Hierarchical forecasting of monthly electricity demand

By

Ignitious Chauke
15001093

submitted in partial fulfilment of the requirements for the degree of Master of Science in e-Science

in the

Department of Mathematical and Computational Sciences
Faculty of Science, Engineering and Agriculture
University of Venda, Thohoyandou, Limpopo
South Africa

Supervisor: **Dr C Sigauke**

Co-Supervisor: **Dr A Bere**

4 April 2022

Abstract

Energy demand forecasting is a vital tool for energy management, maintenance planning, environmental security, and investment decision-making in liberalised energy markets. The mini-dissertation investigates ways to anticipate power usage using hierarchical time series and South African data. Approaches such as top-down, bottom-up, and optimal combination are applied. Top-down forecasting is based on disaggregating total series projections and spreading them down the hierarchy based on historical data proportions. The bottom-up strategy aggregates individual projections at lower levels, whereas the optimal combination methodology optimally combines bottom forecasts. An out-of-sample prediction performance evaluation was performed to assess the models' predicting ability. The best model was chosen using mean absolute percentage error. The top-down technique based on predicted proportions (Top-down forecasted proportions) was superior to the optimal combination and bottom-up approach. To integrate forecasts and build prediction ranges for the proposed models, linear quantile regression, linear regression, simple average, and median were used. The best set of forecasts was picked based on the prediction interval normalised average width. At 95%, the best model based on the prediction interval normalised average width was a simple average.

Keywords: *Bottom-up approach, energy demand, hierarchical forecasting, optimal approach, top-down approaches*

Declaration

I, Ignitious Chauke (15001093), certify that this mini-dissertation for the Master of Science degree in e-Science titled, “Hierarchical forecasting of monthly electricity demand" and the work presented in the project is my own; none of the work displayed in the project has been submitted to another institution or university for a degree. Other writers’ work is acknowledged and cited.

Ignitious Chauke
15001093
4 April 2022

Dedication

To my darling daughter, Xichavo Casey Chauke, who makes all my problems disappear when I look into her inspiring eyes.

Acknowledgements

For me, this mini-dissertation was an excellent piece of work that taught me to respect a large and intriguing topic of study. It took months of hard work and effort, and it would not have been possible without many people's blessings and support. I'd like to begin by thanking Dr. C. Sigauke, my dissertation supervisor, for his guidance, recommendations, and advice. I am grateful for his willingness to commit time to my dissertation and help me complete it. Dr. A. Bere, my co-supervisor, deserves special thanks for his patience and support.

I am also thankful to the University of Venda for admitting me as a student and the National E-Science Postgraduate Teaching and Training Platform (NEPTTP) for providing financial support and introducing me to the field of e-Science. I'd additionally need to explicit my gratitude to my whole circle of relatives for their unwavering aid and patience. So far, your prayers for me have saved me going.

Finally, I would like to thank God for giving me the quality to proceed within the confront of adversity. I've got the opportunity to benefit from your instruction on a day-by-day premise. You were the one who permitted me to finish my research. For my future, I will continue to place my trust in you.

Contents

Abstract i

Declaration ii

Dedication iii

Acknowledgements iv

List of Figures viii

List of Tables x

List of Abbreviations xi

1 Introduction 1

1.1 Background 1
1.2 Problem Statement 3
1.3 Research Aim and Objectives 3
1.3.1 Research Aim 3
1.3.2 Objectives 3
1.4 Significance of the Study 3
1.5 Scope of the Study 4
1.6 Structure of the study 4

2 Literature review 5

2.1 Introduction 5
2.2 An overview of Hierarchical Forecasting 5
2.3 Conclusions from Literature 9

3 Research Data and Methodology **10**
3.1 Introduction . . . 10
3.2 Data . . . 10
3.3 Hierarchical Time Series Structure . . . 12
3.3.1 Cross-sectional Hierarchical Time Series . . . 12
3.4 Approaches to Hierarchical Forecasting . . . 13
3.4.1 The Bottom-up Approach . . . 14
3.4.2 Top-down Approaches Based on Historical Proportions . . . 14
3.4.3 Top-down Approach Based on Forecasted Proportions . . . 15
3.4.4 The Optimal Combination Approach . . . 17
3.5 Forecast Combination . . . 18
3.6 Prediction Intervals . . . 18
3.6.1 Linear Regression . . . 18
3.6.2 Linear Quantile Regression . . . 19
3.6.3 Combining Prediction Limits . . . 19
3.6.4 Evaluation of Prediction Intervals . . . 20
3.7 Evaluation of Forecasts . . . 21
3.8 Conclusion . . . 22

4 Analysis and Empirical Results **23**
4.1 Introduction . . . 23
4.2 Exploratory Data Analysis . . . 23
4.3 Forecasting Hierarchical Electricity Demand Series . . . 28
4.4 Forecasting Accuracy of Models . . . 28
4.4.1 The Individual Models . . . 28
4.4.2 Individual Models and the Forecast Combination Models . . . 29
4.5 Forecasts for Monthly Electricity Demand . . . 29
4.6 Top-down FP Approach Forecasts . . . 31
4.7 Prediction Interval for Top-down FP Forecasts . . . 33
4.7.1 Evaluation of Prediction Intervals . . . 35
4.8 Conclusion . . . 36

5 Conclusions and Future Work **37**
5.1 Introduction . . . 37
5.2 Research Findings . . . 37

5.3 Recommendations 37
5.4 Limitations of the study 38

Appendices **39**

A Monthly electricity demand illustration per province using box plots **40**

B Forecast of the bottom-up approach, two top–down approaches and optimal combination approach **49**

C Some selected R Code **53**

References **64**

List of Figures

3.1 Maps of all districts and provinces of South Africa. 11
3.2 Cross sectional hierarchical tree diagram for South African electricity demand. 12

4.1 Box plots for Gauteng districts monthly electricity demand data from January 2002 to October 2020. 25
4.2 Electricity demand time series plots in Gigawatt-hours (GWh) from January 2002 to October 2020 (Level 0 is the total monthly electricity demand, level 2 is the province's electricity demand and level 3 is the district's electricity demand). 27
4.3 Top-down FP forecasts. 30
4.4 Top-down FP forecast diagnostic plots for monthly electricity demand. 31
4.5 Top-down FP forecast plot of monthly electricity demand from November 2020 to October 2025 superimposed with a fitted smoothing spline trend. 32
4.6 Total forecasts with 95 % prediction interval based on linear regression. 34
4.7 Total forecasts with 95 % prediction interval based on linear quantile regression. 35

A.1 Box plots for Limpopo districts monthly electricity demand data from January 2002 to October 2020. 41
A.2 Box plots for Mpumalanga districts monthly electricity demand data from January 2002 to October 2020. 42
A.3 Box plots for North-West districts monthly electricity demand data from January 2002 to October 2020. 43
A.4 Box plots for Free state districts monthly electricity demand data from January 2002 to October 2020. 44

A.5 Box plots for Kwazulu-Natal districts monthly electricity demand data from January 2002 to October 2020. 45
A.6 Box plots for Western cape districts monthly electricity demand data from January 2002 to October 2020. 46
A.7 Box plots for Northern cape districts monthly electricity demand data from January 2002 to October 2020. 47
A.8 Box plots for Eastern cape districts monthly electricity demand data from January 2002 to October 2020. 48

B.1 bottom-up approach forecasts in Gigawatt-hours (GWh). 49
B.2 Top-down HP1 approach forecasts in Gigawatt-hours (GWh). 50
B.3 Top-down HP2 approach forecasts in Gigawatt-hours (GWh). 51
B.4 Optimal combination approach forecasts in Gigawatt-hours (GWh). . 52

List of Tables

4.1 Monthly electricity demand summarry. 24
4.2 Comparative analysis of the fitted models. 28
4.3 Model comparisons. 29
4.4 Model comparisons. 36

List of Abbreviations

LR	**L**inear **R**egression
CI	**C**onfidence **I**nterval
LQR	**L**inear **Q**uantile **R**egression
MAE	**M**ean **A**bsolute **E**rror
MAPE	**M**ean **A**bsolute **P**ercentage **E**rror
MASE	**M**ean **A**bsolute **S**caled **E**rror
GWt	**G**igawatt **h**ours
PINAW	**P**rediction **I**nterval **N**ormalised **A**verage **W**idth
RMSE	**R**oot **M**ean **S**quare **E**rror
MIDAS	**M**ixed-**D**ata **S**ampling
QCA	**Q**untile**C**ombination **A**pproach
VBEM	**V**ariation**B**ayesian **E**xpection **M**aximisation
EDA	**E**xploratory **D**ata **A**nalysis
ANN	**A**articial **N**eural **N**etwork
GP	**G**aussian **P**rocess
PV	**S**olar **P**hotovoltaic
SVR	**S**upport **V**ector **R**egression
MBE	**M**ean **B**ias **E**rror

Chapter 1

Introduction

1.1 Background

Electricity is regarded as a significant contributor to wealth development and a vital component of economic growth [8]. As a result, electricity demand forecasting is a critical energy system issue. Accurate forecasting can help rational decision-making and management in the power generation and supply sectors while also addressing economic and environmental problems. It can be described as a hierarchical time series forecasting issue with geographical hierarchy aggregation constraints because the sum of the disaggregated time series prediction results should equal the aggregated time series prediction results [26].

Hierarchical time-series are composed of many hierarchical time-series that may be aggregated and disaggregated on several levels based on characteristics such as location, size, and product type [34]. Hierarchical forecasting systems can also deliver predictions for items and respective groups [6]. Electricity demand forecasting has long been demonstrated to be significant in the planning of electricity utilities. The provision of electricity demand can be split into short-term forecasts covering hours to weekly projections, medium-term projections for months to years, and long-term forecasts, from one year to many decades.

Consumption over the entire geographical area can be broken down into multiple sub-regions, which can be broken down further into lower-level regions. Electricity usage in countries, for example, can be separated into provinces, municipalities, and districts. Forecasting electricity demand is difficult due to several factors, including underlying population growth, technology advancements, economic conditions, and current weather conditions [13].

Forecasts in hierarchical data systems to aid decision-making are critical in a variety of applications, including retail and tourism [34]. Depending on the hierarchy's many tasks, hierarchical time series may be constructed in a variety of ways [1]. There have been various debates over the precision of hierarchical forecasting models, but no agreement has been reached on their accuracy [6] [9].

Traditionally, top-down and bottom-up methodologies have been used in hierarchical forecasting. The top-down approach entails projecting a complete series and splitting the predictions into historical or predicted proportions. Bottom-up forecasting anticipates each disaggregate sequence at the lowest hierarchical level, then averages the findings to get higher-level forecasts [15].

Based on the UK home smart meter data, Taieb and Hyndman [31] created a unique probabilistic forecasting approach for a large hierarchy. The method combines an effective mix of forecasts to generate trustworthy and accurate probabilistic projections. The technique also captured the spectrum of distributions in the hierarchy of smart meters. According to the study's findings, we are modelling a smaller set of dependencies rather than the distribution of all the series in the hierarchy as a whole appears to be better to ensure a coherent hierarchy.

This mini-dissertation focuses on using hierarchical forecasting methods to anticipate energy consumption in South Africa, including the top-down approaches, optimal combination strategy, and bottom-up approach. South Africa's energy consumption is among the highest in Africa since the country has one of the world's largest and fastest-growing populations. There has been little research on hierarchical electrical demand forecasting in South Africa, and this dissertation attempts to contribute to this important topic.

1.2 Problem Statement

There are no probabilistic methodologies in the literature on hierarchical forecasting. Random unit interaction models, energy design, operating reserves configuration, value forecasting, and electricity sector commercialism use probabilistic estimates of total device load for power demand. Electrical corporations can use precise forecasting to advocate for expanding energy networks to expand and develop economies. As a result, one of the most important concerns in power management is the demand prognosis for electricity.

1.3 Research Aim and Objectives

1.3.1 Research Aim

This mini-dissertation aims to create a model for forecasting monthly power consumption using hierarchical time series.

1.3.2 Objectives

This study's main objectives are:

- utilize a combination of top-down, bottom-up, and optimal combination methods for medium-term forecasting of monthly electricity demand data,
- use mean absolute percentage error (MAPE), mean absolute error (MAE), mean absolute scaled error (MASE), and root mean square error(RMSE) to compare the performance of various approaches,
- assesses forecast accuracy.

1.4 Significance of the Study

This dissertation is important because forecasting energy demand is essential for utility management, maintenance, and power purchasing and selling. Power demand provision is crucial in the energy business because it provides the foundation for power system planning and operation. It is critical to generate more accurate projections of electricity consumption.

1.5 Scope of the Study

Statistics South Africa provided monthly power demand statistics. To evaluate the model's performance, the accuracy metrics root mean square error (RMSE) and mean absolute error (MAE) were utilised, and the best model was chosen based on the mean absolute percentage error (MAPE). It was compared to find the best model combination of shorter prediction range, top-down, bottom-up prediction intervals, quantile regression mean and linear regression.

1.6 Structure of the study

The remainder of the paper is laid out as follows: The strategies are described in Chapter 3 after a brief survey of the literature on hierarchical forecasting in Chapter 2. The results are described in Chapter 4 and the conclusions are presented in Chapter 5.

Chapter 2

Literature review

2.1 Introduction

This chapter includes an overview of hierarchical time series forecasting and synopsis of research that employed the suggested technique to forecast energy demand.

2.2 An overview of Hierarchical Forecasting

Hierarchical time series forecasting has been the subject of several types of research. Makoni et al. [22] studied the bottom-up, optimum combination, and top-down hierarchical forecasting strategies of foreign visitor arrivals in Zimbabwe. According to the findings of this study, the bottom-up approach provided the most accurate method of calculating foreign visitor arrivals in Zimbabwe. Aggregate series are expected to rise in general, according to predictions.

Athanasopoulos et al.[3] investigated hierarchical projections for domestic tourism in Australia. They looked at five different hierarchical forecasting techniques. The recently proposed top-down approach decomposes top-level predictions based on the predicted proportions of lower-level series, one of two versions of the bottom-up and top-down approaches. According to predicted performance ratings, top-down and optimal combinations based on expected ratios are preferable to bottom-up procedures and are best suited for the tourist class. This approach was then utilised to develop a quantitative projection for Australia's domestic tourism sector.

Hierarchical forecasting approaches have been frequently employed to aid harmonized decision-making by delivering coherent projections at various aggregate levels. Gontijo and Costa [30], investigated Brazil's hourly electricity generation, broken down into subsystems and their energy sources. The hierarchy techniques studied were top-down (TD), optimal reconciliation, and Bottom-up (BU). According to the study findings, the optimum mean performance in the optimal reconciliation models is obtained when the prime predictive windows are used. The south subsystem's energy forecasting was likewise shown to be more inaccurate than the others, emphasizing the necessity for individual modeling for this subsystem.

Wickramasuriya et al. [34], investigated the prediction of hierarchical and group time series using trace reduction. They used a composite prediction technique to generate a set of aggregated and consistent predictions by incorporating information from the entire prediction error covariance matrix. Given fairness, the applied method minimizes the mean square error of consistent predictions aggregated across a set of time series. The proposed strategy is compared against rival techniques using several simulated settings.

Pang et al. [26] propose a new clustering method based on time-series hierarchical electrical prediction. They investigate energy consumption patterns using clustering analyses rather than directly dealing with geographical hierarchy, and they develop a new time series hierarchy on consumption patterns. They discovered that their membership in the relevant consumption patterns is connected with the reconciliation error in the low-level time series. They also performed large-scale experiments on real-time datasets to show that their strategy has the highest predictive accuracy compared to modern methods.

Hyndman et al. [12], proposed a novel hierarchical forecast technique that produces better forecasts than either. Their approach was based on independent forecasting of the complete range across all hierarchies, followed by an optimum combination and a regression model reconciliation of the forecasts. The study's findings demonstrated that their method outperformed the top-down and bottom-up approaches. The proposed method was then shown by forecasting Australia's visitor demand, with data broken down by traveling and geographical area.

All forecasts must be reconciled when developing a time series with hierarchical architecture. Roach [28], created a new hierarchical quantile forecast technique that outperforms a conventional benchmark model. Demand projections were developed using a simulation-based method. To secure the correct reconciliation of all the zonal forecasts and to execute a weighted approach to assure a legitimate total of the bottom-level zonal projections for each of these demand scenarios, the aggregate zonal projections were changed. Discussions were also held on hierarchical time series forecasting and gradients boosting.

Taieb et al. [32], has proposed a unique hierarchical forecasting method for calculating sparse fits while respecting the constraints of aggregation. This issue is characterised as a high-dimensional penalised regression that can be solved fast using the circular coordinate descent approach. In their testing, they also used large-scale hierarchical electricity demand data. The study's findings reveal that their solution outperforms contemporary hierarchical prediction systems in terms of corrected parity and prediction accuracy.

Almeida et al. [2], investigated a hierarchical time series forecasting in electrical grids. This study explored three separate steps: two standard steps, a top-down approach, another step based on a hierarchical data structure, the optimal combination of regressions, and a bottom-up approach. The investigation takes into account short-term forecasts (24-h ahead). Furthermore, the importance of series correlation degrees in increasing forecasting accuracy was emphasized. The study's findings demonstrated that the hierarchical strategy outperforms the bottom-up technique at middle or high levels.

Traditional forecasting methods are often used to predict all hierarchy levels and adjust the estimates to meet this constraint. Roque et al. [29] provided a unique way to predict a large number of hierarchically dependent time series autonomously. An additive Gaussian (GP) process was combined with a hierarchical piecewise linear function to estimate stable, unexplained components of a time series. Each aggregate group in the data hierarchy generates further GPs, which they described as a programmable structure. Two distinct real-world datasets were used to verify the proposed technique, proving its capacity to compete with state-of-the-art methodologies.

Li et al. [20] investigates and compares two commonly used methods for calculating energy output in photovoltaic (PV) systems in Florida: artificial neural networks (ANN) and support vector regression (SVR). Based on reviewed machine learning technology, hierarchical methods were constructed 15 minutes, 1 hour, and 24 hours ago. The production statistics used in this study are based on 15-minute averaged power measurements taken in 2014. To test the model's validity, error statistics such as mean bias error (MBE), mean absolute error (MAE), root mean square error (RMSE), relative MBE (rMBE), mean percentage error (MPE), and relative RMSE were utilised. This research looked into how individual inverter projections may help the whole solar system.

Mijung and Marce [27] developed a probabilistic model with dynamically changing latent variables to describe percentage changes in time series at each layer. Under the new model, they developed the variational Bayesian expectation-maximization (VBEM) approach. They use a sequential implementation of posterior inference in their method, which reduces the processing cost of huge hierarchical time series data. In addition, unlike the traditional EM technique, which produces point estimates of model parameters, their algorithm yields the distribution across the model parameters, revealing which subset of features is responsible for time series proportion changes. The study's simulation findings showed that their strategy beats other methods in terms of prediction.

Quantile combination approach (QCA) in mixed frequency environments was examined by Lima et al. [21]. MIDAS and soft (hard) thresholding approaches were utilized to address the issue of proliferation parameters by estimating quantile retrenchment using mixed frequency data. The proposed method was used to forecast the industrial production index's growth rate. According to the findings, high-frequency information in the QCA enhances predicted accuracy substantially.

Kanda and Veguillas [16] studied data preparation and quantile regression for probabilistic load prediction in the final round of GEFCom 2017. In the GEFCom 2017 final match of hierarchical stochastic load prediction, the quantile regression method created by the R package "quantreg" [18] was applied. In this study, meteorological stations were divided into 11 groups, from which the appropriate stations were selected for each load meter using a boosting process. Strain gauge records were cleaned up and updated in various ways to provide robust quantile estimates. Even though the industrial load meter's equation was used, the variance of the regression equation was minimized by incorporating strategies to reduce forecast instability.

2.3 Conclusions from Literature

This chapter summarizes some of the research that employed the proposed technique to anticipate power consumption. Few studies have been conducted on predicting electricity demand using bottom-up, top-down, or optimal combination approaches. Regarding power usage, quantile regression approaches were also generally overlooked. Some work has been done using two hierarchies [24]. The current study is different because it is based on three hierarchies and considers only cross-sectional hierarchies.

Chapter 3

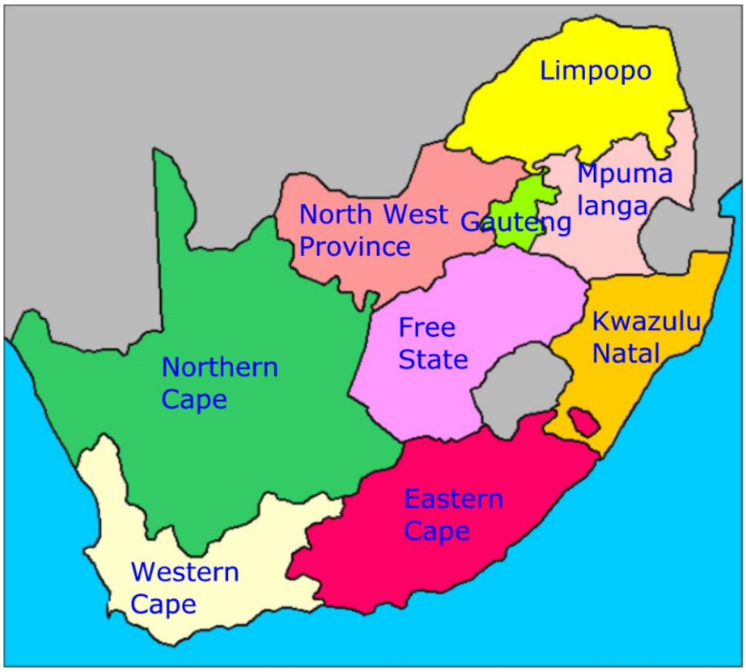
Research Data and Methodology

3.1 Introduction

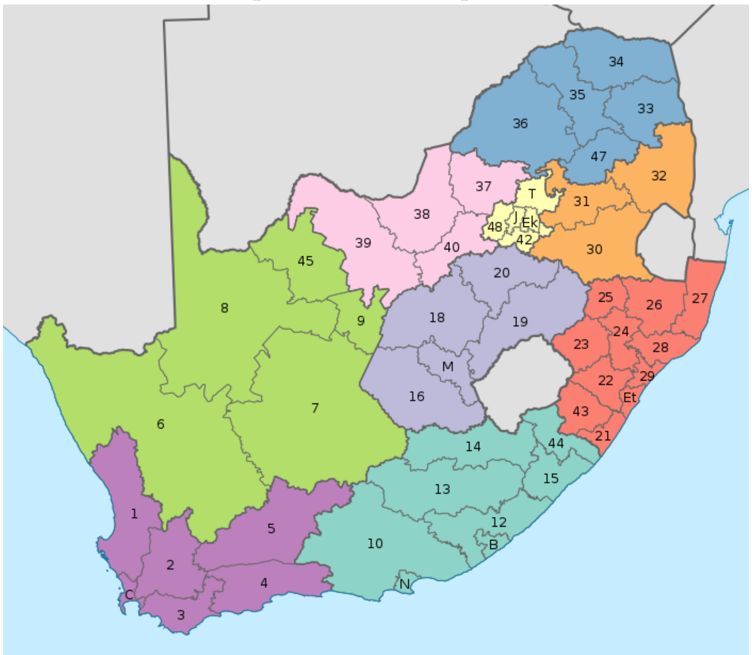
This chapter discusses the tools and strategies utilised in this study to anticipate energy consumption in South Africa. The optimal combination strategy, the Bottom-up approach, the Top-down approach based on historical proportions, the Top-down approach based on proportions expected, and the Bottom-up approach are all covered. This chapter will also cover the methods utilised in model selection and prediction accuracy assessment.

3.2 Data

The mini-dissertation will use monthly electricity demand (GWh) data obtained from Statistics South Africa between January 2002 and October 2020. The data can be studied using two hierarchical methods, namely, cross-sectional and temporal hierarchical time series, but in this mini-dissertation, we will only use cross-sectional hierarchical time series. The data will be disaggregated according to South Africa's nine provinces and fifty-two districts to create a hierarchy of levels 0, 1, and 2. The main reason for selecting this data is that it is traceable and contains some of the local and global events that affected the production and distribution of electricity. Figure 3.1 shows the maps of South Africa indicating all nine provinces (3.1a) and fifty-two districts (3.1b), respectively.



(A) Map of South Africa provinces.



(B) Map of South Africa districts.

FIGURE 3.1: Maps of all districts and provinces of South Africa.

3.3 Hierarchical Time Series Structure

3.3.1 Cross-sectional Hierarchical Time Series

Municipalities and districts can display electricity demand as a hierarchical time series. We propose a three-level hierarchical model of South African power demand based on Hyndman et al. [1] (See Figure 3.2). Level 0 denotes the complete mass series, level 1 the first level of disaggregation, and so on until level K, which contains the most disaggregated series. As a result, Figure 3.2 represents a $K = 2$ level hierarchy.

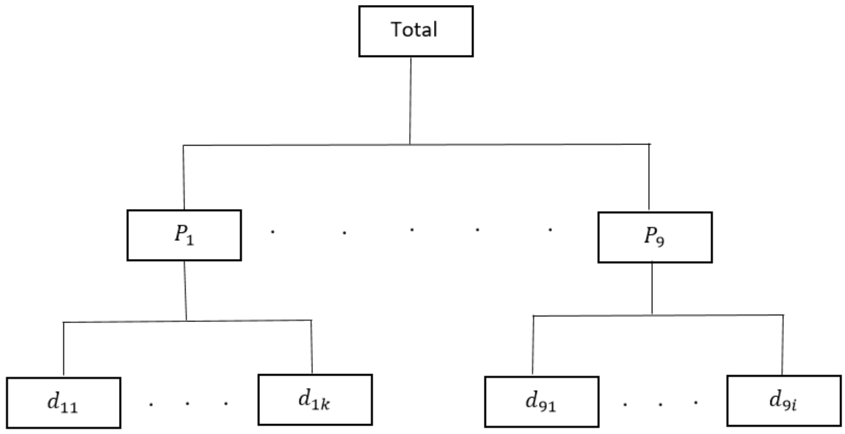


FIGURE 3.2: Cross sectional hierarchical tree diagram for South African electricity demand.

Let $\hat{Y}_{X,t}$ be the t^{th} observation (t=1,...,n) of series $\hat{Y}_{X,t}$ that corresponds to the node X on the hierarchical tree, using the notation of [12]. Where each province and district will be presented, respectively, P_1 =Western Cape, P_9=Limpopo, d_{11}=West Coast District Municipality of Western Cape, d_{1k}=City of Cape Town Metropolitan Municipality of Western Cape, d_{91}= Mopani District Municipality of Limpopo, and d_{9i}= Sekhukhune District Municipality of Limpopo. The individual nodes are denoted by a sequence of letters and numbers, as shown in Figure 3.2. For example, $Y_{P_1,t}$ signifies the t^{th} series observation corresponding to the node P_1 at level 1, $Y_{d_{11},t}$ denotes the t^{th} series observation corresponding to the node d_{11} at level 2, and so on.

3.4 Approaches to Hierarchical Forecasting

To expand the hierarchy rating, we enable the vector $\hat{Y}_{i,t}$ to encompass all levels of electricity demand observations i at time t. At time t, each observation from all series is kept in a column vector described as $\hat{Y}_t = [\hat{Y}_t, \hat{Y}_{1,t}, ..., \hat{Y}_{K,t}]'$. We can now write

$$\hat{Y}_t = S\tilde{Y}_{K,t}, \tag{3.1}$$

where $\hat{Y}_t$ is an n-dimensional vector h-step-ahead forecast for the total electricity demand, $\tilde{Y}_{K,t}$ is an n-dimensional of h-step-ahead forecasts for each of the electricity demand at the bottom-level, and S is a summing matrix of order $n \times n_k$ that aggregates the lowest level series across the hierarchy. For example, consider the hierarchy in Figure 3.2 we have

$$\begin{bmatrix} \hat{Y}_t \\ \hat{Y}_{P_1,t} \\ . \\ . \\ . \\ \hat{Y}_{P_9,t} \\ d_{11,t} \\ . \\ . \\ . \\ d_{9i,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & . & . & . & 1 & 1 & 1 \\ 1 & 1 & 1 & . & . & . & 0 & 0 & 0 \\ . & . & . & & . & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & . & & . & . & . \\ 0 & 0 & 0 & . & . & . & 1 & 1 & 1 \\ & & & & & & & & \\ & & & & I_{52} & & & & \\ & & & & & & & & \end{bmatrix} \begin{bmatrix} d_{11,t} \\ . \\ . \\ . \\ . \\ . \\ . \\ . \\ . \\ d_{9i,t} \end{bmatrix}$$

where I_k denotes an identity matrix of order $k \times k$.

3.4.1 The Bottom-up Approach

The Bottom-up technique is the most often utilised methodology in hierarchical forecasting. This method generates the lowest predictions, which are then aggregated to the highest level using the summing matrix [12]. One can define the Bottom-up approach as:

$$\bar{B} = [0_{nk\times(n-n_k)}|I_{nk}], \tag{3.2}$$

where $0_{i\times j}$ is the $i \times j$ null matrix. The $\bar{B}$ matrix is responsible for collecting the lowest level predictions, which are subsequently aggregated by the S matrix for the full hierarchy to provide the revised provision. The benefits of this technique include the preservation of information and a better understanding of the dynamics of particular series. On the other hand, Bottom-level data may be excessively noisy, making modelling more difficult.

3.4.2 Top-down Approaches Based on Historical Proportions

The top-down technique is the second most commonly used method in hierarchical forecasting. Forecasts are prepared and disaggregated at the highest level of the hierarchy utilising proportions at lower levels of the hierarchy [12]. The approach works well with insufficient counting data. However, it can be difficult to distribute forecasts at low levels. Top-down approaches based on historical proportions can be represented by:

$$\bar{T}_h = [\tilde{q}|0_{nk\times(n_k-1)}], \tag{3.3}$$

where $\tilde{q} = [\tilde{q}_1, \tilde{q}_2 ... \tilde{q}_{mk}]'$ is a collection of proportions for the bottom level series in this example, $\bar{T}_h$'s duty is to disseminate the top-level projections into low-level forecasts. In our mini-dissertation, we will look at two variations of this strategy, both of which performed brilliantly in Gross and Sohl [10]. In the first case,

$$\tilde{q}_{j,t} = \frac{\sum_{t=1}^{n_k} \frac{\tilde{Y}_{(j,t}}{\tilde{Y}_t}}{n_k} \tag{3.4}$$

for $i = 1, ..., n_k$. This will be referred to as Top-down HP1 throughout the study. Each proportion $\tilde{q}_{j,t}$ indicates the average of the historical proportions of the bottom level series $\tilde{Y}_{(j,t)}$ relative to the entire aggregate $\hat{Y}_t$ during the period $t = 1, ..., n_k$; i.e., vector $\tilde{p}$ reflects the average historical proportions. In the second version under consideration,

$$\tilde{q}_{j,t} = \frac{\sum_{t=1}^{n_k} \frac{\tilde{Y}_{(i,t)}}{n_k}}{\sum_{t=1}^{n_k} \frac{\hat{Y}_t}{n_k}} \tag{3.5}$$

for $i = 1, ..., m_k$. Throughout the research, this will be referred to as Top-down HP2. Each $\tilde{q}_{j,t}$ proportion here captures the average historical value of the bottom level series $\tilde{Y}_{(i,t)}$ relative to the average value of the total aggregate $\hat{Y}_t$; i.e., vector $\tilde{q}$ reflects the proportions of the historical averages.

3.4.3 Top-down Approach Based on Forecasted Proportions

We present a top-down way of increasing the historical and statistical character of the levels utilised to break down the high-level predictions to split the share of high-level expectations into predicted proportions for the lower-level series. To demonstrate the intuition of this methodology, we take into account a one-stage hierarchy and a first-stage forecast, which we generate severely for all series. We will calculate the share of each forecast in this level for all individual forecasts at level 1. This procedure will be repeated for each node for a K-level hierarchy, from top to bottom. This will be referred to as top-down FP throughout the study. The biggest disadvantage of this system, which is true of every top-down strategy, is that impartial revised predictions are avoided while bottom projections remain unbiased [12]. As with the previous two top-down techniques,

$$\bar{T}_f = \tilde{q}|0_{nk\times(n_k-1)}], \tag{3.6}$$

where $\tilde{q} = [\tilde{q}_1, \tilde{q}_2, ...\tilde{q}_{nk}]'$ is a collection of proportions for the bottom level series to offer a generic form for the bottom level proportions, we must add a new note. In order to present a general form for the bottom level proportions, we must introduce a new note. Let $\tilde{Y}_{j,n_k}^{(\ell)}$ be the h-step-ahead forecast of the series that corresponding to the node which is ℓ levels above j. Let $\tilde{S}_{j,n_k}(h)$ be the sum of the h-step-ahead predictions below node j which are directly connected to node j. The two notations are going to be integrated. As an example, in Figure 3.2, $\tilde{S}_{AA,n_k}^{(2)}(h) = \tilde{S}_{Total,n_k}(h) =$

$\tilde{Y}_{A,n_k}(h) + \tilde{Y}_{B,n_k}(h) + \tilde{Y}_{C,n_k}(h)$. If we construct h-step-ahead predictions for the series in Figure 3.2, the updated final forecasts travelling down the furthest left branch of the hierarchy will be where $\tilde{q} = [\tilde{q}_1, \tilde{q}_2, ...\tilde{q}_{nk}]'$ is a set of proportions for the lowest level series. A new note is required to provide a generic form for the bottom level proportions.

$$
\begin{aligned}
\tilde{Y}_{A,n_k}(h) &= \left(\frac{\tilde{Y}_{A,n_k}(h)}{\tilde{S}^{(1)}_{A,n_k}(h)} \right) \hat{Y}_{Total,n_k}(h) \\
&= \left(\frac{\tilde{Y}^{(1)}_{AA,n_k}(h)}{\tilde{S}^{(2)}_{AA,n_k}(h)} \right) \hat{Y}_{Total,n_k}(h),
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{Y}_{AA,n_k}(h) &= \left(\frac{\tilde{Y}_{AA,n_k}(h)}{\tilde{S}^{(1)}_{AA,n_k}(h)} \right) \tilde{Y}_{A,n_k}(h) \\
&= \left(\frac{\tilde{Y}^{(1)}_{AA,n_k}(h)}{\tilde{S}^{(1)}_{AA,n_k}(h)} \right) \left(\frac{\tilde{Y}^{(1)}_{AA,n_k}(h)}{\tilde{S}^{(2)}_{AA,n_k}(h)} \right) \hat{Y}_{Total,n_k}(h).
\end{aligned}
$$

Consequently,

$$
\tilde{q}_1 = \left(\frac{\tilde{Y}^{(1)}_{AA,n_k}(h)}{\tilde{S}^{(1)}_{AA,n_k}(h)} \right) \left(\frac{\tilde{Y}^{(1)}_{AA,n_k}(h)}{\tilde{S}^{(2)}_{AA,n_k}(h)} \right).
$$

Similarly, the other proportions are calculated. The overall conclusion is as follows:

$$
\tilde{q}_1 = \prod_{\ell=0}^{K=1} \frac{\tilde{Y}^{\ell}_{j,n_k}(h)}{\tilde{S}^{\ell+1}_{j,n_k}(h)}, \tag{3.7}
$$

for $j = 1, 2, ..., n_K$.

3.4.4 The Optimal Combination Approach

The last hierarchical prediction approach that we investigate is the optimal combination technique. This approach optimally combines the basic forecasts to offer an updated forecast that is as near to the univariate forecasts as feasible while meeting the requirements for higher-level predictions as a sum of the appropriate lower-level forecasts [14]. Furthermore, unlike any other known methodology, this method may provide uncertain forecasts that are consistent at a hierarchical level. The following are the fundamental forecasts in a hierarchy for *h*-step forward:

$$\hat{Y}_{n_k}(h) = S\bar{\beta}_h + \varepsilon_h, \tag{3.8}$$

where

$$\bar{\beta}_h = E[\tilde{Y}_{K,n_k}(h)|\hat{Y}_1, \hat{Y}_2, ...\hat{Y}_{n_k}], \tag{3.9}$$

is the unknown mean of the bottom level K base forecasts, ε_h has zero mean, and covariance matrix V[ε_1]=$\sum_h$. 3.8 then reduces to:

$$\hat{Y}_{n_k}(h) = E[\tilde{Y}_{K,n_k}(h)|\hat{Y}_1, \hat{Y}_2, ...\hat{Y}_{n_k}] + \varepsilon_h. \tag{3.10}$$

The error is anticipated using the forecast error of the lowest level, according to Handyman et al. [14], so $\varepsilon_h \approx S_{\varepsilon_{k,h}}$. The error terms fulfil the aggregate limit as the hierarchical data under this hypothesis. For β_h, the unbiased estimator is:

$$\hat{\beta}_h = (S'S)^{-1}S'\hat{Y}_{n_k}(h). \tag{3.11}$$

As a result, the following estimates have been updated:

$$\hat{Y}_{n_k}(h) = S\hat{B}_h, \tag{3.12}$$

hence

$$\bar{O}_p = (S'S)^{-1}S^{'}. \tag{3.13}$$

3.5 Forecast Combination

Bates and Granger [4] was the first to propose combining predictions. It is especially important when you are unsure about a circumstance or whether the approach is the most accurate, and you want to avoid making big mistakes. It enhances prediction accuracy to the degree that it incorporates valuable and independent data.

Let the total forecasts from the different methods be defined as follows: $y_{Ti}, i =$ 1 (Bottom-up(M1)), 2 (Top-down HP1(M2)), 3 (Top-down HP2(M3)), 4 (Top-down FP (M4)), 5 (Optimal(M5)). The forecasts will be combined using the simple average (M6), median (M7). The average method is given as:

$$\hat{y}_{T(ave)} = \frac{\sum_{i=1}^{M} (\tilde{y}_{T,i})}{M}, \tag{3.14}$$

and the median method as

$$\hat{y}_{T(med)} = \text{Median}\left(\hat{y}_{T1}, \hat{y}_{T2}, \hat{y}_{T3}, \hat{y}_{T4}, \tilde{y}_{T5}\right) \tag{3.15}$$

3.6 Prediction Intervals

In this section we discuss two methods which will used to estimate the prediction intervals. Initially we fit a penalised cubic regression spline given in Equation 3.16.

$$\tau(t) = \sum_{i=1}^{n} (y_i - f(t_i))^2 + \hat{\lambda} \int (f''(t))^2 \, dt \tag{3.16}$$

,
where y_i denotes electricity demand, $\hat{\lambda}$ is a smoothing parameter. We then extract the fitted values and use them in estimating the prediction intervals.

3.6.1 Linear Regression

By adding a linear equation into observational data, linear regression aims to depict the relationship between two variables. The explanatory variable is the first variable, whereas the dependent variable is the second. A definition of the linear regression model is as follows:

$$y_t = \beta_0 + \beta_1 x + \varepsilon_t, \tag{3.17}$$

where y_t are the observed values (electricity demand), β_0 and β_1 are the intercept and parameters, x are the fitted values obtained after fitting the cubic regression spline model described in 3.16 and ε_t is the error term.

3.6.2 Linear Quantile Regression

To integrate forecasts, Quantile Regression (QR) is utilised [25]. Combining or averaging projections from two or more models enhances accuracy while reducing forecasting error variation, according to Koenker and Basset [17]. In this study, we also employ QRA to compute predicted intervals. The QR code is represented as:

$$y_{t,\tau} = \sum_{j=1}^{p} \beta_{j,\tau} x_{tj} + \varepsilon_{t,\tau}; \quad \tau \in (0,1). \tag{3.18}$$

3.18 parameter estimations are obtained by minimising the supplied function in 3.19 as:

$$\tilde{q}_{Y|X}(\tau) = \sum_{t=1}^{n_k} \bar{\rho}_\tau \left(\hat{y}_{t,\tau} - \sum_{j=1}^{p} \hat{\beta}_{j,\tau} x_{tj} \right), \tag{3.19}$$

where $\tilde{q}_{Y|X}(\tau)$ is the extreme conditional quantile function of τ and $\bar{\rho}_\tau(u) = u[\tau - \mathbf{I}(u < 0)]$ is a check function. In this paper, we're interested in estimating extreme conditional quantiles in this work, i.e. for $\tau \in (0,1)$.

3.6.3 Combining Prediction Limits

According to [5][7], combining predictions can improve prediction accuracy. This is then broadened to include both forecast prediction limits. In this section, the GRA and LR combine with a preview interval, and the models' interval is compared.

Robust prediction intervals (PIs) are known to be produced from combining prediction limits from various models [23]. This study shall use the simple average and median methods for combining the prediction limits. The simple average method can be expressed as in 3.20.

$$L_{Av} = \frac{1}{m_k} \sum_{t=1}^{m_k} L_t, \quad U_{Av} = \frac{1}{m_k} \sum_{t=1}^{m_k} U_t \tag{3.20}$$

The median method is known to be less sensitive to outliers. This is given in 3.21 [23]

$$L_{Md} = Median(L_1, \cdots, L_m), \quad U_{Md} = Median(U_1, \cdots, U_m) \tag{3.21}$$

3.6.4 Evaluation of Prediction Intervals

The models used in this study are only a simplification and approximation of the actual electricity demand (patterns). The first index for estimating PI is the prediction interval width (PIW). It is estimated using lower and upper prediction limits and calculated as shown in 3.22.

$$PIW_t = U^{\alpha}(y_t) - L^{\alpha}(y_t) \qquad t = 1, \cdots, m_k \tag{3.22}$$

where $U^{\alpha}(y_t)$ and $L^{\alpha}(y_t)$, denote the upper and lower prediction limits respectively, and α is the nominal confidence.

Various indices, such as the prediction interval coverage probability and the prediction interval normalised average width (PINAW), are used to assess the quality of the PIs. The PINAW is used in this investigation. PINAW shows the model's capacity to collect uncertainty information on interval predictions. It calculates the PIs' average width and is expressed as:

$$\text{PINAW} = \frac{1}{m_k R} \sum_{t=1}^{m_k} (PIW_t), \tag{3.23}$$

where R is the rage of the variable y_t. A smaller PINAW means the PIs are more informative.

3.7 Evaluation of Forecasts

To find the best model for making predictions, the prediction results use mean absolute percent error (MAPE), mean absolute error (MAE), mean absolute scaled error (MASE) and mean square error (RMSE) will be evaluated. The smaller the MAPE, MAE, MASE, and RMSE, the closer the expected return is to the true return and the more fit the model. The formulation of the performance evaluation method is as follows:

Mean absolute percentage error

$$MAPE = \frac{1}{N_k} \sum \frac{[x_t - \hat{x}_t]}{x_t}, \tag{3.24}$$

where x_t are the actual values observed, $\hat{x}_t$ is a predicted value by the model, and N_k is the number of predictions.

Mean absolute error

$$MAE = \frac{1}{N_k} \sum_{t=1}^{N_k} |x_t - \hat{x}_t| \tag{3.25}$$

where x_t are the actual values observed, $\hat{x}_t$ is a predicted value by the model, and N_k is the number of predictions.

Mean absolute scaled error

$$\hat{q}_j = \frac{e_t}{\frac{1}{T-N_k} \sum_{t=N_k+1}^{T} |x_t - x_{t-N_k}|}, \tag{3.26}$$

where N_k is the seasonal period, x_t is the time series of actual observations, and e_t is the forecast error for a specific period. As a result, the average absolute scaled error is simply

$$MASE = mean(|\hat{q}_j|), \tag{3.27}$$

Root mean square error

The Root Mean Square Error (RMSE) is a metric for comparing the model's predicted and observed values.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N_k}(x_t - \hat{x}_t)^2}{N_k}}, \tag{3.28}$$

where $\hat{x}_t$ are predicted values by the model, x_t are the observed values, and N_k is the number of forecasts.

3.8 Conclusion

This chapter began by covering a discussion on variable selection and data processing. It presents the mathematical formulations for the traditional hierarchical time series models, namely bottom-up, optimal combination, and top-down. The chapter further discusses the various formulae for prediction intervals, combining prediction limits and forecast evaluation of prediction intervals.

Chapter 4

Analysis and Empirical Results

4.1 Introduction

Data analysis in this chapter uses the concepts and procedures described in the previous chapter. Fable R [11] is the statistical package that will be used to analyse the data.

4.2 Exploratory Data Analysis

Exploratory data analysis (EDA) is a sort of data analysis that focuses on extracting relevant statistics and other aspects from a dataset [33]. We use data analysis to comprehend a data collection better, confirm basic assumptions, find signaling variables, discover the structure behind them, detect outliers and anomalies, and build models. From 2002 to 2020, the monthly electricity demand (GWh) in South Africa is utilised. To create a hierarchy of levels 0, 1, and 2, the data will be disaggregated by South Africa's nine provinces and fifty-two districts. The descriptive data of all nine (9) provinces' electricity demand are summarised in Table 4.1.

TABLE 4.1: Monthly electricity demand summarry.

Provinces	Min.	1st Qu.	Median	3rd Qu.	Mean	Max.	Skewness	Kurtosis
WC	1464	1828	1894	1883	1960	1202	-0.71	3.88
EC	434	577	640	643	711	850	0.06	2.32
NC	330	422	453	458	497	495	0.15	2.62
FS	748	974	1017	1030	1086	1292	0.31	3.19
KZN	2847	3364	3530	3522	3689	4019	-0.23	2.79
NW	824	1638	1726	1705	1818	1976	-1.42	7.01
GP	3857	4775	5044	5139	5476	6486	0.47	2.67
MP	2077	2611	2710	2692	2794	3011	-0.78	4.19
LP	875	1231	1500	1439	1640	1832	-0.49	2.06

According to Table 4.1 above, the average monthly electricity demand ranges from 453 to 5044, with Gauteng (GP) having the highest mean of 5044. The electricity consumption in Gauteng is the highest, while the Northern Cape (NC) has the lowest number of 330 in one of the periods. The province's highest population may cause the highest electricity demand in Gauteng and the province's lowest population may cause the most insufficient electricity demand. The skewness values of electricity demand in South Africa are positive for the Eastern Cape (EC), Northern Cape (NC), Free State (FS), and Gauteng (GP), indicating that they are positively skewed. Negative for the Western Cape (WC), Kwazulu-Natal (KZN), North West (NW), Mpumalanga (MP), and Limpopo (LP), indicating that their distributions are non-normal. All of the instances in the Western Cape, Free State (FS) North West, and Mpumulanga (MP) have kurtosis larger than three, indicating that their distributions are leptokurtic. This suggests that heavy-tailed distributions can be used to mode the data sets. Eastern Cape, North West, Kwazulu-Natal, Gauteng, and Limpopo have a kurtosis of less than three, indicating platykurtic distribution. This implies that the dataset has lighter tails than a normal distribution.

Konarasinghe and Abeynayake [19] showed that patterns for electricity usage could be revealed using box plots hence what follows below is a graphical representation in the form of a box for monthly electricity demand for districts that are found in Gauteng province only (Others are shown in Figure A.1:A.8). One of the reasons for selecting Gauteng districts is that it is one of the provinces with the highest electricity consumption (As shown in Table 4.1). The box plots for the monthly electricity demand for Gauteng districts are illustrated in Figure 4.1 below, where 42 is Sedibeng West Rand District Municipality, 48 is West Rand District Municipality, Ek is Ekurhuleni Metropolitan Municipality, J and T represent City of Johannesburg Metropolitan Municipality and City of Tshwane Metropolitan Municipality, respectively.

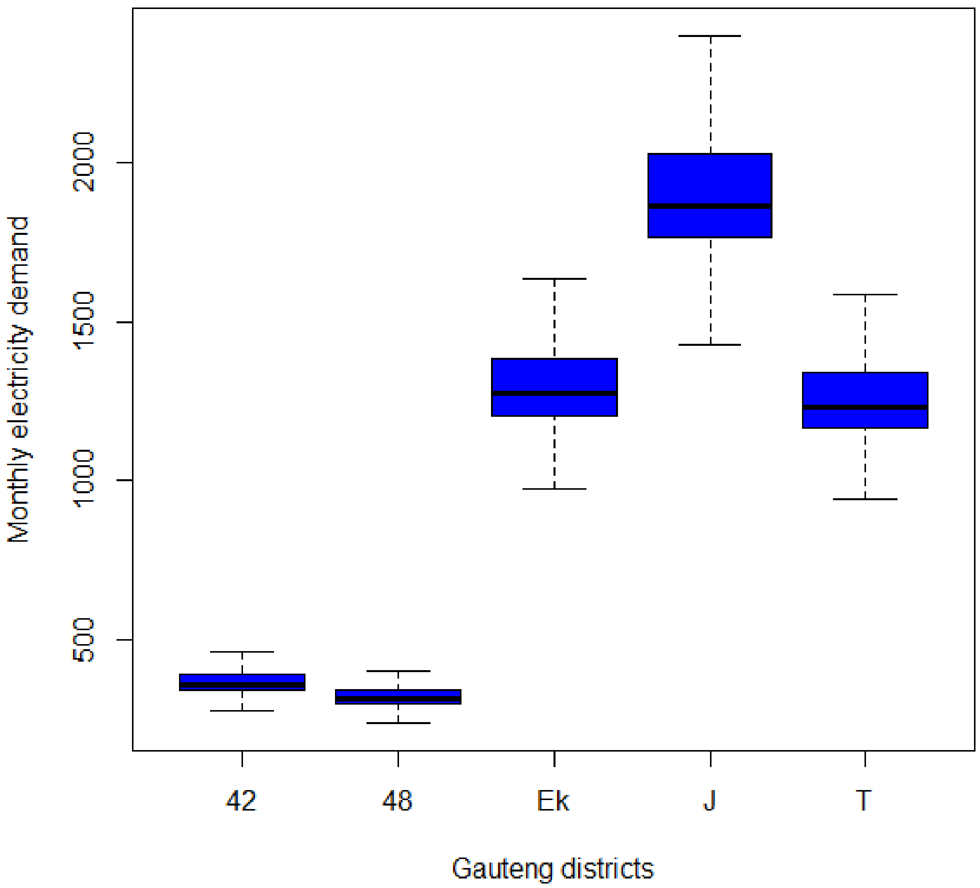


FIGURE 4.1: Box plots for Gauteng districts monthly electricity demand data from January 2002 to October 2020.

The box plots in Figure 4.1 above provide summaries of the minimum, lower quartile, median, higher quartile, and highest power demand in Gauteng province's five districts. The trend in electricity demand can be observed. In appendix, we also showed the monthly electricity demand per district in other provinces.

South African monthly electricity demand data

A time series plot in Figure 4.2 shows the behavior of monthly electricity demand in South Africa for the period January 2002 to October 2020. It can be seen from the level 0 (Total monthly electricity demand) plot that the electricity demand in South Africa is increasing rapidly. The plot also shows some local and global events that affected the electricity demand in 2008 and 2020 (great economic recession and 2020 covid-19). Level 1 and level 2 plots show the monthly electricity demand in South Africa's nine provinces and 52 districts, respectively (See Figure 4.2).

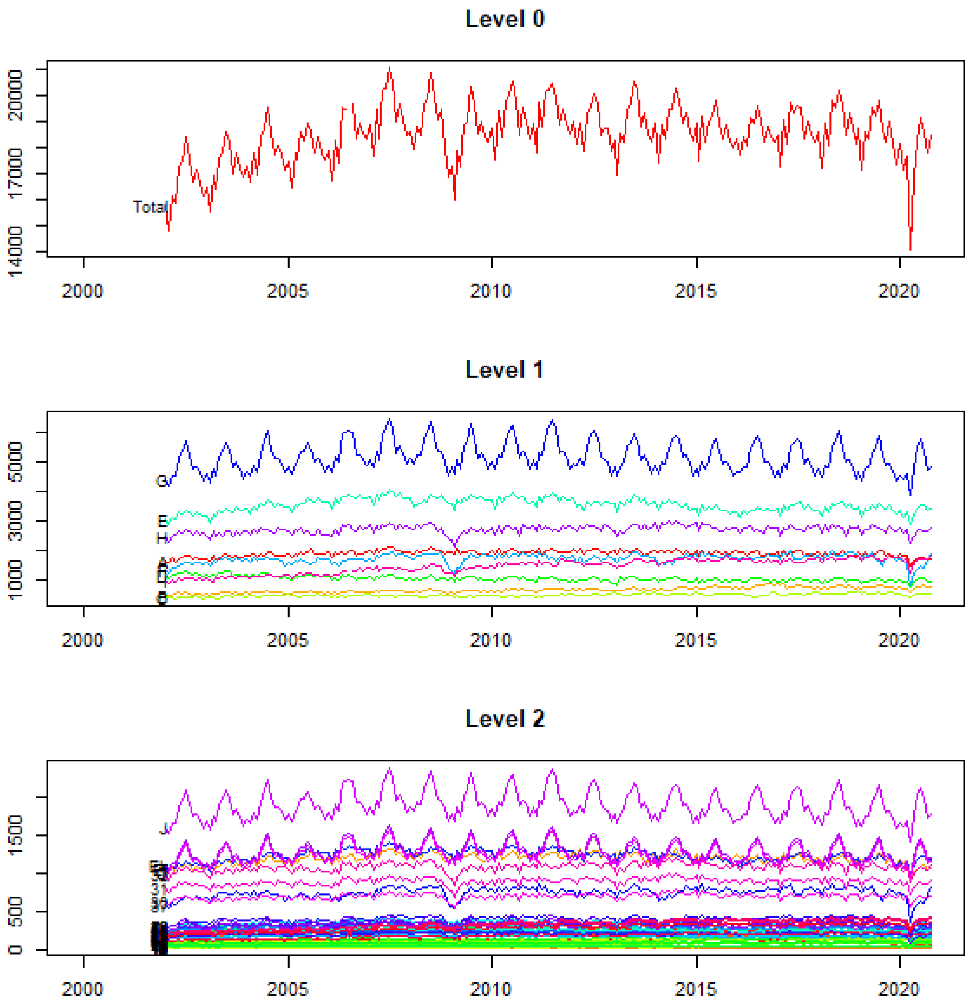


FIGURE 4.2: Electricity demand time series plots in Gigawatt-hours (GWh) from January 2002 to October 2020 (Level 0 is the total monthly electricity demand, level 2 is the province's electricity demand and level 3 is the district's electricity demand).

4.3 Forecasting Hierarchical Electricity Demand Series

Athanasopoulos et al. [3] used exponential smoothing based on an innovative state-space model to estimate Austrian tourism demand. All Level 0, Level 1, and Level 2 monthly power consumption in South Africa will be modelled and forecasted using the same approach in this study. According to [3], the strategies involve additive models that improve forecast accuracy.

4.4 Forecasting Accuracy of Models

4.4.1 The Individual Models

The mean absolute percentage error (MAPE) is used to anticipate accuracy performance at various hierarchy levels based on out-of-sample projections. The MAPE method changed into selected because it is miles famous and easy to comprehend within the literature, and the technique with the bottom MAPE can be decided to be the most successful. The models were trained using data from January 2002 to October 2020, and the models were validated using data from November 2019 to October 2020. The bottom-up method, a top-down method primarily based totally on the common of historic proportions, the proportion of historical averages and due proportions, and the choicest aggregate method, are all used to count on electricity consumption. These strategies are used to reconcile and successfully combine month-to-month electricity calls for predictions. Table 4.2 shows the MAPE, RMSE, MAE, and MASE of each model-averaged across the five methods, which are Bottom-up(M1), Top-down HP1(M2), Top-down HP2(M3), Top-down FP(M4), and Optimal(M5).

TABLE 4.2: Comparative analysis of the fitted models.

	Bottom-up	Top-down HP1	Top-down HP2	Top-down FP	Optimal
RMSE	4432.25	5660.50	5820.84	**4428.09**	4432.25
MAE	3019.58	4522.41	4698.82	**3004.76**	3019.58
MAPE	438.07	686.79	693.63	**436.17**	438.07
MASE	93.01	161.13	163.34	**92.82**	93.01

The Table 4.2 displays the forecasting performance results for each district utilizing the RMSE, MAPE, MAE, and MASE methods. The top-down based on forecasted proportions (Top-down FP) strategy appears to be the overall best performing of the three alternatives because it has the lowest MAPE of 436.17. It is followed by the optimal combination and the bottom-up technique, both of which are equal to 438.07. Thus Top-down FP is regarded to be the overall best performer. This means that this method is capable of producing accurate electricity demand forecasts.

4.4.2 Individual Models and the Forecast Combination Models

The simple average (M6) and median (M7) approaches to integrate the forecasts. Figure 4.3 presents a comparative analysis of the individual models (M1-M5), with the forecast combination models (M6-M7).

TABLE 4.3: Model comparisons.

	M_1	M_2	M_3	M_4	M_5	M_6	M_7
RMSE	4432.25	5660.50	5820.84	4428.09	4432.25	4954.78	4432.25
MAE	3019.58	4522.41	4698.82	3004.76	3019.58	3653.02	3019.58
MAPE	438.07	686.79	693.63	436.17	438.07	538.54	438.07
MASE	93.10	161.01	163.34	92.82	93.10	120.67	93.1

From Table 4.3 above, it can be clearly seen that the Top-down FP (M_4) it is still the best performing model with a MAPE of 436.17, followed by the optimal combination(M_5) and the bottom-up technique (M_1), both of which are equal to 438.07, respectively.

4.5 Forecasts for Monthly Electricity Demand

The Top-down FP approach is used to forecast future out-of-sample monthly electricity demand for the next 60 months across all levels of the hierarchy. Figure 4.3 depicts the forecasts and the original data series graphically.

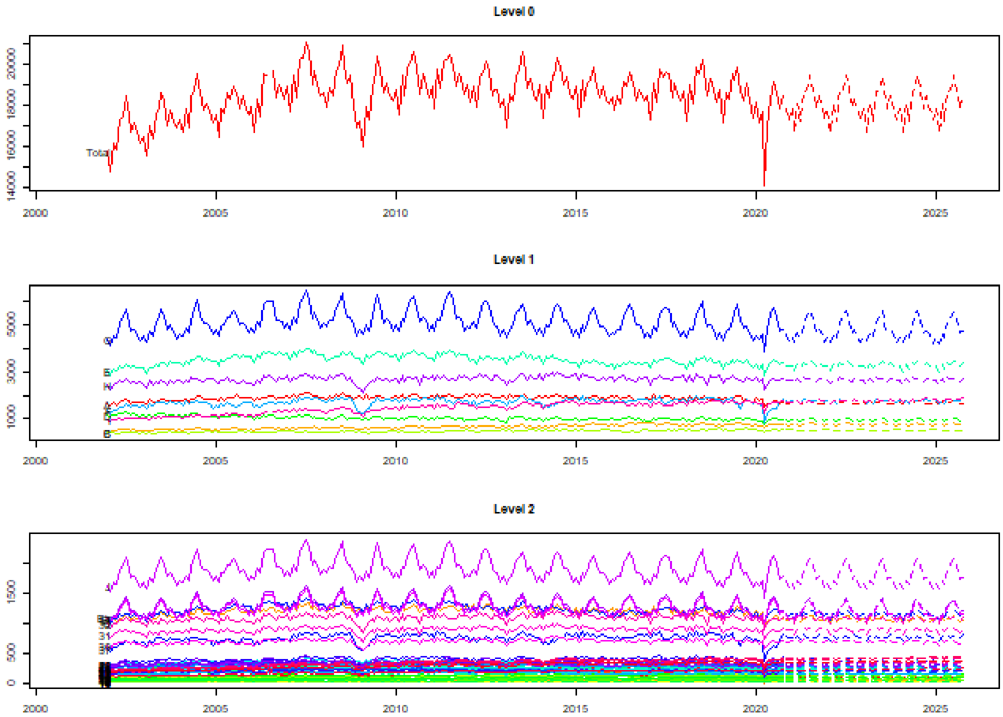


FIGURE 4.3: Top-down FP forecasts.

From Figure 4.3, the top portion indicates South Africa's overall electricity demand. The solid lines on the middle and bottom show historical electricity demand statistics, while the dashed lines represent anticipated electricity demand predictions for each of South Africa's nine provinces and fifty-two districts, respectively.

4.6 Top-down FP Approach Forecasts

Figure 4.4 presents a time series plot of monthly electricity usage anticipated for the next 60 months using a top-down FP approach, replete with density, normal quantile to quantile (QQ), and box plots. The monthly electricity demand forecast trend is depicted in Figure 4.5 using a smoothing spline fitted with an estimated lambda value.

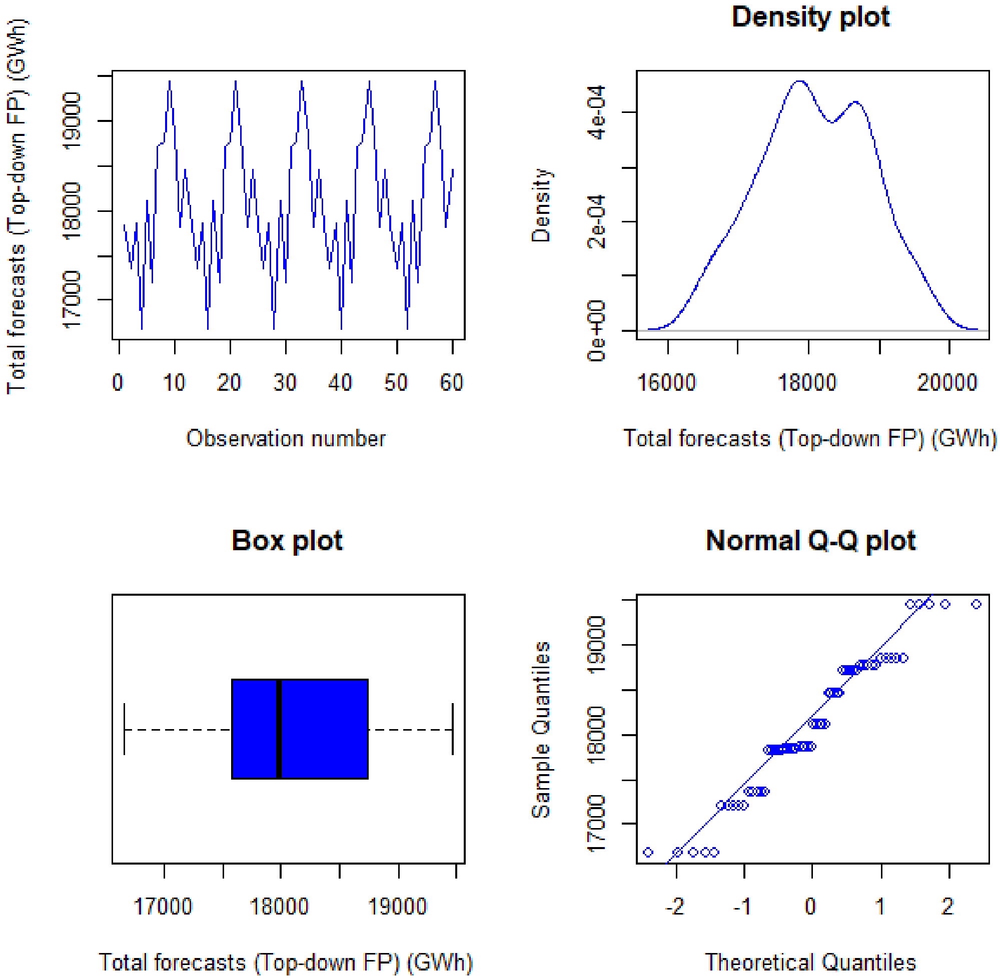


FIGURE 4.4: Top-down FP forecast diagnostic plots for monthly electricity demand.

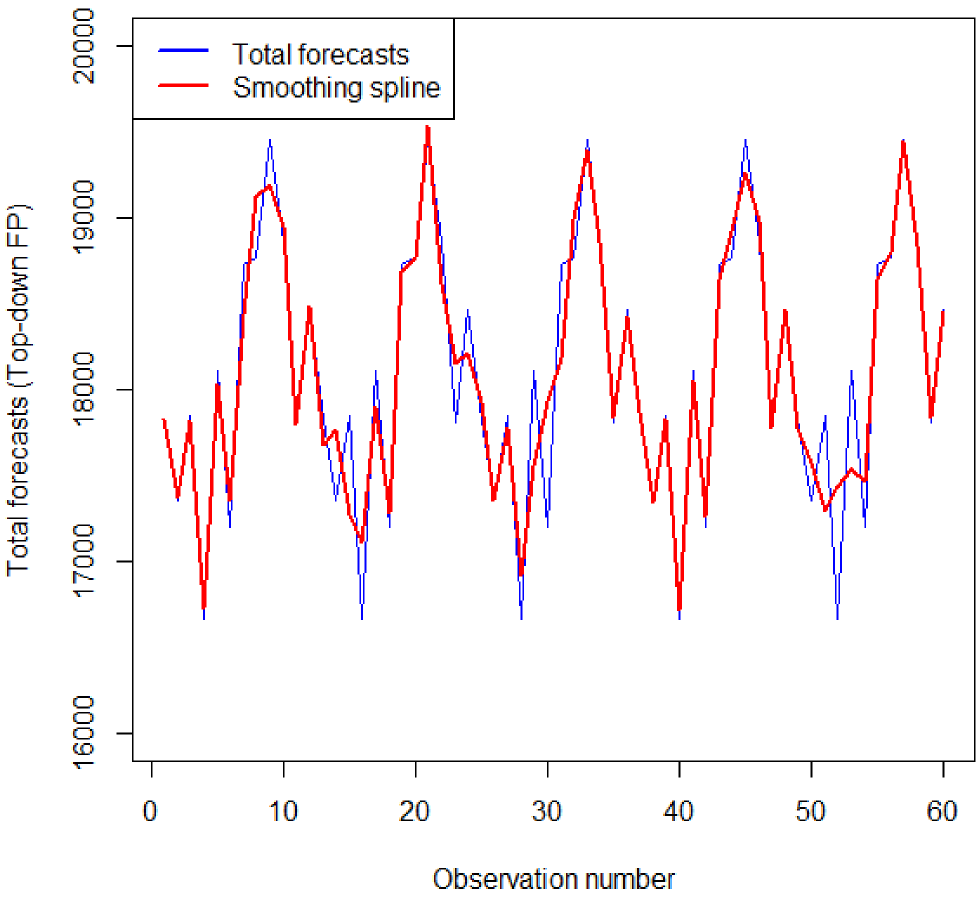


FIGURE 4.5: Top-down FP forecast plot of monthly electricity demand from November 2020 to October 2025 superimposed with a fitted smoothing spline trend.

4.7 Prediction Interval for Top-down FP Forecasts

Prediction intervals are significant because they represent forecast uncertainty. When a prediction interval is calculated, it shows how much uncertainty each forecast has. The penalised cubic smoothing spline function is used to smooth the forecasts. The following is the function:

$$\pi = \sum_{t=1}^{m} (\hat{y}_t - f(t))^2 + \hat{\alpha} \int (f^{''}(t))^2 d_x, \tag{4.1}$$

where α is the smoothing parameter, $f(t)$ is a smoothing spline function at time t obtained from a noisy observation denoted by $\hat{y}_t$, and $f''(t)$ is the smoothing spline function's second derivative at time t. As indicated in Figure 4.5, the $\hat{\alpha}$ value is based on generalised cross validation (GCV) (α = 434842.7).

Intervals of prediction based on linear regression

The upper limit prediction interval is represented by the blue line in Figure 4.6, while the black line shows the current top-down FP method forecast for the next 60 months. Figure 4.6 shows the top limit prediction interval for the 95 percent confidence interval. The trend in the forecast and prediction lines is comparable. We are 95% positive that the projections from the top-down FP method are inside the upper limit since the top-down FP forecast line does not cross the prediction interval.

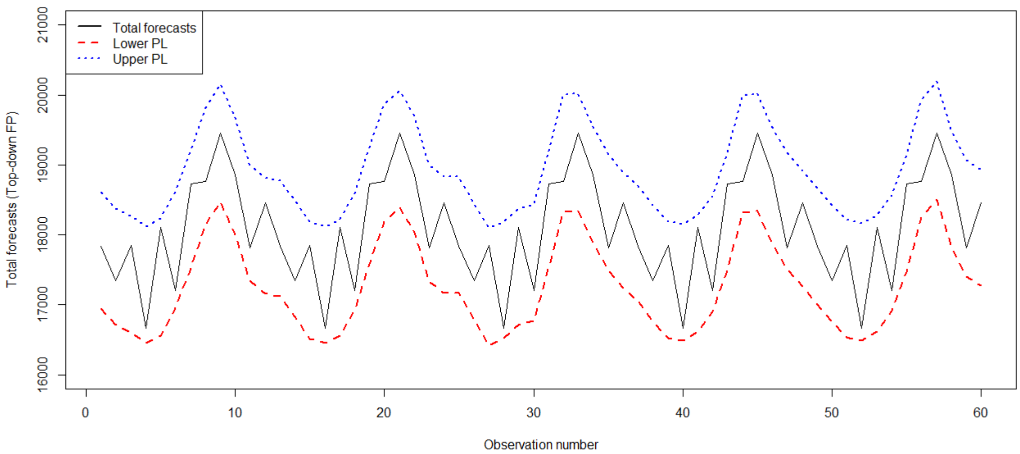


FIGURE 4.6: Total forecasts with 95 % prediction interval based on linear regression.

Prediction intervals based on linear quantile regression

The lower and upper limits prediction intervals for the 95% confidence interval of the quantile regression model are shown in Figure 4.7.

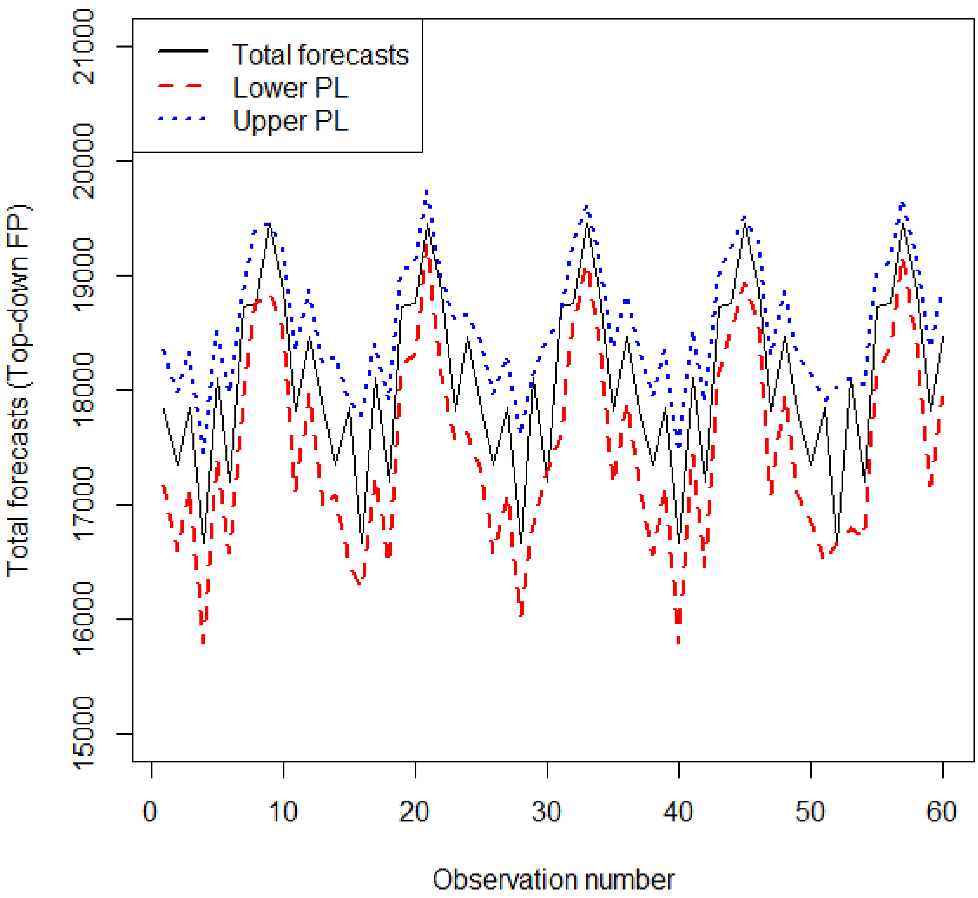


FIGURE 4.7: Total forecasts with 95 % prediction interval based on linear quantile regression.

4.7.1 Evaluation of Prediction Intervals

Model comparisons are shown in Table 4.4, which displays a comparison of the models utilizing PI indices for 95% PINAW. Simple average with 43.20% is the best model based on PINAW at 95%.

TABLE 4.4: Model comparisons.

	LR	LQR	Average	Median
PINAW	67.54 %	63.94 %	44.20%	63.94%

4.8 Conclusion

This chapter proposed monthly electricity demand (GWh) using data obtained from Statistics South Africa for January 2000 and October 2020. Three hierarchical methods were developed. The empirical results showed that the top-down approach produced the best forecast accuracy based on forecasted proportions.

Chapter 5

Conclusions and Future Work

5.1 Introduction

This min-dissertation looked at how top-down, bottom-up, and optimum combination models for energy demand were implemented in South Africa. The models were built using data from Stats SA on monthly power usage in South Africa's provinces. The fitted models were compared using linear regression, simple average, median, and linear quantile regression.

5.2 Research Findings

The purpose of the mini-dissertation was to forecast monthly electricity demand (GWh) between January 2000 and October 2020 using data from Statistics South Africa. This research aims to develop a modelling framework for estimating monthly power consumption based on hierarchical time series models. Monthly electricity consumption forecasting is essential for utility management, maintenance, and power purchase and sale. There have been three hierarchical models established. Based on the MAPE of the three fitted hierarchical models, the top-down technique based on anticipated proportions produced the greatest forecast accuracy.

5.3 Recommendations

Because many members of our society are unaware of South Africa's excessive power use, the government should develop educational programmes on solar energy, wind turbines, generators, and other ecologically acceptable energy sources.

Energy-efficient technology should be pushed, minimising power waste and regulating the amount of electricity used in the country. This study might be useful to system operators, such as decision-makers at power utilities.

5.4 Limitations of the study

This mini-dissertation was based on provincial and district data, which only supplied electricity demand for each province and district, making it hard to establish what causes low or high electricity demand in each province or district. At times, accurately incorporating several sophisticated factors affecting electricity demand into forecasting models might be difficult.

Appendices

Appendix A

Monthly electricity demand illustration per province using box plots

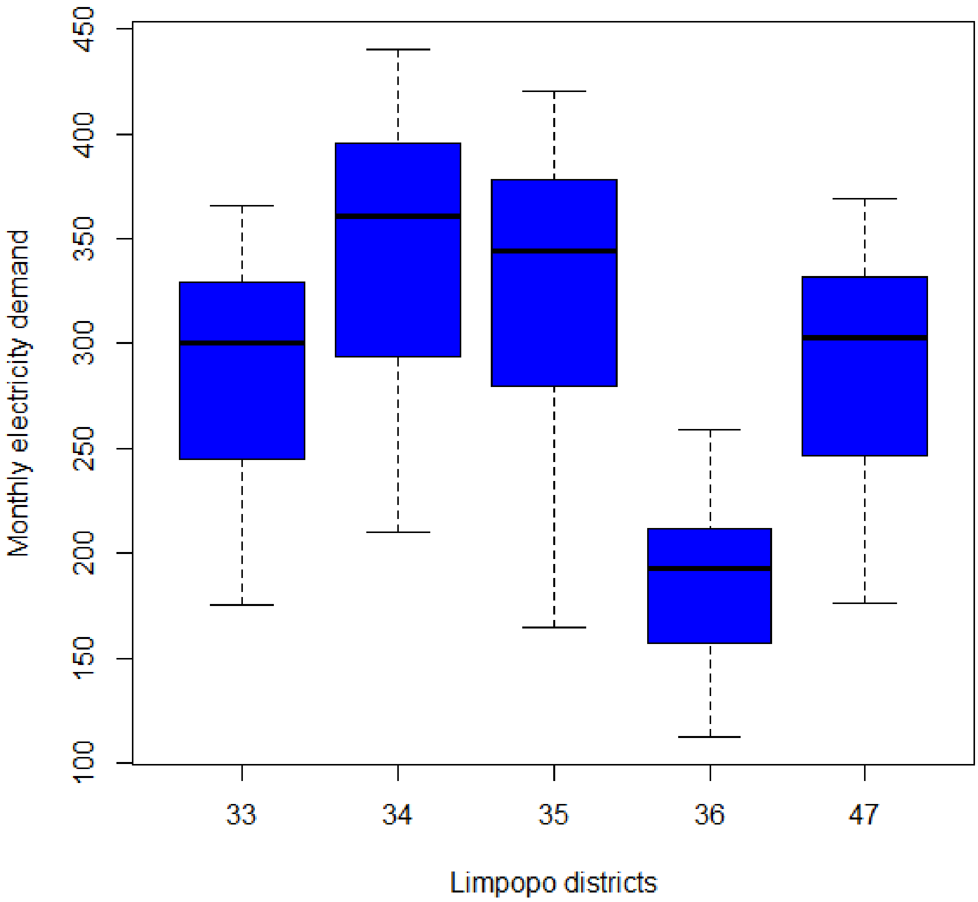


FIGURE A.1: Box plots for Limpopo districts monthly electricity demand data from January 2002 to October 2020.

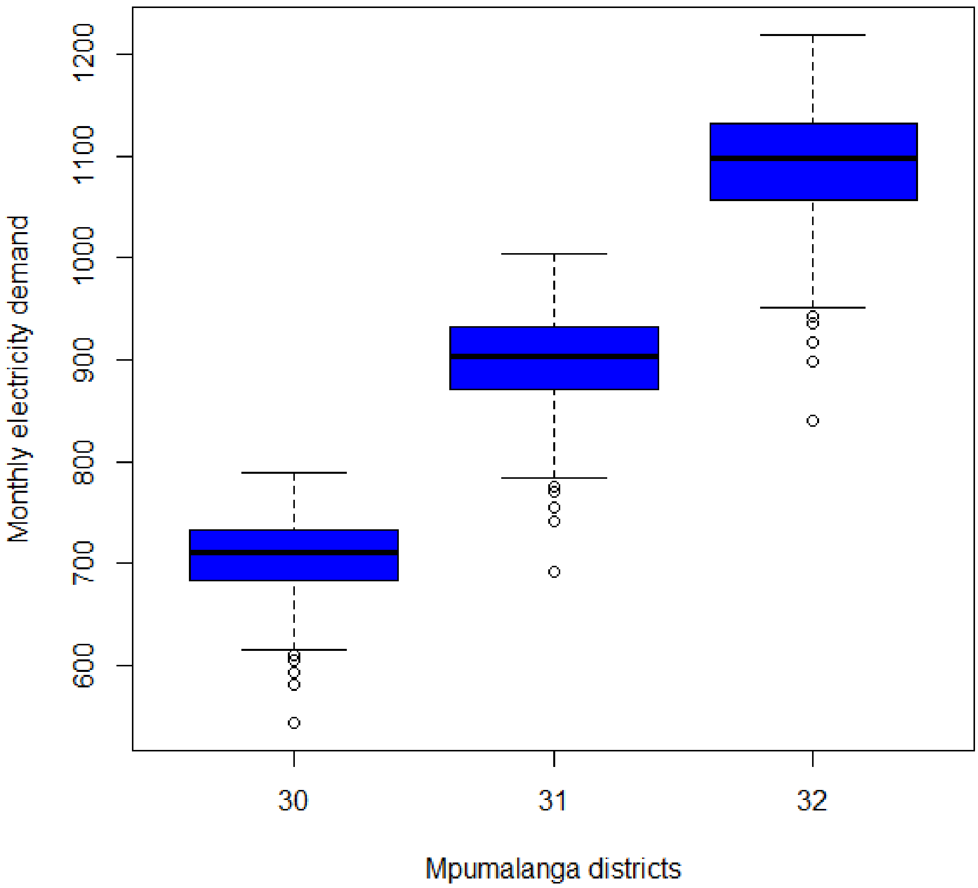


FIGURE A.2: Box plots for Mpumalanga districts monthly electricity demand data from January 2002 to October 2020.

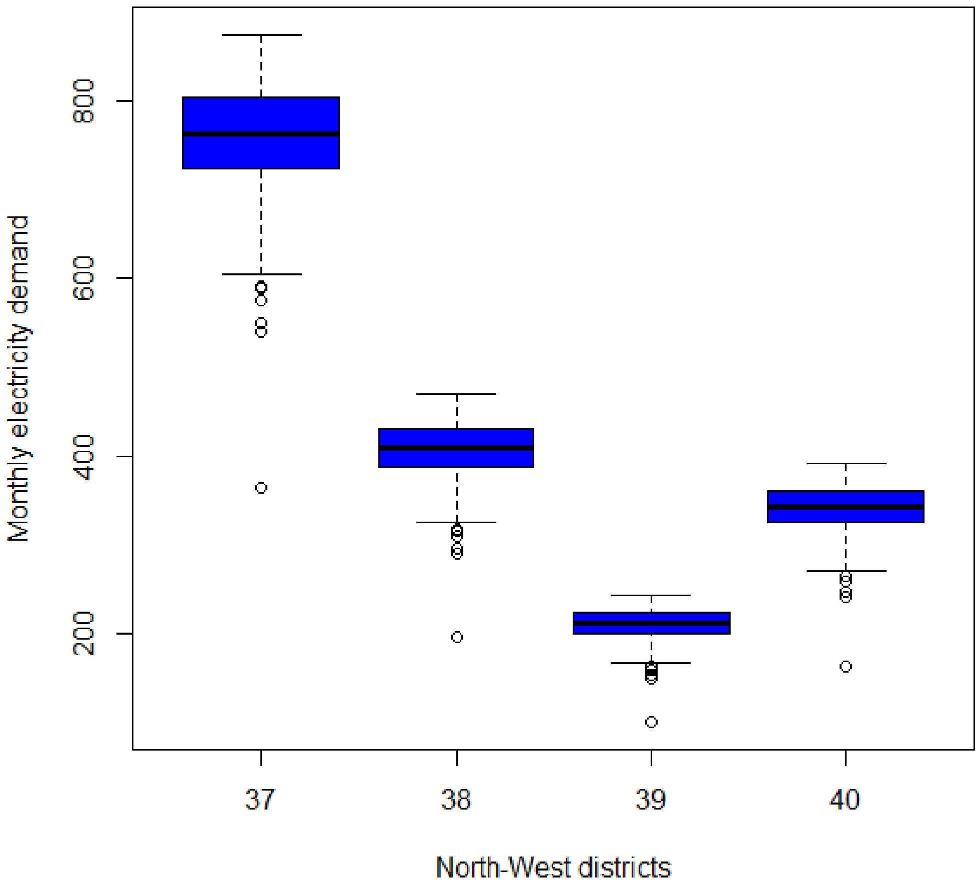


FIGURE A.3: Box plots for North-West districts monthly electricity demand data from January 2002 to October 2020.

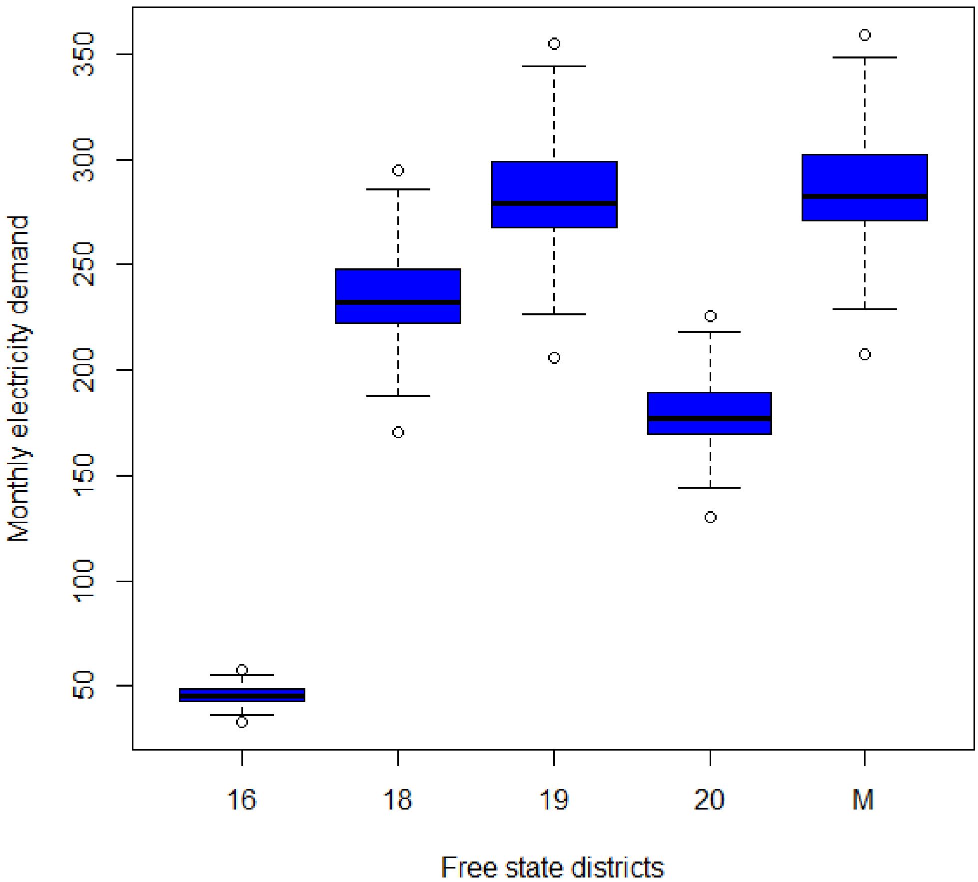


FIGURE A.4: Box plots for Free state districts monthly electricity demand data from January 2002 to October 2020.

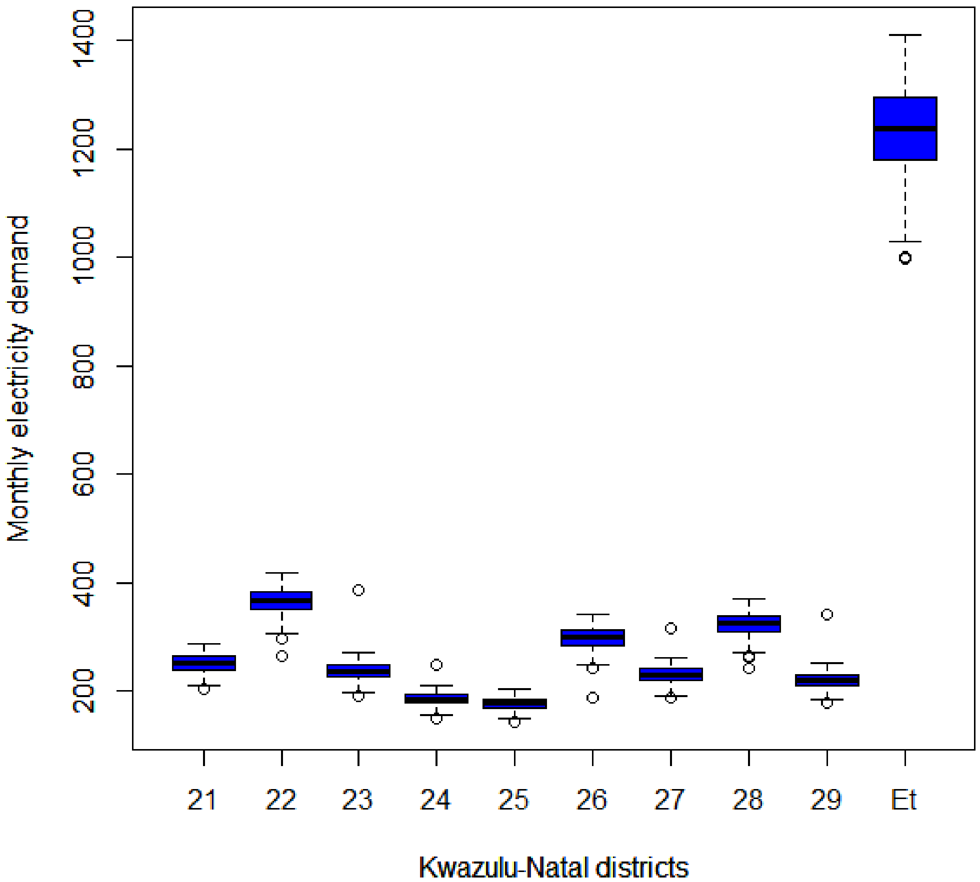


FIGURE A.5: Box plots for Kwazulu-Natal districts monthly electricity demand data from January 2002 to October 2020.

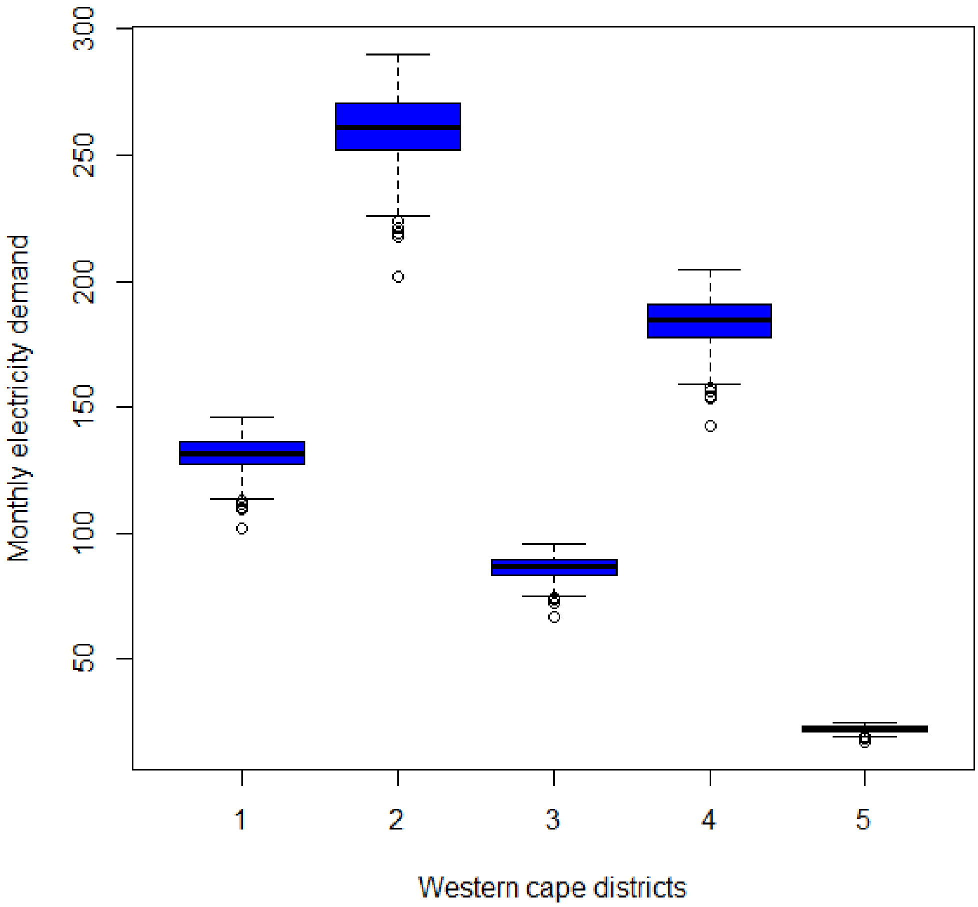


FIGURE A.6: Box plots for Western cape districts monthly electricity demand data from January 2002 to October 2020.

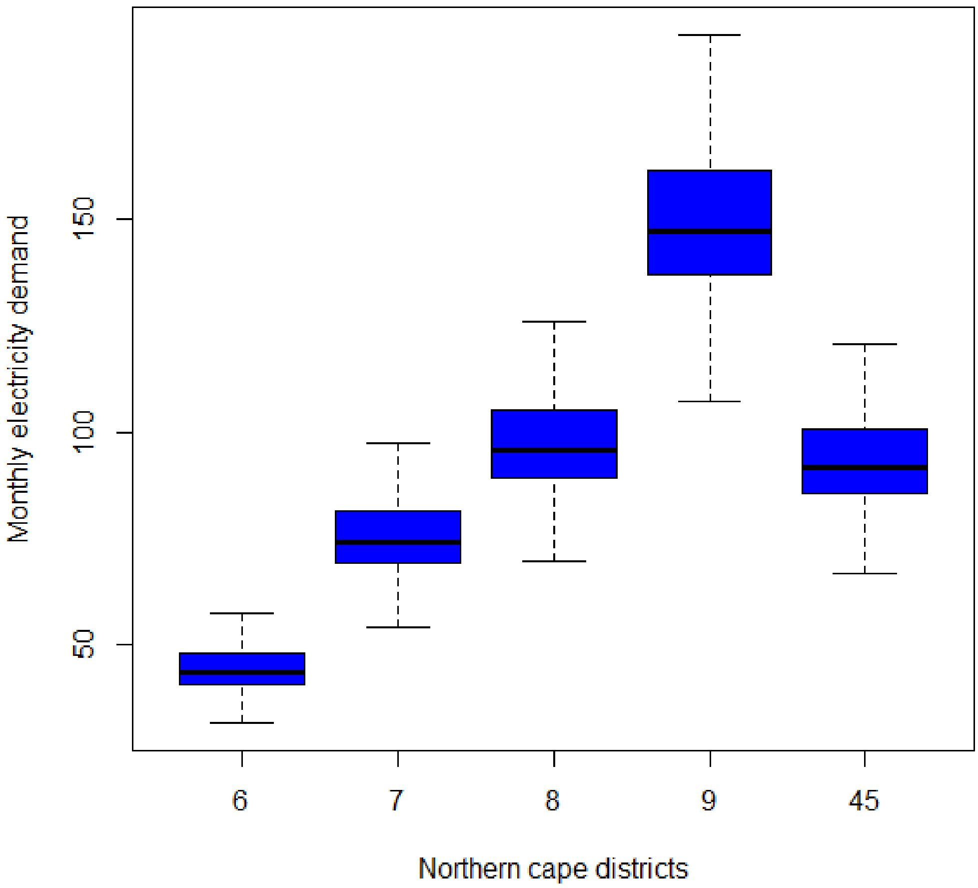


FIGURE A.7: Box plots for Northern cape districts monthly electricity demand data from January 2002 to October 2020.

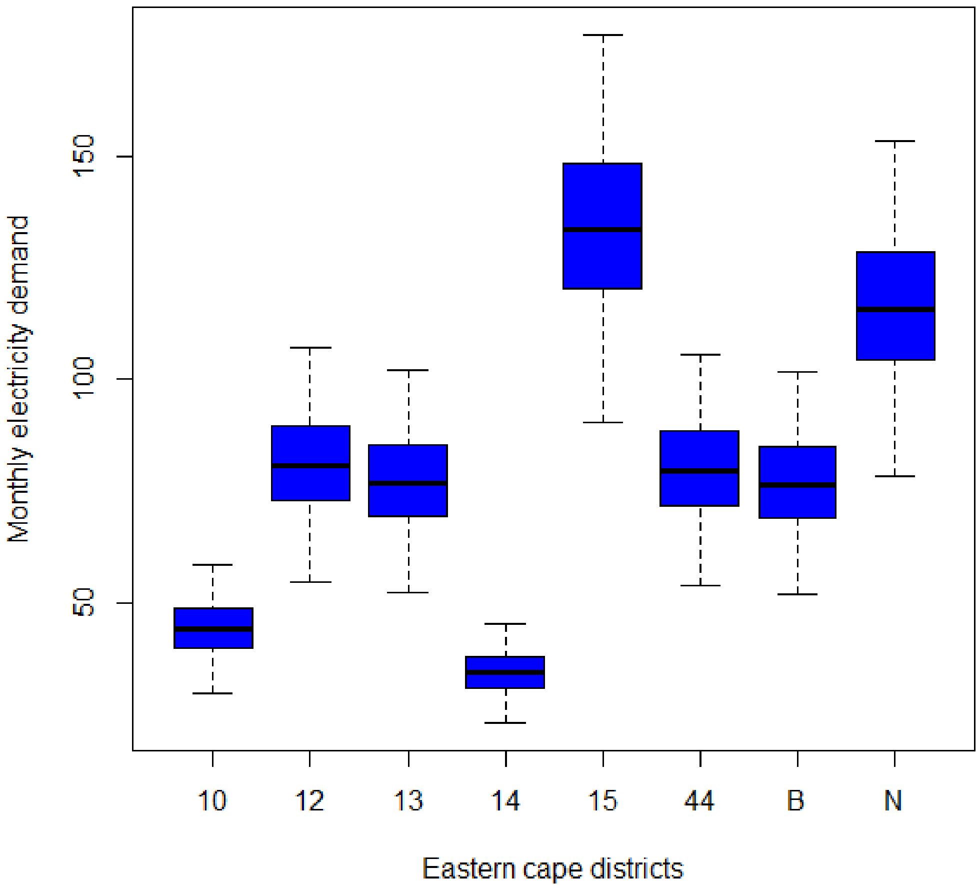


FIGURE A.8: Box plots for Eastern cape districts monthly electricity demand data from January 2002 to October 2020.

Appendix B

Forecast of the bottom-up approach, two top–down approaches and optimal combination approach

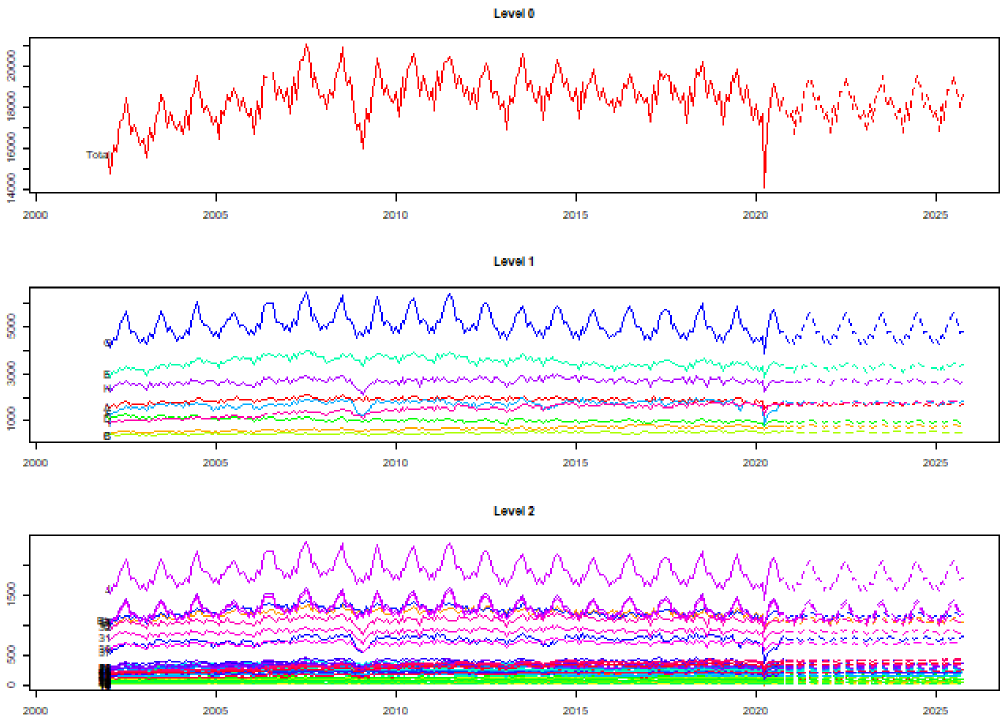


FIGURE B.1: bottom-up approach forecasts in Gigawatt-hours (GWh).

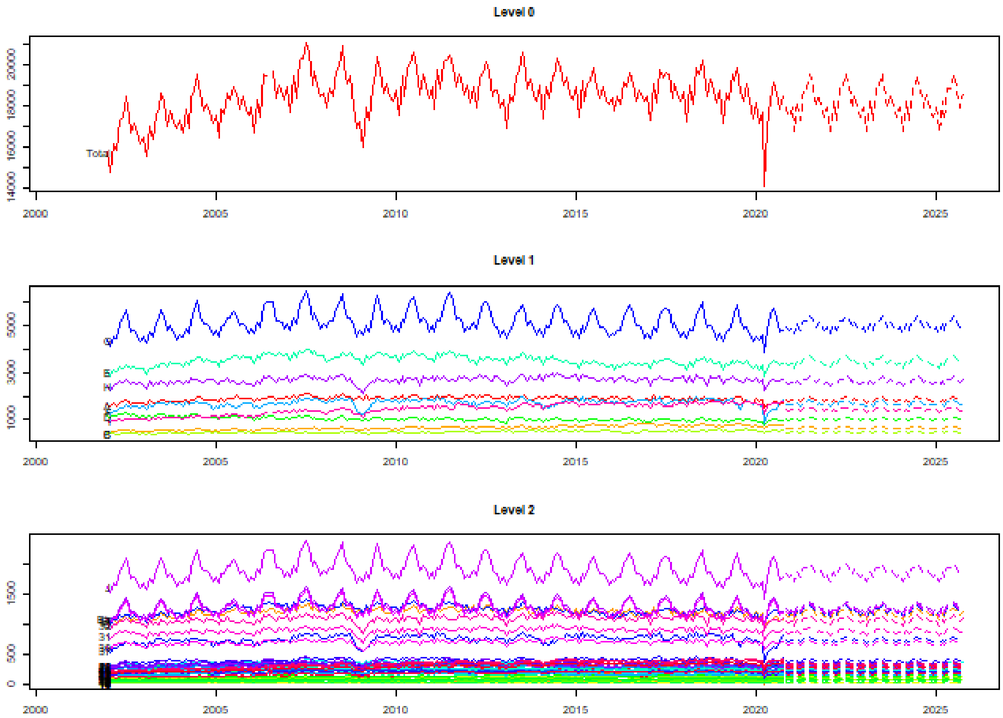


FIGURE B.2: Top-down HP1 approach forecasts in Gigawatt-hours (GWh).

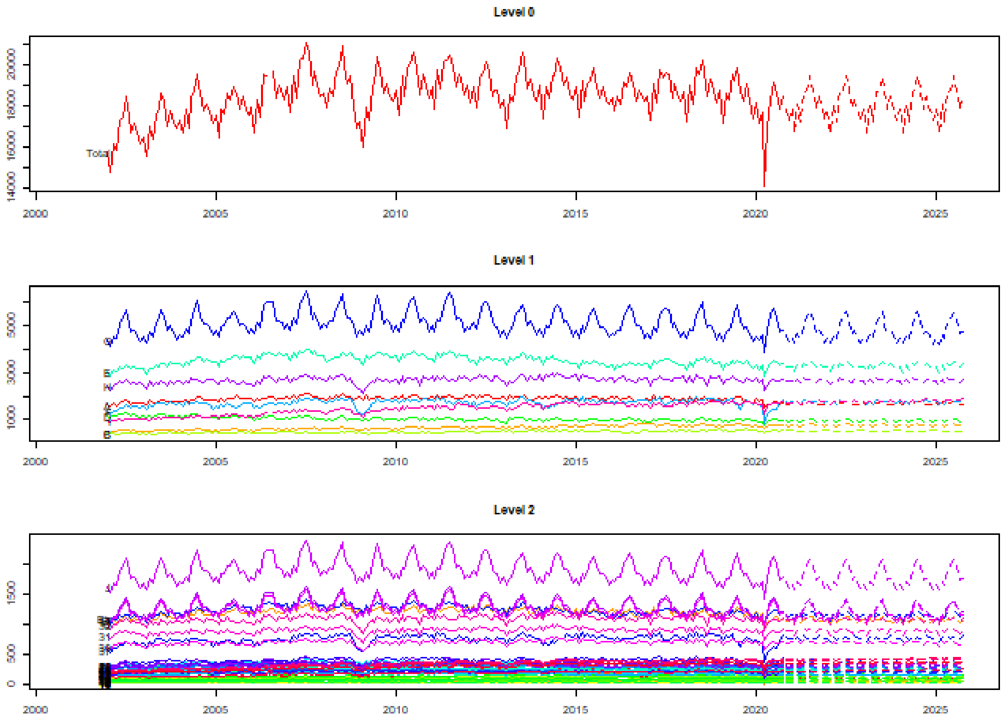


FIGURE B.3: Top-down HP2 approach forecasts in Gigawatt-hours (GWh).

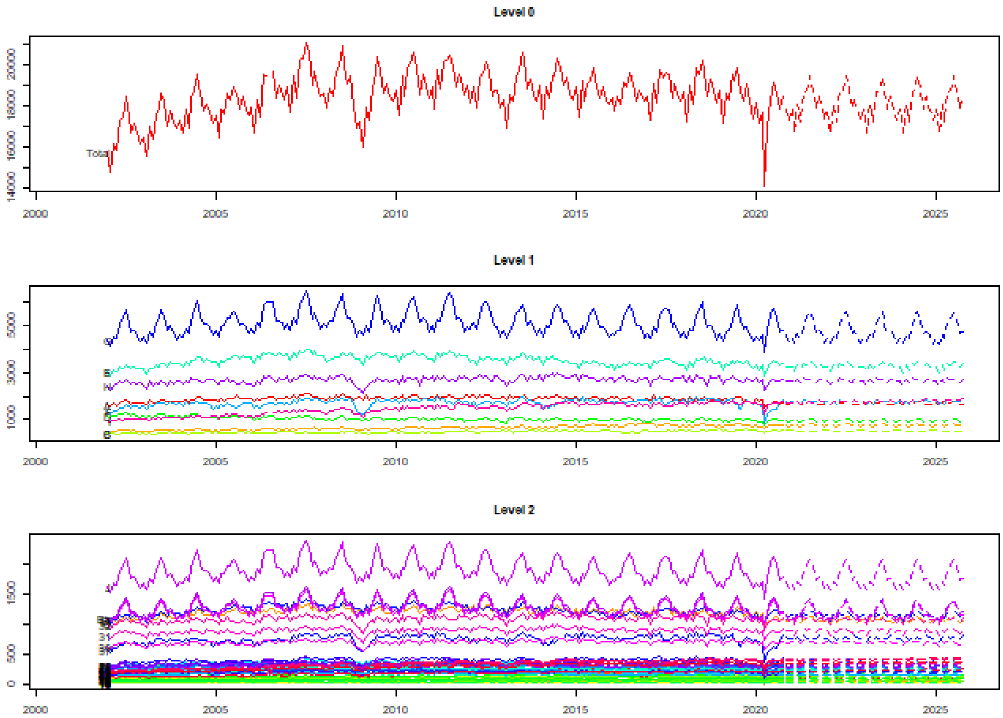


FIGURE B.4: Optimal combination approach forecasts in Gigawatt-hours (GWh).

Appendix C

Some selected R Code

```
#####################################
### IGNITIOUS 2021 MSc mini-dissertation
###################################

attach(data2levels)
head(data2levels)
win.graph()

##HIERACHICAL FORECASTING

library(forecast)
library(hts)
library(thief)
win.graph(width=7,height=5,pointsize=8)

####

#analyticdata
energydata<-as.matrix(data2levels)
energydata
y<-ts(energydata, start=c(2002, 01), end=c(2020,10),
      frequency=12)
y
# Convert the data into an hierarchy
#y<- hts(y, nodes=list(9))
y <- hts(y, nodes=list(9, c(6,8,5,5,10,4,5,3,5)))


#bts <- ts(5 + matrix(sort(rnorm(500)), ncol=5, nrow=100))


```

```r
# Return the hierarchy
y
# Return the labels of the nodes at each level
y$labels
# Return how the nodes are organised
y$nodes

#Ploting a time series plot for all levels
plot(y)

ally <- aggts(y)
ally
somey <- aggts(y, levels = c(0, 2))
somey
S <- smatrix(y)
S
#------------------------
##FORECASTING ACCURACY
#---------------------------

analyticdata<- window(y, start=c(2002,1), end=c(2019,10))
test <- window(y, start=c(2019,11), end=c(2020,10))

forecas <- forecast(analyticdata, h=12, method="bu",
                    fmethod="ets")
forecas
accuracy.gts(forecas, test)

forecas<- forecast(analyticdata, h=12, method="comb",
                   fmethod="ets")
accuracy.gts(forecas, test)
```

```r
forecas<- forecast(analyticdata, h=12, method="tdgsa",
                   fmethod="ets")
accuracy.gts(forecas, test)

forecas<- forecast(analyticdata, h=12, method="tdgsf",
                   fmethod="ets")
accuracy.gts(forecas, test)

forecas<- forecast(analyticdata, h=12, method="tdfp",
                   fmethod="ets")
accuracy.gts(forecas, test)


#################################################################
```

```r
#---------------------------------
##FORECASTING 60 month ahead
## Bottom-up approach
#------------------------------------------

allf <- forecast.gts(y, h=60,method="bu", fmethod="ets")

# Summary of the output
allf

# All-levels forecasts
allts(allf)
# Bottom-level forecasts
allf$bts
# Plot all forecasts
win.graph(width=7,height=5,pointsize=8)
plot(allf)

write.table(allts(allf),"Buforecasts.txt",sep="\t")


#-----------------------------
##FORECASTING 60 month ahead
## Topdown approach
#-------------------------------

allf <- forecast.gts(y, h=60,method="tdgsa", fmethod="ets")

```

```
# Summary of the output
allf

# All-levels forecasts
allts(allf)
# Bottom-level forecasts
allf$bts
# Plot all forecasts
plot(allf)
#library("xlsx")

# Save multiple objects

write.table(allts(allf),file = "Forecast.csv",sep=",")


#-------------------------------
##FORECASTING 60 month ahead
## Topdown approach
#------------------------------------

allf <- forecast.gts(y, h=60,method="tdgsf", fmethod="ets")

# Summary of the output
allf

# All-levels forecasts
allts(allf)
# Bottom-level forecasts
allf$bts
```

```
# Plot all forecasts
win.graph(width=7,height=5,pointsize=8)
plot(allf)
#library("xlsx")

# Save multiple objects

write.table(allts(allf),file = "Forecast1.csv",sep=",")



##FORECASTING 60 month ahead
## Topdown approach

allf <- forecast.gts(y, h=60,method="tdfp", fmethod="ets")

# Summary of the output
allf

# All-levels forecasts
allts(allf)
# Bottom-level forecasts
allf$bts
# Plot all forecasts
win.graph(width=7,height=5,pointsize=8)
plot(allf)
library("xlsx")
```

```
# Save multiple objects

write.table(allts(allf),file = "Forecast.csv",sep=",")




##FORECASTING 60 month ahead
## Optimal combination approach

allf <- forecast.gts(y, h=60,method="comb", fmethod="ets")

# Summary of the output
allf

# All-levels forecasts
allts(allf)
# Bottom-level forecasts
allf$bts
# Plot all forecasts
win.graph(width=7,height=5,pointsize=8)
plot(allf)

# Save multiple objects

write.table(allts(allf),file = "Forecast5.csv",sep=",")
```

```
attach(Top_down_fp)
head(Top_down_fp)
win.graph()
par(mfrow=c(2,2))
TotalTpfo<-ts(Forecast)

plot(TotalTpfo, type = "l", xlab="Observation number",ylab="Total forecast
     col="blue")
plot(density(TotalTpfo),col="blue")
qqnorm(TotalTpfo, col="blue")
qqline(TotalTpfo)
boxplot(TotalTpfo, horizontal = T)

par(mfrow=c(1,1))
plot(TotalTpfo, type = "l", xlab="Observation number",ylab="Total forecast
z = smooth.spline(time(TotalTpfo), TotalTpfo)
z

lines(smooth.spline(time(TotalTpfo),
                    TotalTpfo, spar=0.04),col="red",lwd=2) #spar=0.3792046
Bufitted = fitted((smooth.spline(time(TotalTpfo),
                                 TotalTpfo, spar=0.04)))
Bufitted
#write.table(BUfitted,"~/nolfits.txt",sep="\t")

plot(density(Bufitted),col="blue")
```

```
### PREDICTION INTERVALS METHOD 1 BASED ON LINEAR REGRESSION

data <- cbind(Forecast,Bufitted)
data <- data.frame(data)
mod1 <- lm(Forecast~Bufitted, data=data)
summary(mod1)

f <- cbind(Top_down_fp$Forecast,predict(mod1,interval='prediction'))[1:12
head(f)
write.table(f,"~/predictionint.txt",sep="\t") # 95% prediction intervals

write.table(f,file = "predict.csv",sep=",")


#win.graph()
head(Top_down_fp)
plot(Forecast, type = "l", xlab="Observation number", ylim=c(16000,21000)
     ylab="Total forecasts (Top-down FP)")
lines(lwr,col="red", lty=2, lwd=2)
lines(upr,col="blue", lty=3,lwd=2 )
legend("topleft",col=c("black","red", "blue"), lwd=2,lty=1:3,
        legend=c("Total forecasts", "Lower PL", "Upper PL"))

### PREDICTION INTERVALS METHOD 2 BASED ON
# LinearQUANTILE REGRESSION

library(quantreg)
```

```

win.graph()
plot(TotalTpfo, type = "l", xlab="Observation number",ylab="Total forecas
qr.g = rq(TotalTpfo~ Bufitted,data= Top_down_fp, tau=0.975) #tau = 0.025,

summary.rq(qr.g,se="boot") # can use se = "nid" or se="ker"

lines(qr.g$fit, col="red")

fQRA1 = fitted(qr.g)
write.table(fQRA1,"~/QRA05.txt",sep="\t") #LL0025, QRA05, UL0975
```

References

[1] M. Abolghasemi, R.J. Hyndman, G. Tarr, and C. Bergmeir. "Machine learning applications in time series hierarchical forecasting". In: *arXiv preprint arXiv:1912.00370* (2019).

[2] V. Almeida, R. Ribeiro, and J. Gama. "Hierarchical time series forecast in electrical grids". In: *Information Science and Applications (ICISA) 2016*. Springer, 2016, pp. 995–1005.

[3] G. Athanasopoulos, R.A. Ahmed, and R.J. Hyndman. "Hierarchical forecasts for Australian domestic tourism". In: *International Journal of Forecasting* 25.1 (2009), pp. 146–166.

[4] J.M. Bates and C.W.J. Granger. "The combination of forecasts". In: *Journal of the Operational Research Society* 20.4 (1969), pp. 451–468.

[5] J.M. Bates and C.W.J. Granger. "The combination of forecasts". In: *Journal of the Operational Research Society* 20.4 (1969), pp. 451–468.

[6] G. Fliedner. "Hierarchical forecasting: issues and use guidelines". In: *Industrial Management & Data Systems* (2001).

[7] P. Gaillard and Y. Goude. "Forecasting electricity consumption by aggregating experts; how to design a good set of experts". In: *Modeling and stochastic learning for forecasting in high dimensions*. Springer, 2015, pp. 95–115.

[8] R. Ghedamsi, N. Settou, A. Gouareh, A. Khamouli, N. Saifi, B. Recioui, and B. Dokkar. "Modeling and forecasting energy consumption for residential buildings in Algeria using bottom-up approach". In: *Energy and Buildings* 121 (2016), pp. 309–317.

[9] C.W. Gross and J.E. Sohl. "Disaggregation methods to expedite product line forecasting". In: *Journal of forecasting* 9.3 (1990), pp. 233–254.

[10] C.W. Gross and J.E. Sohl. "Disaggregation methods to expedite product line forecasting". In: *Journal of forecasting* 9.3 (1990), pp. 233–254.

[11] "https://cran.r-project.org/web/packages/fable/fable.pdf". In: ().

[12] R.J. Hyndman, R.A. Ahmed, G. Athanasopoulos, and H.L. Shang. "Optimal combination forecasts for hierarchical time series". In: *Computational statistics & data analysis* 55.9 (2011), pp. 2579–2589.

[13] R.J. Hyndman and G. Athanasopoulos. "Optimally reconciling forecasts in a hierarchy". In: *Foresight: The International Journal of Applied Forecasting* 35 (2014), pp. 42–48.

[14] R.J. Hyndman, G. Athanasopoulos, C. Bergmeir, G. Caceres, L. Chhay, M. O'Hara-Wild, F. Petropoulos, S. Razbash, E. Wang, and F. Yasmeen. "Forecasting functions for time series and linear models". In: (2018).

[15] K.B. Kahn. "Revisiting top-down versus bottom-up forecasting". In: *The Journal of Business Forecasting* 17.2 (1998), p. 14.

[16] I. Kanda and J.M.Q. Veguillas. "Data preprocessing and quantile regression for probabilistic load forecasting in the GEFCom2017 final match". In: *International Journal of Forecasting* 35.4 (2019), pp. 1460–1468.

[17] R. Koenker and K.F. Hallock. "Quantile regression". In: *Journal of economic perspectives* 15.4 (2001), pp. 143–156.

[18] Roger Koenker, Stephen Portnoy, Pin Tian Ng, Achim Zeileis, Philip Grosjean, and Brian D Ripley. "Package 'quantreg'". In: *Reference manual available at R-CRAN: https://cran. rproject. org/web/packages/quantreg/quantreg. pdf* (2018).

[19] W.G.S. Konarasinghe and N.R. Abeynayake. "Modelling stock returns of individual companies of Colombo Stock Exchange". In: *Conference Proceedings of the 1st International Forum for Mathematical Modelling*. 2014, pp. 111–115.

[20] Z. Li, S.M.M. Rahman, R. Vega, and B. Dong. "A hierarchical approach using machine learning methods in solar photovoltaic energy production forecasting". In: *Energies* 9.1 (2016), p. 55.

[21] L.R. Lima, F. Meng, and L. Godeiro. "Quantile forecasting with mixed-frequency data". In: *International Journal of Forecasting* 36.3 (2020), pp. 1149–1162.

[22] T. Makoni, D. Chikobvu, and C. Sigauke. "Hierarchical Forecasting of the Zimbabwe International Tourist Arrivals". In: *Statistics, Optimization & Information Computing* 9.1 (2021), pp. 137–156.

[23] P. Mpfumali, C. Sigauke, A. Bere, and S. Mulaudzi. "Probabilistic solar power forecasting using partially linear additive quantile regression models: An application to South African data". PhD thesis. 2019.

[24] R. Netshiomvani. "Hierarchical forecasting of electricity demand in South Africa". PhD thesis. 2020.

[25] J. Nowotarski, E. Raviv, S. Trück, and R. Weron. "An empirical comparison of alternative schemes for combining electricity spot price forecasts". In: *Energy Economics* 46 (2014), pp. 395–412.

[26] Y. Pang, B. Yao, X. Zhou, Y. Zhang, Y. Xu, and Z. Tan. "Hierarchical Electricity Time Series Forecasting for Integrating Consumption Patterns Analysis and Aggregation Consistency." In: (2018), pp. 3506–3512.

[27] M. Park and M. Nassar. "Variational Bayesian inference for forecasting hierarchical time series". In: *International conference on machine learning (ICML), workshop on divergence methods for probabilistic inference, Beijing*. Citeseer. 2014.

[28] C. Roach. "Reconciled boosted models for GEFCom2017 hierarchical probabilistic load forecasting". In: *International Journal of Forecasting* 35.4 (2019), pp. 1439–1450.

[29] L. Roque, L. Torgo, and C. Soares. "Automatic Hierarchical Time-Series Forecasting Using Gaussian Processes". In: *Engineering Proceedings*. Vol. 5. 1. Multidisciplinary Digital Publishing Institute. 2021, p. 49.

[30] T. Silveira Gontijo and M. Azevedo Costa. "Forecasting hierarchical time series in power generation". In: *Energies* 13.14 (2020), p. 3722.

[31] S.B. Taieb, J.W. Taylor, and R.J. Hyndman. "Hierarchical probabilistic forecasting of electricity demand with smart meter data". In: *Journal of the American Statistical Association* (2020), pp. 1–17.

[32] S.B. Taieb, J. Yu, M. Barreto, and R. Rajagopal. "Regularization in hierarchical time series forecasting with application to electricity smart meter data". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 31. 1. 2017.

[33] J.W. Tukey. *Exploratory data analysis: past, present and future*. Tech. rep. PRINCETON UNIV NJ DEPT OF STATISTICS, 1993.

[34] S.L. Wickramasuriya, G. Athanasopoulos, and R.J. Hyndman. "Forecasting hierarchical and grouped time series through trace minimization". In: *Department of Econometrics and Business Statistics, Monash University* 105 (2015).