Novel data-driven models for forecasting Canadian electricity demand

by

Mohammadreza Makhan

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Supply Chain Management

Asper School of Business

University of Manitoba

Winnipeg

Acknowledgements

I want to express my deep gratitude to my supervisors, Dr. Srimantoorao S. Appadoo and Dr. Yuvraj Gajpal, for their helpful advice and consistent support during my MSc study. Their immense knowledge and experience have been an encouragement for me in my academic research and daily life. My sincere appreciation and gratitude must also go to my MSc committee members: Dr. Aerambamoorthy Thavaneswaran and Dr. Michael Lukie. They generously gave their time to offer me valuable comments toward improving my work. In particular, Dr. Thavaneswaran helped to shed new light on many of my ideas. Without his help this work was not possible to finish. My gratitude extends to the Asper school of business for the funding opportunity to undertake my studies at the Department of supply chain management, University of Manitoba.

Finally, I would like to express my appreciation to my parents and my brother. Without their tremendous encouragement and help in my life, it would be impossible to progress for me in any stage of life.

To Mahsa Amini and all those fighting for freedom in Iran right now.

For “Woman, Life, Freedom”.

abstract

With ever-increasing disruptions in supply chains throughout the world, like pandemics, political conflicts, and trade wars, power generation is becoming one of the major concerns of global and local economies. The recent increasing energy cost in Europe and North America is one of the main consequences and, at the same time, contributes to these disruptions. Therefore, electricity demand forecasting is crucial in power markets to increase cooperation and integration between players in a power grid. This study aims at reviewing the managerial implications of demand forecasting in the electricity supply chain. Also, some recent statistical and machine learning techniques for electricity demand forecasting used in the literature are analyzed and applied to Ontario's historical dataset. A descriptive analysis of electricity demand characteristics in Ontario, post, and pre-pandemic, is conducted. Furthermore, the forecasting performance of methods like dynamic regression, neural network autoregression, and prophet model are discussed and compared. Another contribution of this study is to include fuzzy hourly demand forecasts for a Canadian dataset.

Keywords: electricity demand, demand forecasting, neural networks, dynamic regression

Contents

List of Figures

List of Tables

List of Acronyms

ANN Artificial Neural Networks

DR Dynamic regression

DHR Dynamic Harmonic Regression

HOEP Hourly Ontario Electricity Price

LSTM long short-term memory

ML machine learning

MSE mean squared error

MAE Mean absolute Error

NNETAR Neural Network Autoregression

NNDR Neural networks dynamic regression

RMSE root mean squared error

RMAE root mean absolute error

RNN Recurrent neural network

STLF Short term load forecasting

UC Unit commitment

Chapter 1: Introduction and background

The electricity market has been studied for a long time for its critical influence on the local and global economy. Electricity has some unique characteristics that make it entirely different from other commodities. Electricity is a non-storable commodity, so demand and supply must be matched on a second-to-second basis in every grid node (Kirschen and Strbac, 2018). If the total electricity demand in a transmission system is not met, the whole system can break down, and the recovery process can occur with massive cost and time. It is worth noting that hydro storage systems and batteries for storing large amounts of electricity are developing, yet it is not viable (Kirschen and Strbac, 2018).

Electricity demand is highly inelastic, for example, residential electricity demand will not change significantly with small price changes (Kirschen and Strbac, 2018). Moreover, the high-frequency electricity demand series shows a solid seasonal pattern (hourly, daily, weekly, monthly) (Soliman and Al-Kandari, 2010).

1.1 Electricity supply chain

Traditionally, electricity supply networks in a particular region were operated by one entity. Utility companies were responsible for both generating and distributing electricity to consumers. These utilities owned power generation plants, transmission networks, load centers, and distribution networks for connecting the power plants to all possible customers in a region. Gradually new models, which incorporated more possibilities for competition and thus efficiency was introduced and even realized through the 1990s (Kirschen and Strbac, 2018; Karthikeyan et al., 2013). The electricity supply chain in deregulated markets is generally comprised of four types of nodes (Yi et al., 2017), namely:

1) Generator companies produce electricity with their generation asset at disposal, e.g., hydro turbines, wind turbines, nuclear power plants, solar photovoltaics, etc.
2) Utility companies or retailer companies which, purchase electricity from generator companies to sell to customers.
3) The independent system operator operates the market in which generator companies and utility companies trade in a two-settlement manner and determine the market price.

4) Consumers who purchase electricity from the utility companies for their needs (traditional load of cities and businesses).

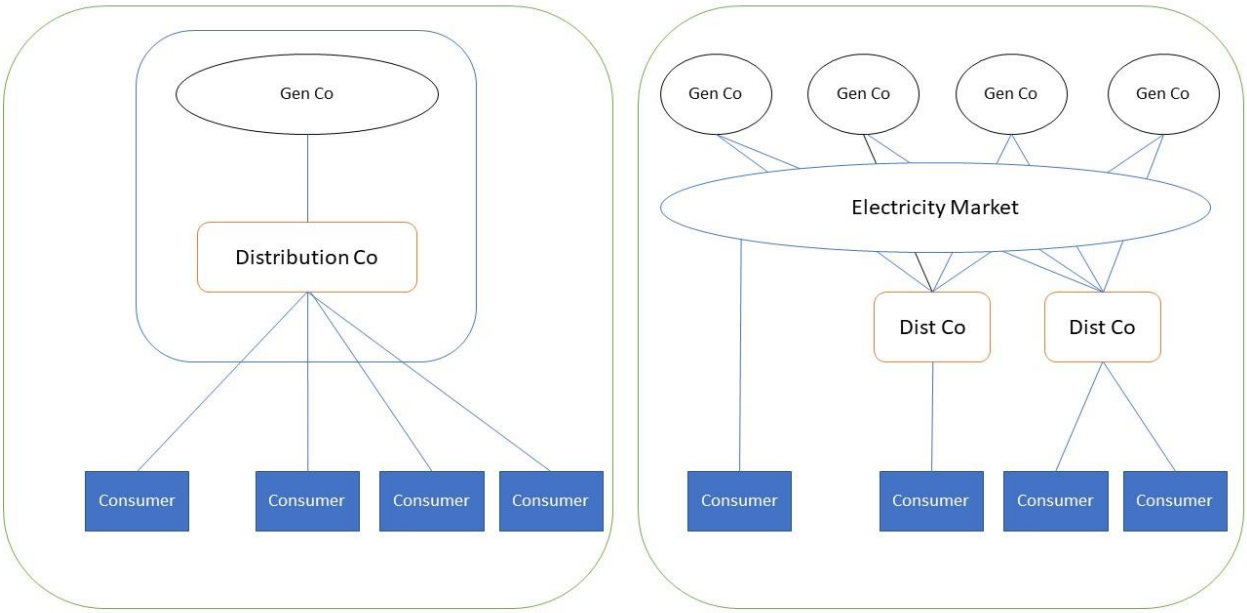


Figure 1.1 Regularized vs deregularized electricity market structures

In both centralized and deregulated market structures, the need for accurate demand forecasts is playing an important role. In centralized systems, the decision-making for operations like power plant scheduling and commitment planning requires very accurate demand forecasts (Lebosta et al., 2018; Banitalebi et al., 2020). In the deregulated structures, each player needs to know the whole system demand so it can decide for its own generation or bidding strategies (Banitalebi et al., 2021).

1.2. Different electricity markets

Three essential markets in the deregulated power supply chains are 1) day-ahead, 2) intra-day, and 3) balancing markets.

Here each of these are going to be explained briefly:

1) Day-ahead markets are the principal mechanism for electricity supply and demand matching. Electricity prices are set in this market, and scheduling for buyers and sellers happens for the next day.

In day-ahead electricity markets, suppliers (power plants) announce their hourly capacity and price for next day to the market operator. Likewise, demand side (distributors) announce their demand and the price they are willing to pay for electricity. These processes are called generation bidding and consumption offering respectively. The mMarket operator receives all information about quantity and price for next day and all players do not know about other players' offers and bids. Then with a clearing algorithm, electricity prices are set for next day. The clearing algorithm working is shown in figure 1.2.

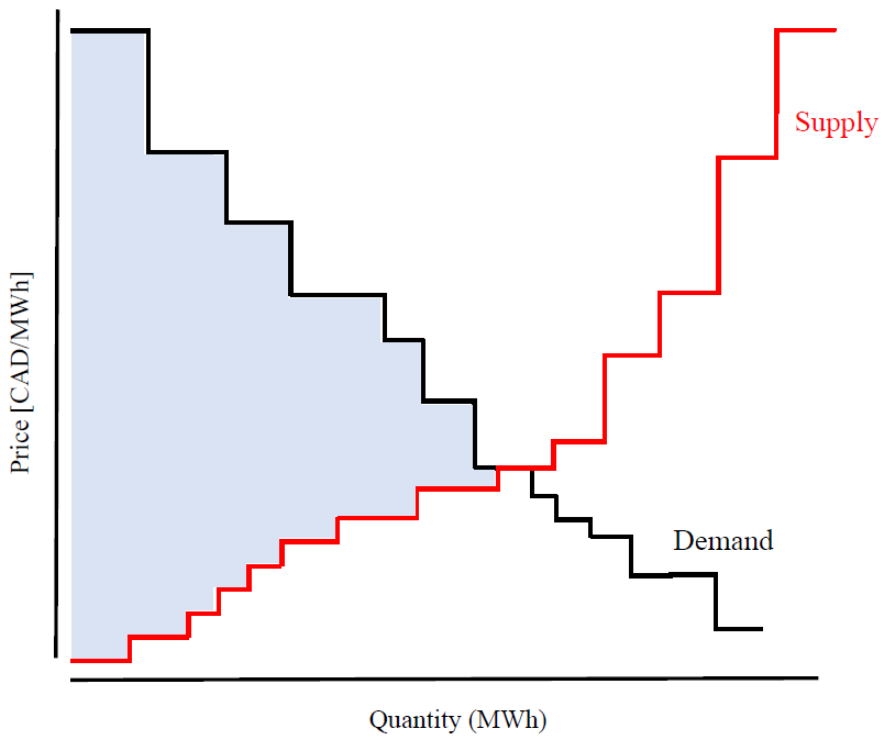


Figure 1.2. Demand and supply clearing in day- ahead electricity markets (Banitalebi ,2021)

In the clearing algorithm, consumption offers (the demand) are sorted in decreasing price order, as illustrated in figure 1.2. Also, the supply side data, i.e., the generation bids are sorted in increasing price order. Then the price for electricity is determined by looking for the crossing point of these two graphs.

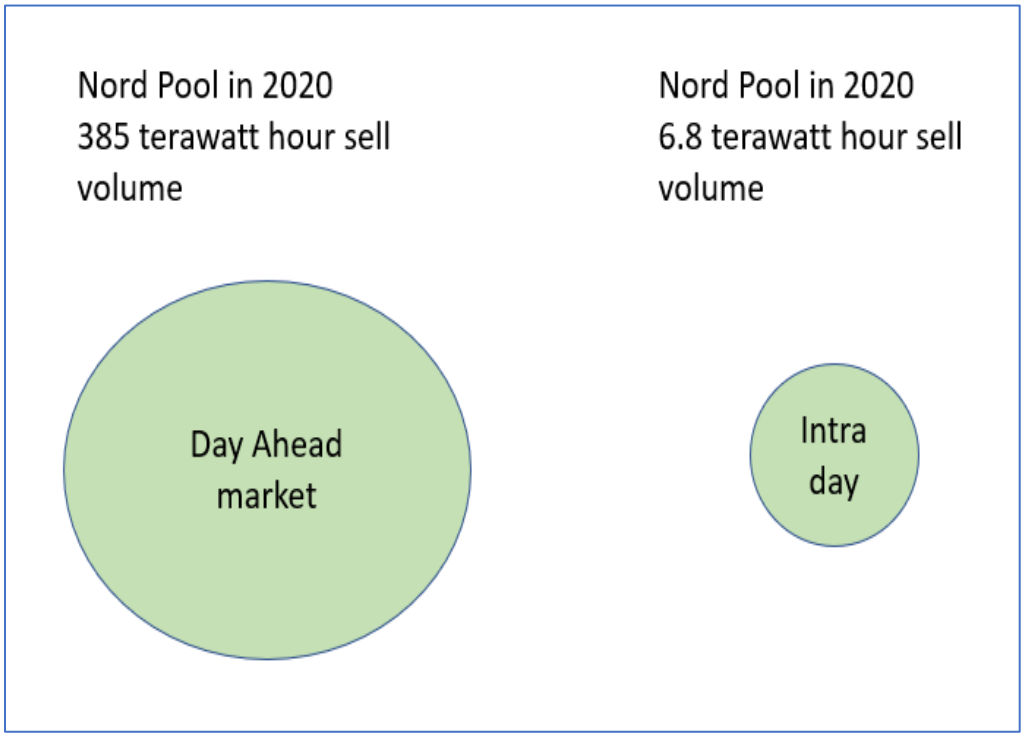


Figure 1.3. Comparison of Day-ahead and intra-day market sell volumes in 2020 NordPool market

2) Intra-day markets enable flexibility for more social welfare (Kattelus, 2021), as it is a continuous trading platform, which acts as a medium between day-ahead and balancing markets. It enables corrections for original schedules, that can happen in instances of wind generation disruptions or emergency plant outages. For instance, if a renewable energy manufacturer gets new information on wind energy predictions, it can adjust its schedule accordingly.
3) Balancing market is a near to real time operation market that assists the system operator to ensure power system balance. Therefore, the balancing market can be considered a tool for system operators to keep the balance between supply and demand.

1.3. Electricity demand forecasting

Demand forecasting horizon varies from hours, days, weeks (short-term load forecasting or STLF) to months and years (medium term load forecasting) and to decades (long term load forecasting) (Mir et al., 2020). Lebosta et al. (2018) implemented a quantile regression model for predicting hourly electricity demand in the peak hours of the day. They used temperature and day of the week as their explanatory variables. The resulting probabilistic forecasts were utilized as an input to a decision-making problem for scheduling power generation units.

A recent trend in load forecasting literature is to provide the probability density function of the load instead of traditional point forecasts. This approach can incorporate the power system's uncertainty caused by introducing renewable energy sources like wind and solar. Yang et al. (2018) used Gaussian process quantile regression to produce one-hour ahead probabilistic forecasts. Temperature is considered the main meteorological feature for STLF in this thesis. Another interesting work is conducted by Banitalebi et al. (2020). The authors propose using a neural network method (multi-layer perceptron) for hourly electricity demand by including more meteorological features like solar irradiance and snowfall, as well as calendar data and historical demand data. Their forecasts are superior to LASSO regression and support vector machine algorithms. One of the most recent works on STLF was a Kalman filter model by Vilmarest and Goude (2021), which performed stronger for predicting electricity demand after Covid-19 disruptive effects. In their r-code competition award-winning paper, Vilmarest and Goude (2021) have used meteorological features like temperature, cloud cover, pressure, and wind speed.

1.4 Motivation

Recently the power grid collapse frequently happens in electric power systems. Once the power grid collapse happens, people's daily life, basic activity and public utility system would be greatly affected by it. For example, the interconnected power grid of the northeastern United States and eastern Canada has gone through a severe large range blackout incident on August 14, 2003. At that time, the electricity supply for about 50 million people was affected. The subway, airport, telecommunication and public transport were all paralyzed by this situation. This power grid collapse gives rise to a 61.8 million kilowatts loss in the cumulative load and billions of dollar loss. Unfortunately, the power grid in the northern regions of India collapsed on July 30, 2012. Shortly after, the power system in the eastern and northeastern region broke down as well on the second day. The basic service, the public transport system, and daily life of more than 600 million people were greatly affected by this power grid collapse event.

Yet there are other new areas, which require more profound electricity demand forecasting techniques. For instance, electric vehicles (EVs) are a novel innovation for controlling climate change. EVs rely on batteries, and batteries can be supplied and charged through different systems like battery swapping stations (BBSs), charging stations (CSs), and central charging stations (CSSs), according to Zhang et al. (2018). Zhang et al. (2018) suggest that a demand response program for pricing electricity at BBSs and CSSs can help EV users purchase electricity at a lower cost during off-peak hours.

Dai et al. (2014) argue that BSSs' load demand is inherently stochastic due to EVs charging random patterns. The random pattern arises from uncertainty in four factors:

1. The number of EVs demanding battery swapping each hour.
2. The beginning of the charging.
3. The distance that EVs aim to travel.
4. The charging time.

Dai et al. (2014) suggest that probabilistic interval forecasts suit BSSs' load forecasting problem. This novel suggestion of replacing point forecasts with probabilistic forecasts is followed in this research in chapters 4 and 5.

1.5 Outline

In this thesis, reviewing the literature in Chapter 2 assesses managerial implications of electricity demand forecasting in different dimensions, including risk management, resilience, and pandemic effects. Then a more thorough examination of recent electricity demand forecasting models follows. Specifically artificial neural network models are discussed in the context of demand forecasting. This examination of artificial neural networks leads to applying a group of these novel models to Ontario's electricity load data in Chapters 3 and 4. The Superiority of the Neural Network Dynamic regression model over other examined methods is shown and discussed. In chapter 4, hourly demand is analyzed and in chapter 5 daily aggregated sum of demand is analyzed and forecasted. Finally, chapter 6 completes the thesis by discussing the limitations and further research directions.

Chapter 2: Literature review

2.1 Risk Management in electricity supply chains

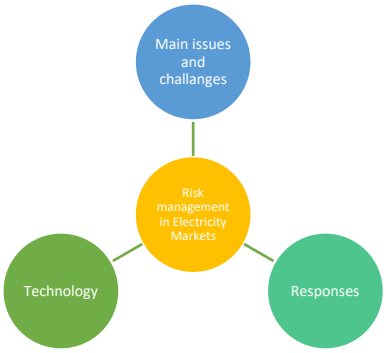


Figure 2.1 Reviewed topics in electricity market risk management

Liu and Wu (2007) have conducted a seminal study in the power market's risk management domain. They have presented a comprehensive risk-management framework for power generation companies that participate in deregulated electricity markets. The authors argue that simulation results imply that generation companies can maximize the profit and minimize risk by taking part in multiple markets. Spot price risk and congestion-charge risk are the two most essential risks a generation company faces. Hedging the risk using futures or swaps, acquiring financial transmission rights, and diversification of markets are among the primary responses proposed by the authors of this paper. They also emphasize the utility of Value at Risk (VaR) as an appropriate tool for risk management purposes in power markets. VaR is a measure for determining how much a provided investment is at risk of loss, given a probability, in a set time, like a day (Thavaneswaran et al., 2019). Also, Banitalebi et al. (2020) have used VaR as a measure to find proper strategies for electricity trading, considering Canadian data.

Applications of supply chain management research in power network optimization is the subject of Wang and Cong's study (2007). They studied how supply chain management essential components like logistics, marketing, and supply differ in electric power supply chain management. The authors conduct their study in the context of climate change as the primary risk challenging China's power network. Extreme weather

conditions bring more uncertainty to electricity demand as a highly perishable commodity, which must be used at the same time as it is generated. This uncertainty calls for more dynamic pricing approaches and the formation of more competitive price mechanisms in the case of China, according to this research. For further information about some market-based policies for power generation climate action, look at US Policy (2020).

Huang et al. (2012) focus on electricity price change risk and its determining factors. They argue bidding strategies of generation companies are one main factor for market price uncertainty in deregulated markets. The bidding strategy of a generation company can be simply put as its price and quantity offering to the market. However, many other factors, like carbon tax policies, encouragement of renewable power production by governments, and implementation of emission trading schemes (ETS), are also crucial to price volatility. ETS is a Cap-and-Trade policy that sets an upper limit on greenhouse gas emissions for a generating company. If a company manages to emit lower than its limit, it can sell its allowances to other entities. Huang et al. (2012) propose a sampling method to assess the price risk. System load uncertainty and wind plant output are considered as main variables determining price risk in their proposed framework for the Australian market.

For putting the supply chain point of view in effect for the power market's risk, Prostean et al.'s paper (2014) is a notable case. The authors consider a project of implementation of a wind power facility. First, a framework for risk analysis in the supply chain is adopted, which was first proposed by Chopra and Sodhi (2004). Then using the Analytic hierarchy process (AHP) methodology and a case study in the Romanian electricity market, the authors suggest that transportation risk and demand risk are the two most crucial factors to be considered by collaborators for designing and purchasing turbines.

Also, renewable energies like wind and solar have an inherent uncertainty, which can cause imbalance risks for producers and the whole grid. Such imbalances can cause substantial financial losses for renewable power generators, as they must pay compensation if they do not meet their committed load. Shamsi and Cuffe (2021) have discussed conventional risk management mechanisms to hedge the imbalance risk for wind power plants. However, they argue these mechanisms are inappropriate for short-term hourly

imbalance costs. Instead, the authors argue that a new market mechanism for electricity called 'prediction market' can address the short-term imbalance cost problem for wind power producers by taking the opposite position. This mechanism works because prediction and day-ahead markets are future markets that rely on wind energy generation results.

2.2 Resilience in electricity supply chains

Supply chain resilience is a rapidly developing research area. In this section, a few outstanding articles published in the last 3 years are examined to make a more suitable understating of the subject.

Hosseini-Motlagh et al. (2020) have shown an overview of a traditional electricity supply network and discussed its pros and cons. One primary problem of traditional electricity supply networks (which are still highly used in many countries) is relying heavily on fossil fuels and their catastrophic consequences on the environment. The authors' proposed alternative is to consider distributed generators (DG) a more resilient and sustainable power network. According to Hosseini-Motlagh et al. (2020), DGs are small power generation plants that create power from renewable sources like photovoltaics or wind. These DGs can be more dispersed in an urban area, decrease the costs of maintaining a massive power plants, and decrease power waste in electricity lines between generators and final consumers. Hosseini-Motlagh et al. (2020) proposes a fuzzy-robust programming model with 3 objectives of decreasing cost, decreasing de-resiliency, and maximizing corporate social responsibility. Resiliency is subdivided into minimizing successive establishment, DG's inadequacy, congestion through electrical lines, and energy dissatisfaction.

Another highly rigorous work in the field of resilient electricity supply chain network design (ESCND) is by Jabbarzadeh et al. (2019). In this article, the authors emphasize exploiting smart-grid concepts for designing a network with more efficiency, resilience, and environment-friendly aspects. Smart grids can have power flow on both sides of an electrical line, and forecasting capabilities come in handy for a more adaptive and sustainable grid. Distributed generation ability is again mentioned as a resiliency measure. The authors define resilience as responding to disruptions like cyberattacks or

natural disasters, which may affect network performance. Also, microgrids can be used to achieve resilient power networks. A local island grid that can work both standalone or in connection with a connected system of grids is called a microgrid (Momoh, 2012). A schematic picture of a microgrid is displayed below.

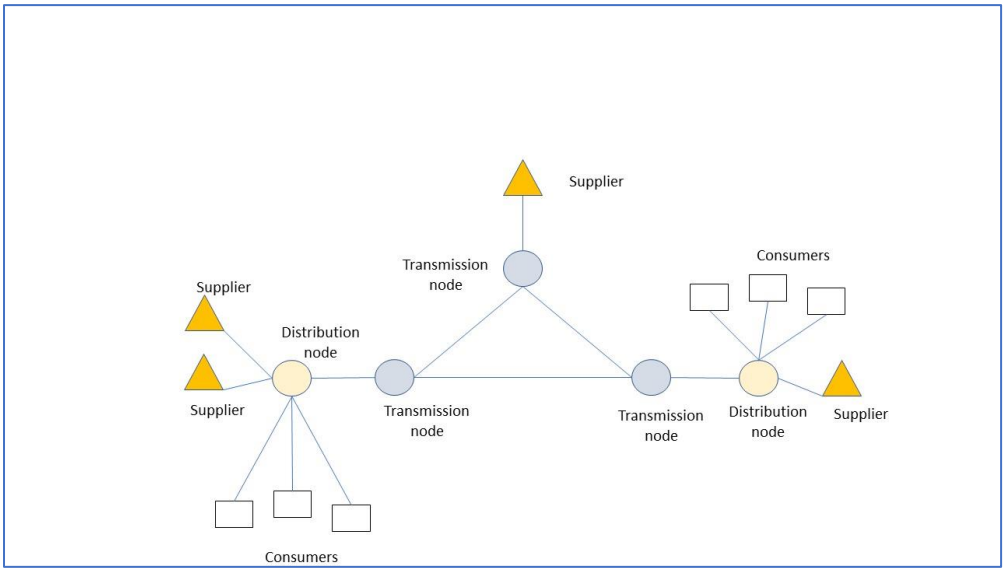


Figure 2.2 A simplified schematic display of a hypothetical microgrid

Zavala-Alcivar et al. (2020) recognize the significance of forecasting capabilities in supply chain resilience. The authors of this article have performed a systematic literature study to enumerate the elements of resilience as a company's strategic capability. Fourteen factors have been identified, and the forecasting powers of a company directly influence 'disruptive environment awareness' as a considerable element. Also, Zavala-Alcivar et al. mention procurement risk and price fluctuations, problems in distribution and logistics, and unforeseen or unstable demand as major risk elements in supply chains. Robust forecasting techniques can help mitigate these risks and is this thesis's main focus. Notably, according to this literature review, only 7 out of 45 articles reviewed were concerned with the electricity sector. Other sectors studied are transport and logistics, agri-food, mining and oil, the petrochemical industry, and health services. Therefore, this literature review has identified that the resilience of electricity supply chains regarding forecasting capabilities is a gap in the literature that warrants further investigation.

Finally, in the case of resilient supply chain designs, Yavari and Zaker (2019) and Yavari and Zaker (2020) are important. Both papers emphasize how integrating two distinct supply networks can yield more sustainability and readiness to face disruptions. In the case of these articles, the dairy industry and electricity generation supply chains are considered to be collaborating. The authors have studied Iran's biggest dairy products producer and how they have managed to diminish power outage risk by sharing information and objectives.

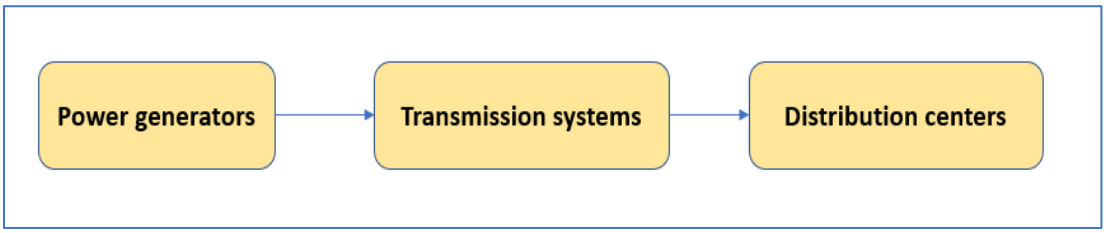


Figure 2.3 A simplified electricity supply chain structure

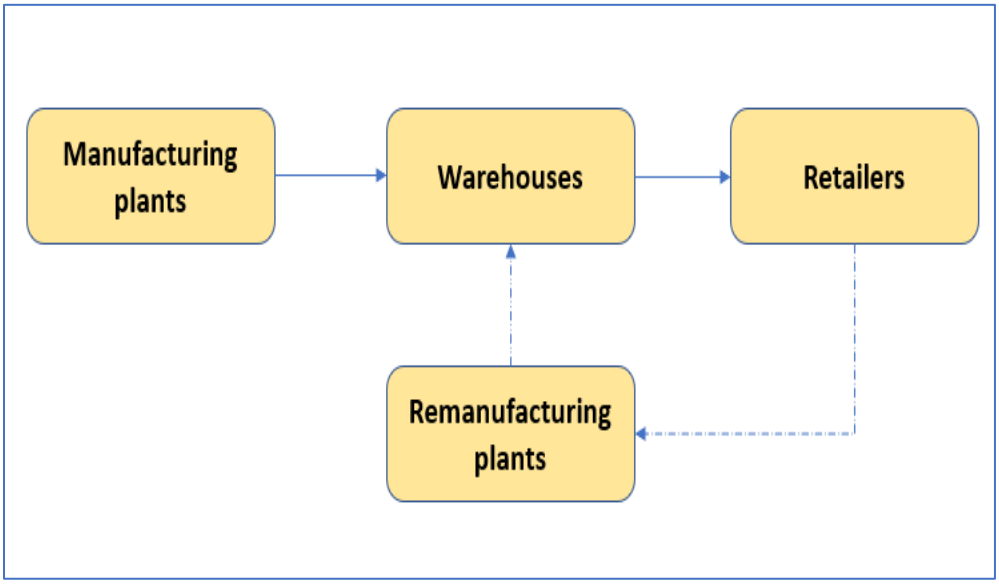


Figure 2.4 A simplified typical supply chain containing reverse logistics

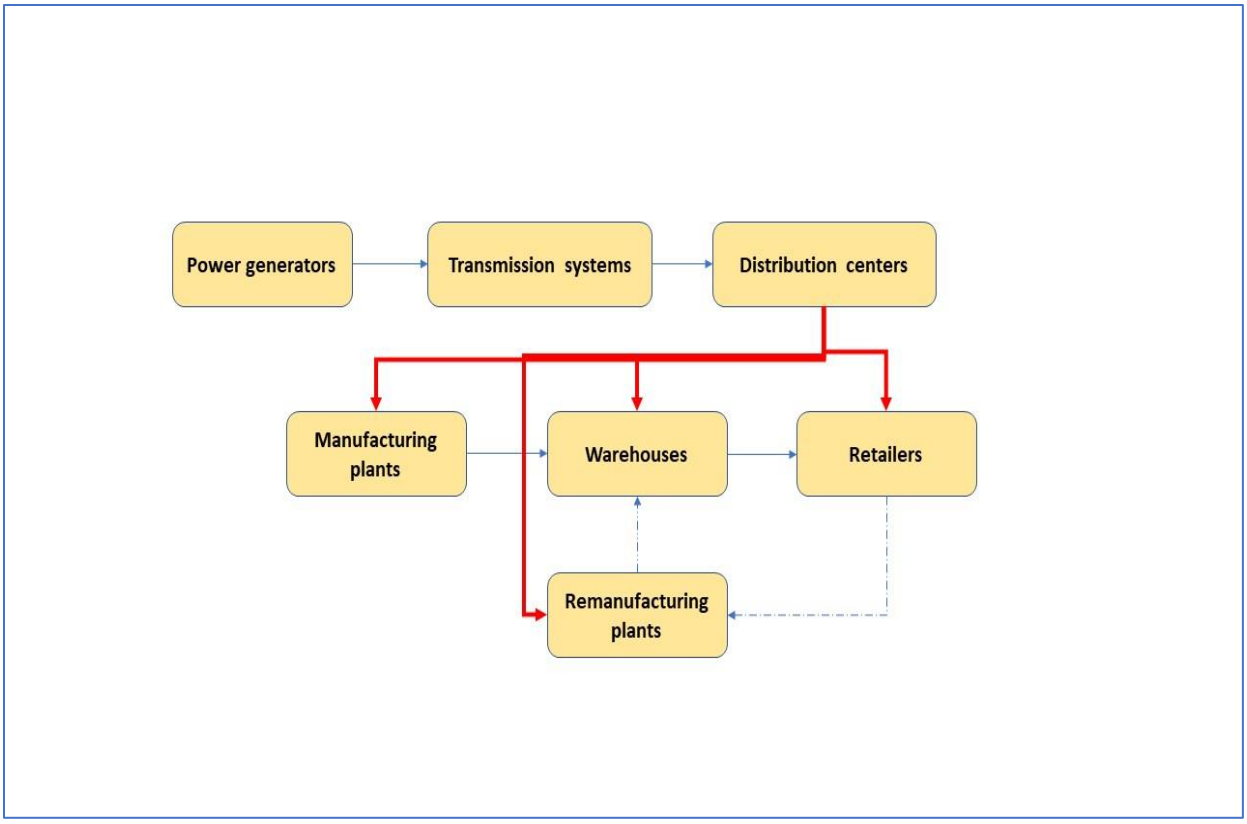


Figure 2.5. combination of power supply chain and reverse logistics in Yavari and Zaker (2020)

2.3 Pandemic effects on electricity supply chains

The Covid-19 pandemic had an immense effect on how supply chains work around the world. Electricity markets were also affected heavily. As a result, the body of research studying the influence of Covid-19 pandemic on electricity markets is growing. Since our dataset includes both pre-Covid and post-Covid era, we take a brief look at a few articles investigating this topic to better understand the research trends here.

Agdas & Barooah (2020) examined U.S. electricity demand in three separate states of California, New York, and Florida, to see if the pandemic altered demand patterns in the power market. The authors cast doubt on the theory that the pandemic reduced electricity demand. The argument is that almost three-fourths of electricity demand in the U.S. belongs to buildings, and the demand for reductions in office and commercial places is

offset by increased demand in the residential sector, as people tended to work from home. Another valuable takeaway from Agdas & Barooah's article is their weather correction procedure for better electricity demand predictions. The fundamental idea of weather correction in this article is that electricity demand has a baseload independent of temperature and two dependent terms for heating and cooling:

$$demand = baseload + \alpha^{Cooling} T^{Cooling} + \alpha^{Heating} T^{Heating}$$

The methodology for predicting electricity demand in this thesis will be to incorporate temperature-related demand in a comparable but more refined manner using data-driven dynamic and dynamic harmonic regression.

However, the case is different in Canada, as Leach et al. (2020) declare that by employing high-frequency electricity data for the two provinces of Ontario and Alberta, at least a 5 percent drop in demand after Covid-19 is perceptible. The authors also call for more data transparency in all regions, as currently only available in an aggregated form, and hourly electricity demand is a highly critical data correlated with economic growth.

Most research in this area suggests that the most significant effect of Covid-19 on electricity demand was not a drop in demand but an increase in demand volatility. For example, Santiago et al. (2020) discuss that lockdowns changed the morning and evening peak patterns leading to more uncertainty for electricity demand in more hours of the day.

2.4 Forecasting methods used in electricity supply chains

2.4.1 Machine learning models

In recent years, with modern machine learning (ML), neural networks, and deep learning algorithms, supply chain demand forecasting has taken an enormous leap. In one of the most seminal works in this area, Carbonneau et al. (2006) advocated using ML techniques in a supply chain demand problem when there is insufficient information about other supply chain participants' demands. They comparing a neural network (NN) models' performance in demand forecasting problems with naïve forecasts, moving averages, and support vector machines (SVM). What they mean by NN is feedforward error back-propagation type networks. In a neural network model, which is a rough simulation of the nervous system, several nodes (or neurons) are organized in multiple layers. Each layer

receives data from the previous layer and passes it to the following layer. By a supervised learning algorithm, desired outputs are given to the network. Therefore, after many iterations, the network approximates the response as close to the ideal prediction. In more advanced NNs like recurrent neural networks (RNNs), nodes in a layer can feed data to other nodes in the same layer. RNNs are highly useful in time series forecasting models.

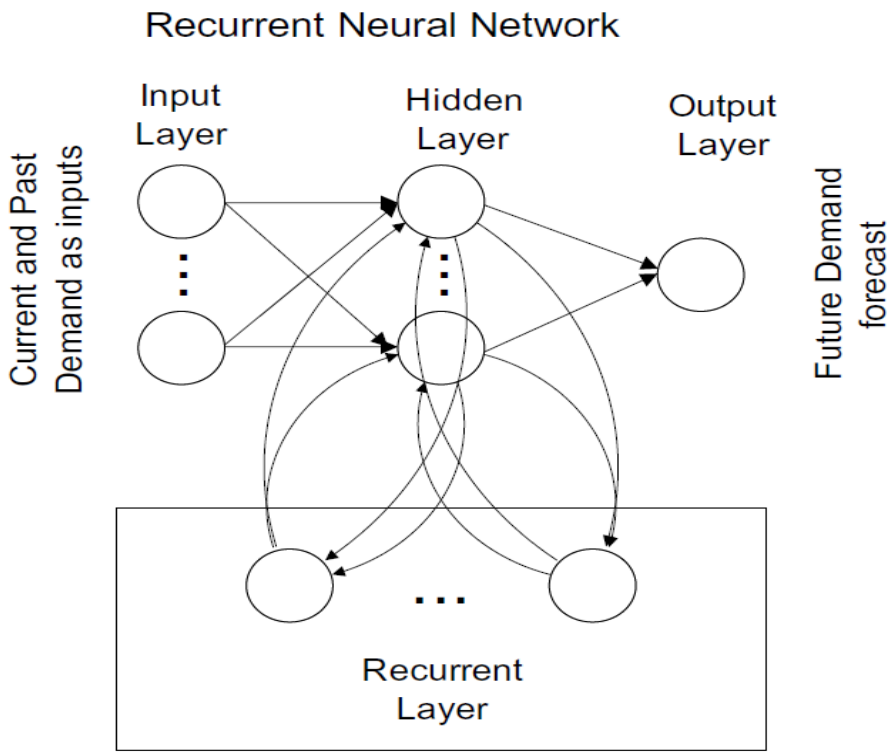


Figure 2.6 Recurrent neural networks schematic

Banitalebi (2021) explains that the performance of an RNN model can get worse as the dimension of inputs gets bigger. A suggested response to this problem is using smoothing functions to reduce the noise in the dataset. For example, Banitalebi (2021) discussed double exponential smoothing (DES) and triple exponential smoothing (TES) algorithms for volatile forecasts in electricity prices in Ontario. DES is suitable when there is a trend in the data, but TES can be used when trends and seasonality exist. Furthermore, optimal parameters can be obtained using the sum of square errors (SSE) metric in both cases of TES and DES.

2.4.2 Neural networks (NN) models in load forecasting

An application of artificial neural networks (ANN) in electricity demand forecasting has an extended history. More sophisticated forecasting techniques like ANN can help supply chain practitioners respond to rapid electricity demand growth over the past years. Yong et al. (2017) applied an improved ANN model for Queensland's load data and reached significant results. Yong et al.'s contribution (2017) is to notice the similarity between each day of the week's demand. This statistical observation was the basis for dividing the demand dataset into seven classes for each day of the week and proposing different ANN models for each class. Queensland load dataset is described as high-frequency, half-hourly data.

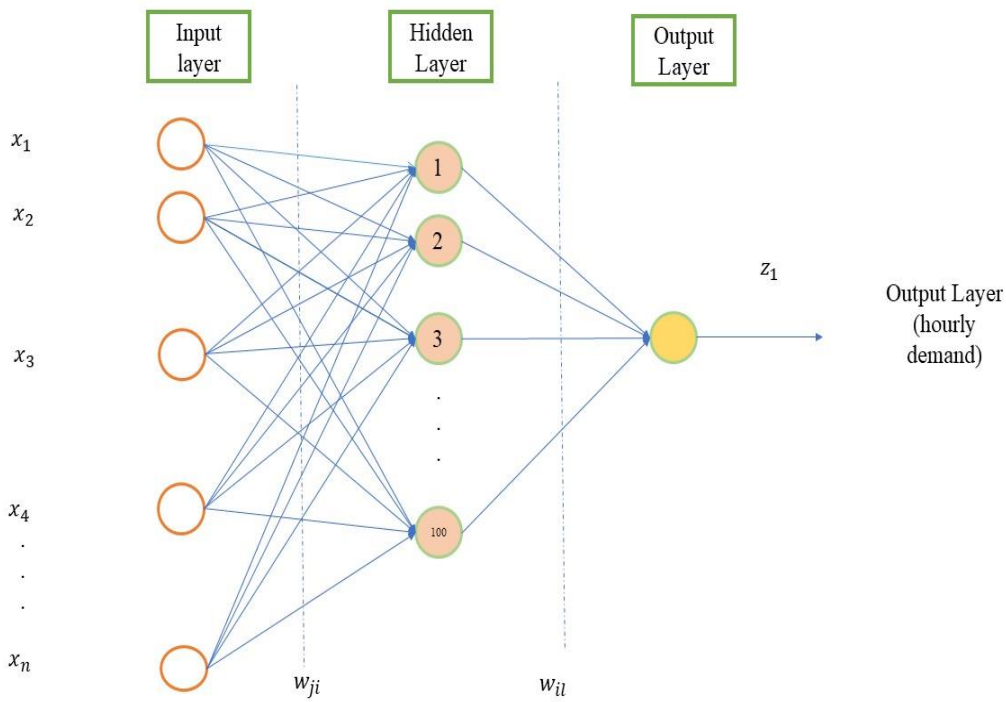


Figure 2.7 ANN model used in Yong et al. (2017)

Yong et al.'s innovative ANN algorithm (2017) modifies how weights in a neural network adjust. They have set the weights (w_{ji} , w_{il}) to be only between -1 and 1 and then in each iteration weights randomly change. If the changed set of weights results in better forecasts, the set is preserved and the next iteration is built upon it. If not, the set of weights

remains unchanged and the process repeats. This method of determining weights is different from methods like back-propagation (BP).

In the following, I explain the Young et al. (2017) ANN algorithm where: x_1 to x_n represent inputs of the ANN model which are hourly electricity demands of lag 1 to lag n. w_{ji} represents weights connecting input node j to hidden layer's node i, and w_{il} is the weight connecting hidden layer's node i to output node. For incorporating non-linearity, sigmoid activation function is used and hidden layer has M=100 nodes (neurons):

$$r_i = \sum_{j=1}^{n} w_{ji}\, x_j + bias_i \;\; , i = 1, \ldots., M$$

$$y_i = \frac{1}{1 + e^{(\beta_i r_i)}}$$

β_i and $bias_i$ are parameters associated with each neuron in the hidden layer. The output of the model can be calculated as:

$$z_l = \sum_{j=1}^{M} w_{il}\, y_i \;, l = 1, \ldots., L$$

Yong et al.'s ANN algorithm (2017)

1: Initialize the parameters randomly

2: randomly create vectors ε_1, ε_2, ε_3

3: modify the parameters $w_{ji} \rightarrow w_{ji} + \varepsilon_1$; $\beta_i \rightarrow \beta_i + \varepsilon_2$; $bias_i \rightarrow bias_i + \varepsilon_3$

4: update the hidden layer r_i according to step 3 modifications

5: update the output z_l according to new r_i and y_i

6: calculate the cost function as C = actual demand$-z_l$

7: if C< satisfaction level, end, otherwise go to step 2

This algorithm tries to capture the growing computational power of computers and reaches better forecasts considering a very large set of weights compared to the BP technique for determining weights, which was not possible when back propagation algorithm was first proposed by Hagan et al. (1996). However, their method only considers point forecast. Interval forecasts are becoming more and more needed to have better and more reliable models. In addition, in this research only day of the week effect on electricity demand is considered while meteorological factors and calendar information have huge impact on load too.

2.4.3 Data driven forecasting models

In many forecasting models, normality is assumed as a condition for model residuals. However, recently a growing literature is suggesting that data-driven t distributions may be more appropriate in many real-world cases. For example, in a foundational article we Thavaneswaran et al. (2020), which discusses a novel data-driven exponential weighted moving average (DDEWMA) forecasting technique for risk in financial markets. The authors propose that estimating the standard deviation (σ) of financial assets return (also called volatility) by estimating the variance (σ^2) and then calculating the square root of it is inefficient. Instead, they propose using sign-correlation (ρ) of return, which can be shown as:

$$\rho = Corr\big(R - m\,, sgn(R - m)\big) =$$

$$\frac{E|R - m|}{2\sigma\sqrt{F(m)(1 - F(m))}}$$

Where R is a random variable with CDF $F(x)$, mean m and variance σ^2. sgn is sign function. The proof for the above formula can be found in Zhu's thesis (2020).

It is obvious that for asymmetric functions,

$$F(m) = \big(1 - F(m)\big) = 0.5$$

So, for asymmetric functions with finite variance (e.g., normal and t)

$$\rho = \frac{E|R - m|}{\sigma}$$

This idea of data-driven models provides an opportunity for estimating the standard deviation directly and more accurately. Thus, in financial markets, volatility can be estimated using DDEWMA as below:

$$\hat{\alpha}_{t+1} = \alpha\widehat{\sigma_t} + (1-\alpha)\frac{|r_t|}{\rho}$$

Where α is the smoothing parameter, which itself can be found by minimizing SSE. This approach to finding the smoothing parameter as well as estimating volatility directly, is called a data-driven approach. Applications of this method is not restricted to financial markets and Hoque et al. (2021) and Banitalebi et al. (2021) have utilized this approach in resilient supply chains, algorithmic trading as well as electricity markets risk forecasting.

2.4.4 Demand forecasting using fuzzy theory

Another new path, which researchers have recently taken is to incorporate fuzzy forecasts. Fuzzy forecasts are appropriate tools when making predictions in an environment of ambiguity and uncertainty. With the recent propagation of disruptions and uncertainties, using fuzzy models in supply chains seems to be more appropriate than ever.

Thavaneswaran et al. (2018) used the concept of fuzzy neuro-volatility models in financial markets risk forecasting. Later in 2022, the idea was further developed in electricity demand forecasting using trapezoidal adaptive fuzzy models (Liang and Thavaneswaran, 2022). In their work, they consider a fuzzy set with a membership function as:

$$A(x) = \begin{cases} \left(\frac{x-a}{b-a}\right)^n & if\ b \le x < b \\ 1 & if\ b \le x < c \\ \left(\frac{d-x}{d-c}\right)^n & if\ c \le x < d \\ 0 & otherwise \end{cases}$$

When $A(a,b,c,d)_n$ and x$\in X$. If $n = 1$ we have commonly known trapezoidal fuzzy set.

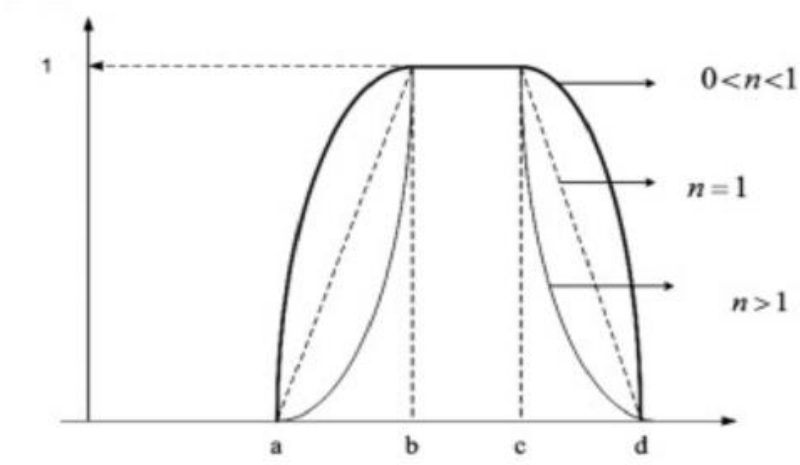


Figure 2.8 Trapezoidal adaptive fuzzy function

In the above diagram, α-cuts (α ∈ [0,1]) are used as horizontal representation of fuzzy sets. In other words, α-cuts are the elements of set X that satisfy the condition of being in vaguely defined set A, with an extent equal or greater than α.

Liang and Thavaneswaran (2022) discuss that in real world cases like demand forecasting, statistical quantile of data can be used instead of values (a, b, c, d). Also, the idea of fuzzy prediction intervals is introduced by the authors.

Chapter 3: Data description

The Independent Electricity System Operator (IESO) website (ieso.ca) offers actual hourly electricity load data. However, high-frequency hourly data from 2017 to 2021 is used for forecasting purposes to obtain the demand forecast. Also, hourly temperature data and other meteorological features in the province of Ontario are accessible through Canada weather stats (weatherstats.ca). The objective is to forecast Ontario day-ahead hourly electricity demand using features like temperature and calendar variables.

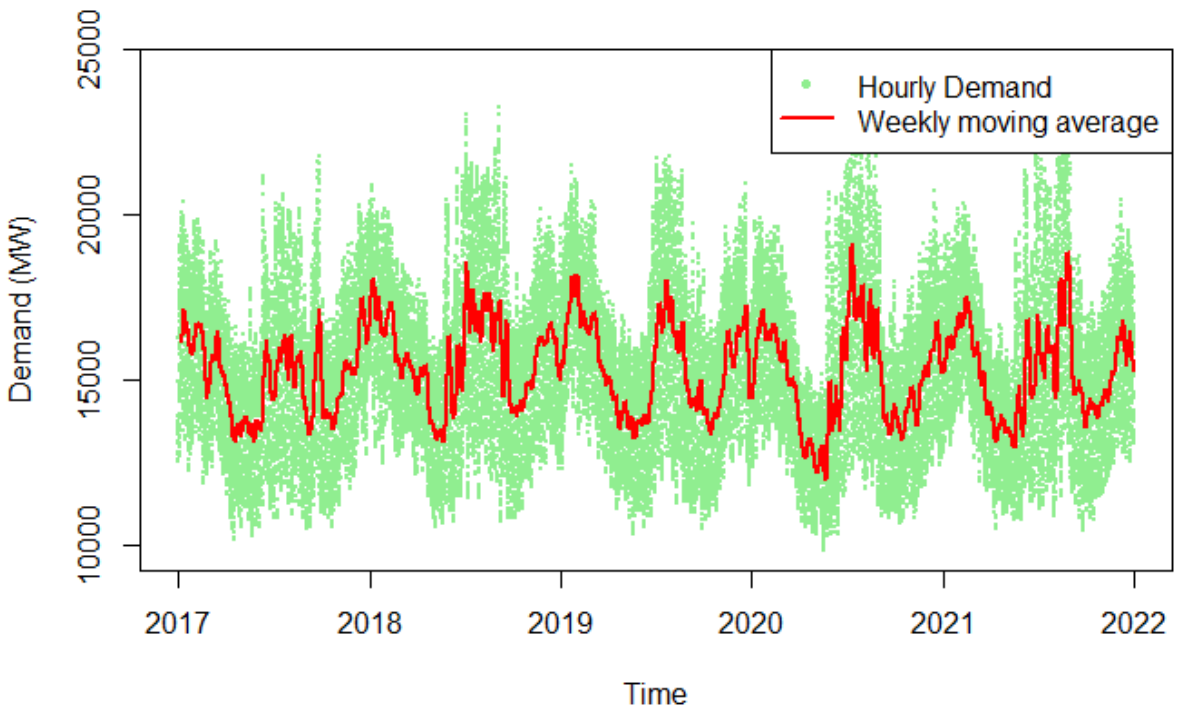


Figure 3.1 Electricity demand time series illustration and simple moving average for last five years in Ontario

It follows from Fig. 3.1 that electricity demand peaks in summer and in winter. A simple moving average (SMA) with a window size of 168 shows the demand behavior in Fig. 3.1. De Vilmarest and Goude (2021) have noticed the abrupt change in electricity demand during the Covid-19 pandemic emergence in their large dataset with almost 35,000 hourly observations and 44 variables. They have associated this change with lockdowns (De Vilmarest & Goude, 2021, p.4). It is worth mentioning that Nouruzi and Fani (2021) have performed an in-depth study to estimate the Covid-19 pandemic's influence on electricity

demand in different countries. Although in Fig.3.1, a cyclical load decrease in spring can be noted, the 2020 spring shows more significant reductions in load.

3.1. Increased uncertainty in demand after Covid-19 pandemic declaration

Fig. 3.2 depicts the increased demand uncertainty by estimating the volatility using the formula below (Thavaneswaran et al., 2020, p.2):

$$\sigma_t = \frac{|d_t - \overline{d_t}|}{\rho}$$

where sign correlation ρ is given as:

$$\rho = Corr(d_t - \overline{d_t}, sgn(d_t - \overline{d_t})$$

and d_t is hourly demand at time t. Thavaneswaran et al. (2020) have used this direct estimation of volatility in financial risk forecasting instead of estimating ${\sigma_t}^2$ and then taking the square root.

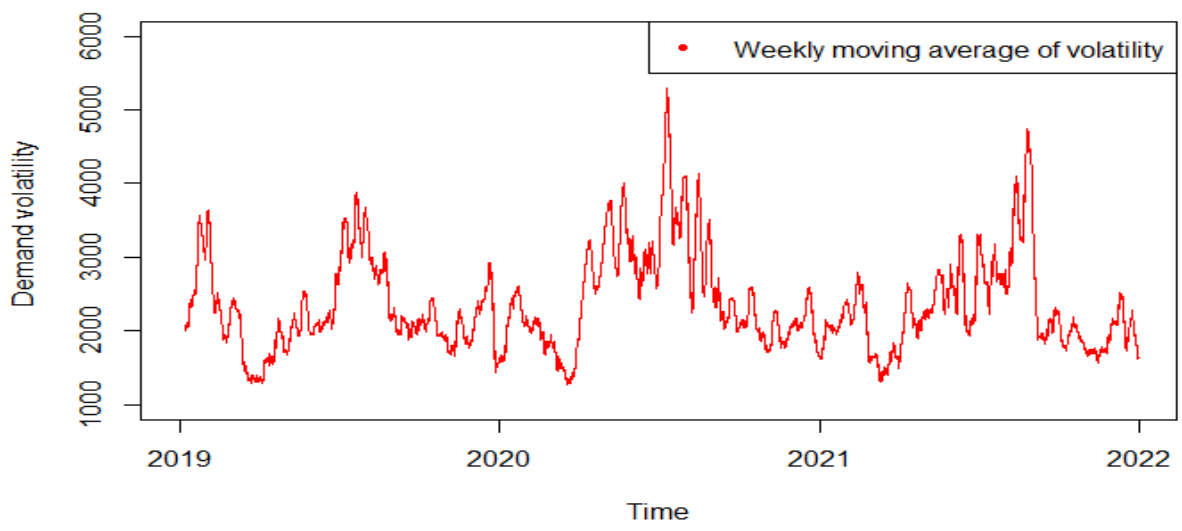


Figure 3.2 Volatility of Electricity Demand in Ontario

Fig. 3.2 clearly shows boosted demand uncertainty (volatility) during the spring and summer of 2020, when most lockdowns and stay-at-home orders were in effect (Ontario newsroom, April 13, 2020). Yet, raised volatility does not necessarily signify a boost in demand. For example, Fig. 3 shows that in March 2020, for each day of the week, total

load demand was smaller than the same day of the week in the March of the previous year (March, 2019). Also, Fig. 3 illustrates two heights of hourly demand during the morning hours (7a.m. to 9 a.m.) as well as evening hours (7a.m. to 9 a.m.) in 2020 (dashed lines) and 2019 (connected lines).

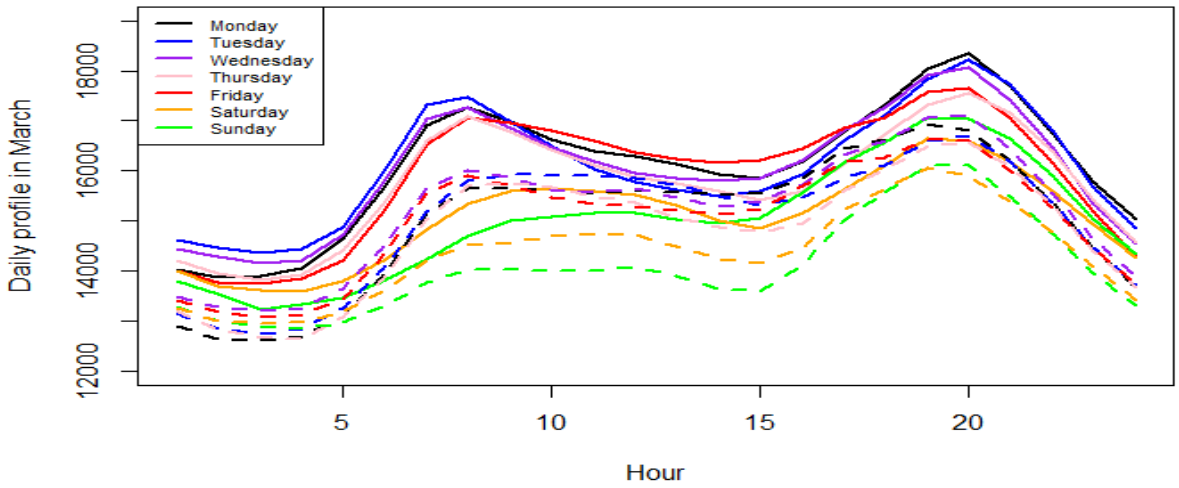


Figure 3.3 Comparison of Electricity Demand before pandemic (2019) and after pandemic (2020)

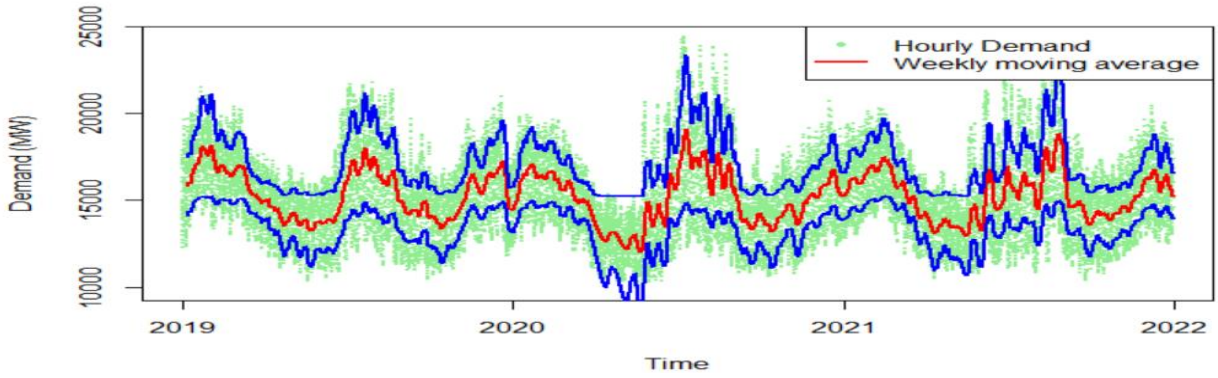


Figure 3.4 Electricity demand trend and its boundaries

In figure 3.4, blue lines depict one standard deviation above and one standard deviation below the moving average demand at each hour. In this figure, it is much better illustrated how demand in spring early summer of 2020 dropped, nevertheless the uncertainty (the gap between the red line and blue lines) has widened.

Table 3.1 presents some basic descriptive information for Ontario hourly demand data.

Table 3.1 Descriptive Statistics of 2020 Ontario Hourly Electricity Demand (MW)

Hourly demand (8760 observations)	Mean	Standard deviation	Min	Q1	Median	Q3	Max
	15052	2493	9831	13180	14719	16727	24446

3.2. Dependency of demand on meteorological variables and other factors

The reliance of demand on meteorological variables, specifically the temperature, is examined in the literature to a great extent. Hyndman and Athanasopoulos (2021) included cooling and heating explanatory features in their dynamic regression model. Cooling and heating variables are functions of the temperature. Also, renewable energy production resources like wind and solar depend highly on temperature and meteorological factors. Renewable energy producers are generally small production units that will consume more electricity from the public grid when their renewable resources are unavailable due to cold weather or lack of wind (De Vilmarest & Goude, 2021).

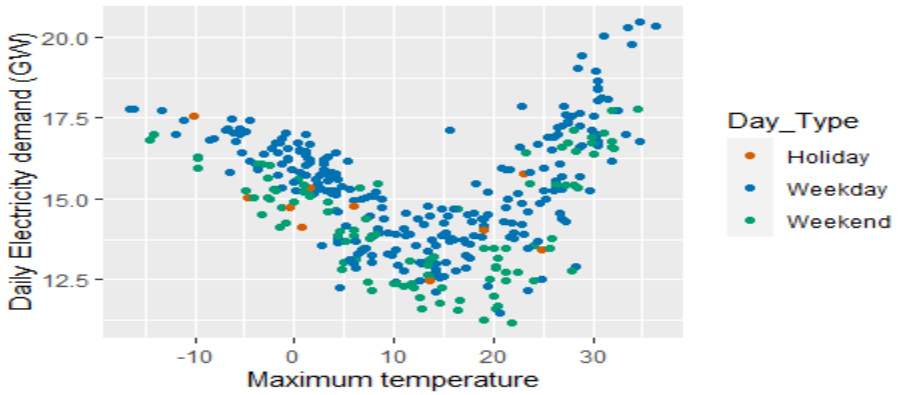


Figure 3.5 Demand highly depends of temperature and day type

Fig. 3.5 shows the effect of maximum daily temperature and day type on electricity demand in 2020. It can be seen that very cold or hot weather increase the electricity demand for heating and cooling respectively. Also, there is more electricity demand during the weekdays than weekends and holidays.

Also, Figure 3.6 explains how each hour temperature and load associate in year 2020. Obviously, working days (blue dots) have a generally higher load than holidays (orange dots).

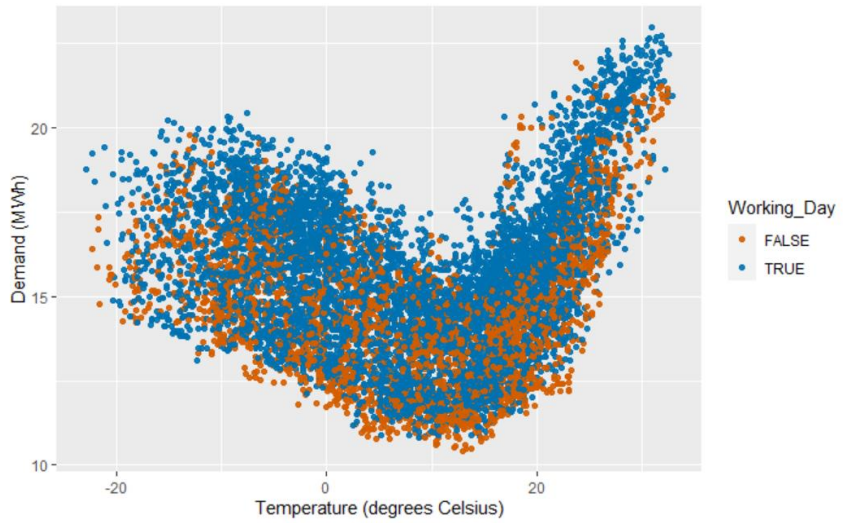


Figure 3.6 Hourly demand and Hourly temperature correlation

If we consider one specific hour of the day such as 3 p.m. and show its demand against its temperature, the same pattern as previous graphs results.

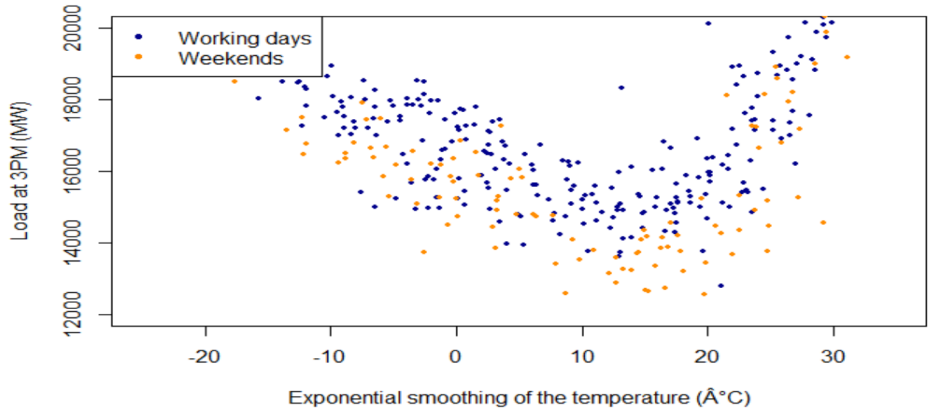


Figure 3.7 Smoothed demand and its dependence on temperature

The side-by-side box plot in Fig. 3.8 shows that hourly electricity demand is higher and more disperse in hot and cold months of 2020.

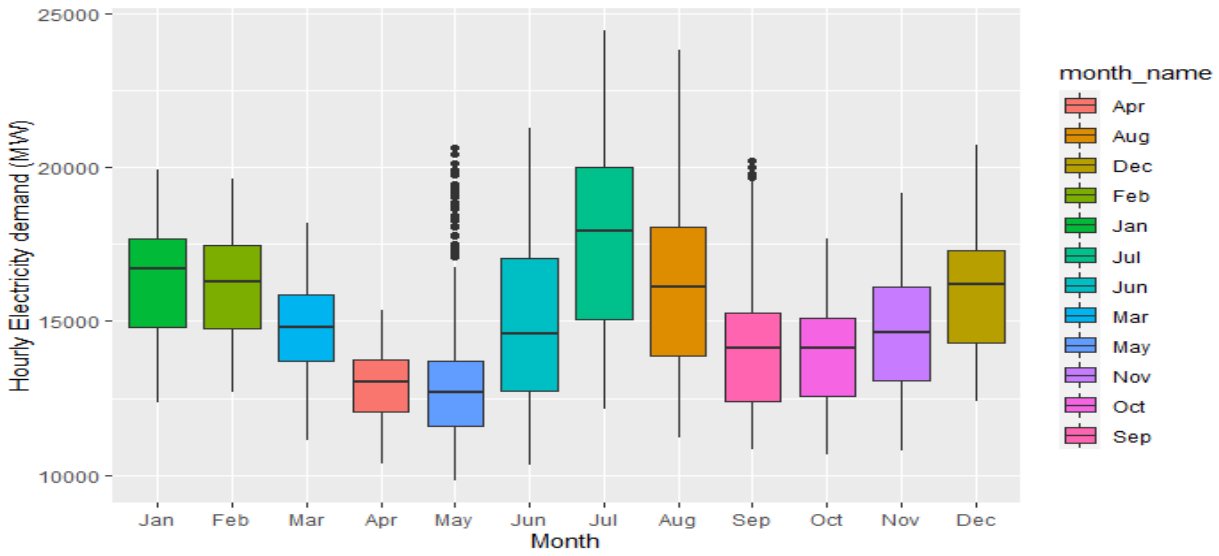


Figure 3.8 Monthly variation in mean and variance of the hourly Demand during 2020

Also, the dependence of hourly load at a particular hour such as 3 pm is examined and illustrated against other variables last day's load, and previous week's load.

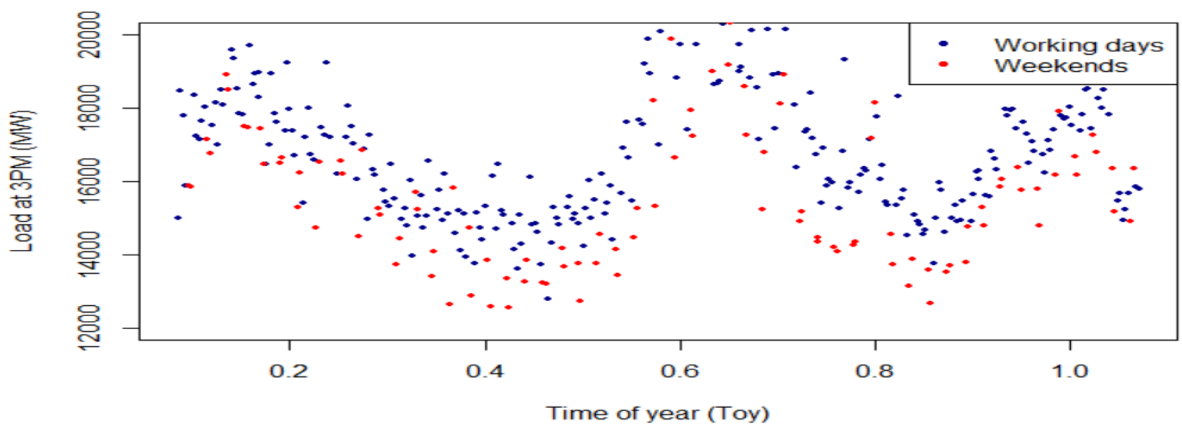


Figure 3. 9 Electricity demand changes throughout a year

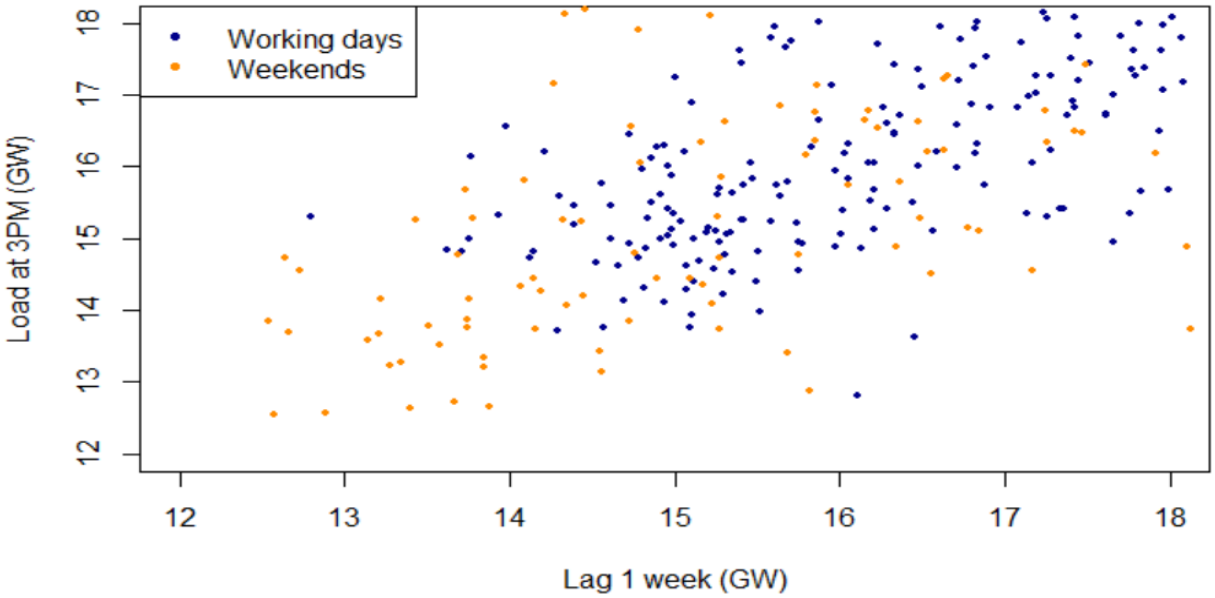


Figure 3. 10 Electricity demand is highly autocorrelated (weekly correlation)

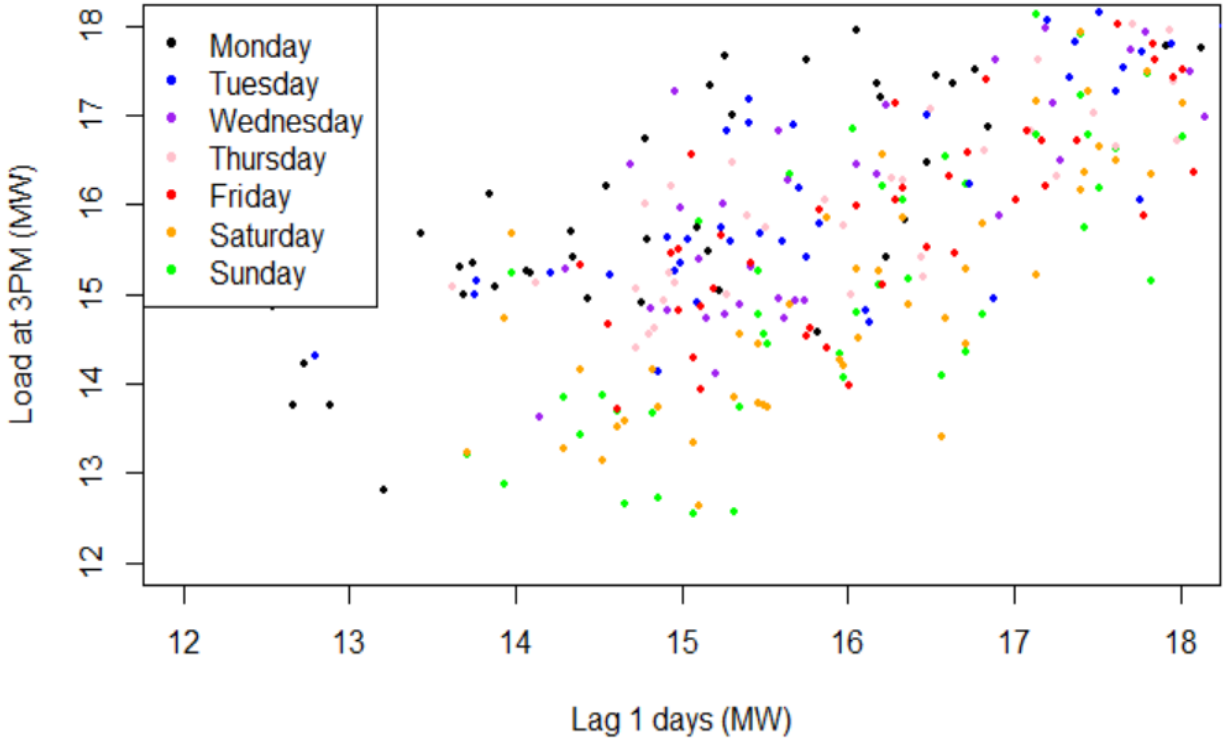


Figure 3. 11 Electricity demand is highly autocorrelated (daily correlation)

Chapter 4: Hourly electricity demand forecasting

4.1 Dynamic regression model

Univariate time series models like ARIMA allow the information of one individual series to be analyzed for forecasting future values. ARIMA models strongly capture existing autocorrelations, However, there may be cases in which the model's inclusion of other related information would result in better forecasting results. Dynamic regression models combine the ability of regression models to include different features with ARIMA models time series analysis strength (Hyndman & Athanasopoulos, 2021).

Consider demand series d_t. Assume there are k different series like $x_{1,t}$, $x_{2,t}$, …, $x_{k,t}$, which can act as predictors for the demand series. It can be written:

$$d_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{k,t} + \eta_t$$

The innovation in this model is the idea that the error terms can be correlated. The error series η_t itself can be modeled as a seasonal ARIMA or any other time series model like generalized autoregressive conditional heteroskedasticity GARCH(1,1). The $\mathrm{ARIMA(p,d,q)(P,D,Q)}_s$ representation of the error terms can be written as

$$\phi_p(B)\Phi_P(L)(1-B^d)(1-L^D)\eta_t = \theta_q(B)\Theta_Q(L)\,\varepsilon_t$$

$$L = B^s$$

Where $\phi_p(B)$ and $\theta_q(B)$ are polynomials of degrees p and q in B (lag operator). $\Phi_P(L)$ and $\Theta_Q(L)$ are polynomials of degrees P and Q in L.

L represents seasonality of the series and ε_t is a white noise series. For example, in this case, observations separated by 24 units are highly correlated. So, for the hourly data, s is considered equal to 24.

The demand forecasts are obtained by fitting dynamic regression models with temperature series and day-type series as explanatory variables. Day-type is a categorical variable that indicates if the day is a working day, holiday, or weekend. Error term η_t is modeled as an $\mathrm{ARIMA(p,d,q)(P,D,Q)}_s$. R package fable is used to provides best fitted model by using

the Akaike information criterion (AIC) criteria. Models are implemented using RStudio v1.4.1106 and fable package v0.3.1 (Mitchell et al., 2021).

For training the model, 4000 data points have been selected from year 2021. The training set for this model contains hourly demand from January 1, 2021 to June 16 of the same year. The period after Covid-19 has been selected, since it is assumed that more volatility in demand can be expected in Covid-19 era.

Table 4.1 Training and testing sets in dynamic regression model

Hourly Electricity demand (GW)	DR model
Training Set	2021-01-01 00:00:00. to 2021-06-16 16:00:00
Testing Set	2021-06-16 17:00:00. to 2021-06-18 16:00:00

As suggested by Hyndman and Athanasopoulos (2021), temperature, temperature square, and day type are considered as explanatory variables of the regression model. Also, two other variables of heating and cooling are put into the model for the sake of better forecasts.

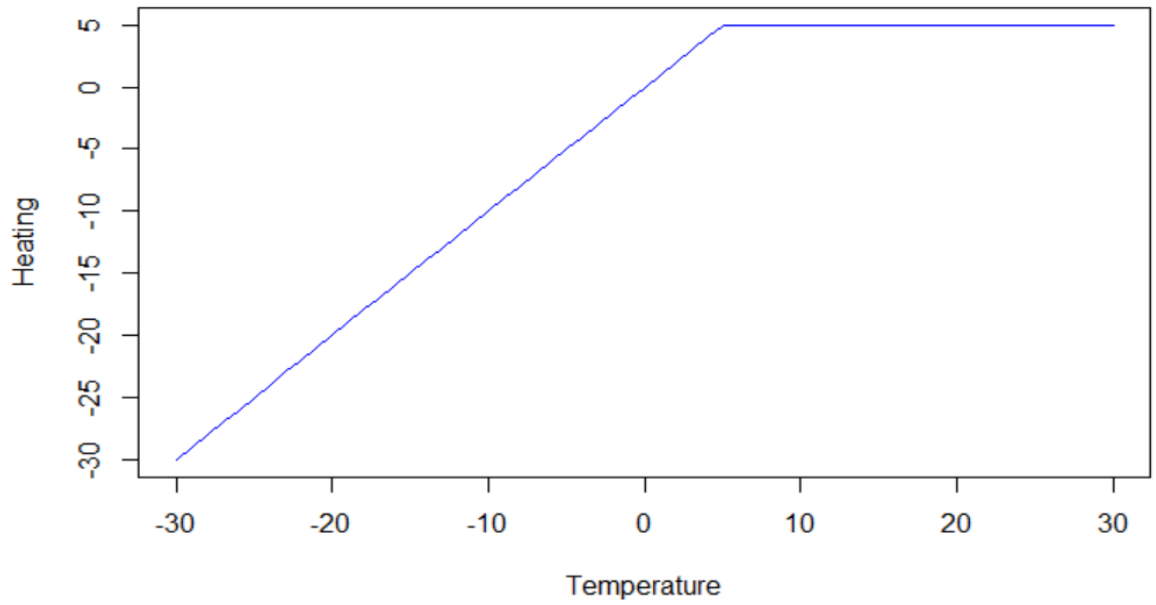
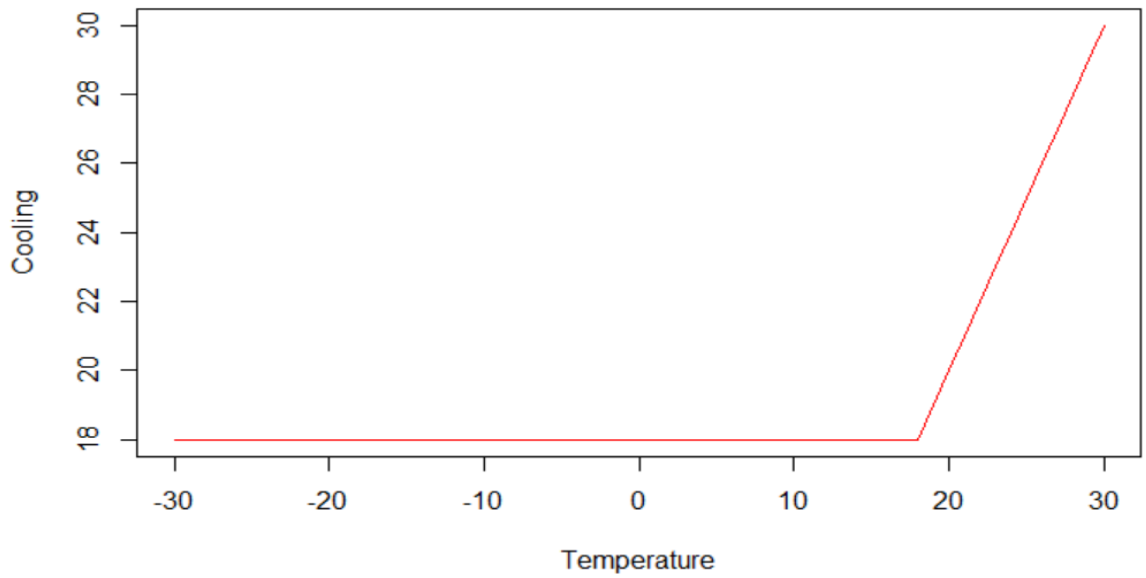


Figure 4.1 Cooling and heating functions graph, used in dynamic regression model



The resulting model would be as below:

```
## Series: Ontario.Demand
## Model: LM w/ ARIMA(1,0,3)(2,1,0)[24] errors
##
## Coefficients:
##          ar1     ma1   ma2     ma3     sar1     sar2
##       0.9238  0.5872  0.30  0.0833  -0.5058  -0.2570
## s.e.  0.0069  0.0172  0.02  0.0169   0.0157   0.0154
##       temperature  I(temperature^2)  heating  Cooling
##           -0.0055             1e-04  -0.0010   0.0107
## s.e.       0.0069             2e-04   0.0103   0.0095
##       Working_DayTRUE
##                0.0144
## s.e.           0.0208
##
## sigma^2 estimated as 0.03807:  log likelihood=856
## AIC=-1688   AICc=-1688   BIC=-1612
```

The fitted model is a seasonal ARIAM with 24 seasonality. It is completely expected to have such a seasonality as we have daily data and patterns of demand are repeated each

day. Residuals of the model show normal distribution behavior (sign correlation $\rho = 0.743$ for residuals).

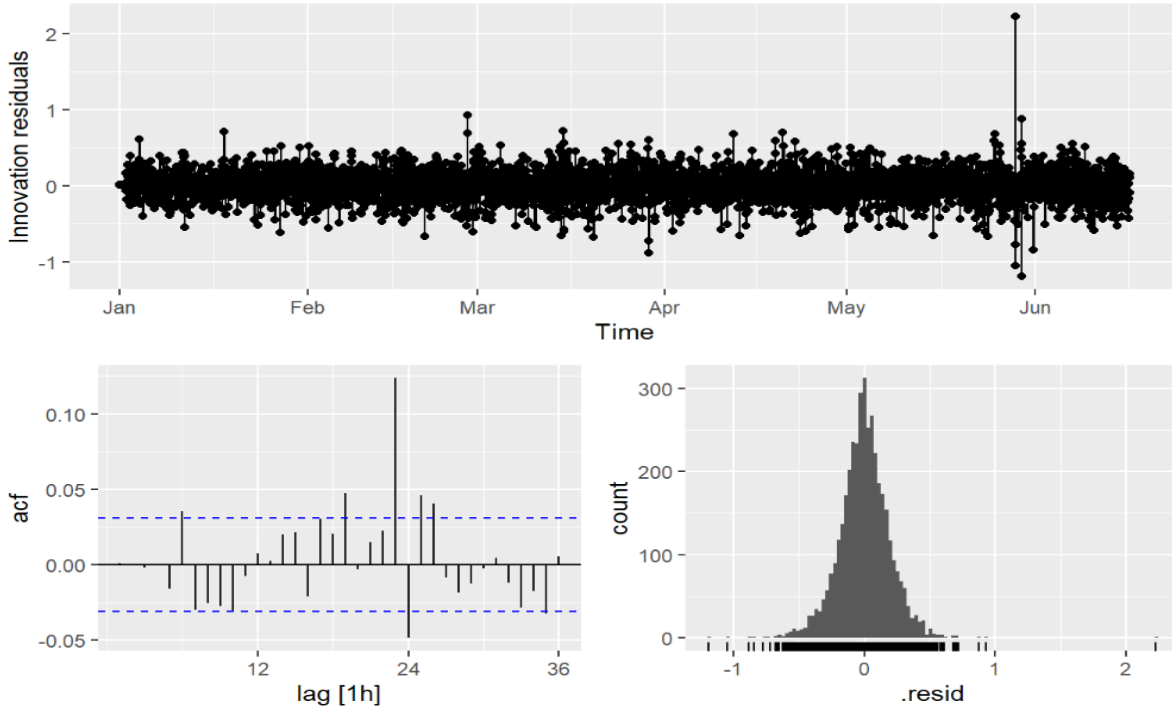


Figure 4.2 Residual analysis of dynamic regression model

However, the model presented in this chapter may be improved by incorporating a suggestion by Vilmarest & Goude (2021). Their innovation is to create separate models for the demand for each hour of the day. They have observed that electricity demand during different hours of the day show high changes in average and variance. This observation is also valid for Ontario's dataset. The below figure shows the hourly electricity demand dispersion from March 2002 to December 2021. During the early hours of the morning (3 a.m. to 6 a.m.) average demand and variance of the demand are relatively low. However, in the afternoon and early evening hours, maybe due to lighting needs, demand average and volatility increase significantly. For example, 5 p.m. shows highest demand and highest demand dispersion for Ontario's dataset.

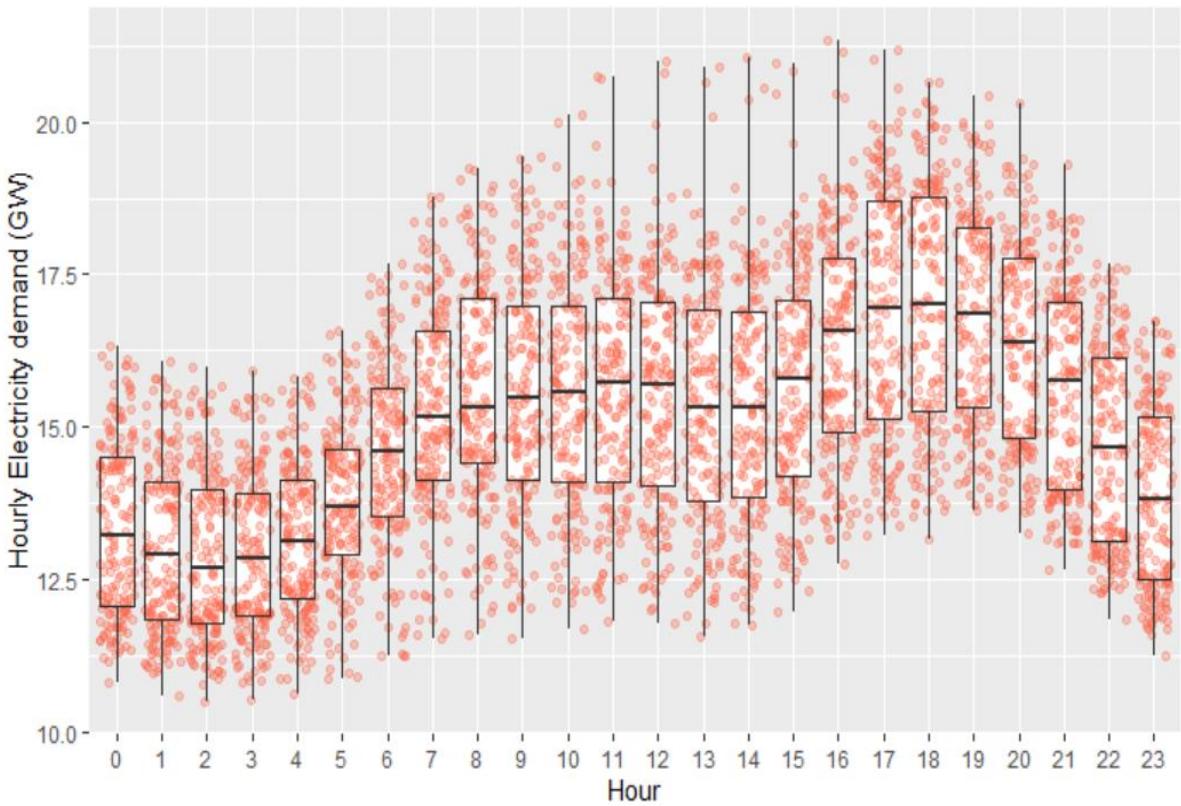


Figure 4.3 Electricity demand changes throughout a day

Furthermore, Vilmarest and Goude (2021) have used many other meteorological variables like wind speed and cloud cover for forecasting electricity demand in their proposed methodologies. For an extended dynamic regression model, additional meteorological features of the province of Ontario have been gathered in this research. A full sample of these added features can be observed in Table 4.2.

Table 4.2 Extended features for forecasting electricity demand

Time <S3: POSIXct>	Dema... <dbl>	Temperature <dbl>	Pressure <dbl>	Cloud_cover <int>	wind_d <int>	Wind_speed <int>	Humidity <int>	Location <chr>	Day_Type <chr>
2021-06-17 17:00:00	17.1	26.1	101.1	3	25	14	23	Ottawa	Weekday
2021-06-18 17:00:00	16.8	17.4	100.6	8	17	18	96	Ottawa	Weekday
2021-06-19 17:00:00	16.3	25.9	100.3	3	28	21	40	Ottawa	Weekend
2021-06-20 17:00:00	18.5	26.9	100.5	7	24	13	39	Ottawa	Weekend
2021-06-21 17:00:00	18.3	28.8	99.3	5	20	34	50	Ottawa	Weekday
2021-06-22 17:00:00	15.0	15.6	101.0	6	32	8	50	Ottawa	Weekday

Considering all these extensions, a dynamic regression model for electricity demand forecasting is proposed here, using all features explained above and for a specific hour of the day. For example, the model below has been created for 5 p.m., as it is the hour with the highest electricity demand for the day.

```
Series: Demand
Model: LM w/ ARIMA(2,1,2)(2,0,0)[7] errors

Coefficients:
        ar1     ar2     ma1     ma2    sar1     sar2  Temperature  Pressure  Cloud_cover  Wind_speed   wind_d
      0.562  0.0598  -0.624  -0.188  -0.165  -0.0470      -0.0138    0.0351       0.0539      0.0123  -0.0036
s.e.  0.518  0.4146   0.506   0.453   0.103   0.0903       0.0133    0.0868       0.0216      0.0058   0.0070
      Day_Type == "Weekday"TRUE  intercept
                          0.651    -0.0051
s.e.                      0.108     0.0249

sigma^2 estimated as 0.6013:  log likelihood=-186
AIC=400   AICc=403   BIC=443
```

The resulting forecasts indicate improvement for the next first two days, but the model's performance declines after that.

Table 4.3 Comparison of different dynamic regression models for Ontario electricity demand

	Dynamic regression model (without added features and without separation of hours)	**Dynamic regression model (with added features, for 5 pm) next two days**	**Dynamic regression model (with added features, for 5 pm) next ten days**
RMSE	0.33	0.07	1.36
MAE	0.21	0.06	1.14

Forecasts are depicted for the next ten days in the figures below (all demands are in GW).

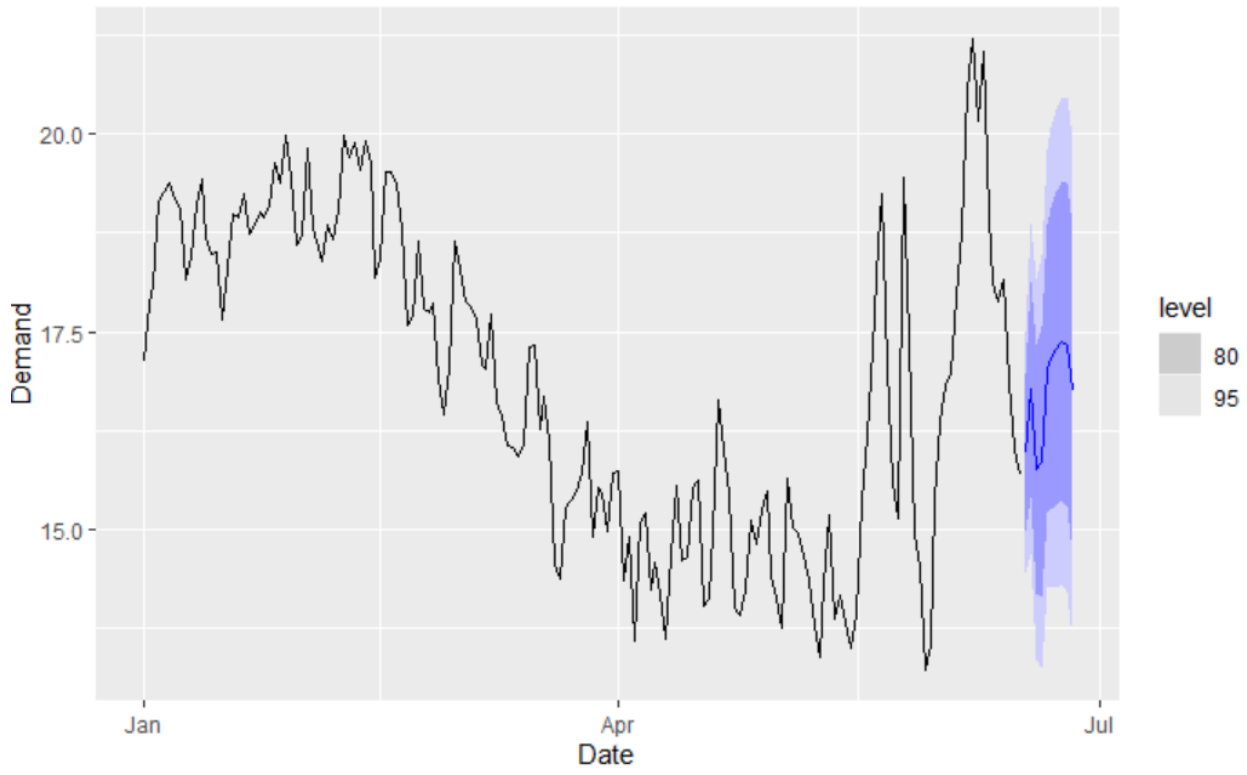


Figure 4.4 Electricity demand forecasts obtained using dynamic regression

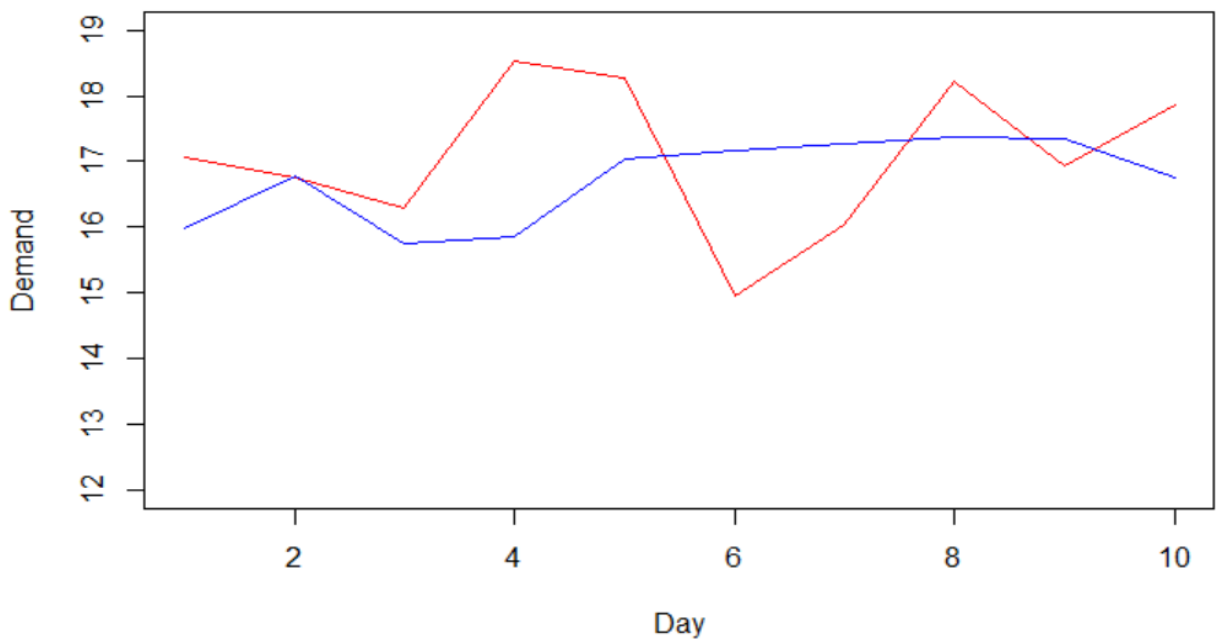


Figure 4.5 Electricity demand forecast for next 10 days

Furthermore, fuzzy interval forecasts are obtained using the adaptive nonlinear fuzzy model suggested in Liang and Thavaneswaran (2022), where values a, b, c, and d in fuzzy membership function are represented as 1st, 25th, 75th and 99th quantiles of the dynamic regression residuals. Fuzzy forecasts of the model for next ten days can be calculated using

$$[\hat{d}_{n+h} + a + \alpha^{\frac{1}{m}}(b-a), \hat{d}_{n+h} + d - \alpha^{\frac{1}{n}}(d-c)]$$

Where $\hat{d}_{n+h}$ is the point forecasts from the dynamic regression model, and α represents the α-cuts of the residuals.

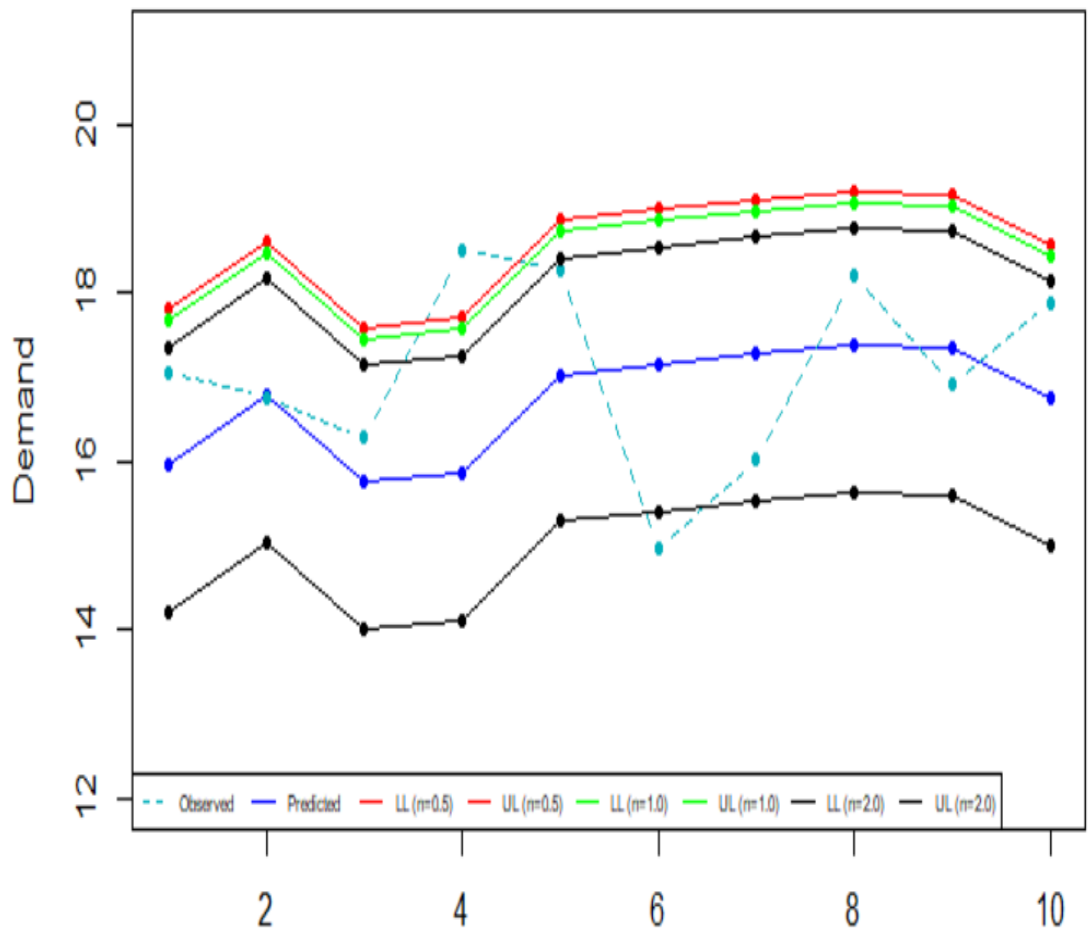


Figure 4.6 Fuzzy forecasts for electricity demand

4.2. Dynamic harmonic regression (DHR)

Recently there has been a growing interest in using dynamic harmonic regression for electricity demand forecasting (Aronsen & Gravem, 2021; Permata & Prastyo, 2022). In addition, Young et al. (1999) introduced dynamic harmonic regression that provides a better forecasting model for processes with seasonality. The idea is that Fourier terms like the formula below (with different K values for weekly or monthly seasonality) can accommodate the cyclical and seasonal behaviors of the processes better than commonly-used seasonal ARIMA models (Young et al., 1999, p.3).

$$S_t = \sum_{i=1}^{K} \left\{ a_i \cos \frac{2\pi i t}{K} + b_i \sin \frac{2\pi i t}{K} \right\}$$

S_t is the seasonal component of demand series d_t . d_t is assumed to be comprised of trend, seasonality and error components:

$$S_t = T_t + S_t + e_t$$

Hyndman and Athanasopoulos (2021) show that SARIMA models do not perform properly for models with lengthy periods of seasonality like weekly seasonality for hourly data (s = 7*24 = 168).

Table 4.4 Training and testing sets for DHR model

Hourly Electricity demand (GW)	DHR model
Training Set	2021-01-01 00:00:00. to 2021-06-16 16:00:00
Testing Set	2021-06-16 17:00:00. to 2021-06-18 16:00:00

The same variables are incorporated here as in the DR model to observe if the results improve after running the dynamic harmonic regression.

```
## Series: Ontario.Demand
## Model: LM w/ ARIMA(2,0,3) errors
##
## Coefficients:
##          ar1    ar2    ma1   ma2     ma3  temperature
##        0.622  0.307  1.063  0.51  0.0927      -0.0176
## s.e.   0.668  0.632  0.671  0.50  0.1870       0.0056
##        Cooling  heating  Working_DayTRUE
##         0.0402  -0.0137          -0.0151
## s.e.    0.0101   0.0076           0.0286
```

Residuals of the model do not indicate promising improvements when compared to dynamic regression models (sign correlation $\rho = 0.743$ for residuals). When running a DHR model for the above-mentioned training data set, about 15 minutes was required on a Core i7 Windows laptop. However, the same data provides a dynamic regression model with better residuals in about 2 minutes. While the same dataset would provide a dynamic regression model with better residuals in about 2 minutes.

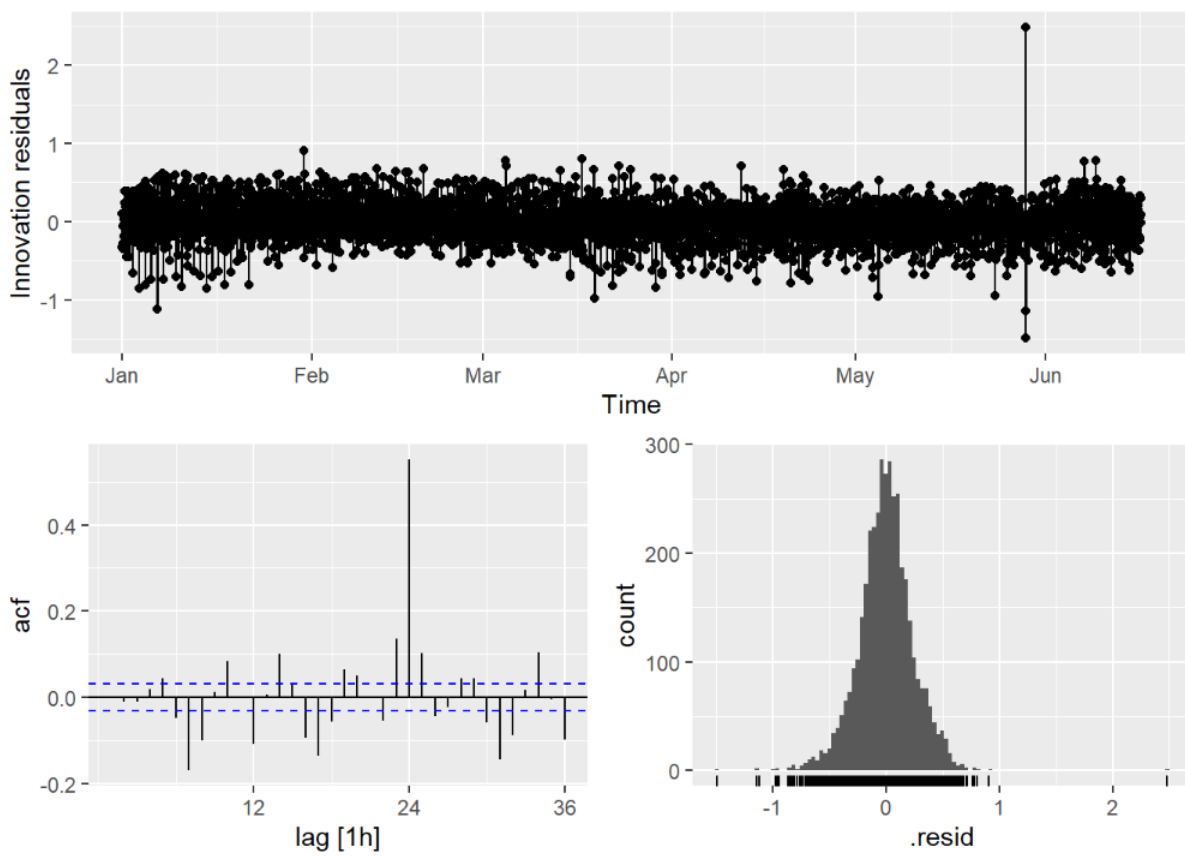


Figure 4.7 Residual analysis for DHR

4.3 NNETAR

Neural network autoregression (NNETAR) is a univariate model that uses a series' previous p lags for forecasting. It is distinct from $AR(p)$ models because the weights given to each lag are obtained through running a feed-forward neural network (Hyndman & Athanasopoulos, 2021). NNETAR is an automated R function in package fable (v0.3.1; Mitchell et al., 2021), which determines the number of lag (p) and number of nodes (K) in a hidden layer for a NNETAR (p,K) model. Model selection is made based on information criteria AIC criteria. Figure 4.8 represents how a feedforward NNETAR is modeled.

Table 4.5 Training and testing set for NNETAR

Hourly Electricity demand (GW)	NNETAR model
Training Set	2021-01-01 00:00:00. to 2021-06-16 16:00:00
Testing Set	2021-06-16 17:00:00. to 2021-06-18 16:00:00

The resulting model is a NNETAR (28,1,14) [24], a seasonal model with daily seasonality (each 24 hours), that considers last 28 demand lags as inputs for the neural networks model as well as last days demand at the same hour. It includes one hidden layer of 14 neurons fully connected to input nodes. This model may be represented in the figure below.

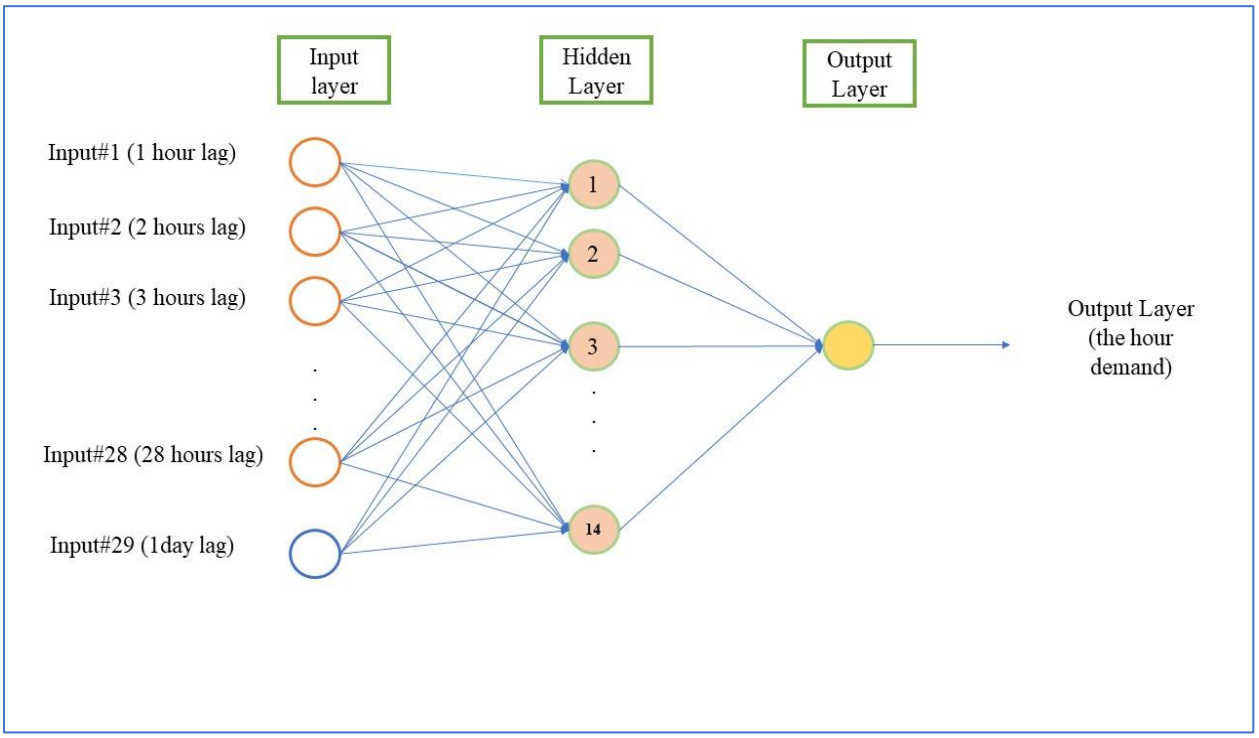


Figure 4.8 NNETAR model used for forecasting Ontario electricity demand

4.4 Prophet

Prophet is a forecasting package developed by Facebook that has been well received among analysts and forecasters with different backgrounds (Taylor & Letham, 2018, p.23). Taylor and Letham (2018, p.5) show superior forecasting performance of prophet compared to automatic ARIMA, exponential smoothing, and naïve approaches, specifically for series with long-term seasonality.

Prophet decomposes time series into 4 components. A piecewise-linear trend (G_t), seasonality modeled with Fourier terms (S_t), holiday effects as dummy variables (H_t), and white noise error term (e_t).

$$d_t = G_t + S_t + H_t + e_t$$

Table 4.6. Training and testing set for Prophet

Hourly Electricity demand (GW)	Prophet model
Training Set	2021-01-01 00:00:00. to 2021-06-16 16:00:00
Testing Set	2021-06-16 17:00:00. to 2021-06-18 16:00:00

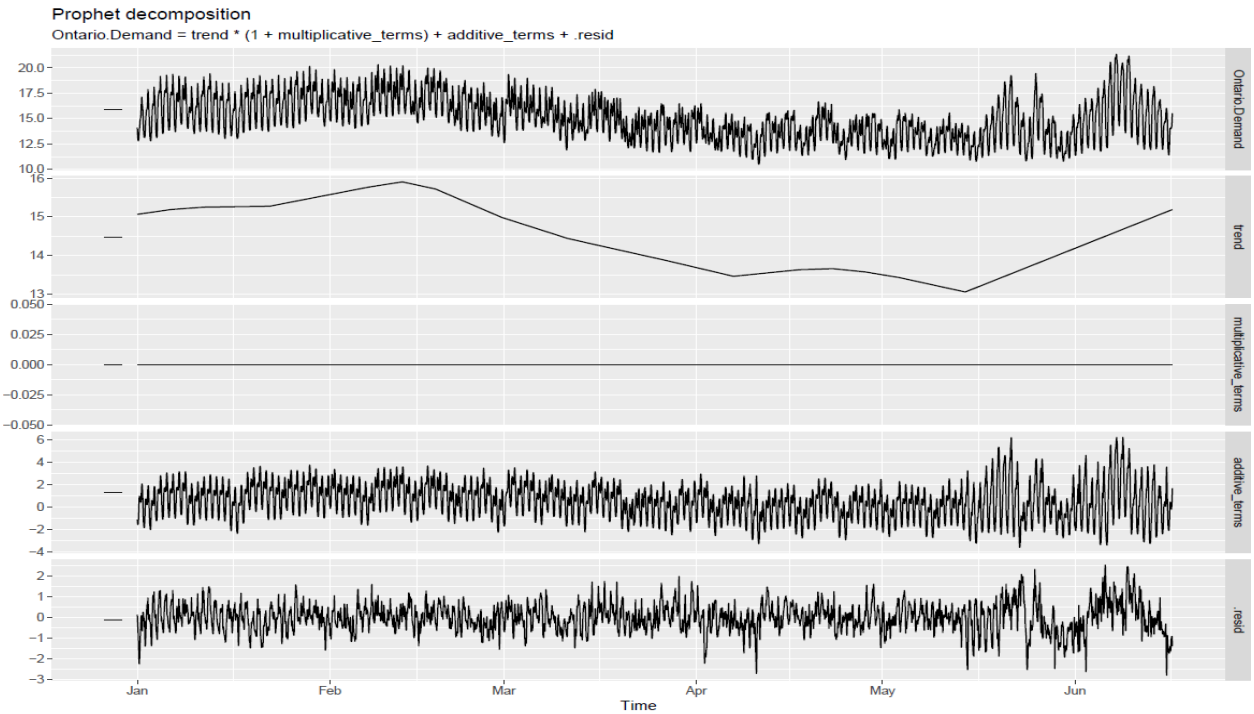


Figure 4.9 Prophet model decomposition

Prophet model is very easy to use and almost completely automated. It runs very quickly, but the results are not as promising as dynamic regression model. The residuals of the model, depicted below, show large standard deviation in residuals compared to the dynamic regression model.

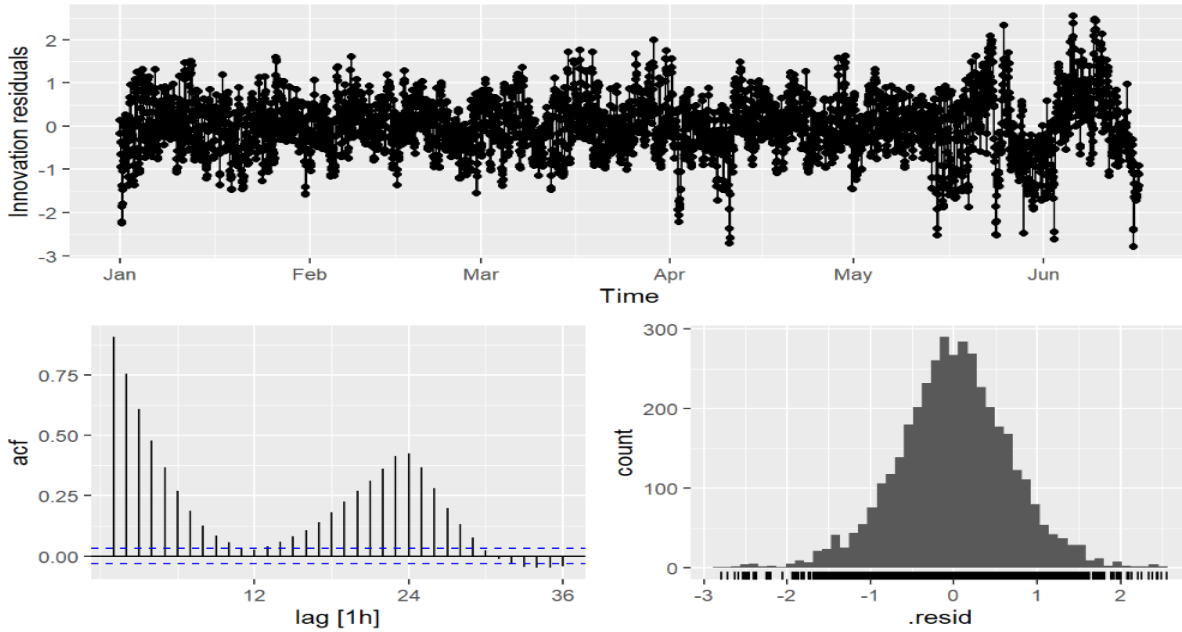


Figure 4.10 Prophet model residuals analysis

4.5 Discussion of the results

This part compares applied models for forecasting hourly electricity demand based on two error measures: root mean square error (RMSE) and mean absolute error (MAE). Each model had the same input for electricity demand data from January 1, 2021 to June 16,2021 (4000 data points). Also, dynamic regression, dynamic harmonic regression, and prophet models utilize temperature and day-type data as inputs. Nevertheless, seasonal ARIMA and NNETAR rely solely on load data and do not include other components in their forecasting model (univariate).

Following Hyndman and Athanasopoulos (2021), a combined forecast approach is used in this thesis to combine four of the forecasts, i.e., dynamic regression, dynamic harmonic regression, prophet and NNETAR. Thus, we identify a combined forecast as

$$combination = \frac{(Dynamic\ regression + Dynamic\ Harmonic\ Regression + Seasonal\ ARIMA + prophet)}{4}$$

The newly popular prophet (Taylor & Letham, 2018) model is fairly fast in processing the data compared to NNETAR, However, it follows from Table 3.8 that RMSE for the prophet is large compared to the other models.

Table 4.7. Comparison of forecast performance for different models

Model	RMSE	MAE
Dynamic Regression	0.329	0.205
NN	0.376	0.233
Seasonal ARIMA	0.411	0.353
Dynamic Harmonic Regression	0.978	0.882
Prophet	1.789	1.752
Combined forecast	0.592	0.552

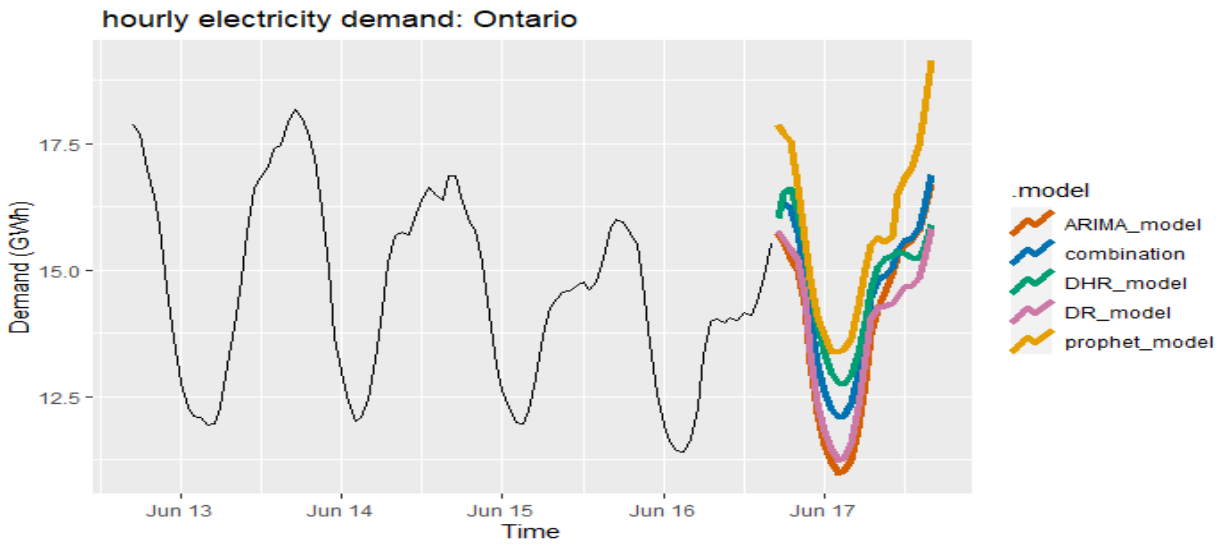


Figure 4.11. Next 24 hours forecast of hourly demand

Dynamic regression model-based predictions are more suitable (according to MSE and MAE) than dynamic harmonic regression-based forecasts in this study. It suggests substituting seasonal ARIMA autoregression models with Fourier terms does not lead to better outcomes.

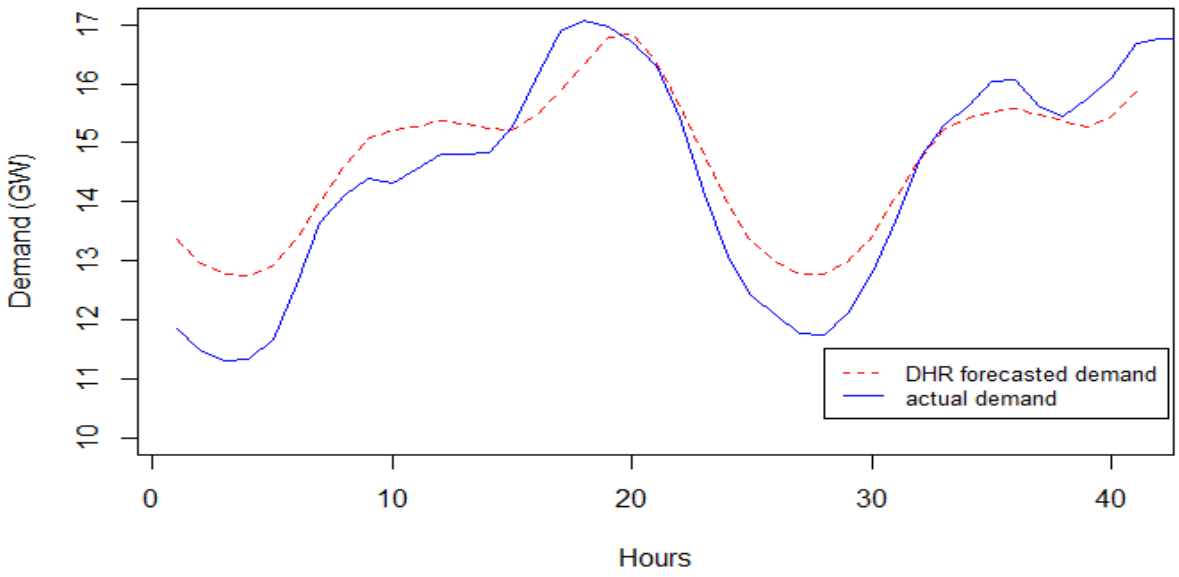


Figure4.12 Dynamic Harmonic Regression forecasts

Also, NNETAR forecast outperforms the forecasts obtained by using the benchmark model (seasonal ARIMA). Resulted NNETAR model has 14 hidden layers and looks back to 28 lags as well as a seasonal lag (lag 24). While, optimal seasonal ARIMA model by using r package fable is ARIMA (4,0,0) (0,1,1) [24].

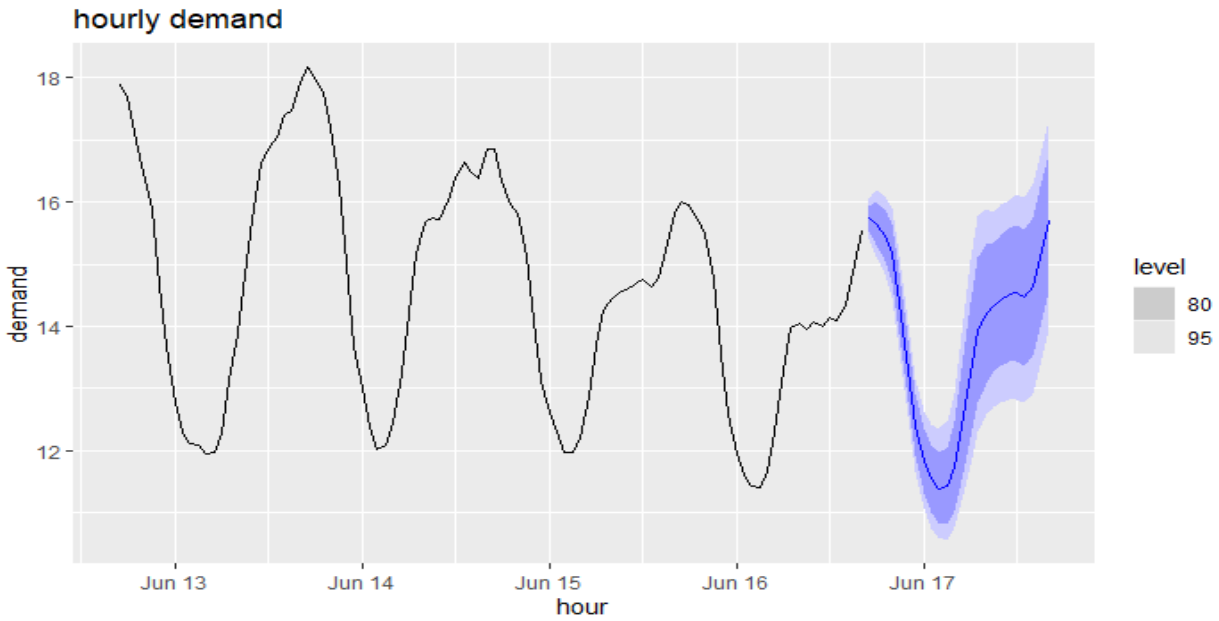


Figure 4.13. NNETAR model forecast for next 24 hours

Chapter 5: Daily demand forecasting

Daily electricity demand forecasting is about forecasting the total demand of one or multiple days ahead for demand of a certain region or province. According to Mirasgedis et al. (2006) daily demand forecasts are becoming increasingly crucial as climate change dictates that governments and policy makers pay more attention to mid-term and long-term planning for development of power infrastructures. Mirasgedis et al. (2006) considered Greek hourly electricity demand data and for the sake of daily aggregation, they have used sum of 24 hours demand for each day. Also, they have considered the monthly demand data by aggregating all days demand of a month. They have also gathered meteorological factors like temperature and relative humidity for forecasting purposes.

5.1 Daily data analysis

Having in mind the suggestions of the research by Mirasgedis et al. (2006) and the book by Hyndman and Athanasopoulos (2021), in this thesis the Ontario dataset is aggregated at a daily level by adding all 24 hour demands in MW and then dividing by 1000 to make the GW unit.

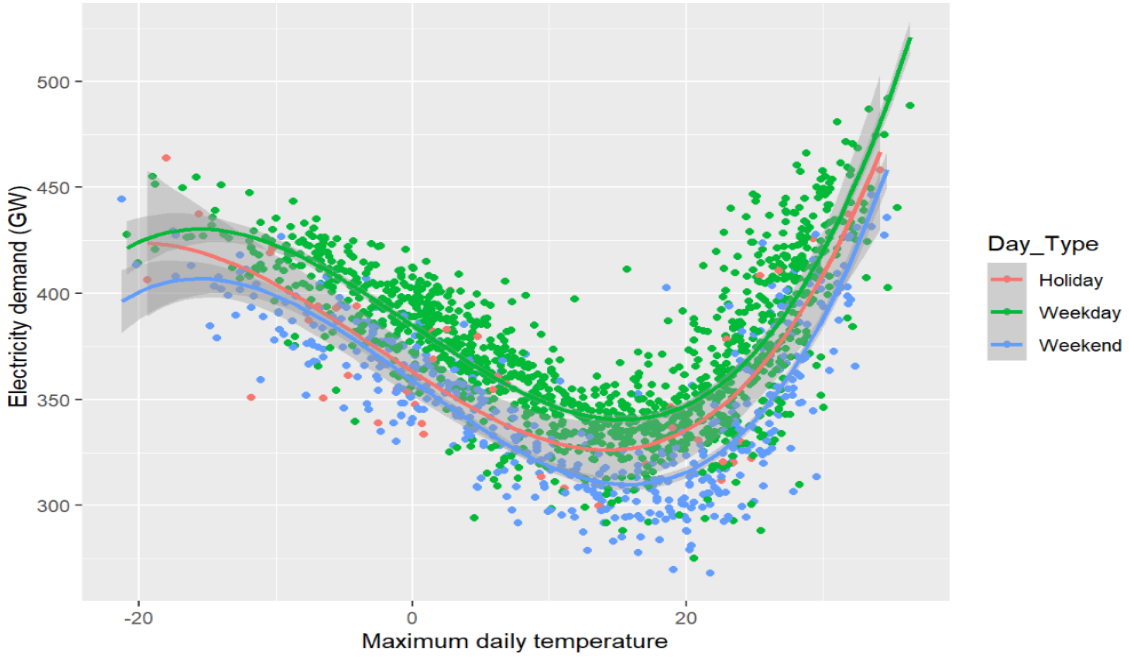


Figure 5.1 Relationship between daily electricity demand and maximum temperature

In figure 5.1, the relationship between demand (GW) and maximum daily temperature is depicted. The demand data used for this graph is from 2017 to 2021, so it includes both pre-pandemic and post-pandemic historical data. The maximum temperature is used because in upcoming forecasting methods like dynamic regression and NNETAR, maximum temperature shows better results that other aggregation forms like mean temperature or minimum temperature. Also, it is worth considering that holidays and weekends have different relationship. Generally, weekdays demand for electricity (green dots and green fitted curve) are higher than in weekend demands (blue dots and blue fitted curve) and holidays (red dots and red fitted curve). I have included the relationship between demand and temperature from the Greek dataset (Mirasgedis et al., 2006, pg. 210), to emphasize on the general non-linear relationship between demand and temperature in other datasets as well. However, obviously temperature ranges differ in Ontario and Greece.

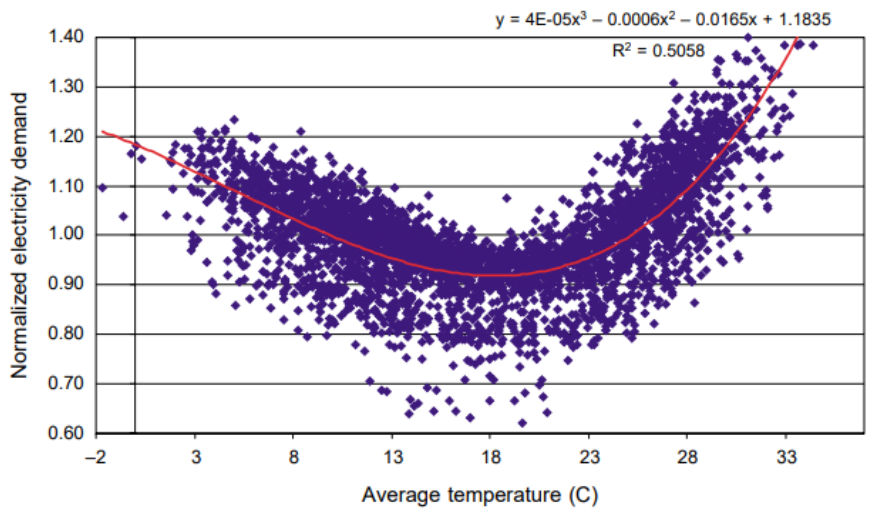


Figure 5.2 Electricity demand and temperature relationship in Greek dataset from Mirasgedis et al. (2006)

In Figure 5.3, daily demand (GW) and the maximum temperature (°C) series are illustrated together to better grasp the impact of temperature on demand in each season of the year.

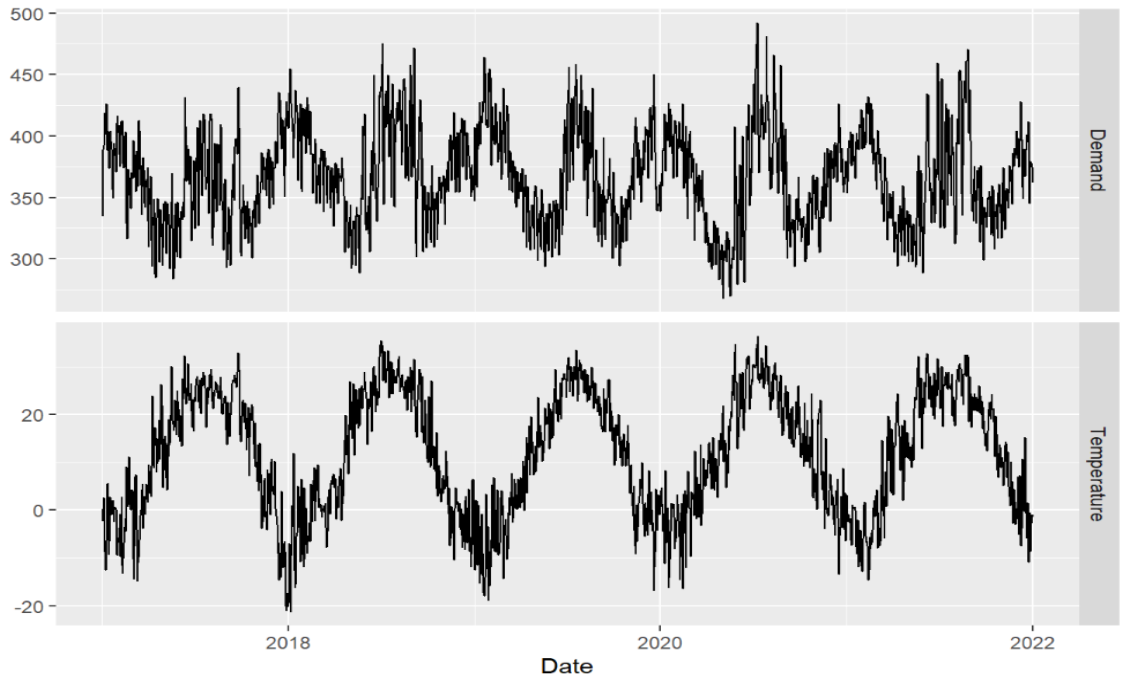


Figure 5.3 Daily demand and daily maximum temperature series from 2017 to 2018

Also, to gather the daily demand data at a monthly level, we can infer from figure 5.4 that winter (December to March) and summer months (July to September) Have more demand and more demand variation as well. The daily demand is again from Januray 2017 to December 2021.

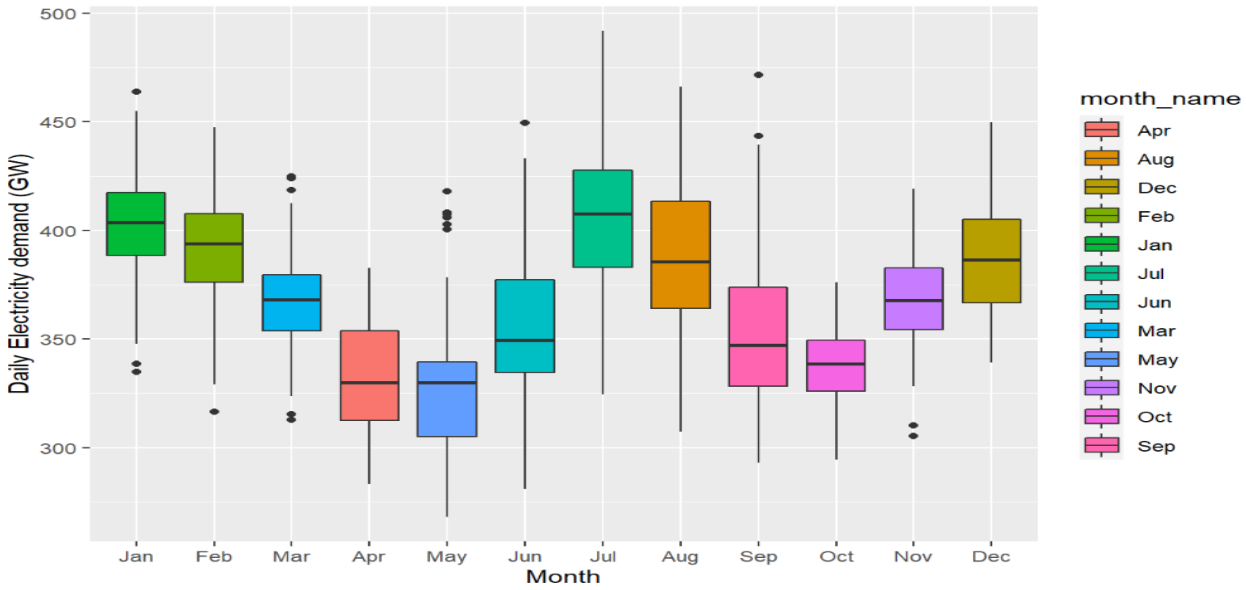


Figure 5.4 Variation of daily demand in different months of the year

For the sake of finding the best degree for temperature in polynomial form for forecasting daily demand, multiple polynomial regression models with degrees from 1 to 5 had been run and their forecasting performance was compared. I observed that

$$\text{Demand} \sim (\text{Temperature}^3)$$

has the best forecasting performance regarding RMSE and R^2 . Figure 6.5 depicts the best fitted polynomial function for daily demand.

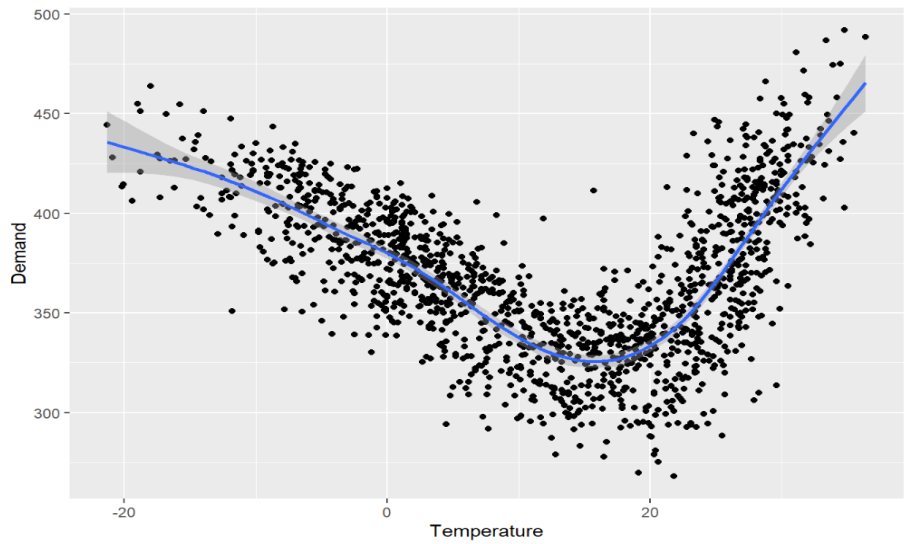


Figure 5.5 Smoothing function for daily demand data

Figure 5.6 shows the forecasting perfomacne of the polynomial smoothing model Demand ~ a + b (Temperature)+ c (Temperature^2) + d (Temperature^3) .

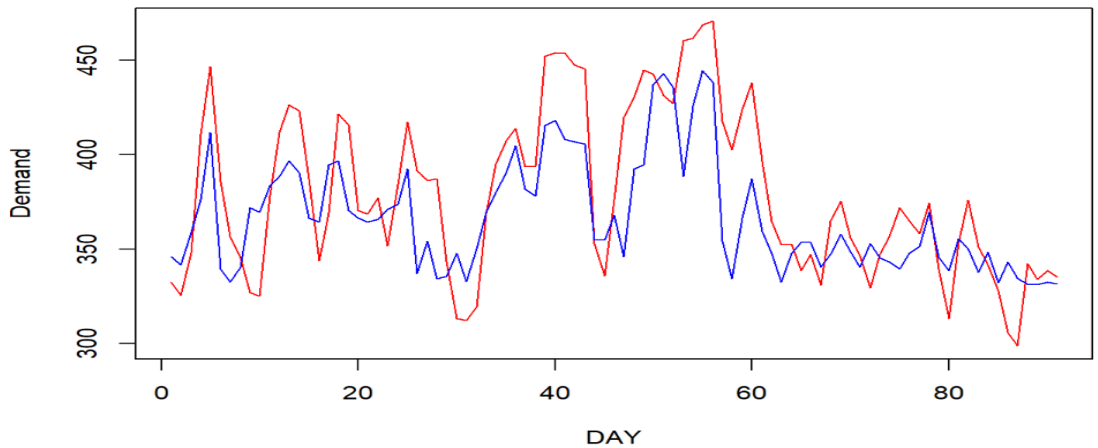


Figure 5.6 Forecasting performance of simple polynomial smoothing model

5.2 Dynamic regression for daily demand

As pointed out in the previous section of this chapter, the 3rd degree for the polynomial regression on temperature is chosen. Also, day type as categorical variable is used for the dynamic regression model.

The fitted dynamic regression model is as follows:

```
## Series: Demand
## Model: LM w/ ARIMA(0,1,4)(2,0,0)[7] errors
##
## Coefficients:
##           ma1      ma2      ma3      ma4    sar1    sar2  Temperature
##       -0.2052  -0.1999  -0.1472  -0.1362  0.0149  0.0712      -2.0647
## s.e.   0.0309   0.0268   0.0266   0.0259  0.0281  0.0272       0.1143
##       I(Temperature^2)  I(Temperature^3)  Day_Type == "Weekday"TRUE
##                 0.0155            0.0023                    22.4707
## s.e.            0.0061            0.0002                     0.8344
##
## sigma^2 estimated as 180.9:  log likelihood=-5873.66
## AIC=11769.32   AICc=11769.5   BIC=11827.49
```

Temperature ($x_{1,t}$) is the signifying maximum daily temperature and day type($x_{2,t}$) is a categorical variable with two possible values: One when day is a weekday, and zero when the day is a holiday or weekend day.

Table 5.1. Training and testing set for daily DR

Daily Electricity demand (GW)	DR model
Training Set	2017-07-01 00. to 2021-06-30
Testing Set	2021-07-01 to 2021-09-30

If we consider demand as daily demand at day t as D_t then,

$$D_t = -2.0647x_{1,t} + 0155{x_{1,t}}^2 + 0.0023{x_{1,t}}^3 + 22.4707x_{2,t} + \eta_t$$

and

$$(1 - 0.0149L - 0.0712L^2)(1 - \mathrm{B})\eta_t$$
$$= (1 + 0.2025B + 0.1999B^2 + .1472B^3 + 0.1362B^4)\ \varepsilon_t$$

$$\varepsilon_t \sim NID(0,180.9)$$

$$L = B^7$$

It is assumed that if the model is adequate, then the ε_t would be a normally distributed white noise series. Figure 5.7 depicts ε_t series and its autocorrelation function (ACF) as well as its distribution. The heavy tails of the distribution and high autocreation in lags 21 and some other lags can lead to the conclusion that residuals are not necessarily normally distributed.

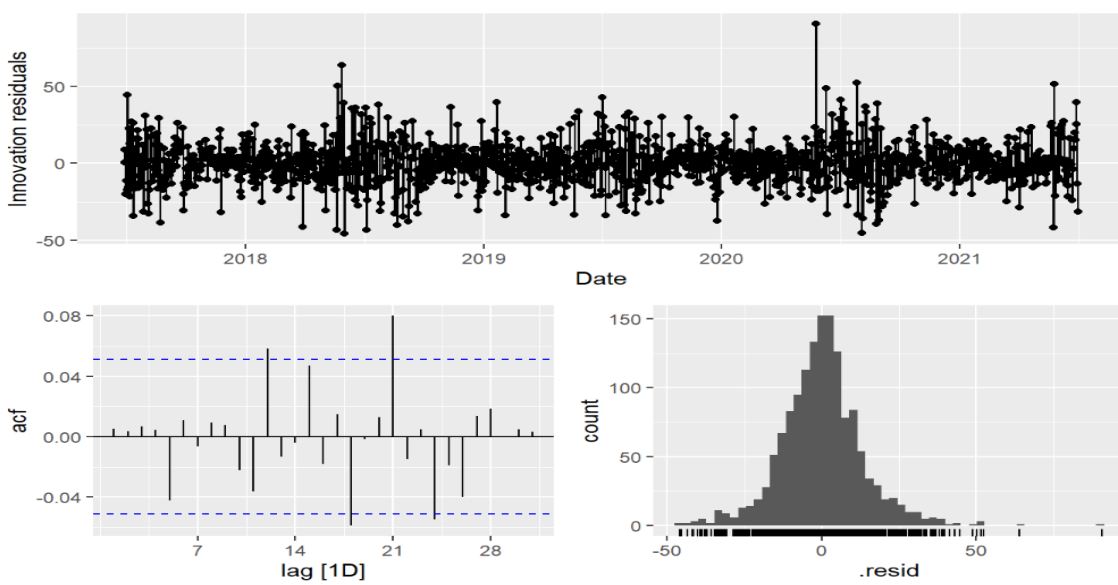


Figure 5.7 Residuals analysis for the dynamic regression model for daily demand

Figure 5.8 shows predictions for elecricty demand using the dynamic regression model, assuming that normality assumption is true.

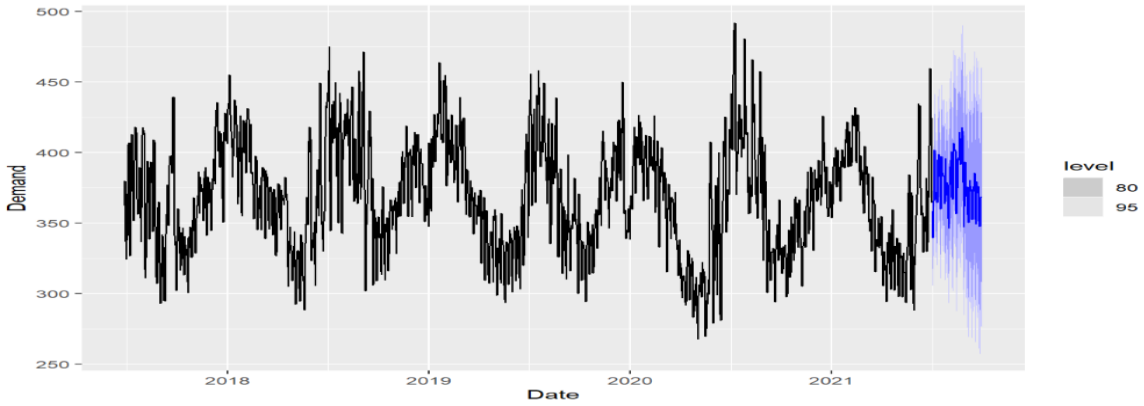


Figure 5.8 Daily forecasts of dynamic regression model

Also, a more detailed depiction of the dynamic regression model forecasts are represented in Figure 5.9. The blue line represents the actual demand (GW) and the red line represents daily forecasts for the next three months.

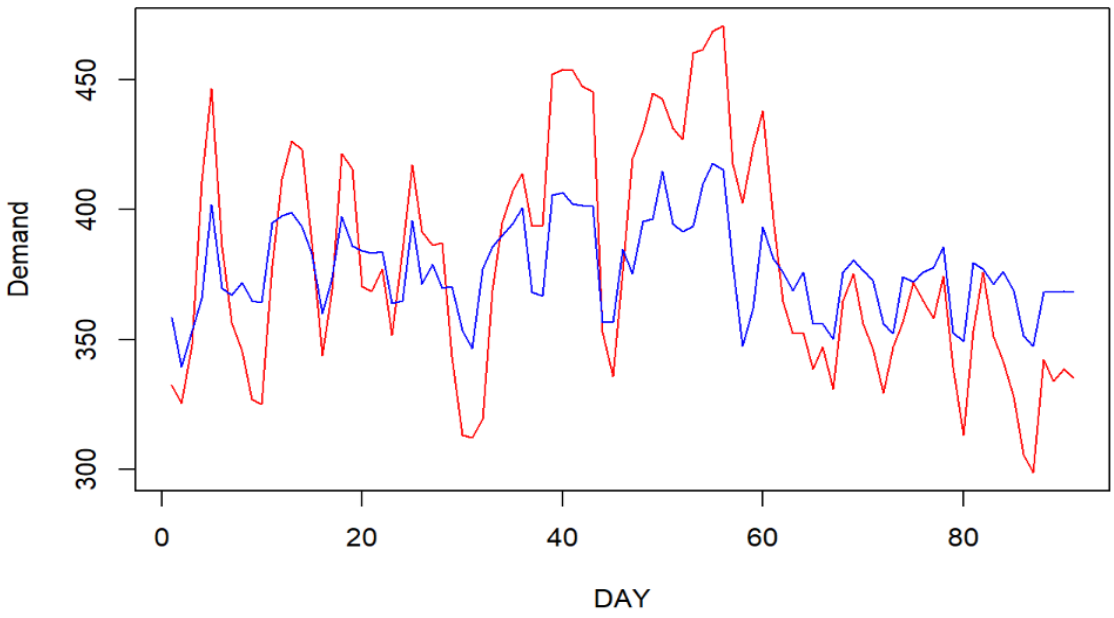


Figure 5.9 another depiction of demand forecasts for next three months (blue: forecasts, red: actual demand)

However, we can come up with better prediction intervals (PIs) for the model according to Liang and Thavaneswaran (2022). Conducting a detailed residual analysis,

we can infer that residuals distribution for the dynamic regression model is a student's t with degrees of freedom $\upsilon = 4.682$.

Liang and Thavaneswaran (2022) describe a data-driven method for calculating PIs when the normality assumption is not acceptable. Using the sign correlation concept, sign correlation of the residuals $e_1, e_2, e_3, \dots, e_n$, when the sample mean is $\bar{e}$ is:

$$\hat{\rho}_e = Corr\big(e_t - \bar{e}\,, sgn(e_t - \bar{e})\big)$$

Form Liang and Thavaneswaran (2022), It is known that if e_t follows the normal distribution, then $\hat{\rho} = \sqrt{2/\pi} = 0.7918$. But if e_t has the characteristics of student t distribution, $\hat{\rho}$ is going to be smaller than $\sqrt{2/\pi} = 0.7918$. In case the degrees of freedom (υ) of the associated student's t distribution can be computed by:

$$2\sqrt{\upsilon - 2} = \hat{\rho}_e(\upsilon - 1)Beta[\frac{\upsilon}{2}, \frac{1}{2}]$$

where $Beta[x, y]$ is Beta function. These concepts and calculations are highly useful for producing better data-driven PIs for electricity demand instead of normal PIs depicted in Figure 5.8.

After calculating the student's t, $\upsilon = 4.682$. The PIs with significance level of $100(1 - p)\%$ are calculated as:

$$\hat{d}_{n+h} \pm t_{\frac{p}{2},\upsilon}\hat{\sigma}$$

$$h = 1,2,3, \dots l$$

when l is the prediction horizion. For example, if $p = 0.05$ and $\upsilon = 5$, then $t_{\frac{p}{2},\upsilon} = 2.57$

Bootstrap prediction intervals

Bootstrapping is another method which can be used to come up with prediction intervals for l future forecasts. The bootstrapping method assumes that the training set and testing set have the same errors. By sampling l residuals from the training data

$(e_1, e_2, e_3, \dots, e_l)$, and adding those values to point forecasts $\hat{d}_{n+1}, \hat{d}_{n+2}, \hat{d}_{n+3}, \dots, \hat{d}_{n+l}$, a simulated string of l future observations are gathered. When this procedure is repeated a large number of times $s = 1000$, s realization of future observations is gathered. Now, with calculating the percentiles of each simulated step realization, PIs are obtained. For example, 0.5th percentile and 99.5th percentile of simulated future observation $\hat{d}_{n+1}$, constitute the 99% PI for one step ahead forecast using bootstrapping technique.

Figure 5.10 depicts and compares PIs for the dynamic regression model forecast obtained by the student's t distribution (green lines for upper bound and lower bound) and bootstrapping method (purple lines for upper bound and lower bound). The blue line represents the daily forecasts obtained from the dynamic regression model shown in the previous section and the red line shows actual demand values. 95 % Bootstrapping PIs are captured from the resampling $s = 1000$ times from training set and t PIs are gathered using $\upsilon = 4.682$, $p = 0.05$ and $\hat{\sigma} = 13.49$.

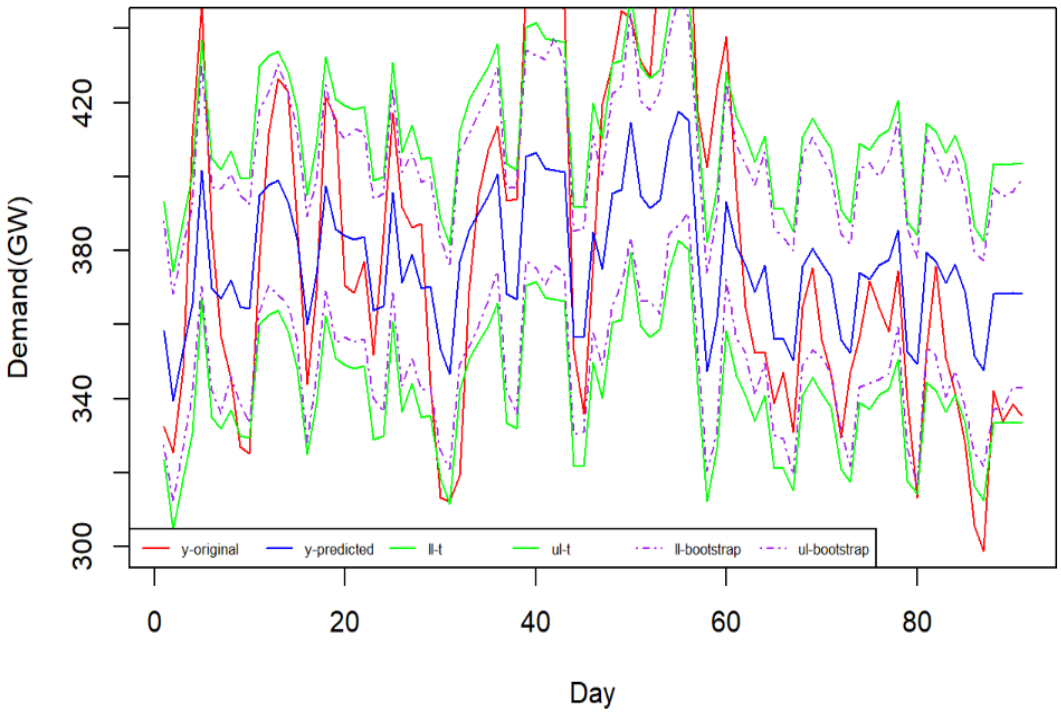


Figure 5.10 Bootstrap and student's t data-driven PIs for daily electricity demand

5.3 NNDR for daily demand forecasting

Feed-forward neural networks (NNs) are among the most popular methods for approximating nonlinear functions. NNs consist of an input layer for including the input variables, hidden layer(s) and an output variable. A hidden layer includes elements called neurons, which receive the previous layer's information (can be the input layer of another hidden layer) and then this information is passed to the newly processed information to the next layer (can be the output layer or another hidden layer). The process which transforms the input into output is called supervised learning. With a sufficient number of neurons, the process can learn useful features of input for capturing the signal of the output layer.

In a neural networks dynamic regression (NNDR) model for electricity demand forecasting, the input layer consists of p lags of demand data $(d_{t-1}, d_{t-2}, \dots, d_{t-p})$ and n features for forecasting demand like temperature $(x_{t,1})$ and day-type $(x_{t,2})$. The output layer is the target value, which is the demand for time t, d_t.

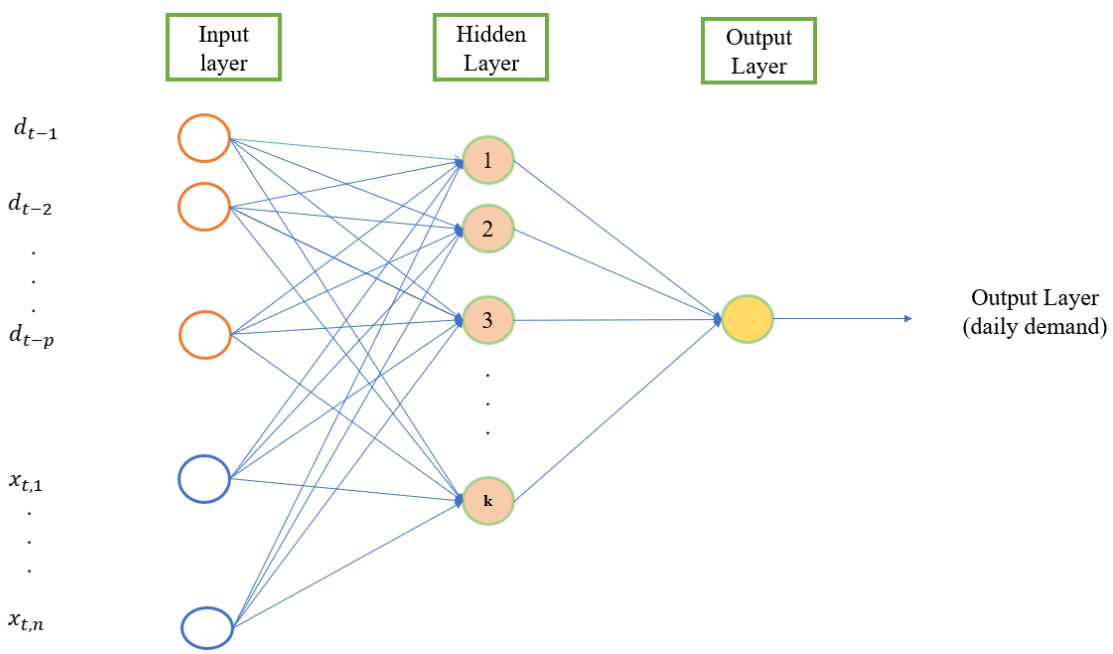


Figure 5.11 NNDR model for daily electricity demand forecasting

Since each node in the hidden layer is a function of the previous layer, from layer 1 to layer 2 we can have:

$$z^2 = \omega_{l0}^1 + \sum_{j=1}^{p} \omega_{lj}^1 y_{t-j} + \sum_{j=p+1}^{p+n} \omega_{lj}^1 x_{t,j-p}$$

And

$$a_l^2 = g^2(z^2)$$

g^2 is a non-linear transformation for the layers function z^2.

The transition from layer $k-1$ to layer k can be shown as:

$$z^k = \omega_{l0}^{k-1} + \sum_{j=1}^{p_{k-1}} \omega_{lj}^{k-1} a_{lj}^{k-1}$$

$$a_l^k = g^k(z^k)$$

Table 5.2. Training and testing set for daily NNDR

Daily Electricity demand (GW)	NNDR model
Training Set	2017-07-01 00. to 2021-06-30
Testing Set	2021-07-01 to 2021-09-30

This thesis examines the forecasts of a feed-forward NNDR with a hidden layer to improve the forecast of a dynamic regression model discussed in the previous section. The variable day-type has three levels (weekdays, weekends, and holidays), and the day-type variable is converted to a single numeric variable first. This can be done by designTreatmentsZ() and prepare() functions available in R package vtrea (Mount & Zumel, 2021)

The next step is to fit an $NNDR(p, P, k)_m$ model to obtain demand forecasts, and the function nnetar() is fed with temperature and level-coded day-type features. In the NNDR model, p and P are the AR orders of the non-seasonal and seasonal parts, and k is the number of nodes in the hidden layer. The fitted NNDR model contains 16 neurons in the hidden layer with the last 29 observations $d_{t-1}, d_{t-2}, \dots, d_{t-29}$ used as inputs for forecasting the target d_t. The decay rate is set to 0.1 for regularization in order to

compensate for model complexity. Note that the histogram and ACF plots in Figure 5.14 indicate that the residuals are autocorrelated with heavy tails. Figure 5.12 and Figure 5.13 visualize point forecasts of electricity demand using NNDR models for July 2021 to September 2021.

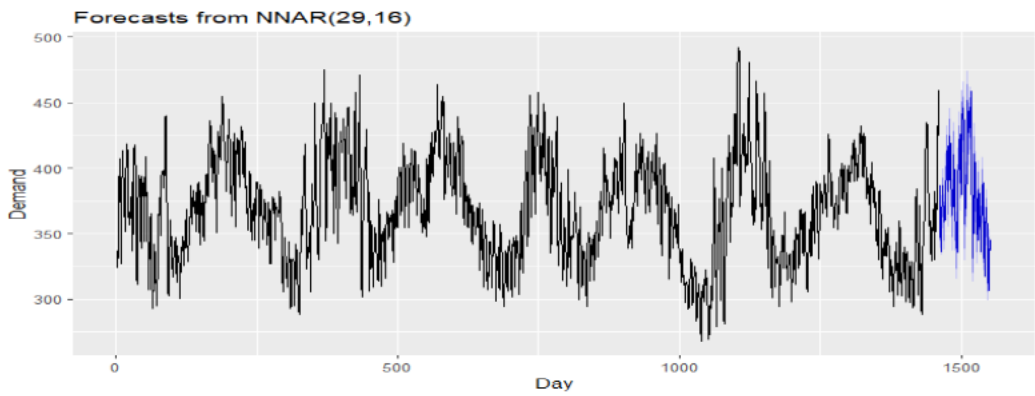


Figure 5.12 NNDR model results for daily demand forecasting

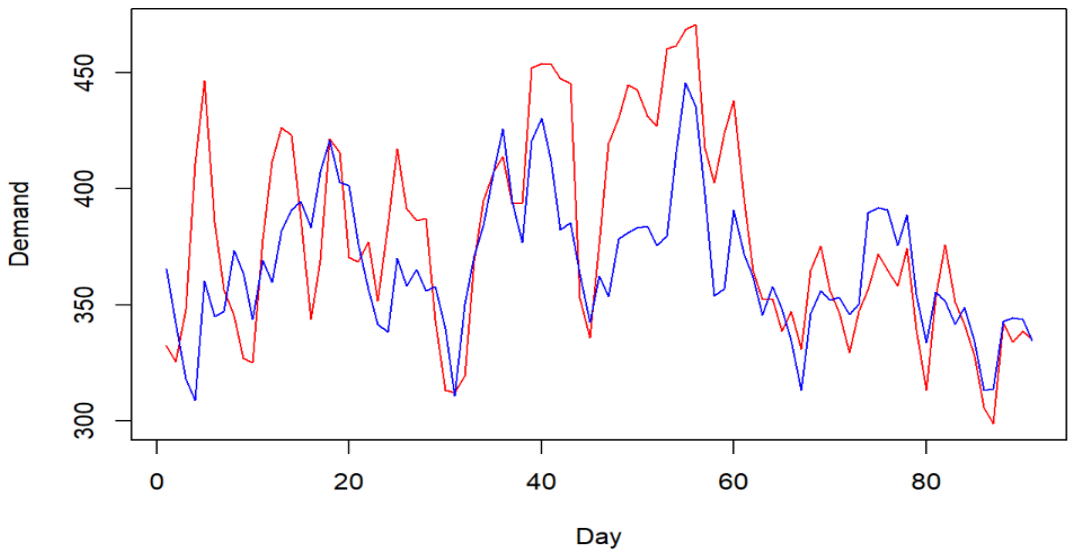


Figure 5.13 Another depiction of NNDR demand forecasts for next three months (blue: forecasts, red: actual demand)

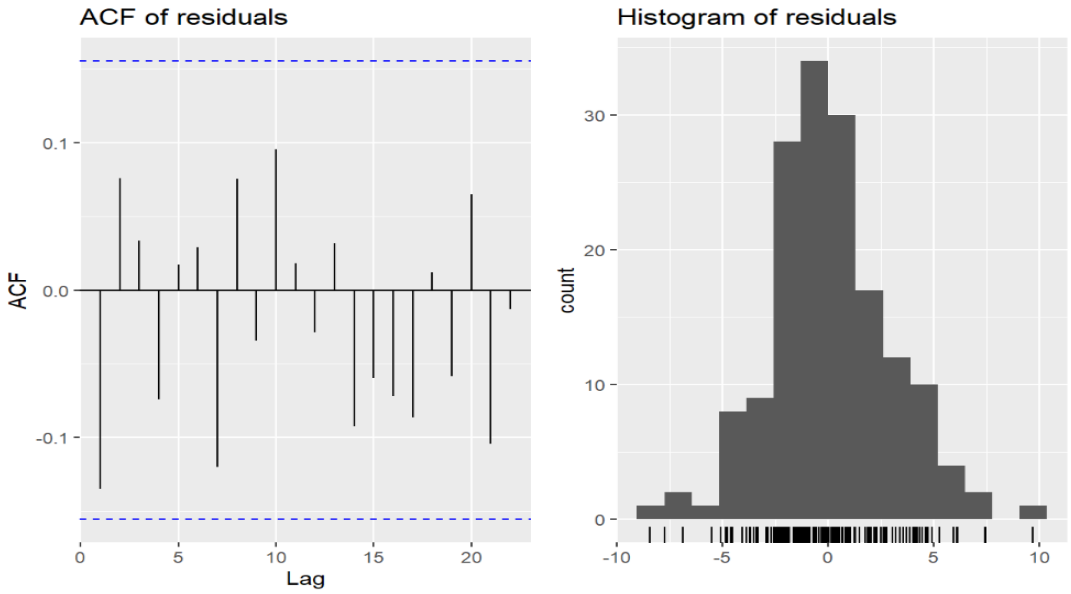


Figure 5.14 Residuals of NNDR model

By default, nnetar() uses information from the in-sample residuals for the innovations used in PIs. The residuals can be arbitrarily small depending on the complexity of the model. As a reference, the standard deviation of the nnetar() innovations is roughly 4 times smaller than the residuals from ARIMA(). Therefore, we obtain the innovations from cross validation (R CVar() function). The innovations from the cross validation follow a heavy-tailed Student t distribution with the estimated $\upsilon = 5.5185$.

95% probabilistic Student t PIs are computed with $\upsilon = 5.5185$, $p = 0.05$ and $\hat{\sigma} = 0.0744$. 95% bootstrap PIs are calculated based on s = 1000 times of resampling from the residuals. Student t PIs (green lines) and bootstrap PIs (dashed purple lines) are shown in Figure 5.15.

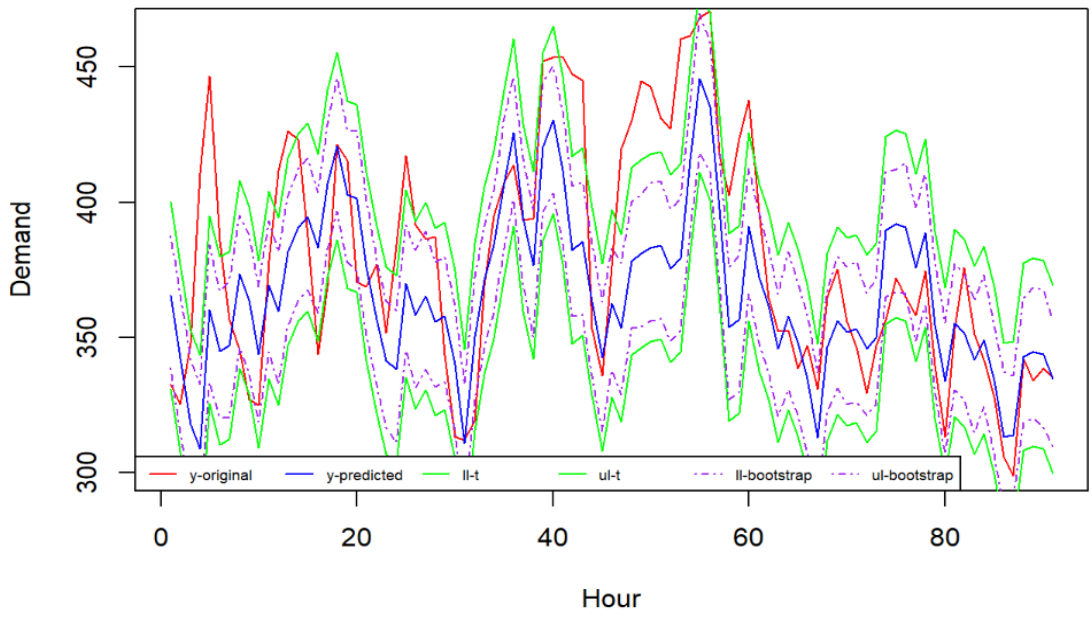


Figure 5.15 Bootstrap and student's t data-driven PIs for daily electricity demand

5.4 Prophet model for daily demand

Finally, comparing the results of the dynamic regression and NNDR models, the same training and testing sets have been applied to the automated prophet model as an acceptable benchmark in the industry. The Prophet model has an interface in the fable Package, which this study uses to forecast daily electricity demand. Temperature and day type are the input variables. Figure 5.17 depicts point forecasts for the testing period along with the training data. The blue lines denote forecasts of electricity demands from July 2021 to September 2021. A more detailed depiction of prophet model forecasts are represented in Figure 5.18. The blue line represents the actual demand (GW) and the red line represents daily forecasts for the next three months.

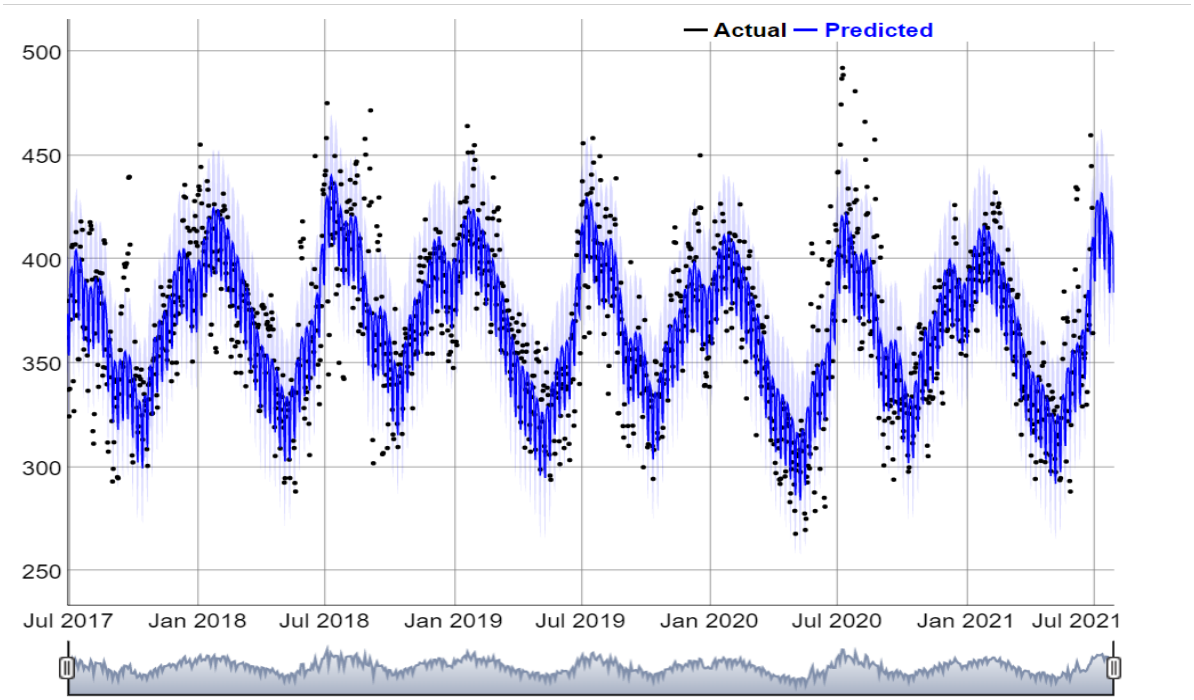


Figure 5.16 Fitted prophet model for daily forecasts

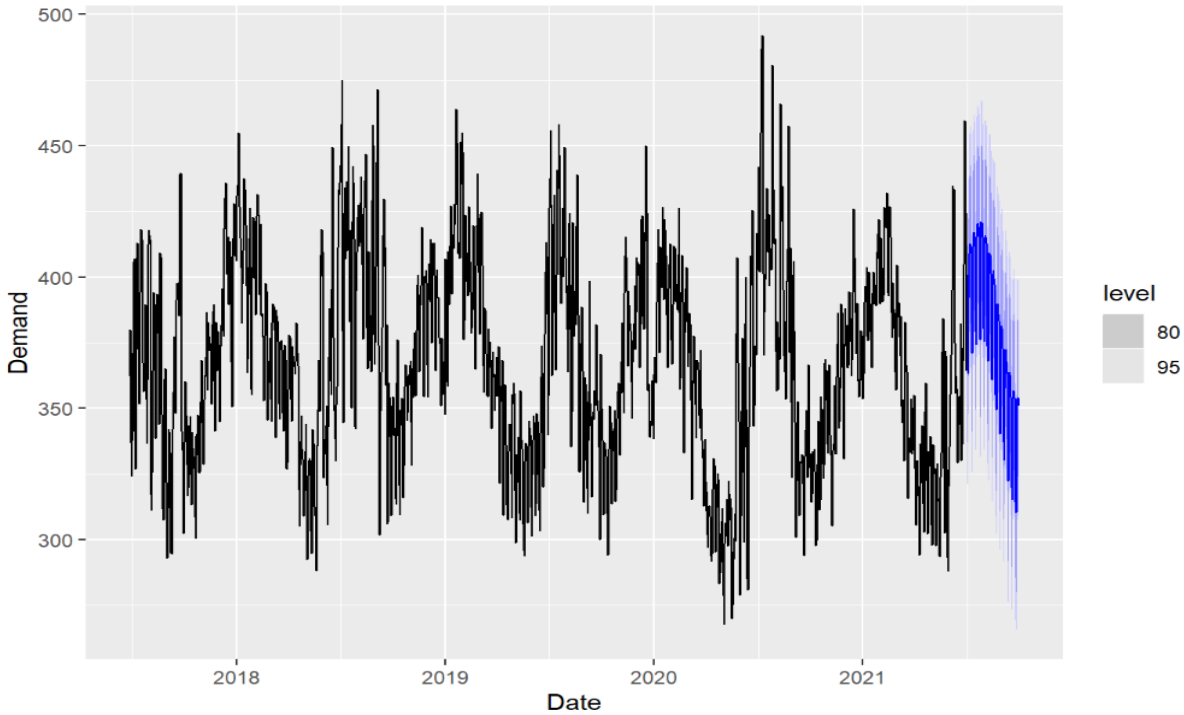


Figure 5.17 Prophet model results for daily demand forecasting

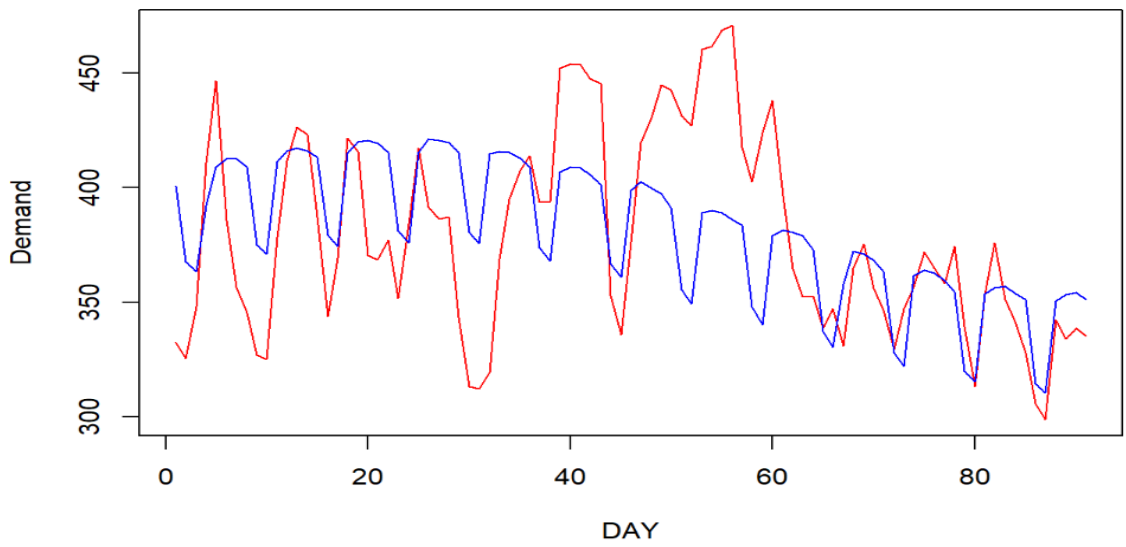


Figure 5.18 Another depiction of prophet demand forecasts for next three months (blue: forecasts, red: actual demand

5.5 Discussion

When comparing the performance of the three models discussed in this chapter, different metrics for point forecasts and PIs have been devised. Testing set mean square error (TRMSE) and Testing set mean absolute error (TMAE) are tools to compare the performance of the models regarding their point forecasts. On the other hand, coverage probability is a metric measuring the number of times that actual values of target variable (daily demand) fall inside the prediction interval. In other words, it is a measure for quantifying how successful prediction intervals are in capturing the target variable.

Table 5.3. Daily forecast models performance comparison

Method	**TRMSE**	**TMAE**	**Coverage probability**	
			Student t PI	**Bootstrap PI**
DR	32.0015	28.4963	59.78%	51.09%
NNDR	24.7315	18.3434	83.70%	81.52%
prophet	38.6067	30.5683	59.78%	55.43%

The NNDR model dominates the other models in every category. Its TRMSE and TMAE are lower than both other models and coverage probabilities (both bootstrap and t) are higher. Dynamic regression performs better than prophet in point forecasts, however,

its high standard deviation leads to the weakest coverage probabilities. Yet the prophet model can be considered an acceptable demand forecasting method due to its ease of use.

Chapter 6: Conclusion

In this research, the issue of electricity demand forecasting is reviewed and investigated in several layers and forms. The significance of electricity supply chains and their tremendous impact on other sectors of the economy has been introduced. Without proper management of a power network, huge costs are incurred to the whole economy and the well-being of citizens. Blackouts and waste of energy resources are two crucial catastrophic implications of weak power demand forecasting in both centralized and decentralized supply chains. Any player in the electricity supply chain, whether producers, distributors, or consumers, benefit from accurate demand forecasts.

Electricity demand forecasting can play a huge role in developing more sustainable and resilient supply chains. For instance, with the gradual transition of road transport from fossil fuel vehicles to electric vehicles (EVs), electricity pricing during peak hours and off-peak hours has become more crucial for EV owners. Many renewable energy generations which happen during off-peak hours are non-dispatchable, because the supply surpasses the demand at that specific hour. However, battery swapping stations (BSSs) with strong forecasting tools can purchase the energy at these hours with a much-diminished price. This research mostly focuses on analyzing and discussing these strong data-driven forecasting tools. To name a few of these forecasting models, dynamic regression (DR), dynamic harmonic regression (DHR), neural networks dynamic regression (NNDR), and prophet are utilized in this research.

Contributions:

Short-term load forecasting (STLF) is a highly crucial topic for research. There are many applications for STLF in operations research, like scheduling and commitment planning in power generation facilities, especially during peak hours. STFL usually ranges for half-hours, hours, and days ahead of consumption prediction. For instance, unit commitment (UC) is an optimization problem for determining which generation units in a power generation company should work and what resources should be used. Such planning enables power generation companies to meet the demand with the lowest possible cost and to diminish fuel consumption.

This research sheds light on how electricity supply chain research links the decision-making body of knowledge with actual problems of real-world like threats and sustainability issues of power markets. This research also investigates how with better forecasts, these problems can be addressed. Then in the numerical section of the thesis, which contains Canadian data, the emphasis is to evaluate the predictive power of specific recent innovations in the field.

It is evident from the province of Ontario's historical data that multiple factors impact the forecasting performance of varied statistical and machine learning techniques. Some of these factors are calendar variables, meteorological aspects, and global and national economic disruptions like the Covid-19 pandemic. Also, the performance of different models differs as further forecasting horizons are chosen.

Another contribution of this research is to incorporate fuzzy forecasts for responding to inherent uncertainty and ambiguity of energy markets. With the current increase of disruptions and uncertainties, utilizing fuzzy models in supply chains seems to be more appropriate than ever. Trapezoidal fuzzy numbers and adaptive fuzzy sets for demand forecasting introduced and fuzzy forecast intervals developed and provided in this thesis.

With the numerical analysis of the Canadian dataset for Ontario, it can be suggested that dynamic regression model has the best forecasting performance for hourly data. While, for daily aggregation of data, neural networks dynamic regression (NNDR) provides stronger forecasts with better coverage probability.

Further research suggestions:

For further research, probabilistic load forecasts of this research can be combined with a UC optimization problem for a power generation company. The company operates multiple units with diverse resources like nuclear, coal-fired, gas, hydroelectric, and pumped storage. Each unit has a lower limit and upper limit of power output. The UC problem determines which units to operate for each hour of a day and upcoming days forward from 1 to n. It can be translated to a mixed-integer linear programming (MILP). The primary decision variable will be a 0-1 variable determining if a unit will be on or off at a specific hour or a particular day.

In the UC problem, load forecast for specific day hours can be supplied into a balance constraint for units. A novel research approach that can be undertaken from this thesis is to incorporate stochastic programming instead of linear programming. This research already recognizes the importance of probabilistic forecasts instead of point load forecasts. However, incorporating this probabilistic approach into a UC problem is a new challenge worth investigating.

Another attractive research area in combining load forecasting and operation research is economic dispatch. Similar to the UC problem, economic dispatch tries to minimize generation costs. However, in the economic dispatch, fuel and efficiency of units are considered, while in UC problem, maintenance and start-up costs of a unit are included too.

For further research, conducting the same methods on other available Canadian provinces datasets can help investigate the variability and risk associated with power markets in provinces like Alberta or Manitoba. However, finding reliable high-frequency data for all other regions (like what is used in this research) is challenging.

Also, for another research direction, the price volatility of Ontario can be retrieved and analyzed through the proposed models of this research. As another research direction, the combined application of demand forecasting and mathematical programming is enticing. Problems like plant scheduling and resource allocation in generation units can be combined with the forecasting results of a proposed model in this research.

References

Agdas, D., & Barooah, P. (2020). Impact of the COVID-19 pandemic on the US electricity demand and supply: an early view from data. *IEEE Access*, *8*, 151523-151534

Banitalebi, B., Appadoo, S. S., & Thavaneswaran, A. (2020). Data Driven Approach for Reduced Value at Risk Forecasts in Renewable Power Supply Systems. In *2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp. 1-6). IEEE

Banitalebi, B. (2021). *Probabilistic forecasts of day-ahead electricity prices in a highly volatile electricity market* [Master's thesis, University of Manitoba]

Bedi, J., & Toshniwal, D. (2019). Deep learning framework to forecast electricity demand. *Applied energy*, *238*, 1312-1326

B. Efron and T. J. Hastie (2017), *Computer age statistical inference: Algorithms, evidence, and data science.* Cambridge University Press.

Carvalho, M., Bandeira de Mello Delgado, D., de Lima, K. M., de Camargo Cancela, M., dos Siqueira, C. A., & de Souza, D. L. B. (2021). Effects of the COVID-19 pandemic on the Brazilian electricity consumption patterns. *International Journal of Energy Research*, *45*(2), 3358-3364

Chae, Y. T., Horesh, R., Hwang, Y., & Lee, Y. M. (2016). Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy and Buildings*, *111*, 184-194

Chopra, S., & Sodhi, M. S. (2004). Supply-chain breakdown. *MIT Sloan management*

review, *46*(1), 53-61

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research, 184*(3), 1140-1154

Dai, Q., Cai, T., Duan, S., & Zhao, F. (2014). Stochastic modeling and forecasting of load demand for electric bus battery-swap station. *IEEE Transactions on Power Delivery*, *29*(4), 1909-1917.

J. De Vilmarest and Y. Goude. (2022) State-Space Models for Online Post-Covid Electricity Load Forecasting Competition, *IEEE Open Access Journal of Power and Energy*,192-201, doi: 10.1109/OAJPE.2022.3141883.

Eseye, A. T., Lehtonen, M., Tukia, T., Uimonen, S., & Millar, R. J. (2019). Machine learning based integrated feature selection approach for improved electricity demand forecasting in decentralized energy systems. *IEEE Access*, 7, 91463-91475

Guo, C., Ge, Q., Jiang, H., Yao, G., & Hua, Q. (2020). Maximum power demand prediction using Fbprophet with adaptive Kalman filtering. *IEEE Access*, *8*, 19236-19247

Halbrügge, S., Schott, P., Weibelzahl, M., Buhl, H. U., Fridgen, G., & Schöpf, M. (2021). How did the German and other European electricity systems react to the COVID-19 pandemic? *Applied Energy*, *285*, 116370

Huang, J., Xue, Y., Dong, Z. Y., & Wong, K. P. (2012). An efficient probabilistic assessment method for electricity market risk management. *IEEE Transactions on Power Systems*, *27*(3), 1485-1493

Hagan, M. T., Demuth, H. B., and Beale, M. (1996). *Neural network design*. Pws Pub.

Hoque, M. E., Thavaneswaran, A., Appadoo, S. S., Thulasiram, R. K., & Banitalebi, B. (2021, July). A Novel Dynamic Demand Forecasting Model for Resilient Supply Chains using Machine Learning. *In 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 218-227). IEEE.

Hoque, M. E., Thavaneswaran, A., Paseka, A., & Thulasiram, R. K. (2021, July). An Algorithmic Multiple Trading Strategy Using Data-Driven Random Weights Innovation Volatility. *In 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 1760-1765). IEEE.

Jabbarzadeh, A., Fahimnia, B., & Rastegar, S. (2017). Green and resilient design of electricity supply chain networks: a multiobjective robust optimization approach. *IEEE Transactions on Engineering Management*, *66*(1), 52-72.

Liu, M., & Wu, F. F. (2007). Risk management in a competitive electricity market. *International Journal of Electrical Power & Energy Systems*, *29*(9), 690-697.

Leach, A., Rivers, N., & Shaffer, B. (2020). Canadian electricity markets during the COVID-19 pandemic: An initial assessment. *Canadian Public Policy*, *46*(S2), S145-S159

Liang, Y., & Thavaneswaran, A. (2022, June). Long Term Interval Forecasts of Demand using Data-Driven Dynamic Regression Models. *In 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 250-259). IEEE.

Mirasgedis, S., Sarafidis, Y., Georgopoulou, E., Lalas, D. P., Moschovits, M., Karagiannis, F., & Papakonstantinou, D. (2006). Models for mid-term electricity

demand forecasting incorporating weather influences. *Energy*, *31*(2-3), 208-227

Hewamalage, C. Bergmeir and K. Bandara (2021). Recurrent neural networks for time series forecasting: Current status and future directions, *International Journal of Forecasting*, *37*(1) (388-427).

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, *2*(5), 359-366.

Mount, J., & Zumel, N. (2017). Vtreat: A statistically sound 'data. frame'processor/conditioner.

Mayer, K., & Trück, S. (2018). Electricity markets around the world. *Journal of Commodity Markets*, *9*, 77-100.

Kattelus, M. (2021, July 21). Understanding how the power market creates social welfare and utilises flexibility is essential for the implementation of a carbon-free future *FortUM* https://www.fortum.com/about-us/forthedoers-blog/power-markets-social-welfare.

Manowska, A. (2020). Using the LSTM network to forecast the demand for electricity in Poland. *Applied Sciences*, *10*(23), 8455.

Pierrot, A., & Goude, Y. (2011). Short-term electricity load forecasting with generalized additive models. *Proceedings of ISAP power*, *2011*.

POLICY, U. (2020). Market Mechanisms: Options for Climate Policy. *Center for Climate and Energy Solutions.*

Prostean, G., Badea, A., Vasar, C., & Octavian, P. (2014). Risk variables in wind power supply chain. *Procedia-Social and Behavioral Sciences*, *124*, 124-132.

Román-Portabales, A., López-Nores, M., & Pazos-Arias, J. J. (2021). Systematic review of electricity demand forecast using ANN-based machine learning algorithms. *Sensors*, *21*(13), 4544.

Shayeghi, H., & Ghasemi, A. (2013). Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme. *Energy Conversion and Management*, *74*, 482-491.

Shamsi, M., & Cuffe, P. (2021). A prediction market trading strategy to hedge financial risks of wind power producers in electricity markets. *IEEE Transactions on Power Systems*, *36*(5), 4513-4523

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In *2018 17th IEEE international conference on machine learning and applications (ICMLA)* (pp. 1394-1401). IEEE.

Santiago, I., Moreno-Munoz, A., Quintero-Jiménez, P., Garcia-Torres, F., & Gonzalez-Redondo, M. J. (2021). Electricity demand during pandemic times: The case of the COVID-19 in Spain. *Energy policy*, *148*, 111964

Takeda, H., Tamura, Y., & Sato, S. (2016). Using the ensemble Kalman filter for electricity load forecasting and analysis. *Energy*, *104*, 184-198.

Thavaneswaran, A., Thulasiram, R. K., Zhu, Z., Hoque, M. E., & Ravishanker, N. (2019, July). Fuzzy value-at-risk forecasts using a novel data-driven neuro volatility predictive model. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)* (Vol. 2, pp. 221-226). IEEE

Thavaneswaran, A., Thulasiram, R. K., Zhu, Z., Hoque, M. E., & Ravishanker, N.

(2019, July). Fuzzy value-at-risk forecasts using a novel data-driven neuro volatility predictive model. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)* (Vol. 2, pp. 221-226). IEEE.

Thavaneswaran, A., Thulasiram, R. K., Hoque, M. E., & Appadoo, S. S. (2021, December). Data-Driven Fuzzy Demand Forecasting Models for Resilient Supply Chains. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 01-08). IEEE.

Thavaneswaran, A., Paseka, A., & Frank, J. (2020). Generalized value at risk forecasting. Communications in Statistics-Theory and Methods, 49(20), 4988-499

Vallera, A. M., Nunes, P. M., & Brito, M. C. (2021). Why we need battery swapping technology. *Energy Policy*, *157*, 112481.

Wang, X. H., & Cong, R. G. (2012). Electric power supply chain management addressing climate change. *Procedia Engineering*, *29*, 749-753

Wood, S. N., Goude, Y., & Shaw, S. (2015). Generalized additive models for large datasets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *64*(1), 139-155.

Yavari, M., & Zaker, H. (2020). Designing a resilient-green closed loop supply chain network for perishable products by considering disruption in both supply chain and power networks. *Computers & Chemical Engineering*, *134*, 106680.

Yong, B., Xu, Z., Shen, J., Chen, H., Tian, Y., & Zhou, Q. (2017, January). Neural network model with Monte Carlo algorithm for electricity demand forecasting in Queensland. In *Proceedings of the Australasian Computer*

Science Week Multiconference (pp. 1-7).

Zavala-Alcívar, A., Verdecho, M. J., & Alfaro-Saiz, J. J. (2020). A conceptual framework to manage resilience and increase sustainability in the supply chain. *Sustainability*, *12*(16), 6300.

Zhang, L., Zhou, S., An, J., & Kang, Q. (2019). Demand-side management optimization in electric vehicles battery swapping service. *IEEE Access*, *7*, 95224-95232.

Zhu, Z. (2020). *Dynamic data science applications in finance* (Master's thesis).

List of publications and conference presentations:

Journal Publications:

Appadoo, S. S., Makhan, M., & Kumar, A. (2021). Commentary on "Utilizing Linguistic Picture Fuzzy Aggregation Operators for Multiple-Attribute Decision-Making Problems". *International Journal of Fuzzy Systems, 23*(6), 1914-1917.

Appadoo, S. S., Makhan, M., & Kumar, A. (2022). A note on "Dealer using a new trapezoidal cubic hesitant fuzzy TOPSIS method and application to group decision-making program". *Soft Computing*, 1-3.

Appadoo, S. S., Makhan, M., & Kumar, A. (2022). A note on "Pythagorean uncertain linguistic hesitant fuzzy weighted averaging operator and its application in financial group decision making". *Soft Computing, 26*(17), 8841-8843.

Conference proceedings and presentations:

Makhan, M., Thavaneswaran, A., Appadoo, S. S., & Gajpal, Y. (2022, June 3-7). "Application of novel forecasting methods for Ontario electricity demand" *Administrative Sciences Association of Canada (ASAC)*.

Bowala ,S., Makhan, M., Liang, Y., Thavaneswaran, A., & Appadoo, S. S. (2022, September 18-20). "Superiority of the Neural Network Dynamic Regression Models for Ontario Electricity Demand Forecasting*" Canadian Conference on Electrical and Computer Engineering (CCECE).*

Appadoo, S. S., Gajpal, Y., & Makhan, M. (2022, November 19-21). "Mellin Transform and Possibilistic Assisted MCDM Model with Application in Health Care Management." *Decision sciences institute 53rd annual*

CONFERENCE

Appadoo, S. S., Gajpal, Y., & Makhan, M. (2022, August 11-14). "A Covid-19 Vaccines Selection Model in the presence of Arbitrariness, Imprecision, and Randomness In the Model Parameters." *XVI International Conference on Applied Research in Engineering and Management Sciences*

Bowala ,S., Makhan, M., Liang, Y., Thavaneswaran, A., & Appadoo, S. S. (2022, December 4 to 7). "Daily and Hourly Fuzzy Forecasts of Electricity Demand Using Neural Network Dynamic Regression Models" *Symposium Series on Computational Intelligence (SSCI)*. **(Accepted for poster presentation)**

Appadoo, S. S. Makhan ,M., & Gajpal, Y. (2021, April 30 – May 5). Covid-19 Vaccine distribution optimization. Production and Operations Management (POMS) Conference **(Presented Online)**