Peak Electricity Demand Density Forecasting with Macroeconomic Data

By:

Wei Xin, Tan

Supervisor:

Dr. Kaveh Rajab Khalilpour

Year:

2019

This thesis is presented in partial fulfillment of the requirements for the degree of Master of Data Science at Monash University, Faculty of Information Technology.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signed

31/5/2019

Date

Faculty of Information Technology
Monash University
Caulfield, Australia

Acknowledgement

I would like to thank Dr. Kaveh Rajab Khalilpour for encouraging and guiding me throughout the thesis writing process. I also thank my family for the unparalleled love and support.

Abstract

Energy forecasting is widely used in both commercial and non-commercial setting. In today's rising awareness of climate change, it has becoming more relevant. In recent year, many researches are focusing on short-term energy forecasting due to the availability of higher resolution data. However, long-term energy forecasting remains an important aspect, government entities are using it to evaluate environmental impact, and energy companies are using it for infrastructure planning.

Peak load density forecasting will be the focus of this thesis. Peak load density forecast can be produced by modeling the short-term effect and the long-term effect. This thesis aims to improve the long-term effect with additional macroeconomic indicators. In order to accommodate for more variables and to adopt a more holistic approach, stepwise feature selection and out-of-sample model selection will be used.

The case study in this thesis is the peak residential electricity load density forecasting in South Australia, and the benchmark methodology for this thesis will be Monash Electricity Forecasting Model (MEFM). Although the empirical results in the finding section showed that the proposed methodology for the long-term effect can indeed improve the peak electric demand density forecasting, but there are still some issue to be resolved to make the methodology more robust.

Contents

1 Introduction **1**

1.1 Brief Introduction . . . 1

1.2 Aim and Scope . . . 2

1.3 Research Goal . . . 2

1.4 Research Statement . . . 3

2 Literature Review **3**

2.1 Definition . . . 3

2.1.1 Energy Forecasting . . . 3

2.1.2 Forecasting Horizon and Coverage . . . 4

2.1.3 Microeconomic Indicator . . . 4

2.1.4 Macroeconomic Indicators . . . 5

2.1.5 Temperature Effect . . . 5

2.1.6 Calendar Effect . . . 6

2.1.7 Probabilistic Forecast . . . 7

2.1.8 Hierarchical Probabilistic Forecasting . . . 8

2.1.9 Point to Probabilistic Estimate with Scenarios Generation . . . 8

2.1.10 Temperature Scenario Generation . . . 9

2.1.11 Residual Simulation . . . 9

2.1.12 Evaluation Metric for Probabilistic Forecasting . . . 10

2.2 Forecasting Models . . . 10

2.2.1 Time Series Models . . . 10

2.2.2 Regression Models . . . 11

2.3 Energy Forecasting Landscape . . . 12

2.3.1 Historical View of Energy Forecasting . . . 12

2.3.2 Short-Term versus Long-Term . . . 13

2.3.3 Industrial Practice . . . 14

2.3.4 Motivation for Holistic Approach . . . 15

3 Research Method **16**
3.1 Overview 16
3.2 Short-Term Effect 18
3.2.1 Data Description 18
3.2.2 Data Source 19
3.2.3 Modeling 19
3.2.4 Forecasting 20
3.3 Long-Term Effect 21
3.3.1 Data Description 21
3.3.2 Data Source 22
3.3.3 Data Pre-processing 22
3.3.4 Modelling 23
3.3.5 Forecasting 25
3.4 Peak Demand Density 26

4 Findings **26**
4.1 Long-Term Effect 26
4.1.1 Feature Selection 26
4.1.2 Model Selection 27
4.2 Peak Load Demand Density Forecasting 29
4.3 Peak Load Demand Point Forecasting 32

5 Discussion **33**
5.1 Feature Selection 33
5.2 Model Selection 33
5.3 Possible Application 34
5.4 Implementation Code 34

6 Limitation **34**

7 Conclusion **35**

References **37**

List of Figures

1 Non-Linear Relationship between Residential Load Demand and Temperature 6
2 Process Flow for Residential Consumption Forecasting in AEMO, image extracted from AEMO (2019). 15
3 Overall Methodology 17
4 Overall Modelling Process for Long-Term Effect 25
5 Forecast and Fit for Base Model and Proposed Model 29
6 Weekly Density Plot of Weekly Peak Load Demand 30
7 Annual Density Plot of Weekly Peak Load Demand 31

List of Tables

1 Table of the Data Sources 22
2 Table of the Top 10 Most Appeared Macroeconomic Indicators 27
3 Table Summary of Coefficients for the Final Model 28
4 Pinball Loss for Base Model and Proposed Model 32
5 MAE for Base Model and Proposed Model 32

1. Introduction

1.1. Brief Introduction

Energy Forecasting is used by different entities for different purposes, these entities include government bodies, international organizations, business corporations, and researchers. Business corporations used energy forecasting for their operation needs and business planning. Other entities use it to forecast future environmental impact, economic impact, climate change and et cetera. In the mix of rising awareness of climate change, data-driven business models and big data era, energy forecasting has becoming more and more relevant.

The use case of energy forecasting can basically split into two contexts: business and non-business. In the business context, energy companies use short-term forecasting to fulfil their operational needs, and long-term forecasting for reliability planning, rate planning, infrastructure planning and more (T. Hong & Fan, 2016). Other businesses that are heavily rely on energy consumption might do energy forecasting as part of their business plan as well. Therefore, energy forecasting remains an important element in the business world.

Energy forecasting is not only used by energy companies, non-business entities like policy makers also rely on them to evaluate environmental and economic impact. In recent Intergovernmental Panel on Climate Change special report (IPCC, 2018), it stated the global warming might reach 1.5C between year 2030 and 2052 with the current pace. The report warned that there will be irreversible damages and risks to our ecosystem if the 1.5C mark is passed (IPCC, 2018). Climate change is mainly caused by the greenhouse gasses in the atmosphere, and most of these greenhouse gasses are coming from the energy industry. Energy industry emitted 72 percent of the greenhouse gases globally in year 2013 (C2ES, 2018). Furthermore, the global energy demand will continue to growth by 50 percent from 2017 to 2040 (IEA, 2018). A better energy forecasting will not directly alleviate these problems, but it can provide a better tools for organizations to gauge the progress and impact, thus allow them to make better decisions and planning.

1.2. Aim and Scope

As mentioned earlier, energy forecasting remains an important tool for many entities, and the aim of this thesis is to improve energy forecasting. Because energy forecasting has many components in it, the literature review on next section will help us explore the research gaps, and subsequently frame the research goal. This thesis is an applied research rather than a fundamental research, so there will be no attempt to improve or reinvent the underlying modelling theory.

1.3. Research Goal

Below is the list of key ideas that shaped our research goal, and they will be discussed more in-depth in the literature review.

- Electric production is the biggest greenhouse gas emitter in the energy industry.
- Although macroeconomic indicators such as price, GSP, population have been used in many forecasting models, it lacks discussion on why and how they are selected.
- Many recent researches in energy forecasting are focusing on short-term forecasting, and it is understandable due to the availability of higher resolution data.
- Maximum and minimum demand forecasting is one of the crucial energy forecasting for the energy companies.
- In real-world, energy forecasting is more complicated, and involves several models to produce forecasts. Therefore, we should try to understand the entire framework instead of focusing on a particular model.
- In the rise of big data era, adopting holistic approach could improve the forecasting accuracy.
- It is important to include the prediction interval or probabilistic density for a better understanding of the forecasts and its associated uncertainty.

For abovementioned reasons, the goal for this research is to improve energy forecasting in the area of probabilistic peak electricity demand with a more holistic approach towards macroeconomic indicators. More specifically, this thesis will propose a framework to improve peak electricity demand forecasting by including more macroeconomic data, and hence the title ‘Peak electricity demand density forecasting with macroeconomic data’.

1.4. Research Statement

The research statement/question would be ”Does using more macroeconomic indicators improve peak electricity demand density forecast?”. Hence, the thesis will rigorously testify the research statement, and determine whether the proposed framework and data can actually produce a better forecasting result.

2. Literature Review

The literature review will be split into three parts:

- The first part and second part will be discussing on the concepts and forecasting models respectively, which are prerequisite for understanding the energy forecasting research. Therefore, more advanced readers could overlook those that they are familiar with.
- The Third part will be discussing on the past and present of relevant researches and industrial practices in energy forecasting, while at the same time exploring the limitations and research gaps.

2.1. Definition

2.1.1. Energy Forecasting

Energy forecasting is a broad term. There are different types of energy, and also different aspects of the energy that can be forecasted. Types of energy includes solar energy, wind power, electricity, fossil fuels, and the aspects of energy can be load, peak, consumption, production, price and et cetera. A few popular subdomains of energy forecasting are load forecasting, electric price forecasting, wind power forecasting and solar power forecasting.

2.1.2. Forecasting Horizon and Coverage

In energy forecasting, there are different forecasting horizons: short-term, medium-term and long-term. Medium-term is between two weeks and three years, less than that would be short-term and beyond that would be long-term (T. Hong et al., 2016). It is important for forecaster to identify this in the early stage before building the model, because it will shape how the data is gather and processed. For example, building short-term load forecast requires higher resolution data such as half – hourly load demand. However, some energy forecasts such as peak density load forecast, utilizes both short- and long-term data to build the model.

Coverage is another aspect that forecaster should consider beforehand. Similar to the forecasting horizon, it would shape the data processing and model selection. The coverage can be ranging from a single residential household to the entire nation. For example, annual industrial load forecast for South Australia is a medium-term forecast that covers only the industrial electric consumption in South Australia. Moreover, residential and industrial forecasts are typically modelled separately due to their dissimilarity in relationship with the explanatory variables, and also the explanatory variables used.

2.1.3. Microeconomic Indicator

Microeconomic indicators concern with the local outlook of a specific company or industry. Some of the microeconomic indicators are supply and demand, market competition, elasticity, opportunity cost, technological improvement and tax rate. However, these indicators vary across different industry, different industry will have different indicators. In the context of energy forecasting, the microeconomic indicators can be viewed as predictors variables for the short-term forecast. It is usually temperature data such as hourly temperatures, cooling degree day and heating degree days. Calendar effects such as workday, weekends and holidays are often included by practitioner for short-term forecasting as well.

2.1.4. Macroeconomic Indicators

Macroeconomic indicators are those indicators that give us an overview of the economic. For instance, gross domestic product (GDP), unemployment rate, interest rate, consumer price index (CPI), population and et cetera. However, in order to aptly adding macroeconomic indicators into to the forecasting model, some knowledges on the composition, calculation and assumption of the macroeconomic indicators are essential. Nevertheless, it is known that macroeconomic indicators like population, gross state product could improve the medium to long-term electric load forecast. For example, in order to forecast 2 years ahead for the load demand in the Victoria state in Australia, gross state product, population of Victoria state, and price of utilities could be incorporated in the forecasting model.

Apart from using the macroeconomic indicators directly, practitioners could explore the ideas of further extracting information from these macroeconomic indicators using methods such as cross-impact analysis, computable general equilibrium (CGE), index decomposition analysis to generate scenarios, then use it as an input features for the forecasting model. This modelling method is commonly seen in other data analysis tasks, for example, in text classification, topic modelling is used to generate the posteriors for each topic, and the posteriors will be used as new or extra input feature for a strong classifier like Support Vector Machine or Gradient Boosting Tree. Nonetheless, as more features are adding to the forecasting model, feature selection technics such as subset selection, regularization, dimension reduction should be considered to prevent overfitting.

2.1.5. Temperature Effect

In residential electric load forecast, temperature data is an important variable. Huge amount of residential electric consumption is used for heating and cooling. Furthermore, heating and cooling require relatively huge amount of electricity, and these activities are hugely influenced by the weather. Therefore, temperature data are used extensively in the residential load forecasting, it can be seen in all of the Global Energy Forecasting Competition (GEFCom), Australian Energy Market Operator (AEMO) methodology paper, and so on. However, the relationship between temperature and load consumption is non-linear as shown in Figure 1 below, because of low temperature leads to heating, and high temperature leads to cooling. In order to better model the temperature effect, different representations such as lagged temperatures, cooling degree days, heating degree days, temperatures from multiple locations, and et cetera are used.

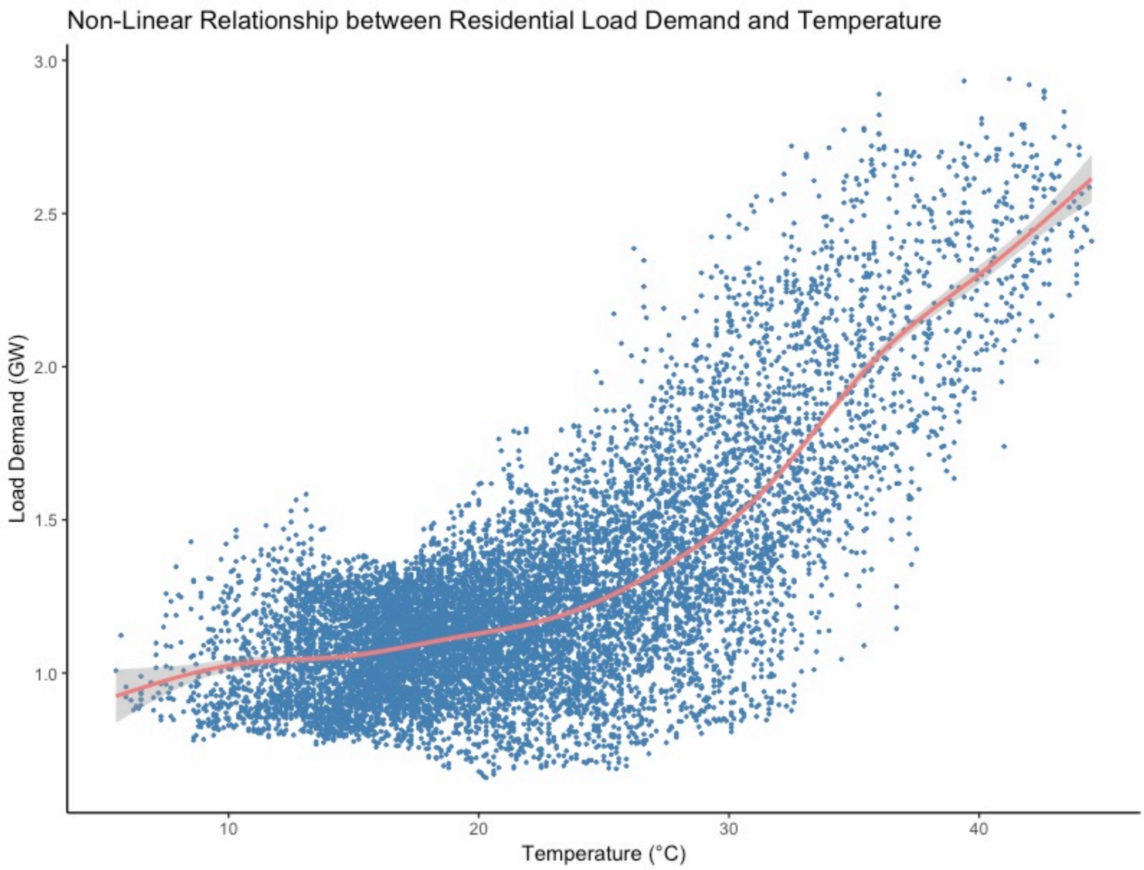

Figure 1: *Non-Linear Relationship between Residential Load Demand and Temperature*

2.1.6. Calendar Effect

In energy forecast, it is important to take into consideration of calendar effect, as it has substantially impact on the accuracy (Florian, 2018). Particularly holiday effect, most electric forecaster will incorporate it into their model, especially short- to medium- term forecasting. Apart from the holiday effect, other calendar effect like weekend effect, seasonal effect and intra-day effect are being used to improve the forecasting model (Lusis, Khalilpour, Andrew, & Liebman, 2017). Weekend effect can be observed in the industry and commercial sectors due to lower electric consumption during the weekend (Lusis et al., 2017). As for residential, intra-day effect is peak during evening due to people returning home from work (Lusis et al., 2017). However, the variance and correlation of these calendar effect variables must be evaluated beforehand to prevent redundant features to the underlying forecasting model, and causing optimization issue. Moreover, in Lusis et al. (2017) paper, their experiment demonstrated the combination of calendar effect, load forecast granularity, forecasting model will have different effect on the accuracy for day-ahead household load forecast, therefore, practitioners need to pay more attention when applying these calendar effects.

2.1.7. Probabilistic Forecast

According to T. Hong et al. (2016), the greatest uncertainties that energy forecasters will face in future are climate variability, electrics vehicles, wind and solar power generation, energy efficiency and demand response. Furthermore, the uncertainty of energy forecasting becomes greater when the coverage is smaller. For example, the uncertainty of predicting the energy consumption of a particular housing area is greater than predicting the nation energy consumption. It is similar to predicting stock market, the uncertainty of predicting the stock indices is smaller than predicting an individual stock price.

When there is huge uncertainty involved, having probabilistic forecasting makes more sense, even though is not as satisfying as having a point estimate. The reason is when the future outcome is uncertain, it better to provide a range of values, instead of one single value. Therefore, in the ever-increasing dynamicity in energy industry, being able to capture the uncertainty is a plus for many decision makers. In energy production, the supply is almost never produced at the equilibrium, for that reason, there is a need for energy companies to know the probabilistic outcome, so that they can supply energy at a low power outage rate.

The key characteristic of probabilistic models is the prediction produced by the model is a prediction interval instead of a single value like point estimate model. Not to confuse with confidence interval, prediction interval is referring to the prediction, and confidence interval is referring to the indicators (T. Hong & Fan, 2016). However, in real world, the outcome is often a single point, therefore, one of the goals for probabilistic forecasting model is to increase the sharpness of the prediction interval while subject to calibration (Gneiting, Balabdaoui, & Raftery, 2007).

2.1.8. Hierarchical Probabilistic Forecasting

Hierarchical forecasting aggregates or disaggregates data to produce forecasting at different level. For instance, in order to generate forecasts for nation GDP, we first generate forecasts for the lower level components, and subsequently aggregate them. There are different ways of producing the point forecasts from different hierarchy: bottom-up approach, top-down approach, middle-out approach (Athanasopoulos et al., 2019). It can also be extended to produce probabilistic forecasts, depending on whether it is parametric or non-parametric, it can be derived analytically or bootstrapped from the errors (Athanasopoulos et al., 2019).

In the energy forecasting competition, the latest GEFcom2017 is competed on hierarchical probabilistic load forecasting, it is deemed to be their most challenging competition to date (T. Hong, Xie, & Black, 2019). The winning method was combining forecasts from quantile regression and generalized additive model (T. Hong et al., 2019).

2.1.9. Point to Probabilistic Estimate with Scenarios Generation

Another common approach to generate prediction interval is to use scenarios generation. The idea is drawing multiple temperature samples from past data through bootstrapping, and use these samples to generate temperature scenarios according to the quantiles, then pass each quantile to a point forecasting model to get the prediction interval. For example, team ADADA who ranked top 5 in GEFCom2014 in load forecasting, used temperature stimulation model to generate a range of temperature scenarios, and feed them into a generalized additive model to derive a sequence of quantiles (Dordonnat, Pichavant, & Pierrot, 2016).

This method allowed many statistical models such as Multi-linear Regression models, Semi-parametric Additive models, Artificial Neural Network, Fuzzy Regression models, Support Vector Machines, and Gradient Boosting Machines to be able extend to probabilistic models (T. Hong & Fan, 2016). As a result, it opens up many statistical learning algorithms for researchers and practitioners to explore and develop for probabilistic forecasting. However, this approach raised an interesting research question: can improving the point estimate in turn improves the probabilistic estimate (T. Hong & Fan, 2016). Nevertheless, the prerequisite of this method is being able to generate appropriate scenarios at the first place.

2.1.10. Temperature Scenario Generation

As discussed in the earlier section, probabilistic forecast can be simulated through scenario generation. According to Xie and Hong (2018), there are main four methods from probabilistic load forecasting literature for generating temperature scenarios: fixed date, shifted-date, bootstrap and surrogate. For fixed day method, if 5 years is chosen as the sample size, the temperatures will be generated by drawing samples that are on the same day for the past 5 years. For example, if we want to generate one week ahead temperatures for 1st November to 7th November, we randomly draw temperatures from the 5 years data that are on 1st November to 7th November. Whereas, for shifted day method, it aims to improve upon the fixed day method by taking into consideration of the days for each year might not be the same, so there is need to shift the days to match its similarity. As for bootstrap method, a year with 365 days is divided by the number of bins, and samples are drawn from each equal size bin. There are also more advance bootstrapping methods such as double seasonal block bootstrap, which used by Hyndman and Fan (2010). The final method which is surrogate method simulates a new temperatures series through Fourier transformation. However, the first three methods are more popular in the energy industry. Additionally, according to Xie and Hong (2018), the ideal sample size is 10 years with 16 shifted days which yielded the lowest quantile score.

2.1.11. Residual Simulation

Residual simulation is another way to capture the uncertainty and improve the forecasting model. This method was used in GEFCom2014, and it was one of the winning solutions (Xie & Hong, 2016). Furthermore, in Hyndman and Fan (2010) paper, they demonstrated how simulated residuals can be bootstrapped, and subsequently incorporating it into the forecasting model to improve prediction. However, bootstrapping method assumes the residuals are normally distributed, and in Xie, Hong, Laing, and Kang (2017) paper, it testified that simulating residuals with normality assumption indeed could improve the accuracy.

2.1.12. Evaluation Metric for Probabilistic Forecasting

As mentioned before, the sharpness of a probabilistic forecasting model shows how good the model can predict, and we need a way to quantify it. Pinball loss function and Winkler score are considered to be the more comprehensive evaluation metric for probabilistic forecasting, they are able to capture the sharpness of the probabilistic model. However, pinball loss function is more widely use among researchers, it is even used as the scoring benchmark for GEFCom2014 and GEFCom2017 (T. Hong & Fan, 2016; T. Hong et al., 2019).

2.2. Forecasting Models

2.2.1. Time Series Models

Time series models are widely used for forecasting in numerous fields such as econometrics, quantitative finance, meteorology and more. Exponential Smoothing, Autoregressive Integrated Moving Average (ARIMA) are the two popular time series models. These models extrapolate information like trend, level, seasonality, autocorrelation, error, from the past observations to generate point forecasts and prediction interval. There are a several ways for time series models to generate prediction interval: bootstrapped residuals, scenario generation or mathematically derived the forecast standard deviation (Hyndman & Athanasopoulos, 2018). Thus, these time series models only require the data of the dependent variable, and user does not need to gather or forecast any other predictor variable data. Nonetheless, since these models do not use any predictor variable, user would not be able to infer the causality or correlation of the data, which makes it less useful in some cases. Moreover, temperature-based models like regression models can usually outperform time series models (D. T. Hong, 2018).

2.2.2. Regression Models

A few regression models that are commonly used in energy forecasting will be discussed in this section. Firstly, linear regression is perhaps the most ubiquitous model in the modeling world, hence, there is no introduction required. Nevertheless, linear regression remains a powerful tool in energy forecasting. For example, the external driving factors in load forecasting is often modeled using linear regression.

Quantile regression was first introduced by Koenker and Bassett in 1987, and it is often introduced by comparing it with linear regression. Nevertheless, quantile regression can be viewed as an extension of linear regression, because it was invented around the limitation of linear regression which is unable to extend the prediction of the response variable beyond the central location (Koenker, 2015). The central tendency that these two models are trying to measure is different. Instead of optimizing the parameters to estimate the mean of the response variable, quantile regression optimizes the parameters to estimate the median and the quantiles of the response variable. It is argued that estimating mean alone does not provide a complete picture of a single distribution, hence, quantile regression is known to be more robust because it provides a more comprehensive picture of the distribution (Koenker, 2015). It is robust also because of its insensitivity to outliers. The way quantile regression optimizes the parameters is also different from linear regression. Linear regression uses least squares method, and it is often solved using closed form solution or gradient descent. Whereas, for quantile regression, the parameters are optimized through linear programming, in particularly, simplex or barrier method (Koenker, 2015). Despite being more robust, quantile regression has some disadvantages: it required more data and it is computationally more expensive (Rodriguez & Yao, 2017). Although quantile regression is not as popular as linear regression, but in eclectic load forecasting, it has been the winning method for two of the latest GEFComs (T. Hong et al., 2019; Xie & Hong, 2016). In nutshell, quantile regression it is a promising candidate for energy forecasting.

Spline regression is a non-linear method, it is an extension of steps functions and polynomial regression (James, Witten, Hastie, & Tibshirani, 2013). Furthermore, spline regression can be extended to smoothing spline, by optimizing error function with smoothing constraints. Lastly, generalized additive models (GAM) allows us to formulate an additive model using predictors with different relationships mentioned above: linear and non-linear (James et al., 2013). Non-linear models are crucial for modelling the residential load with temperature data because of their non-linear relationship as mentioned in the earlier section.

Dynamic regression combines both regression model and time-series model. In regression models, the residuals are assumed to be normally distributed with zero mean, in dynamic regression, the residuals will be modeled using ARIMA. Dynamic regression is used by Hyndman and Fan (2015) to model the short-term effect of the peak load density forecasting, the combination used is spline regression, and autoregressive of first order for the residuals.

2.3. Energy Forecasting Landscape

2.3.1. Historical View of Energy Forecasting

The earliest energy forecasting method is just counting the number of lightbulbs, because lightbulb was the only use case for energy (T. Hong et al., 2014). After the industrial revolution, there are more demand for energy including air-conditioning, mobile computing devices, and et cetera. With more variables adding into the equation, energy forecasting becomes more complex and uncertain. Without high accuracy, forecasting would be pointless. Thus, forecasting methods has since developed from a simple trending method to statistical or AI-based model (T. Hong et al., 2014). The shift is made possible by the improvement of computing power coupled with huge data availability. However, the caveat in using these advance AI-based models such as Artificial Neural Network is interpretability, these models often refer as black-box algorithms. Although, these models can perform well in some cases, but it does not meet some of the business needs for interpretability (T. Hong et al., 2014). For instance, some of the decision plans such as rate cases, they are required to be published and archived for shareholders, regulatory commissions and other stakeholders to review, hence, the interpretability and transparency are mandatory (T. Hong et al., 2014). Another important milestone for energy forecasting is shifting from point estimate to probabilistic estimate, using probability theory to capture the uncertainty (T. Hong et al., 2016).

Despite years of research, there is no one best model for energy forecasting (Wang et al., 2018). Different energy forecasting faces different challenges. For electric price forecasting, programs such as demand response or rebate must be accounted by the model (T. Hong et al., 2014). For renewable energy forecasting such as solar and wind energy, they are highly correlated to the meteorology, hence, their forecast accuracy is hugely depended on the accuracy of the weather forecast (T. Hong et al., 2014). Despite the challenges, our improved understanding of the atmospheric physics and the availability of the data allowed us to better forecast the weather, and in turn, a better energy forecast. However, due to the chaotic nature of the dynamical system like our Earth's atmosphere, which sensitive to the initial condition and subject to butterfly effect, the further we forecast the lower the accuracy of the forecast.

Today, the adoption of smart grid give rise some of the challenges and advantages for energy forecasting. With smart grid, energy providers can better manage the demand and supply to reduce operational costs (Smartgrid.gov, 2018). Moreover, smart grid allowed many possibilities, for instant, integration of energy from different sources such as solar energy and wind energy, demand response program, smart meter, and et cetera (Smartgrid.gov, 2018). The smart meter in smart grid system produces stream data which can provide consumers and producers timelier information (Smartgrid.gov, 2018). In essence, this fundamental engineering change of energy infrastructure increases the forecasting complication, but at the same time, generated huge data for forecaster to create better forecasting models (T. Hong & Fan, 2016).

2.3.2. Short-Term versus Long-Term

Short-term forecasting has been the focus for the past decades, and it has gone through some critical breakthrough which allowed by technology and the availability of higher resolution data (T. Hong et al., 2016). However, long-term forecasting has gained some tractions due to aging infrastructure and design, therefore, energy companies need a longer forecast for making long-term business plans (T. Hong et al., 2016). According to T. Hong, Wilson, and Xie (2014), long-term probabilistic load forecasting can be built using normalized low-resolution data and macroeconomic indicators.

2.3.3. Industrial Practice

In previous sections, we only discussed about the overview of some energy forecasting methods, this part we will look into how energy companies actually do their energy forecasting. Australian Energy Market Operator (AEMO), who is responsible for operating Australia's largest gas and electricity markets and power systems, on February 2019 published a paper detailing how they produce electric demand forecast. The paper discussed three major forecasts: business annual consumption, residential annual consumption and maximum and minimum demand. Research papers are usually scoped to focus on improving one particular area, whereas the paper from AEMO is more comprehensive, detailing all the processes involved in building the forecasts. This allowed us to get bigger picture of how energy forecasting works. For instance, below flow chart shows us how AEMO generates residential consumption forecast. Although, most of the relationships are modelled using linear regression with ordinary least square method, but AEMO has considered a very extensive list of factors that could affect the residential consumption like heating degree days, cooling degree days, gas-to-electricity fuel switching, solar PV rebound effect, and et cetera (AEMO, 2019).

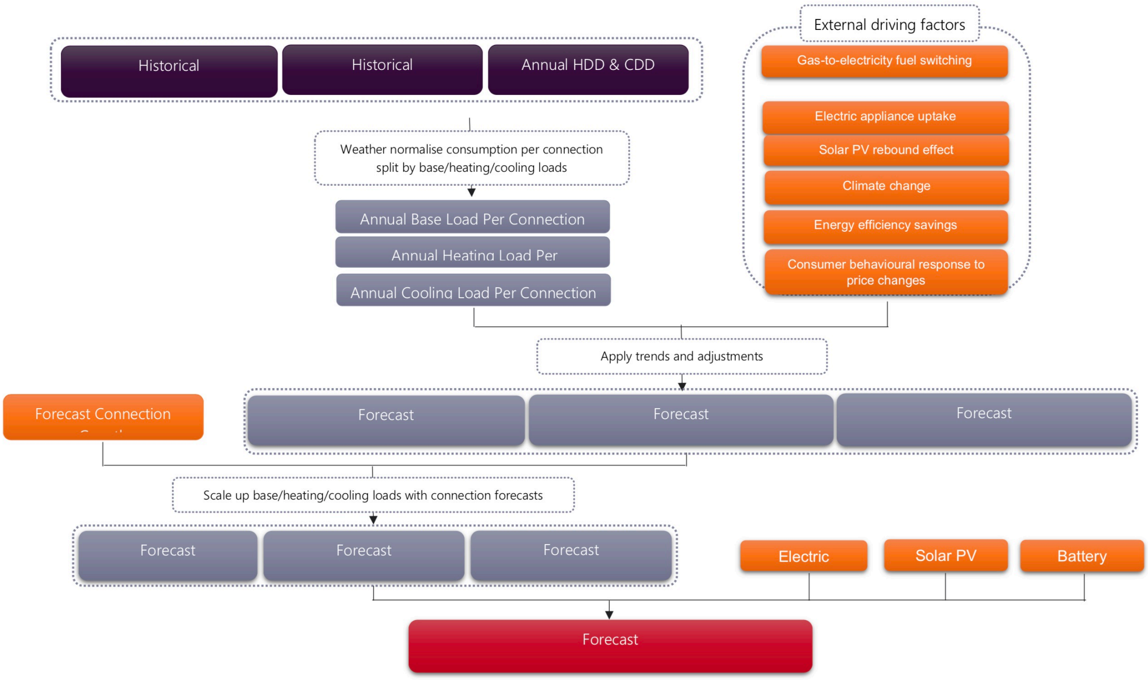

Figure 2: *Process Flow for Residential Consumption Forecasting in AEMO, image extracted from AEMO (2019).*

2.3.4. Motivation for Holistic Approach

In the previous section, AEMO's sophisticated forecasting framework showcased their well understanding of the energy system, and is able to identify some contributing factors for consumptions, because some of the factors are not obvious, especially for the uninitiated. However, in short-run, this framework would work, but in long-run the contributing factors might change, and they will have to revise their forecasting framework again like they did in the past. Hence, instead of always instilling our domain knowledge for forecasting, I propose we remove some of the heuristic part of the framework and replace it with a more holistic approach. In AEMO case, it would be the 'external driving factors' part.

It might seem to be a counterintuitive way of modelling, but the advance in computing power allowed for a more computational search method. Moreover, when we embed our human knowledge into a system, it sometimes inhibits the system to learn a certain pattern or relationship, and this is illustrated in many AI research problems (Sutton, 2019). For example, Deepmind's AlphaZero, an AI system that plays go and chess game, surpassed its predecessor that train on matches played by professional players, and it did it by just playing with itself (DeepMind, 2019).

These AI systems can sometimes teach us new or unconventional ways of playing the games. By applying these algorithms to real-world problems, we are in hope to discover new knowledge like we did in these games. However, these AI research problems usually have access to all the data and variables of their environment, and this cannot be said to many real-world problems. Hence, I am by no means suggest that holistic approach is the solution for all problems, at least for now. But it is still worthy of a try and hope there is something new to be discovered as there are more and more data being collected each day.

3. Research Method

3.1. Overview

The methodology for this thesis is heavily based on Monash Electricity Forecasting Model (MEFM) (Hyndman & Fan, 2010, 2015). One of the reason that MEFM is adopted in this thesis is their comprehensiveness in forecast peak demand density, and each of the details required for generating the forecast are well-documented. Besides, it is one of the well formulated and state-of-the-art methodology for generating peak density load forecast. Moreover, it is also adopted by AEMO. So when the term 'base model' is mentioned in this thesis, it will be referring to MEFM. In MEFM framework, two effects will be modelled separately: long-term effect and short-term effect. The long-term effect includes the annual or seasonal effect that is driven my macroeconomic parameter such as population and GSP. On the other hand, the short-term effect is the half-hourly normalized load demand that is driven by temperature and calendar effect. The relationship of the long-term and short-term effects are defined as the formula below:

$$\begin{aligned} y_{t,p} &= y^*_{t,p} \times \bar{y_i} \\ log(y_{t,p}) &= log(y^*_{t,p}) + log(\bar{y_i}) \end{aligned} \tag{1}$$

- $y_{t,p}$ is the half-hourly load demand for time t and period p
- y^* is the half-hourly normalized load demand (half-hourly demand/annual demand)
- $\bar{y}_i$ is the annual average load demand where period t falls

The proposed changes is only on the long-term effect. There will be more macroeconomic data added to model the long-term effect. Therefore, each of the data description and modeling steps will discussed separately for long-term effect and short-term effect. The overall methodology is summarized in Figure 3, which consists four different stages, and each of the stages will be discussed in the following sections.

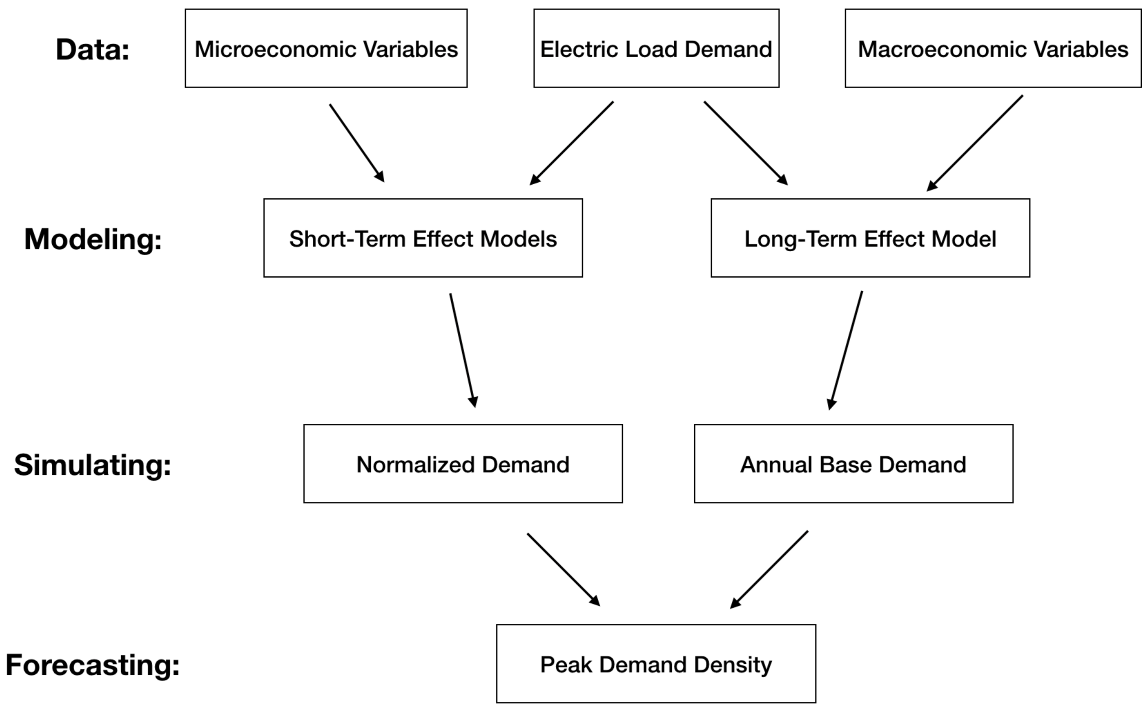

Figure 3: *Overall Methodology*

The load demand is this thesis will be referring to electric load, and only the residential load will be modelled. In practise, industrial load will be modelled separately because its driving factors are different from residential load.

3.2. Short-Term Effect

3.2.1. Data Description

The target variable for the short-term effect is half-hourly electric load demand. For each day, it will have 48 data points. While for each 'year', it only considers the summer period which is from October until next year March (6 months or approximately 183 days). The reason is the peak load demands usually occurs in the summer. The total time spam of the data is 15 years from 2000 to 2014. Nonetheless, 14 years of the data will be used for training the model, and 1 year for performance evaluation. In total there are 124,848 data points, and it covers only the electric load demand in South Australia. In addition, the industrial portion has already been excluded from the electric load demand.

The primarily predictor variables for modelling the short-term effect will be temperature data. In this case, temperature data from two locations will be used. However, in order to model their various relationships with load demand, a few more variables are extrapolated from the temperature data. All the temperature variables that are going to be used in the modelling stage are listed below:

- Current average half-hourly temperature between two locations (1 variables)
- Current temperature difference between two locations (1 variables)
- Previous 4 half-hourly temperatures (4 variables)
- Previous 3 days temperatures of the same period (3 variables)
- Previous day temperature difference between two locations of the same period (1 variables)
- Previous 3 half-hourly temperature difference between two locations (3 variables)
- Average and minimum temperatures for the past 24 hours (2 variables)

Lastly, calendar variables, specifically holiday indicator, time of the year (continuous variable of 0 to 182) and day of the week, will be used to capture the calendar effect. Overall, there are 17 predictor variables (3 Calendar effect, 14 temperature effect) will be used to build the short-term effect model.

3.2.2. Data Source

The half-hourly load demand data can be acquired form the AEMO' official website, however, the load demand values are aggregated by states, and there is no segregation between industrial load and residential load. For the temperature data, it can be requested or download from Bureau of Meteorology (BOM). There is also a R package called 'bomrang', which is an API for extracting the weather data from BOM by specifying the weather station id. However, these data are from Australian organizations, so the coverage is only within Australia.

3.2.3. Modeling

For short-term forecasting, the allegedly better model such as quantile regression can be used to model the short-term effect, but in this case, it is not the objective of this research, so the modelling process will remain unchanged as MEFM in order to control the experiment environment. In order to account for intra-day calendar effect, each half hourly period will be modelled separately, so there will be 48 models in total to capture the short-term effect. Moreover, the relationship between temperature and load demand is non-linear, thus natural spline is used to model their relationship. While the calendar effects are modelled using linear relationship. The 48 models will be generalized additive models that adding both temperature effects and calendars effects together for each half hourly periods. As a result of creating new temperature variables like previous hour or day temperature, there are empty values in some rows simply because there is no previous temperature, and these rows will be ignored during train time. Before optimizing the models, the half - hourly load demands will be log transformed and normalized against annual average load demand. Below is the equation for the 48 short-term effect models.

$$log(y^*_{t,p}) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + e_t \tag{2}$$

- $h_p(t)$ is the linear and additive function of calendar effect
- $f_p(w_{1,t}, w_{2,t})$ is the temperature effect of two locations that will be modelled in natural spline function

- e_t is the residuals to be simulated

3.2.4. Forecasting

In order to generate the forecasts and prediction interval for the normalized half-hourly load demand, the temperature data and residuals will be simulated. The temperature variables are assumed to be the only driver of uncertainty, while the calendar variables are known beforehand.

For the reason of preserving time series seasonality, trend and autocorrelation, the seasonal block bootstrap method will be used (Hyndman & Fan, 2015). Due to two types of seasonality: annually and daily, double season block is used instead of single season block. Moreover, in order to counter for unrealistic large spikes in temperature, variable block length will be implemented (Hyndman & Fan, 2015). With this bootstrapping method, temperature data generated are more akin to the real world weather patterns such as heatwave or consecutive cooling days in the summer.

As for simulating the residuals, first the medians of the half-hourly residuals in 35 block days will be extracted from the fitted models. Then the medians will be modelled using autoregressive of first order AR(1), and subsequently the AR model will be used to simulate future values with normal errors $\mathcal{N} \sim$ (0,1) (Hyndman & Fan, 2015). Finally, these simulated medians will be added into the bootstrapped residuals which are generated using the centered fitted values from the medians and the season block bootstrapping. The reason for this process is due to is the residuals of the fitted half-hourly models depicted some serial correlations during the residual analysis.

In summary, the temperatures are simulated from double season block bootstrapping with block days of 20 and variability of 5 days. Furthermore, for each of these blocks, they will be drawn from different year. Subsequently, these simulated temperatures will be fed into the 48 short-term models trained earlier to predict the normalized load demand values. Finally, the residuals will be added as shown in equation 2 above. This process will repeat for 1,000 times in order for generating the probabilistic distribution later. Thus, in total there are 8,784,000 ($1,000 \times 183$ $(days\ per\ year) \times 48$ $(periods\ per\ day)$) of half-hourly normalized load demands that are being simulated.

3.3. Long-Term Effect

3.3.1. Data Description

For the proposed framework to model the long-term effect on annual load demand, below variables will be added.

1. 117 Australian Macroeconomic data are collected Australian Bureau of Statistics (ABS) and the Reserve Bank of Australia (RBA) and compiled for /cite study.

2. Five variables for measuring technological progress:

   - Supercomputer Power (FLOPS)
   - Microprocessor Clock Speed (Hertz)
   - Transistors Per Microprocessor (transistors per chip)
   - Computing Efficiency (Watts per MIPS)

   However, some of these data have missing data between years, so it was imputed using Kalman Smoothing on structural time series models.

3. Six other macroeconomic indicators includes :

   - Gross Domestic Product (GDP) of Australia, America, China and the World
   - Australian Economic Productivity
   - South Australia Internet Subscribers Count

In comparison, the macroeconomic indicators for the base model are gross state product chain volume estimate (millions of dollars), residential price index of electric consumption (cents per kWh) and cooling degree days with a threshold of 18.5 degrees Celsius, and these indicators are for South Australia.

As for the target value, it will be the average load demand of the year or season, and in this thesis it will be refer as ’annual base load demand’. It is calculated by averaging the half-hourly load demands by number of days in a year, and the ’year’ as mentioned above is 183 days.

The time spam for the macroeconomic data is also from 2000-2014, but since the frequency of the data is annually, there is only 15 data points in total for each indicators .

3.3.2. Data Source

The data come from four sources, and they are summarized in Table 1 below:

Source	**Macroeconomic Indicator - 133 in total**
Australian Macro Database (AMD)	117 Australian Macroeconomic Indicators
Australian Bureau of Statistics (ABS)	South Australia Internet Subscribers Count
MEFM R Packages	South Australia GSP South Australia Residential Electric Price Index South Australia Cooling Degree Days South Australia Population South Australia Total Price of Electric
Our World in Data	Australia GDP America GDP China GDP World GDP Australia Economic Productivity Supercomputer Power Microprocessor Clock Speed Transistors Per Microprocessor Computing Efficiency Microprocessor Clock Speed

Table 1: *Table of the Data Sources*

3.3.3. Data Pre-processing

Before going into modelling, there are some pre-processing steps required. The first step is dealing with the missing values by imputation or removal, which is a rudimentary step for data analysis. The second step is to check the correlations between all variables, and remove those high correlated data. The reasons are these variables are not giving us enough additional information, and it might also cause optimization problems such as multicollinearity. Hence, in this thesis, the correlation threshold is set to 0.9. The third step is to remove data that have near-zero variance which would not be useful for any data analysis task. Lastly, the data needed to be scaled and centered, this is a crucial steps especially for doing regression. The data will be split into 3 sets: train set, test set and validation set. The train and test set will be used in pre-processing, modelling and model selection. Whereas the validation set will be an unobserved data to any of the models, and it will solely be used for model evaluation in section 4. In this case, the train set has 11 data points, test set has 3 data points and validation set has 1 data point, and each data point represent one year.

3.3.4. Modelling

Due to the nature of forecasting, it is not a good idea to use models that generate forecast based on all the predictors variables. The reason is if we were to forecast ahead, all the predictors variables must be made available. If they are not, they must be forecasted, and ultimately increasing the uncertainty. Therefore, only models that have the properties of feature selection will be considered, for instance, lasso regression and linear regression with stepwise selection.

Although the idea is to be holistic, but to assess all the possible combinations 2^{133} is still beyond current mobile computing technology. Hence backward stepwise features selection based on Akaike Information Criterion (AIC) will be used. Furthermore, to compensate for not considering all the possible combinations with backward stepwise method, the features will be randomly shuffled before performing the backward stepwise selection. This process will be repeated 300 times to get a good amount of random combinations for the backward stepwise selection. Linear regression with least square method will be the underlying model that model the relationship between the annual base load demand and the predictors variables. At this stage, only the train set is used.

In order to take advantage of the ever-increasing computing power, these models will be trained on parallel. Since each of these models does not need any information from each other, they can be trained on parallel. However, the parallelism is done on CPU, thus the scalability is not as prominent as using GPU. Besides, the GPU might not be able to efficiently perform the serial optimization process of training a single linear regression model.

Since there are many variables, it is not difficult to find a combination that can fit the model perfectly, but it does not guarantee that the model will be able to forecast well. For this reason, the test set will be used for model selection. The 300 fitted models will be used to generate three years ahead forecasts (length of the test set), and its point estimate will be used for performance evaluation. The error function for evaluating the performance between the forecasts and the actual values is mean absolute error (MAE) instead of mean square error (RMSE). The square term of RMSE has the effect of inflating residuals that are bigger, but it would not be effective in this case, because the annual load measure is a relatively small number. The annual base load demand is around 1.3, and the difference between the forecasts and the actual is often between -1 and 1, hence applying square term to these numbers would actually make them smaller. The three years test set will supposedly help us choose a model that fit well and forecast well. After using MAE to choose the model, the model will be refitted with both train and test set. For the remaining discussion, the term 'proposed model' will be referring to the above modelling framework, and it also summarized as below Figure 4.

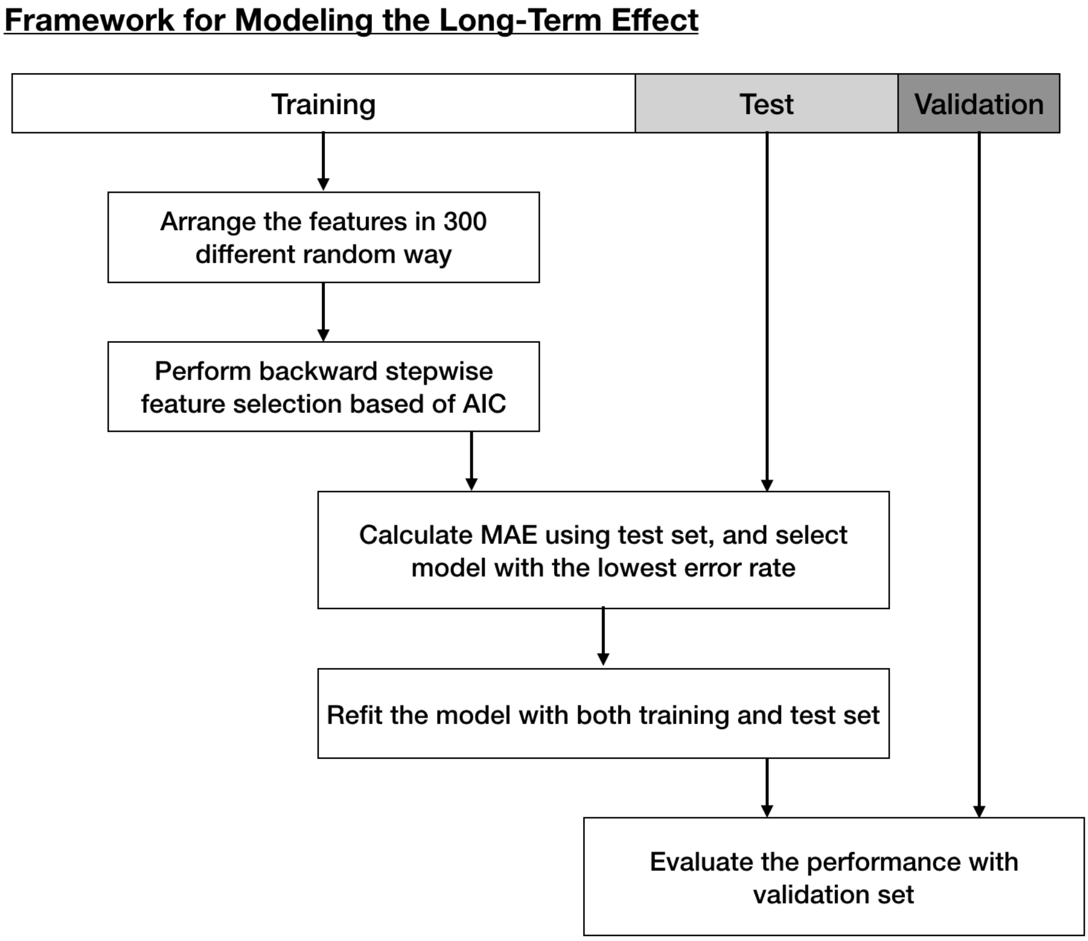

Figure 4: *Overall Modelling Process for Long-Term Effect*

3.3.5. Forecasting

Before forming the final load demand, 1,000 annual base load demands are randomly drawn from a Gaussian distribution. The mean of Gaussian distribution is the point forecast of the long-term model. While the variance is derived analytically with the confidence level of 95.45 %, and calculated using the fitted values and variance of the long-term model. This process is to account for the forecasting uncertainty. Then each of these 1,000 annual base load demands will replicated 8,784 (183 $(days\ per\ year) \times 48$ $(periods\ per\ day)$) times to match the half-hourly frequency of a year.

3.4. Peak Demand Density

Now that both normalized load demand and annual load demand are simulated in a matching frequency, they can be multiplied to produce the final load demands as per equation 1. As for the peak demand density, it will be inferred from these simulated load demands in according to the percentile. There are two types of peak demand: weekly peak and annual peak. So if the annual peak demand density is to be generated, a single highest value of the load demand will be extracted from each of the 1000 simulated time series of load demand. Whereas for weekly peak density demand, there will be one peak value per week, and total up to 26 values (183 days/7) per year. Hence, 26 values will be extracted for each of the 1000 simulated time series of load demand. By taking the percentiles of these peak values, we can then generate the full density.

4. Findings

In this section, the result from the short-term effect will not be evaluated since there are no changes has being made to it. The results from long-term effect and how its accuracy has translated to the peak density forecasting will be discussed separately in the following sections.

4.1. Long-Term Effect

4.1.1. Feature Selection

During the 300 runs of feature selection through backward stepwise, the residuals of the fitted models are all near zero and the AIC values are around -660. This means now there are 300 models that are perfectly fitted on the train set. Each of these models will have around ten variables, and the top ten most appeared variables are listed below, and the description of these variables are from Jiang et al. (2017).

Coefficient Name	Appearance (out of 300)	Description (Notes: Each macroeconomic indicators are for Australia)
HStarts..TAS	46	Housing; Tasmania; Number of Commitments
Emp.PtPer	37	Employed – Part-time; Persons
COMMP	36	Index of Commodity Prices; All items; AUD, Index, 2013/14=100
Cons.FurEqu	36	CVM (Chain Volume Measure) Household Consumption; Furnishings and household equipment
Cons.RenDwe	36	CVM Household Consumption; Rent and Other dwelling services
Emp.HoursPt	36	Aggregate Monthly Hours Worked (Part-time); Persons
IP.TCO	35	Industrial Production of Textile, Clothing and Other manufacturing
BM	34	Broad Money
Emp.JobVac	34	Job Vacancies
IP.Min	34	Industrial Production of Mining excluding Exploration and Mining support services

Table 2: *Table of the Top 10 Most Appeared Macroeconomic Indicators*

4.1.2. Model Selection

The final model for the long-term effect is selected from the 300 models based on the lowest MAE. The variables and coefficients estimate of the model are shown in the following Table 3. Moreover, the final model selected has good fitness statistics. For instance, the adjusted R-squared is 0.9196, and except for ”COMMP” variable, the other variables have significant t-statistic value.

Coefficient Name	**Coefficient Estimate**	**Description (Notes: Each macroeconomic indicators are for Australia)**
(Intercept)	1.310410304	Forecast when all the predictors values are 0
X10.yr.T.bond	-0.019353858	10 years Australian Government
HCE.AVCE	-0.032482679	Price Index; Audio, Visual and Computing Equipment
M3	-0.019276428	Money Supply; M3
HSexfinVal	-0.140023940	Housing; Total excluding refinancing of established dwellings – Value
Cons.Alc	0.043622866	CVM Household Consumption; Alcoholic Beverages
IP.ME	-0.089102385	Industrial Production of Machinery and Equipment
Cons.Food	-0.048589460	CVM Household Consumption; Food
COMMP	-0.004038426	Index of Commodity Prices; All items; AUD, Index, 2013/14=100
HStarts..TAS	0.068384031	Housing; Tasmania; Number of Commitments

Table 3: *Table Summary of Coefficients for the Final Model*

In order to compare the performance, the validation set will be used now, and none of the models has ever observed this data point in any of the previous stages. The validation set has only one data point, so the forecast is one step ahead. As it shown in following line plot, the proposed methods can indeed produce a slightly more accurate forecast than the base model used in MEFM which is also trained on both the train and test set.

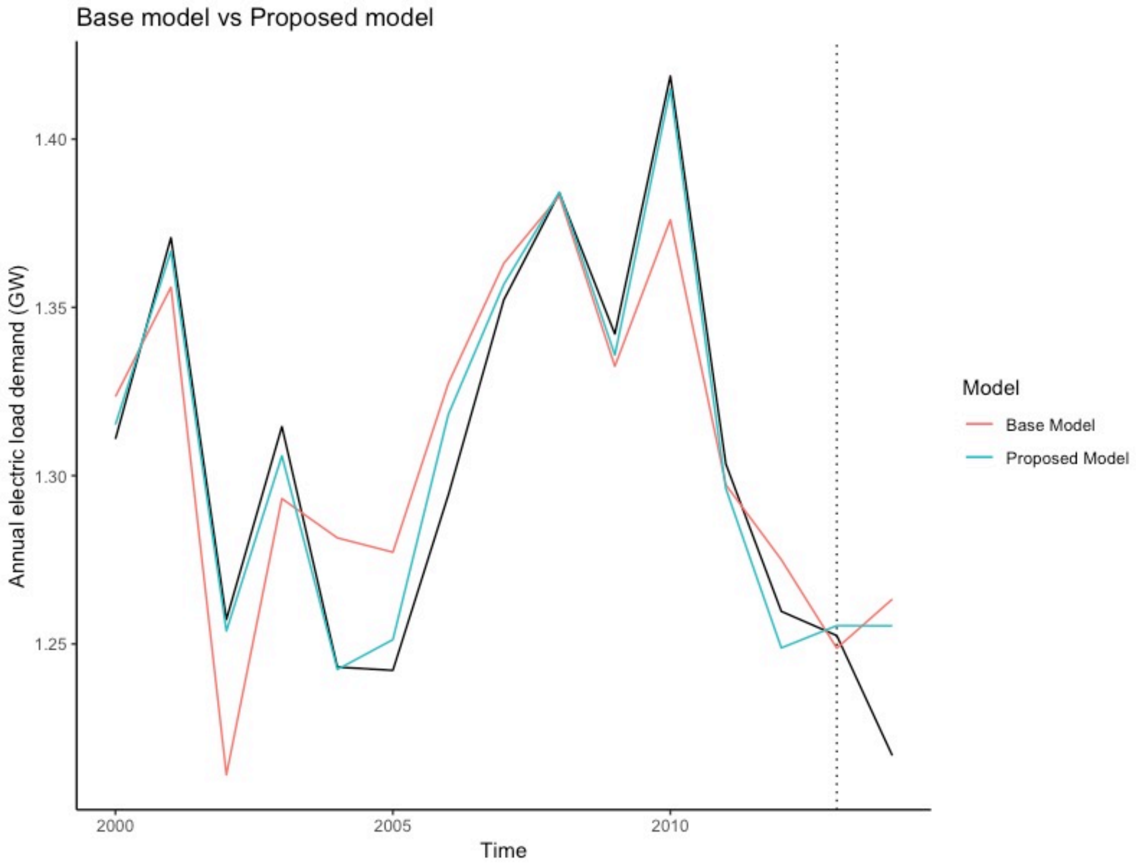

Figure 5: *Forecast and Fit for Base Model and Proposed Model*

The dotted line in the Figure 5 separates the plot into two areas: the left area is the fitted values, and the right area is the forecasts. In terms of the fitted values, the proposed model outperform the base model, as it fitted the actual more closely. As for the forecast, the base model is forecasting into a different direction as the actual annual base load demand. Although the proposed model does not forecast exactly as the actual base load demand, but it is marginally closer compared to the base model.

4.2. Peak Load Demand Density Forecasting

In order to see how the slightly improved accuracy in the long-term effect model translates to the final peak density forecasting, ex-post evaluation will be used. It is ex-post because all the macroeconomic indicators are made available for forecasting the annual base load demand. Hence, the only difference would be how well these models can generate the peak density giving the same data. Three peak density will be generated: peak density with actual annual base demand, peak density with forecasted annual demand from the base model, and peak density with forecasted annual demand from the proposed model.

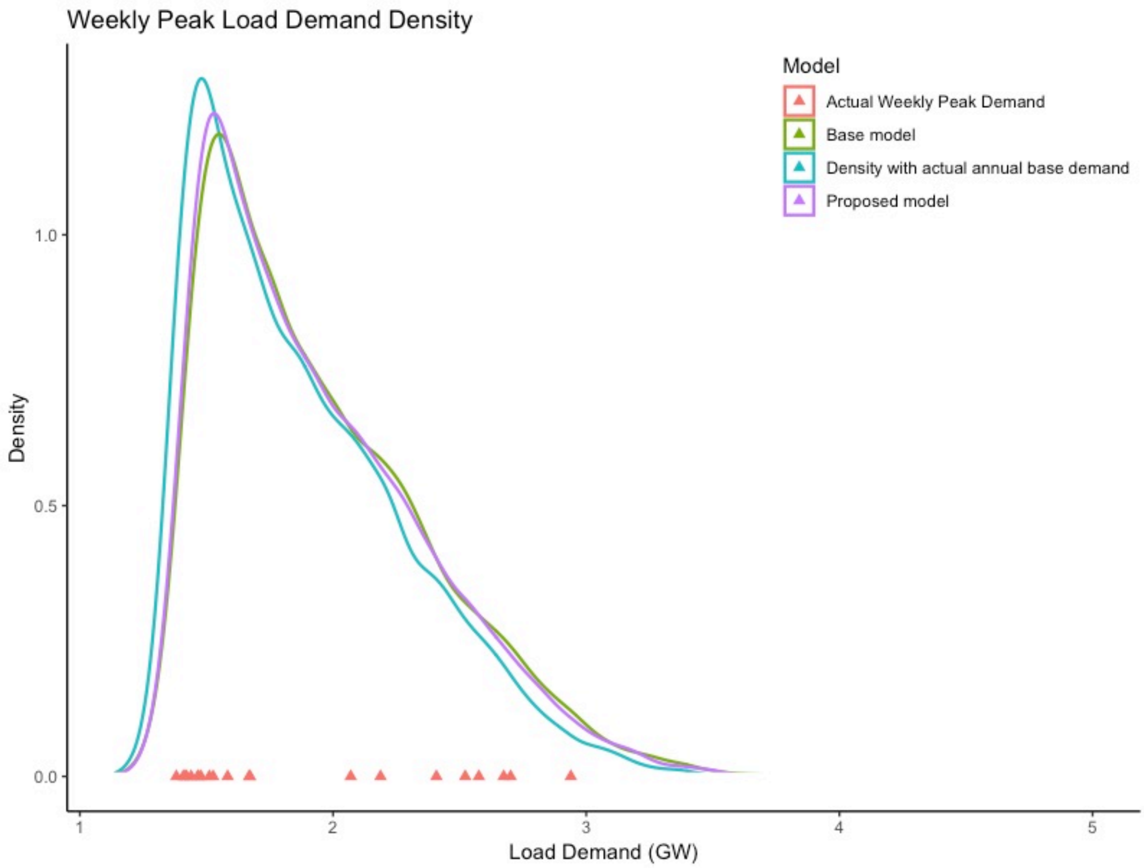

Figure 6: *Weekly Density Plot of Weekly Peak Load Demand*

The Figure 6 above depicts the density of weekly peak load demands. The weekly peak load demands are drawn from the 1000 simulated load demands earlier, and the density is generated according to the percentiles. From the figure, it shows that for some percentiles the proposed model is closer to the actual density. Nevertheless, all three of the distributions are very close to each other, and they covered the actual weekly peak demands pretty well.

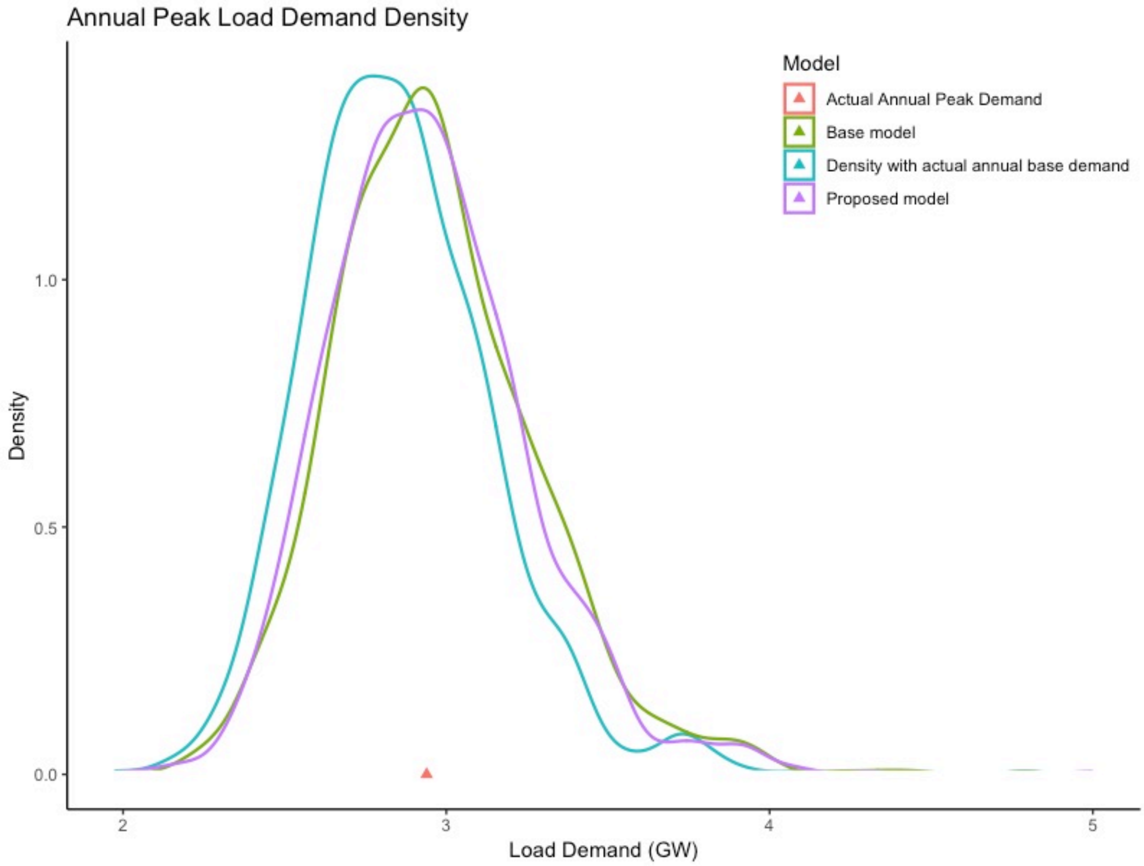

Figure 7: *Annual Density Plot of Weekly Peak Load Demand*

As for the annual peak load demand density, the three distributions are also very close to each other, and the actual annual peak demand falls into the middle of the distribution. As it shows in both earlier figures, even with the base model, these density distributions are pretty well generated.

Since it is hard to tell from the figures which model is closer to the peak density generated using actual annual base demand, error measure function will be used. As discussed in the literature review, pinball loss function is a good way to measure the error rate from a probabilistic distribution. The pinball loss function is formulated as below:

$$\begin{aligned} L_\tau(y,z) &= (y-z)\tau \qquad & if\ y \geq z \\ &= (z-y)(1-\tau) & if\ z > y \end{aligned} \tag{3}$$

- $L_\tau(y,z)$ is the pinball loss function for percentile value τ

- τ is the percentile, for example, 0.01,0.99
- y is actual value of percentile value τ
- z is forecasted value of percentile value τ

The τ values that used in this case is from 0.01 to 0.99, and increment by 0.01, thus in total there are 99 τ values. After calculating the pinball losses, the error rate for each percentile will be summed up to get a final error rate. The final results are presented in Table 4 below, and it shows that the proposed model is indeed closer to the peak density with the actual annual base demand than the base model.

Method	**Pinball Loss (Weekly)**	**Pinball Loss (Annually)**
Base Model	7.262329	11.08008
Proposed Model	5.942139	9.118248

Table 4: *Pinball Loss for Base Model and Proposed Model*

Although there are random processes involved in the forecasting and simulation, but the effect on the error rate are negligible. In every run the error rates are consistently around the same numbers.

4.3. Peak Load Demand Point Forecasting

For point forecasting, the mean of the simulated peak load demand values will be the point forecast. It is the same for weekly and annually. The error function is again MAE. The errors are measure against the actual peak values of the validation set. Below Table 5 summarized the error rate, and again the proposed model has lower error rate than the base model in both weekly and annually

Method	**MAE (Weekly)**	**MAE (Annually)**
Base Model	0.02902172	0.6393038
Proposed Model	0.00862434	0.6362365

Table 5: *MAE for Base Model and Proposed Model*

5. Discussion

5.1. Feature Selection

From the previous Table 2, it shows that the most selected feature is the number of housing finance commitments in Tasmania. Even if the number of randomized features arrangement is change from 300 to 600 and 1200, it is the only indicator that stay consistently at the first spot. This seems to be an abnormal choice, even so when there is already a housing finance commitments indicator for South Australia, because the load demands to be forecasted in this case is in South Australia. As discussed earlier in the literature review, part of the reason to adopt holistic approach is to find interesting choices from the algorithm, and this could be an interesting phenomena to investigate.

The coefficients of the final selected model are shown in Table 3, and for some of the coefficient estimate like price index of audio, visual and computing equipment, it is possible to infer their linear relationship with the load demand. For instance, the price of computing equipment increase causes drop in purchase, and hence, lower usage of load consumption in the household. But others like household consumption on alcoholic beverages or food are harder to draw a direct connection. However, when there is no complete measures or indicators of a system, correlation without causality can still be useful for forecasting.

5.2. Model Selection

Although the proposed model outperformed the base model in many different error metrics, but there is a caveat. The number of randomized arrangements that chose this particular model is 300, and within the 300 this model has the lowest MAE on test set. However, if number is significantly different from 300, a different model will be chosen, and that might not even perform as well as the base model. Hence, this could implied that the model selected is purely coincidence or more holistic searches are causing overfitting. Without any theoretical support of the latter, the former seems to be a more plausible explanation. Nevertheless, for the model to be able to fit well on the training set, forecast well on test set (3 steps ahead), and ultimately perform well on validation set is not an easy feat, hence it would be a very lucky coincidence.

5.3. Possible Application

In order to claim a methodology works, it must produce similar decent result in different experiment. In this case, it would be testing the methodology in different location or different time frame.

The proposed framework showcased the possibility of modelling the long-term effect without any manual intervention like handpicking the variables. Moreover, the selected model performed relatively well albeit the incomprehensible choice on feature selection. Thus, researchers or practitioner can use this thesis as a reference to explore the a holistic way of modelling the long-term effect of peak demand density in a different location or time frame. However, the prerequisite is having enough of the macroeconomic indicators like AMD dataset for the model to be able to search for a good fit.

5.4. Implementation Code

Large part of this thesis is possible because Hyndman and Fan (2010) had made their R codes for MEFM publicly available. Therefore, to continue their spirit and to make this research reproducible, the R codes for thesis has also been made available in Github : `https://github.com/ptwx/finalthesis` .

The proposed framework is designed to automatically perform feature selection, model selection and model fitting, and it has built into a R function. Therefore, the user only needs to pass in the data and number of randomized features arrangements. As a result of that, it is relatively easy to apply the code to another dataset.

6. Limitation

As discussed earlier, there are problems in the proposed framework like the questionable feature selection that needed to be further investigated. For example, are these problems caused by the underlying data that are being used or it might be the method for feature selection is incorrect. Finally like any real-world forecasting, it will be subject to changes in variables selection and estimates, due to the dynamicity of our environment.

As for the overall framework for peak density forecasting, it is a relatively well researched area except for the long-term effect. As we can see from Figure 5, neither the MEFM long-term effect model nor the proposed model forecasted the same direction as the actual annual average demand. Hence below is a list of suggestions to consider for modelling the long-term effect. Since it is only my inexperience suggestions, it might be obvious or naive for the experienced.

- Taking advantage of the unreasonable effectiveness of data by increasing the width and depth of the predictor variables. However, structural change must be considered when adding more data points.
- Explore multiple years forecasts instead of one single year.
- Explore different feature selection method. For instance, using different information criterion and different stepwise or search method.
- Explore different model selection criteria and method. For instance, different error function or different regression model.
- Explore different kind of economic indicators such as endogenous technological change by Paul M. Romer.
- Challenge the linear assumption of the relationship between target and predictor variable.

7. Conclusion

In summary, this thesis demonstrated using stepwise feature selection coupled with out-of-sample model selection allowed for more economic indicators to be modelled. Moreover, the AMD dataset is valuable as it provided many Australian macroeconomic indicators. Although the MEFM framework for peak load demand density forecasting is very well researched, the long-term effect model still can be improved.

Finally, it is time to address the research question of ”Does using more macroeconomic indicators improve peak electricity demand density forecasting?”. Regardless of the contradicting analysis in the findings and discussion, the answer would be yes as showcased in the findings with the empirical results, but the methodology still need refinements to be robust enough for further application in a different case study.

References

AEMO. (2019). *Electricity demand forecasting methodology information paper.* Retrieved 2019-02-28, from https://www.aemo.com.au/-/media/Files/Electricity/NEM/Planning_and_Forecasting/NEM_ESOO/2018/Electricity-Demand-Forecasting-Methodology-Information-Paper.pdf

Athanasopoulos, G., Gamakumara, P., Panagiotelis, A., Hyndman, R. J., Affan, M., et al. (2019). *Hierarchical forecasting* (Tech. Rep.). Monash University, Department of Econometrics and Business Statistics.

C2ES. (2018). *Global emissions.* Retrieved 2018-08-31, from https://www.c2es.org/content/international-emissions/

DeepMind. (2019). *Alphago zero: Learning from scratch.* Retrieved 2017-10-18, from https://deepmind.com/blog/alphago-zero-learning-scratch/

Dordonnat, V., Pichavant, A., & Pierrot, A. (2016). Gefcom2014 probabilistic electric load forecasting using time series and semi-parametric regression models. *International Journal of Forecasting*, *32*(3), 1005–1011.

Florian, Z. (2018). Modeling public holidays in load forecasting: a german case study. *Journal of Modern Power Systems and Clean Energy*, *6*(2), 191–207.

Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *69*(2), 243–268.

Hong, D. T. (2018). *Temperature-based models vs. time series models.* Retrieved 2018-12-02, from http://blog.drhongtao.com/2018/12/temperature-based-models-vs-time-series-models.html

Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32*(3), 914–938.

Hong, T., et al. (2014). Energy forecasting: Past, present, and future. *Foresight: The International Journal of Applied Forecasting*(32), 43–48.

Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., & Hyndman, R. J. (2016). *Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond.*

Elsevier.

Hong, T., Wilson, J., & Xie, J. (2014). Long term probabilistic load forecasting and normalization with hourly information. *IEEE Transactions on Smart Grid*, *5*(1), 456–462.

Hong, T., Xie, J., & Black, J. (2019). Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting. *International Journal of Forecasting*.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Hyndman, R. J., & Fan, S. (2010). Density forecasting for long-term peak electricity demand. *IEEE Transactions on Power Systems*, *25*(2), 1142–1153.

Hyndman, R. J., & Fan, S. (2015). *Monash electricity forecasting model.* Report for Australian Energy Market Operator (AEMO). http://robjhyndman. com ....

IEA. (2018). *World energy outlook 2017.* Retrieved 2018-08-25, from `https://www.iea.org/weo2017/`

IPCC. (2018). *an ipcc special report on the impacts of global warming of 1.5 °c above pre-industrial levels and related global greenhouse gas emission pathways.* Retrieved 2018-08-25, from `http://www.ipcc.ch/report/sr15/`

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Jiang, B., Athanasopoulos, G., Hyndman, R. J., Panagiotelis, A., Vahid, F., et al. (2017). Macroeconomic forecasting for australia using a large number of predictors. *Monash Econometrics and Business Statistics Working Papers*, *2*, 17.

Koenker, R. (2015). Quantile regression.

Lusis, P., Khalilpour, K. R., Andrew, L., & Liebman, A. (2017). Short-term residential load forecasting: Impact of calendar effects and forecast granularity. *Applied Energy*, *205*, 654–669.

Rodriguez, R. N., & Yao, Y. (2017). Five things you should know about quantile regression. In *Proceedings of the sas global forum 2017 conference. cary, nc: Sas institute inc. http://support. sas. com/resources/papers/proceedings17/sas525-2017. pdf.*

Smartgrid.gov. (2018). *What is the smart grid?* Retrieved 2018-08-25, from `https://www.smartgrid.gov/the\_smart\_grid/smart\_grid.html`

Sutton, R. (2019). *The bitter lesson.* Retrieved 2019-03-19, from `http://www.incompleteideas.net/IncIdeas/BitterLesson.html`

Wang, Y., Zhang, N., Tan, Y., Hong, T., Kirschen, D. S., & Kang, C. (2018). Combining probabilistic load forecasts. *IEEE Transactions on Smart Grid*.

Xie, J., & Hong, T. (2016). Gefcom2014 probabilistic electric load forecasting: An integrated solution with forecast combination and residual simulation. *International Journal of Forecasting*, *32*(3), 1012–1016.

Xie, J., & Hong, T. (2018). Temperature scenario generation for probabilistic load forecasting. *IEEE Transactions on Smart Grid*, *9*(3), 1680–1687.

Xie, J., Hong, T., Laing, T., & Kang, C. (2017). On normality assumption in residual simulation for probabilistic load forecasting. *IEEE Transactions on Smart Grid*, *8*(3), 1046–1053.