DEVELOPMENT AND EVALUATION OF DATA-DRIVEN MODELS FOR ELECTRICITY DEMAND FORECASTING IN QUEENSLAND, AUSTRALIA

A Thesis Submitted by

TOBIAS KUMIE

BEng

For the Award of

Master of Science Research (MSCR)

2020

ABSTRACT

Queensland (QLD) is the second largest state in Australia, with a growing demand for electricity, but existing studies appear to lack their ability to accurately model the consumer demand for electricity. In this Master of Science Research (MSCR) thesis, two kinds of hybrid forecasting models were developed by integrating the Extreme Learning Machines (ELM) with a Markov Chain Monte Carlo (MCMC) algorithm based bivariate copula model (ELM-MCMC) and also, a conditional bivariate copula model to probabilistically forecast the electricity demand (*D*). The study has incorporated statistically significant lagged electricity price (*PR*) datasets as a non-linear regression covariate into the final *D*-forecasting model.

In the first objective of the MSCR thesis, the ELM model was trained using statistically significant historical electricity demand at (*t*–1) timesteps for the state of Queensland used as a predictor variable, derived from Partial Autocorrelation Functions (PACF). This represented historical usage patterns in the electricity demand datasets used to forecast the future usage. It was then tested against current electricity demand (*D*(*t*)) to forecast the future *D* values. The output (*i.e.*, simulated and observed tested *D* values) from the independent test dataset of the ELM model was used as the input for the MCMC-based copula model to derive the best copula model and to further improve forecasting accuracy. This involved the adoption of twenty-six copulas (*e.g.*, Gaussian, *t*, Clayton, Gumble, Frank, etc.) and enabled us to also rank the best copulas based on the Akaike Information Criterion *(AIC*), Bayesian Information Criterion (*BIC*), and Maximum Likelihood (*MaxL*) to establish the dependence of historical *D* with the current and future *D* values. The results for the ELM-MCMC copula-based model outperformed both of its counterpart models (*i.e.* MCMC copula-based model and the standalone ELM model) based on vigorous statistical performance metrics. For 6 and 12-hours timescales, the MCMC-Fischer-Hinzmann copula yielded the highest Legates and McCabe Index (L_M) (0.98 and 0.98), and lowest error terms including root mean square error (*RMSE*) (285.480 and 534.090), relative root mean square error (*RRMSE*) (0.348 and 0.320%), mean absolute error (*MAE*) (262.241 and 490.661 MW), relative mean absolute error (*RMAE*) (0.336 and 0.309 %), *AIC* (-63136.102 and -34727.466), *BIC* (-63125.530 and -34718.279), and Max_L (51570.051 and 17365.733), respectively. Similarly, for the daily timescale, the ELM-MCMC-Cuadras-Auge copula outclassed its counterpart models by

displaying L_M (0.98), *MSE* (482703.8 MW), *RMSE* (694.769 MW), *RRMSE* (0.220 %), *MAE* (638.365 MW), *RMAE* (0.208 %), *AIC* (-14514.312), *BIC* (-14510.412), and Max_L (7258.156).

These present results indicated that the hybrid ELM-MCMC copula-based model had an excellent performance, evidenced by attaining less than 10% *RRMSE* and *RMAE*, and Legates McCabe value close to unity. This is further supported by better model fits as denoted by lower *AIC* and *BIC* values and small residual error between observed and predicted data as indicated in higher Max_L values for the respective timescales.

In another phase of this study, we explored the ability of both local and global optimization techniques in achieving the best parameter estimate for the 26 copulas. It has shown that the global MCMC optimization method delivers accurate parameter estimates for 6 and 12-hours timescales whilst presenting information on the posterior distribution by computing uncertainty range of parameter values within a Bayesian framework. The local method appeared to provide better estimates of copula parameters for the daily timescale of *D*-forecasting.

In the second objective of the MSCR thesis, this study has developed a conditional bivariate copula model to probabilistically forecast electricity demand by incorporating the significant lagged electricity price (*PR)* from the Australian Energy Market Operator (AEMO) as a covariate into the final *D*-forecasting model. The use of energy price data to predict the energy demand is an important contribution given the relationships between these variables are well established. This objective resulted in the bivariate BB7 and BB8 copulas as being ranked highly for the probabilistic forecasting of *D* at a timescale of 30 minutes, 1-hour, and daily. The conditional exceedance probability of electricity demand greater than 7000 MW, 14000 MW, and 360000 MW for 30-minutes, 1-hour, and daily timescales given their respective prices greater than AU$25/MWh, AU$60/MWh, and AU$165/MWh predicted to be 20%, 30%, and 50% respectively. Similarly, the conditional non-exceedance probability of electricity demand greater than 7000 MW, 14000 MW, and 360000 MW for 30-minutes, 1-hour, and daily timescales given their respective prices greater than AU$25/MWh, AU$60/MWh, and AU$165/MWh was predicted to be 80%, 72%, and 70% respectively.

When benchmarked with literature, the proposed research methodologies for objective 1 (*i.e*., projection of demand based on antecedent behaviour) and objective 2 (*i.e*., projection of demand based on antecedent energy price data) appear to be versatile tools possessing a robust predictive capability for forecasting *D* in Queensland, Australia. Hence, this research project is the first to develop and test these novel techniques, especially using price as regression covariate to forecast demand to achieve high forecasting accuracy, when the models are applied for multiple forecasting horizons of 30-minutes, 1-hour, 6-hourly, 12-hourly, and daily. It is noted that these timescales are relevant for stakeholders (*e.g.*, energy utilities) to develop decision systems for better energy security, and can potentially be adopted in real power grid operations to ensure stability, cost reduction and improved efficiency whilst granting consumer satisfaction.

In summary, the novel energy demand modelling techniques presented here can help address research gaps in electricity usage monitoring sector by making a significant contribution towards improved forecasting accuracy of energy demand. While this study has currently been limited to Queensland, the research findings are immensely useful for energy experts in the National Energy Markets elsewhere including supporting the work of AEMO, Energex and other companies to enhance their energy forecasting and monitoring skills. These can assist in informed decisions and addressing the growing challenges within electricity industry, through improving energy demand and price monitoring, consumer satisfaction and maximized profitability endeavours of energy companies.

CERTIFICATION OF THESIS

This Thesis is entirely the work of TOBIAS KUMIE except where otherwise acknowledged. The work is original and has not previously been submitted for any other award, except where acknowledged.

Principal Supervisor : Associate Professor Ravinesh C. Deo

Associate Supervisor : Dr Thong Nguyen

Associate Supervisor : Dr Nawin Raj

Associate Supervisor : Dr Mumtaz Ali at Deakin University

Student and supervisors signatures of endorsement are held at the University.

ACKNOWLEDGEMENTS

I wish to thank and acknowledge various people for their distinguished contributions toward my project.

Firstly, I would like to express my deepest gratitude to Associate Professor Ravinesh Deo for his tireless effort in providing much needed support in ensuring my application for the Australian Awards Scholarship was successful. I also thank him highly for his endless commitment, courage, and motivation injected into my scholarly life to ensure the successful completion of my master's program. Furthermore, I deeply thank him for his expert advice and support rendered in formulating my project and his continuous guidance throughout the entire project.

I would also like to extend my sincere gratitude to my Associate Supervisor, Dr Thong Nguyen for his concepts in developing the bivariate copula models and his guidance on the same. On the same note, I also thank my Associate Supervisor, Dr Nawin Raj, for providing useful feedback and editing the thesis. Moreover, I greatly thank my External Supervisor, Dr Mumtaz Ali for his expert advice and direction on the development of the Monte Carlo Markov Chain copula-based model, and valuable comments throughout the project.

In addition, I duly appreciate and acknowledge the Australian Department of Foreign Affairs and Trade through the Australian Awards Scholarship scheme for funding this study and the University of Southern Queensland for accepting me in securing the placement under the school of sciences, which collectively made this study feasible.

I also acknowledge and extend my gratitude to the Australian Energy Market Operator for accessing its aggregated electricity demand data for the state of Queensland, Australia.

Furthermore, I extend my thanks to Shirley Munks for proof editing the thesis in terms of grammar, punctuation, and sentence structure. I also acknowledge and appreciate Karin Hoskin for her professional service in copyediting my thesis.

Moreover, I extend my gratitude to the Advanced Data Analytics: Environmental Modelling and Simulation Research group for communicating research ideas, particularly in advanced data analysis technique via email collaborations and personal discussions.

Finally, this research would not have been possible without the continuous support of both my immediate and extended families and friends in Papua New Guinea.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AEMO	Australian Energy Market Operator
AI	Artificial Intelligence
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
BIC	Bayesian Information Criterion
BMA	Bayesian Model Averaging
CCF	Cross-correlation Function
CDF	Cumulative Distribution Function
D	Electricity load (demand)
ECMWF	European Centre for Medium Range Weather Forecast
ELM	Extreme Learning Machine
E_{NS}	Nash–Sutcliffe Coefficient
GDP	Gross Domestic Product
GP	Genetic Programming
ICEEMDAN	Improved version of Empirical Mode Decomposition with Adaptive Noise
L_M	Legates and McCabe's Index
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MARS	Multivariate Adaptive Regression Splines
MCMC	Monte Carlo Markov Chain
ML	Machine Learning
MLR	Multiple Linear Regression

MSE	Mean Square Error
MvCAT	Multivariate Copula Adaptive Toolbox
MW	Megawatts
NEM	National Electricity Market
PACF	Partial Autocorrelation Function
PIT	Probability Integral Transform (PIT)
PR	Electricity Price
QLD	Queensland
r	Correlation Coefficient
RMAE	Relative Mean Absolute Error
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
SDG	Sustainable Development Goal
SILO	Scientific Information for Landowners
SLFN	Single Layer Feed-Forward Neural Network
STLF	Short-Term Load Forecasting
SVR	Support Vector Regression
UN	United Nations
WI	Willmott's Index of Agreement
WT	Wavelet Transformation

CHAPTER 1: INTRODUCTION

1.1 Background

In this contemporary world, electricity is increasingly used as a necessity for human advancement and survival. Electricity demand forecasting is a vital tool for effective energy management, operation, and policy planning in the energy sector. Innovative data-driven models are powerful tools that possess accurate forecasting capability, hence, they are required by electricity operators and other stakeholders along the energy value chain to make informed decisions for improved performance, increased reliability, reduced cost, and above all, avoid energy crisis. Such models are important platforms for enhancing key operational parameters of the National Electricity Market (NEM) including the determination of electricity generation capacity, price regulation and the introduction of new policies to reflect changes and make room for upcoming expansion of electricity distribution networks.

Electricity is an essential commodity in contemporary world and is vital for socio-economic and infrastructure development to ensure economic growth and success in any country. Electricity load (predominantly called demand) (abbreviated as *D* in this thesis and measured in megawatts, MW) forecasting is essential for enriching and upgrading system deficiencies in ensuring adequate generation and consistent electricity supply to its distribution networks, which in turn can positively impact development outcomes (Toman & Jemelkova 2003). The global electricity markets have confirmed forecasting irregularities and uncertainties present in traditional *D*-forecasting techniques, which have led to substantial financial losses for electric utility companies (Bunn & Farmer 1985; Haida & Muto 1994; Fan, S. & Chen, L. N. 2006). Therefore, they have emphasised and embarked on researchers to strategize and develop robust forecasting models that hold high accuracy to deliver reliable forecasting solutions. This also helps to meet the United Nations (UN) Sustainable Development Goal (SDG) 7 (Assembly 2015), which facilitates smooth transition from non-renewable to renewable energy sources. This is done to mitigate climate change treats whilst providing a safe and vibrant energy systems to meet consumer demands in real-time. It is projected that during the imminent shift to sustainable energy system, electricity consumption will increase (Akay & Atak 2007; Campbell 2018) to ensure tranquility during this transition phase. To achieve this smooth transition, new and robust energy models are needed for precise capacity planning as old *D*-

forecasting methods are reported to have exhibited inaccurate results when dealing with complex datasets of various predictors.

1.2 Classification of electricity demand forecasting method

Generally, electricity demand forecasting models are classified as either physical or data-driven models. Physical models are driven by physics and use mathematical equations to solve problems. They require initial (boundary) conditions to force the model to work and can be non-linear and complex. In contrast, Artificial Intelligence models are cheap and easy to develop and do not require initial conditions. They can learn from historical features of the data using computer algorithms and are purely dependent on data, hence are called "data-driven models". These models use data analytics techniques where big data are examined to reveal hidden patterns, trends, periodicity and "cause and effect" information necessary to predict target variables. They do not require mathematical equations (except a transfer function to analyse patterns in datasets) nor model initialization and are self-adaptive and highly automated to provide real-time solutions.

Energy forecasting is usually done using qualitative or quantitative methods, depending on the degree of statistical or mathematical forecasting models. The qualitative forecasting technique is widely applied by forecasters in forward planning of electricity load generation and distribution. It involves an expert knowledge system where historical electricity consumption data across the spectrum of the electricity sector is analysed to reach common consensus to forecast future electricity load requirements. This includes the Delphi method, curve fitting and advanced comparative analysis (Dalkey & Helmer 1963; Suganthi & Jagadeesan 1992; Haida & Muto 1994). Conversely, quantitative forecasting methods are categorized as deterministic or data driven. The deterministic methods are governed by mathematical equations while data-driven methods purely use past data to train computer systems using a set of soft algorithms to learn the prevailing relationships, trends, seasonality, and cause and effect to rapidly produce logical and accurate results for future projections. Data-driven models (predominantly called machine learning (ML) models) use historical datasets to teach computer which emulate the human thinking and produce logical and accurate forecasting solutions (Wang & Ramsay 1998). Therefore, they have been used extensively (Florens et al. 2007; Xydas et al. 2016) in the contemporary world due to their powerful computational ability (Suganthi & Samuel 2012) in generating precise forecasting results for complex systems (Haida & Muto 1994) such as time-series modelling.

1.3 Classifications for electricity demand forecasting horizons

Electricity demand forecasting is usually classified into time intervals based on the duration of the forecast range. This includes short-term, medium-term, and long-term forecasting. Short-term forecasts commonly range from half an-hour to a week and are vital for the daily routine planning and management of the utility systems to ensure stability and continuity. Likewise, the medium-term forecasts commonly range from weeks to a year. They are used as decision support tools for the timely procurement of field consumables (for example, fuel and spare parts), planning for preventive maintenance of utility systems, and assessing sales and income. Long-term forecasts are usually greater than a year and are essential for long-term planning of infrastructure development for the utility system to cater for future expansion.

Different modelling techniques have been used globally for short, medium, and long-term *D* and price *(PR)* forecasting depending on the duration of forecasting (Al-Alawi & Islam 1996; Mamun & Nagasaka 2004; Chui et al. 2009; Zhang et al. 2013; Papadopoulos & Karakatsanis 2015; Bello et al. 2016; Ziel & Steinert 2018; Cao et al. 2019; da Silva et al. 2019).

1.4 Complexity of electricity demand forecasting

The forecasting of *D* is a multivariate issue comprising many exogenous features such as the local GDP, population, and climate variables. As such, it cannot be appropriately modelled in time series using a bivariate normal distribution (Alexander 2004) with simple statistical methods. However, this problem can be addressed using copulas, which are advanced statistical methods. Multivariate normality can be accommodated by capturing essentials features of the *D* datasets such as asymmetry, non-linear dependence, and heavy-tail behaviour using ranked Spearman or Kendall tau coefficients. Further, this enables independent modelling of marginal selections and dependence structure. In previous studies, He et al. (2017) used copula for short-term *D* probability density forecasting. Similarly, Grégoire et al. (2008) used copulas to model price in energy trades. Therefore, this project firstly develops and tests a hybrid ELM-MCMC copula-based model to achieve robust forecasting results. Furthermore, it also investigates and evaluates the joint predictive distribution of *D* for probabilistic forecasting by utilizing conditional bivariate copula whereby the electricity price (*PR*) is used as a covariate predictor.

1.5 Forecasting model for the national electricity market in Australia and its limitation

The national electricity market (NEM) in Australia was established in December 1998 and operates the world's biggest integrated electricity systems comprising five states, namely Tasmania, South Australia, Queensland, Victoria, and New South Wales (Clements et al. 2016). Electricity in these five states is supplied from generators (sellers), which are managed by Energex and the Australian Energy Market Operator (AEMO) whereby they organize trade between customers and sellers. In particular, the AEMO adopts semi-parametric-type additive models for forecasting *D* in these respective regions. However, the emergence of technological advancement in the energy sector, infrastructure development, climate change, and economic interventions have downplayed the forecasting capability of the existing model to forecast *D* accurately. Therefore, there is a need for the development of robust models for accurate *D* projection to meet the growing consumer demands in real-time whilst optimizing operation and maximizing profitability.

This Master of Science Research Thesis is centred on developing models for short-term electricity demand forecasting, which is an essential service in infrastructural and economic development for the prosperity of any nation.

The timely and adequate provision of electricity supply by any utility company is vital for economic growth and industrialization in any county (Morimoto & Hope 2004; Altinay & Karagol 2005; Nwankwo & Njogo 2013). Lacking such supply can deter economic development, hence hindering the prosperity of any nation (Stern et al. 2019). This study is focused on Queensland (the second largest state in Australia) where demand for electricity consumption has experienced significant growth in the recent past (2014 - 2015) consuming 24.4% of the national consumption (Ball et al. 2016). Despite Queensland having high demand for electricity consumption, the literature shows that it lacks accurate *D*-forecasting models. A few studies have been conducted recently to address this issue. This includes a study by Wu et al. (2019) which amalgamated ensemble empirical mode decomposition, ELM, and a grasshopper optimization algorithm for daily *D*-forecasting for five states in Australia, particularly New South Wales, QLD, South Australia, Victoria, and Tasmania which exhibited the highest efficiency by means of generating the lowest error terms against four comparative models of ensemble empirical mode decomposition. Also, another study by Benaouda et al. (2006) applied a wavelet-based non-linear multiscale decomposition model for daily *D* forecasting in New South Wales which performed much better than traditional forecasting methods such as multilayer perceptron and recurrent neural network. However, the limitation of these studies is their inability to predict abnormal events such as extreme weather and energy fluctuations.

Therefore, this study tests a novel hybrid copula-based model (extreme learning machine integrated with Markov Chain Monte Carlo, ELM-MCMC) for improved accuracy in *D*-forecasting and compares its performance against the standalone ELM model and MCMC copula-based model. This study is the first to apply this advanced hybrid copula-based model as a novel technique which has never been explored and used anywhere for *D*-forecasting. Its test results present accurate *D*-forecasting solutions which energy utility companies such as AEMO and Energex could adopt to make intelligent business decisions. This would assist in improving forecasting efficiency by delivering accurate and timely response to meet consumer demand through process optimization whilst maximizing organizational profitability.

1.6 Research Problem

Precise and real-time information on *D* is vital for electricity operators in a competitive electricity market (*e.g*., Australia) as it can present critical information for energy policy formulation and reforms, and assist in making informed decisions on capacity planning to optimize operations and minimize the cost. This is achieved by the development and use of intelligent *D*-forecasting models. The accuracy level of *D* is of utmost importance for electricity operators as *D* is sensitive to slight errors and can account for millions of dollars of financial losses if there is a one percent increase in error value (Haida & Muto 1994; Fan, S. & Chen, L. 2006). It can also cause operational and financial instability in the energy sector and even jeopardise the cohesiveness of electricity markets (Erdogdu 2007). Thus, there is an apparent need for the NEM to have reliable and intelligent predictive models for precise *D*-forecasting to maintain a vibrant and competitive electricity markets as well as contributing to enabling the economic prosperity of the nation in terms of the provision of efficient electricity systems.

Despite QLD being the second largest state in Australia, research into *D*-forecasting has been lacking. A few studies (Al-Musaylh et al. 2018a, 2019; Wu et al. 2019) have been conducted recently in QLD that have addressed the forecasting uncertainties associated with projecting *D*. These studies employed MARS, SVR, ARIMA, WT, B-ANN models that displayed accurate results. However, as stated by Nti et al. (2020), further research interventions are required for *D*-forecasting as it is a complex task that is influenced by multiple predictors such as GDP, demographic, temperature, and technological variables (Nasr et al. 2002; Mirasgedis et al. 2006; Zareipour et al. 2006a; Odhiambo 2009; Suganthi & Samuel 2012; Hu et al. 2013). Therefore, it demands comprehensive research that entails all factors to develop and adopt a robust predictive model for *D*-forecasting in QLD.

1.7 Research Questions

This Master of Science Research (MSCR) thesis entirely based on addressing the three research questions:

1) Can a hybrid predictive model based on Extreme Learning Machine and Markov Chain Monte Carlo (ELM-MCMC) approaches help improve predictive performance in forecasting short-term (*i.e.*, 6-hours, 12-hours, and daily) electricity demand for the State of Queensland?

2) What is the most appropriate, advanced copula-based statistical model that can provide probabilistic forecasting of short-term electricity demand for multiple timescales (*i.e.*, 30-minutes, 1-hour, and daily) for the State of Queensland?

3) What are the differences between global (*i.e.*, Bayesian inference) and local (maximum likelihood) methods for estimating copula parameters for research question one when used to develop electricity demand forecasting models?

The research questions (1 - 3) were addressed by employing both electricity demand and price datasets for Queensland for the period 2017 – 2019.

1.8 Research Aim and Objectives

The primary objective of this study is to develop robust predictive models to improve the short-term (6-hours, 12-hours, and daily) *D*-forecast for the QLD region using historical data of *D* and to compare their performance against existing models for improvements. The three objectives of this project are as follows:

1. To develop hybrid ELM-MCMC based statistical bivariate copula models using the statistically significant lag, which is commonly the first lag (t-1) of historical demand as input for short-term forecasting of *D* in QLD.

2. To advance a multivariate copula approach for probabilistic forecasting of short-term *D* in QLD by applying the antecedent significant lags (t-1) of the average electricity price (*PR*) as a covariate predictor at multiple forecasting horizons (30-minutes, 1-hour, and daily).

3. To evaluate the differences between global (Bayesian inference) and local (maximum likelihood) methods for estimating copula parameters.

To fulfill these objectives, historical *D* data for QLD are used in this study to develop and use the novel hybrid ELM-MCMC model for *D* forecasting in QLD, the second largest state in Australia where there is growing demand for electricity consumption. The hybrid ELM-MCMC model incorporates two or more algorithms utilizing the merits of each model to solve challenges, rather than a single learning model. It is generally accurate since data features are pre-processed and revealed more clearly for forecasting. Also, this study separately investigates and evaluates the conditional bivariate copula model using *PR* as a covariate predictor to improve the short-term forecasting accuracy of *D* at various timescales (6-hours, 12-hours, and daily). It is evident from existing literature that the proposed method has not previously been exploited for probabilistic *D*-forecasting. Consequently, this forecasting technique is tested and evaluated as a novel data intelligent model to reduce forecasting inaccuracies experienced by the NEM.

These objectives are focused on answering the research questions and addressing the research gaps in forecasting inaccuracies for *D* in QLD, thereby providing better decision-making tool in terms of modelling energy demand and pricing.

1.9 Organization of Thesis

This Master's thesis has been written and compiled with the following chapters.

Chapter 1 provides an introduction to the study, with background information pertaining to the importance of *D*-forecasting globally as well as in QLD, Australia. It also defines the types of forecasting models used for *D*-forecasting.

Chapter 2 presents the literature review, detailing existing studies of *D*- forecasting in QLD and their limitations and covers the various forecasting models that have been applied for *D*-forecasting with their corresponding results. It also defines the research problem associated with *D*-forecasting in QLD and provides a novel forecasting technique as a new approach to addressing *D*-forecasting issues in QLD

Chapter 3 provides the theoretical framework of the different models used in this study. It includes the mathematical equations and theorems that were applied to derive the respective forecasting models such as the Artificial Neural Network (ANN), Extreme Learning Machine (ELM), and Markov Chain Monte Carlo (MCMC) models in conjunction with the application of copula-based models.

Chapter 4 outlines the novel hybrid ELM-MCMC copula-based model for forecasting *D*. It also defines the data source, the study site, and the methodology that was applied to develop and test the ELM- MCMC copula-based model. The results for various timescales (6-hours, 12-hours, and daily) for this model are also displayed in this chapter.

Chapter 5 discusses the methods that were used to develop the conditional bivariate copula models to make a probabilistic forecast of *D*. It also displays the results for these models across various timescales (30-minutes, 1-hour, and daily).

Chapter 6 provides the conclusion of the entire research thesis based on the research findings and points out the research limitations and opportunities for future research.

CHAPTER 2: LITERATURE REVIEW

2.1 Importance of accurate predictive models for *D*-forecasting in QLD

Supplying a specific electricity load to meet customers' need in a specified timeframe is paramount. Reliable forecast models present a vital decision support tools for utility operators to plan and manage resource requirements for smooth and sustainable energy systems to serve consumers on time and when required. Without such models, the electricity supply will be under - or over- supplied, which destabilizes the daily and business activities undertaken by electricity consumers and incurs excessive costs for the utility operator, respectively. As inferred by Haida and Muto (1994), and Fan, S. and Chen, L. (2006), small percentage increases in projection can cause million-dollar losses for utility companies. Therefore, it is very important for utility operators to have a reliable model to accurately predict *D* to appropriate meet consumer demand, as well as avoiding excessive production costs by not producing excess quantities.

Prudent and accurate forecasting is fundamental for efficient planning in any organization. In the case of the electricity industry, developing a reliable forecast model is crucial for accurately projecting *D* to appropriately meet future demand. Since a loss of electricity is inevitable in any power utility system, *D*-forecast is necessary for maintaining the required supply of load at any given time. Since it is impossible to achieve an accurate forecast model as the future is unforeseeable, it is always advisable to test model performance using various predictor variables as well as having sufficient length of data for optimized performance.

2.2 Factors affecting electricity demand in Australia

The factors influencing *D* are volatile and complex in nature, comprising non-stationary and non-linear variables such as local GDP, meteorological features, social and demographic influences (Nasr *et al*. 2002; Mirasgedis *et al.* 2006; Zareipour *et al*. 2006a; Odhiambo 2009; Suganthi & Samuel 2012; Hu *et al*. 2013). Because of this, *D* is also volatile at any timescale, which subsequently tends to be challenging for utility providers to optimize management and operational inputs to maximize the efficiency of the electricity system.

D-forecasting is a multivariate task that is influenced by various interconnected factors such as changes wind and solar radiation patterns causing instability and uncertainty in energy provision for both the wind and solar industries. Also, local climatic features (air temperature, and humidity) affect *D* requirements as residents use more electrical energy by means of heaters and air conditioning units to maintain their body temperatures during winter and summer seasons, when local temperatures are much more extreme. *D* is also impacted by the local population density in any region as more people use more electrical energy to fulfill their daily and economic needs. Moreover, *D* is also dependent on GDP as any region which is exposed to industrialization, requires large quantities of energy to meet its economic and technological needs to survive and prosper. Some seasonal effects which seldom affect *D* are festive and religious events, public holidays, and sporting events (Al-Alawi & Islam 1996).

Energex and AEMO are the leading managers of NEM the in Australia. They organize the market and trade between customers and sellers (generators). Since *D* is affected by various interconnected predictors, the AEMO may not be able to predict *D* sufficiently for QLD during the imminent rise in *D* expected by the Australian Electricity Market from 2020 to 2030 (Brinsmead et al. 2014). This includes a likely increase in energy demand by the solar and wind energy sectors during the transition from conventional to renewable energy, as well as fluctuation in air temperature during various climatic seasons (*e.g.* winter and summer), compelling users to consume more energy to maintain their normal body temperature (Al-Musaylh *et al*. 2019). It also entails the prospective transformation and participation of electricity trading in the local market and the growing demand for intelligent systems for precise and real-time forecasting in competitive electricity markets.

The AEMO currently uses semi-parametric-type additive models (Hyndman & Fan 2010) that incorporate various predictors. These predictors are mainly the gross domestic product (GDP), air temperature, population census, and seasonal anomaly which are processed and analysed statistically to forecast the load requirements for all Australian states (Ball *et al*. 2016). However, since the semi-parametric-type additive model is outdated and holds high uncertainty due to the impacts of modernization and climate change, there is a need to explore new forecasting techniques to accurately forecast *D* to meet the growing load demand, which this study addresses.

It is crucial that Australian utility providers use precise and real-time information on *D* in a competitive electricity market as it can present valuable information for energy policy formulation and reforms, and can further assist in making informed decisions on capacity planning for process optimization and subsequent cost reduction. Such information will guide and maintain cohesive and sustainable policy planning and the associated energy markets (Erdogdu 2007). Hence, there is an apparent need for the NEM to adopt a robust predictive model with which intelligent business decisions can be made to address forecasting inaccuracies appropriately whilst meeting customers' demand, maximizing profit and being competitive in the electricity market.

2.3 Rationale for machine learning techniques in *D*-forecasting

Machine learning (ML) is a sub-divisional field of artificial intelligence (AI). It uses data to feed machines for self-training and learning to uncover hidden trends and relationships and discover new insights to assist in making informed decisions for the future in various applications including *D*-forecasting. The ML technique is perceived as a powerful tool for data analytics that mimics human intelligence to manipulate and analyse large data to compute fast, consistent, and accurate estimations. However, optimizing management and operational inputs to maximize the efficiency of the electricity system is difficult as *D* is volatile at any time scale and is influenced by complex networks of non-stationary and non-linear variables such as local GDP, meteorological features, demographic, and social influences (Nasr *et al*. 2002; Mirasgedis *et al*. 2006; Zareipour *et al*. 2006a; Odhiambo 2009; Suganthi & Samuel 2012; Hu *et al*. 2013). Therefore, it is difficult to simulate *D* precisely using conventional computing methods given the complexity of historical explanatory variables which are mostly big data. However, ML perfectly address this issue as they can utilize complex data logically and compute fast, accurate, and consistent solutions (Wang & Ramsay 1998). Because of this, they have been increasingly used in the contemporary world in various engineering applications as reported in Florens *et al*. (2007) and Xydas *et al*. (2016). Therefore, ML techniques are required to accurately address forecasting uncertainties experienced by the National Electricity Market in developing robust forecasting models of high precision to assist in making intelligent business decisions.

2.4 Standalone machine learning techniques used for *D*-forecasting

Machine learning techniques have been widely applied for predictive forecasting in various engineering applications (Salcedo-Sanz et al. 2014) as they hold promising data analytical tools that can solve time-series data (*e.g. D*) using numerous predictors. Some common standalone ML models that have been extensively used by the NEM for *D*-forecasting include the Multivariate Adaptive Regressive Spline (MARS) (Sigauke & Chikobvu 2010), knowledge-based expert systems (Singh et al. 2013), genetic algorithms (Ali et al. 2018b), Support Vector Regression (SVR), Artificial Neural Network (ANN) (Sozen & Akcayol 2004), Autoregressive Integrated Moving Average (ARIMA) (Contreras et al. 2003), Wavelength Transformation (WT), Bootstrapping Procedure (B) and Bayesian Model Averaging (BMA).

The SVR model can solve regression problems with many predictor variables. It is highly automated which automatically tunes the model parameters to achieve the optimal model combination. It operates by combining uniform and systematic networks by way of feature extractions (Smola & Schölkopf 1998; Yu et al. 2006). Its programming functions are operated on the conceptual framework of structural risk minimization and are intended to minimize overfitting data by limiting the imminent error during training. This technique has been commonly practised for various forecasting applications to provide practical solutions for the real-world problems by delivering accurate forecasting data to assist planners and energy experts to make informed decisions to minimize operation cost, optimize operation, and maximize profitability.

In Istanbul (Turkey), the SVR model was applied in conjunction with the Radial Basis Kernel Function (RBF) to predict daily *D,* obtaining a mean absolute percentage error *(MAPE)* of 3.67% between the observed and predicted values (Türkay & Demren 2011). Also, the SVR model was used for forecasting *D* in the eastern part of Saudi Arabia which generated accurate results when benchmarked against the Autoregressive (AR) model. Moreover, Sivapragasam and Liong (2005) employed various SVR models for regional projection of stream flow in Taiwan and achieved accurate results compared to the ANN model. In a recent study, Al-Musaylh et al. (2018a), used significant lagged *D* to feed the SVR model and attained precise *D*-forecasts for daily horizons by achieving a highest Willmott's Index of Agreement (*WI*) of 0.980 and lowest *MAE* of 162.363 MW against comparative MARS and ARIMA models.

MARS is a non-parametric regression technique that is flexible in processing various data and can provide fast results. It operates by fitting relationships of various linear regression models involving different predictors. However, its application in *D*-forecasting was limited as it was intentionally developed for piecewise function (Sharda et al. 2008; Deo et al. 2016). A study by Sigauke and Chikobvu (2010) used a MARS model to forecast *D* in South Africa and generated smaller *RMSE* when benchmarked against piecewise regression models. Also, Al-Musaylh et al. (2018a) reported that their MARS model forecasted accurate *D* for 30-minutes and 1-hour timescales by exhibiting the highest *WI* of 0.993 and 0.990 and lowest *MAE* of 45.363 and 86.502 MW, respectively.

The ARIMA model has also delivered satisfactory results for *PR*-forecasting (Contreras et al. 2003). Another study on *PR*-forecasting was conducted by Zareipour et al. (2006b) in Turkey by amalgamating the ARIMA model with a cointegration technique. Based on the comparative analysis in this study, the projection was more than 40% when benchmarked against the target values of the ARIMA model. Also, a study by Al-Musaylh et al. (2018a) indicated that the ARIMA model performs well given the shorter training dataset as it minimizes the degree of fluctuation thereby inhibiting cumulative errors compared to a larger dataset. In this study, SVR and MARS models were also tested using significant lagged demand as the predictor variable for forecasting *D* in the state of Queensland at different timescales.

ANN is a type of AI technique that the resembles human brain and can solve complicated problems involving various variables of non-linear relationships in time series. It has been highly adopted in forecasting short-term *D*-forecasting involving complex data due to its self-learning ability. This includes studies involving an adaptive ANN model to determine *D*-requirements for residential buildings (Yang et al. 2005a, 2005b) to adapt to periodic changes such as the emergence of new predictors data for *D* such as climatic and operational data with changing patterns. In this study, the model parameters were automatically updated to suit the features of the incoming data for real-time *D*-forecasting. Other studies involving ANN for short-term *D*-forecasting have also been undertaken in many locations using various methodologies (Papalexopoulos et al. 1994; Khotanzad et al. 1997; Chen et al. 2001; Taylor & Buizza 2002; Hsu & Chen 2003; Pai & Hong 2005; Sheikh & Unde 2012).

2.5 Studies on *D*-forecasting using climate-based predictors

According to the literature review, a few research studies have been conducted on *D*-forecasting using climatic variables as predictors. This includes a study by Mirasgedis et al. (2006) which employed solar radiation, wind speed, humidity, and air temperature to forecast *D* for a site in Greece. This study reported that air temperature and humidity are two fundamental parameters that influence *D*-forecasting. In another study conducted in South Africa by Lebotsa et al. (2018), they used air temperature, calendar anomaly, and time lagged demand to model *D* using a partial addictive quantile regression model and attained improved results. Similarly, studies in Australia (Deo & Şahin 2015; Deo & Sahin 2017) on solar radiation and precipitation also showed that climatic variables directly influence the solar radiation modelling. In a recent study by Al-Musaylh *et al*. (2019), they used climate-based predictors extracted from SILO and ECMWF reanalysis parameters to develop and test a hybrid Bootstrapping-ANN model and obtained accurate *D*-forecasting results for eight locations in southeast Queensland. Therefore, to build robust forecast models for both solar radiation and *D*-forecasting, climate variables are essential predictors. This would enable the delivery of accurate *D*-forecasting data which is crucial for energy experts, power planners, and utility operators to make informed decisions in addressing practical *D*-forecasting issues.

2.6 Studies on *D*-forecasting using hybrid machine learning models

The hybrid method combines two or more forecasting techniques to solve complex problems, of a non-linear and stochastically volatile nature, in a timely manner to significantly improve forecasting errors (Singh *et al*. 2013). Such models take advantage of the combined merits of both models under consideration, thereby further strengthening the predictive performance of *D*-forecasting at multiple timescales. For instance, Khan et al. (2019) applied empirical mode decomposition (EMD) to remove noise and randomness in *D* data (as noise caused inaccurate forecasting) before feeding the ELM model (EMD-ELM) and determined the daily, weekly, and monthly *D* of a building. The simulated results for this study were much better than those of the standalone Convolution Neural Network (CNN), Long Short-Term Memory (LSTM), and ELM. Similarly, Marwala and Twala (2017) used an ensemble of Optimally-Pruned integrated with Basic Extreme Learning Machines (OP-ELM) to forecast *D* against other traditional methods (*e.g.* ANN and SVR). The results outperformed both the ANN and SVR models.

Furthermore, a study by Clements et al. (2016) used 30 minutes data from NEM and forecasted 12-hourly *D* by the integrating a multiple equation time series method (recursive system) to seize the intra-day *D* relationship. This was reported to have outperformed the three standard benchmarking models utilized by AEMO by 33% against *MAPE* criteria. Additionally, Al-Musaylh et al. (2019) developed and tested a hybrid ANN model to forecast *D* for eight station in south east Queensland using the statistically lagged cross-correlation of *D* with climate-based predictors. In this study, they first used the bootstrapping (B) method to minimize uncertainties by means of rigorous resampling with substitution, prior to feeding the ANN model (*i.e.* B-ANN) and attained accurate *D*-forecasts for multiple timescales. Moreover, in another study by Al-Musaylh et al. (2018b), they first applied an improved version of the empirical model decomposition with adaptive noise (ICEEMDAN) tool to decompose non-stationary time-series data to address frequency components associated with *D* data prior to model development. This followed tuning of the weights and biases from *D* features using particle swarm optimization (PSO) that was approximated at the global optimum before feeding the SVR model for the final *D*-forecast. They developed a hybrid ICEEMDAN-PSO-SVR model which projected accurate *D* results in real-time for practical applications in the energy sector.

The above literature affirms that model hybridization with adequate predictor variables improves performance in delivering accurate forecasting results. In recent years, research has been focused on advancing energy markets, optimizing generation and distribution operations, minimizing expenditure, and maximizing profitability (Marzband et al. 2015; Marzband et al. 2017). Uncertainty about *D* requirements impacts the energy provision by its utility systems. Therefore, it is vital for utility operators and forecasters to have robust forecasting models to plan adequately for *D* requirements of various consumers to avoid destabilizing energy systems as well as incurring extra costs in operational expenditure. This can be achieved by integrating appropriate ML techniques with a wide range of *D* predictor variables to develop hybrid models for reliable forecasting.

2.7 Copula-statistical models in forecasting applications

Copulas are powerful mathematical equations that can link numerous time-independent variables regardless of their marginal distribution (Joe 1997; Nelsen 2003; Nguyen-Huy et al. 2017; Nguyen-Huy et al. 2018). They adopt a global optimization method to discover accurate copula parameters based on the highest marginal and joint distribution for probabilistic forecasting. In statistics and probability theory, a copula is a multivariate probability distribution for which the marginal distribution of each variable, coming from any distribution function is uniform. Multivariate data can be modelled through more advanced techniques like copulas, which are used to conduct multivariate analyses. Copula statistical models are very useful in probability statistics in studying scale free measures of dependence and in constructing families of bivariate and multivariate distributions, which have non-linear dependence. In addition, copulas can measure dependence for heavy-tail distributions and can allow the dependence structure to be modelled independently from the marginal selection (Nguyen-Huy *et al*. 2017). Copulas have been extensively used for predicting the dependence structure of water and climate-based variables (Salvadori 2003; Salvadori & De Michele 2004; De Michele et al. 2005). Also, the multivariate copulas have been applied in various studies for drought monitoring (Hao & AghaKouchak 2013), analysis of flood hazard (Jongman et al. 2014), rainfall prediction (Li et al. 2013; Jongman et al. 2014; Vernieuwe et al. 2015), and agricultural science (Fousekis & Grigoriadis 2017), among others.

Mainly, a copula is a joint distribution of random variables $x_1, x_2 \ \ldots\ldots x_p$ each of which is marginally uniformly distributed as $x_{(0,1)}$ as introduced by Sklar (1959). Ranked Spearman and Kendall tau correlation coefficients are utilized in copula statistical models to model non-linear and joint dependence of bivariate and multivariate datasets for probabilistic forecasting.

Basically, copulas are advanced data analytical tools, which have recently emerged and are perceived by researchers as useful tools for accurate crop and weather forecasting. They hold promising results, as reported in studies by Ali et al. (2018b); (Ali et al. 2018a) and Ali et al. (2020). However, copulas have not been used for *D*-forecasting anywhere.

2.8 Hybrid copula-based models in forecasting applications

The hybrid copula-based models commonly integrate ML models with a copula-statistical model to achieve accurate forecasting for various applications. The MCMC copula-based model is generally the feature selection method, which creates distributions for inputs against target copula parameters, in which they are ranked in the highest order of copula model performance. For example, Ali *et al*. (2018a) developed and used a hybridized online sequential ELM-MCMC based model which demonstrated superior performance in forecasting rainfall for various locations in Pakistan. In another study by the same authors (Ali *et a*l. 2018b) on cotton yield prediction in Pakistan, the performance of the GP-MCMC based copula model was demonstrated to be superior to standalone MCMC and GP models. These studies confirm that the integration of various models improves predictive accuracy leading to better estimation than standalone models. The above studies involved hybrid copula-based models for forecasting rainfall and cotton yield production. However, it is evident from the literature that a study of *D*-forecasting in Queensland, or indeed anywhere in the world, using a hybrid copula-based model is yet to be undertaken. Therefore, the hybrid copula-based model and conditional bivariate copula models are the novel techniques targeted to improve accuracy in *D*-forecasting.

CHAPTER 3: MATERIALS AND METHOD

3.1 Study Area

This study is focused on the entire Queensland region (20.92º S, 142.70º E) (refer to **Figure 1**). It is the second largest state in Australia with a total area of 1,730,648 km² and a population of 5.13 million (ABS 2020). It has experienced significant growth in demand for electricity in the recent past (2014 - 2015), consuming 24.4% of the national consumption (Ball et al. 2016). However, studies for *D*-forecasting have been lacking and this study addresses this need.

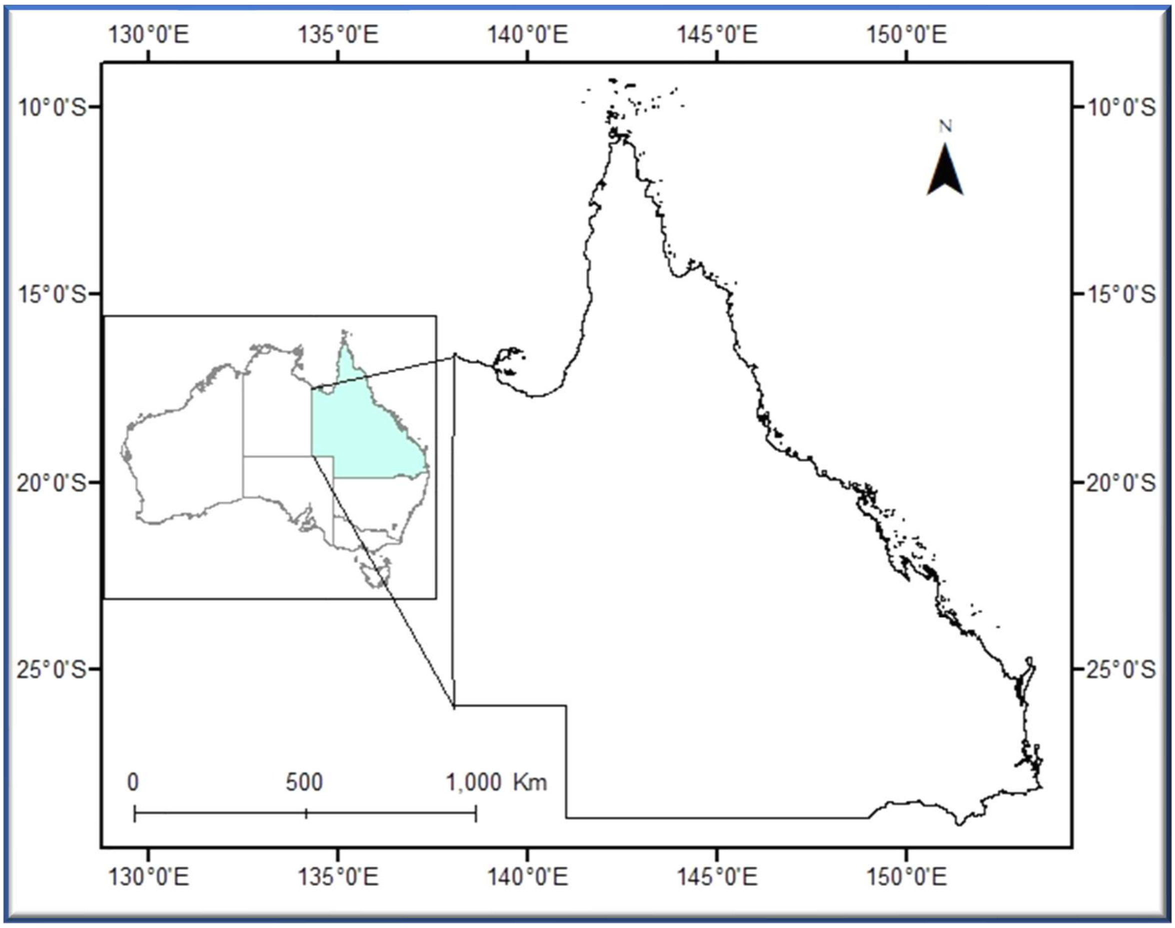


Figure 1: *Study map of Queensland region*

3.2 Data

3.2.1 Electricity demand data and pattern

The *D* dataset for the entire state of Queensland that was used in this thesis was downloaded from AEMO's website (AEMO 2020), covering 01-January-2017 to 31-December-2019. This dataset was partitioned with data from 01-January 2017 to 31-December 2018 used as training sets (see **Table 1**), while those ranging from 01-January 2019 to 31-December 2019 were used used as testing sets (see **Table 2**). This dataset is of high temporal resolution, featuring a 30 -minutes timescale, and has been recently used in a study conducted by Al-Musaylh et al. (2018a) using the ML method. However, the data not been used in copula models as pursued in the current study. The selected data was cleaned and sorted from extraneous features, including missing values, using MATLAB and R software. This process was followed by performing arithmetic averaging and summing was then performed to achieve average price and total electricity demand respectively for each timescale (*e.g*. 6-hours, 12-hours, daily), and descriptive statistics for the dataset were computed as displayed in **Tables 1** and **2** for the training and testing phases respectively. The relationship between *D* and *PR* was also analysed by means of time-series plots as seen in **Figure 2**. Finally, the linear relationship between the predictor (price) and predictand (demand) was determined using scatter plots which then generated the value of the coefficient of determination (r^2) for the assessment of linear dependence. Generally, values of r^2 of 0.5 and above imply a strong correlation between the target and objective variables. This means only the price variable can be used to predict demand. In this study, r^2 was less than 0.5 so time-lagged demand was used to forecast future *D*. The time-lagged demand was computed from the partial autocorrelation function (PACF) in the MATLAB software.

Table 1: *Descriptive statistics of aggregated electricity demand (megawatts, MW) for Queensland State for training set (2017 - 2018).*

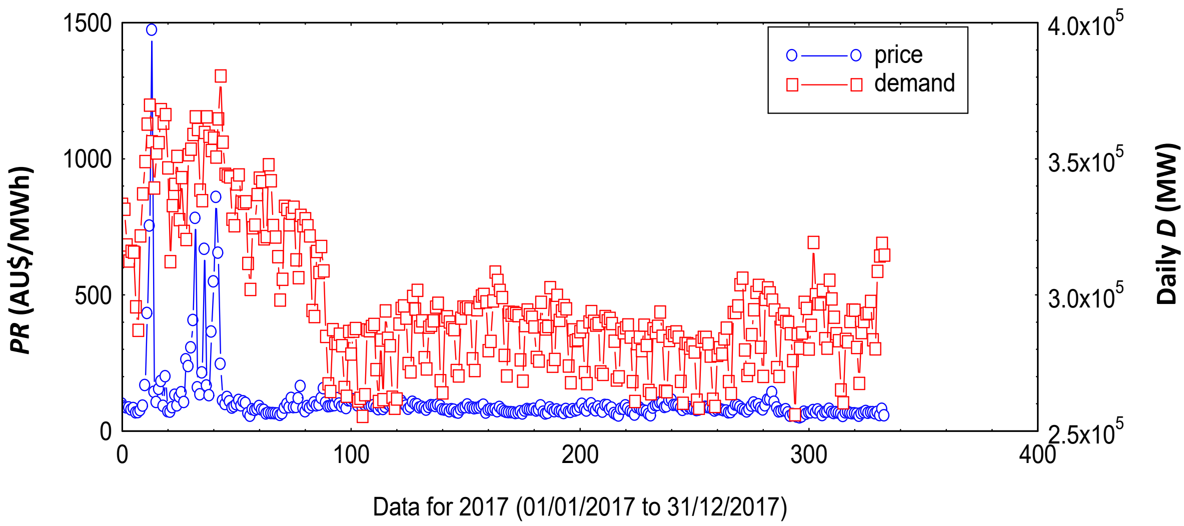
Forecast Period	Data Period	Minimum (MW)	Maximum (MW)	Mean (MW)	Standard Deviation (MW)	Skewness	Kurtosis
6 hours	01-January 2017 to 31-December 2018	58495.01	110573.9	74820.66	9451.04	0.50	-0.09
12 hours	01-January 2017 to 31-December 2018	117487	214327.6	149641.3	16806.01	0.88	0.74
Daily	01-January 2017 to 31-December 2018	253867.6	383711	299282.7	24008.75	0.84	0.46

Table 2: *Descriptive statistics of aggregated electricity demand (megawatts, MW) for Queensland State for testing set (2019).*

Forecast Period	Data Period	Minimum (MW)	Maximum (MW)	Mean (MW)	Standard Deviation (MW)	Skewness	Kurtosis
6 hours	01-January 2019 to 31-December 2019	56339.28	109281.3	74566.22	9620.33	0.58	-0.20
12 hours	01-January 2019 to 31-December 2019	117762.6	216812	149132.4	17400.41	0.88	0.59
Daily	01-January 2019 to 31-December 2019	252230.4	381046.8	298264.9	24996.72	0.65	-0.06

Figure 2 portrays the time series plot of the aggregated *D* data for the entire Queensland region. The stochastic features that are associated with the *D* dataset are attributed to the behaviour of electricity consumers at any one time. The *D* was very high in the beginning of 2017 that had caused spikes in corresponding *PR*. However, in the progressive months, the *PR* was not responsive to changes in *D* that was likely caused by consumer preferences, less population, and low-income rates for consumers. It is also shown that the relationship between *D* and *PR* for 2018 and 2019 has progressively improved. These suggested that it would not be feasible to train models using 2017 data and test them using 2018 and 2019 data because 2017 had some very peculiar features that will not be captured in test set (2019). This is further supported by an increase in standard deviation for each timescale in each year and the dissimilar values of skewness and kurtosis (*i.e.*, 0.65 and -0.06 respectively) for daily timescale as opposed to the 6 and 12-hours timescale in **Tables 1** and **2**.

The 12-hours and daily timescales have positive kurtosis of 0.74 and 0.46 respectively for the training set. Similarly, the testing set have the kurtosis of 0.59 for 12-hours timescale. These positive kurtoses indicated that distribution of *D* data have tail dependence as well as outliers. All the timescales in both training and testing sets have positive skewness. This implies that most data are congregated at the left tail of the distribution while the right tail have longer distribution, meaning the mode is less than both mean and median.



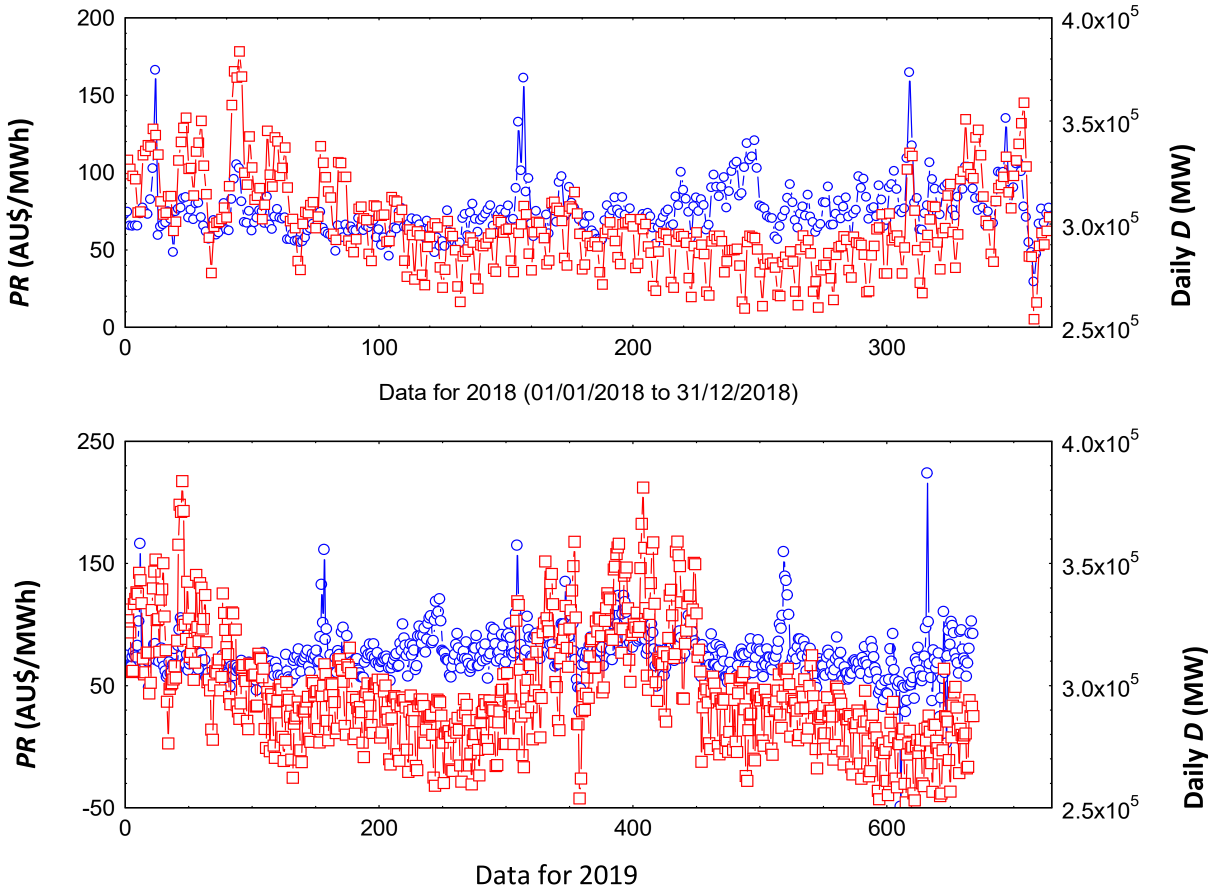


Figure 2: *Time-series plot of electricity demand and price for daily intervals for 2017, 2018, and 2019 for the whole study period of 01-January-2017 to 31-December-2019.*

Figure 3 shows that every yearly data (2017 - 2019) has different correlation between *D* and *PR*. Therefore, it is not feasible to treat or develop models to behave the same way for different years. Of all the years, 2019 had the best correlation between *D* and *PR* with the corresponding coefficient of determination of 0.321.

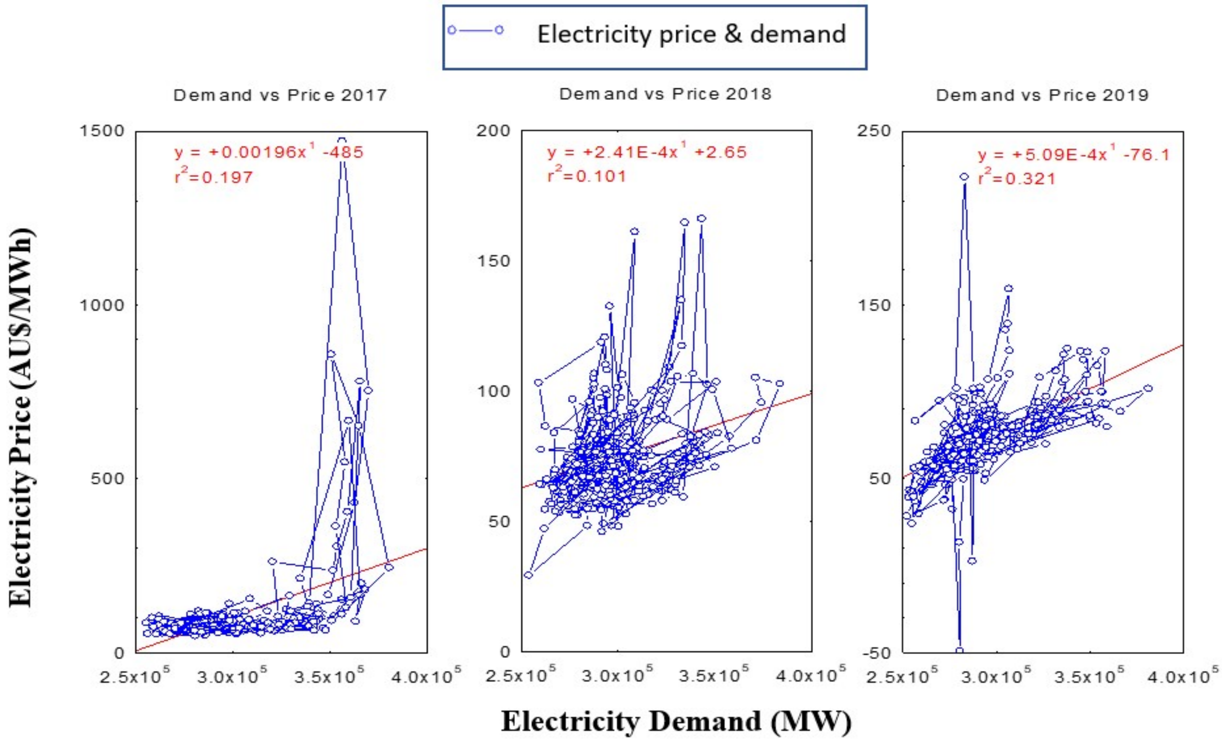


Figure 3: *Relationship between electricity demand and price for 30-minutes intervals for 2017, 2018, and 2019*

3.2.2 Electricity price data

The price of electricity (*PR*) is charged in Australian dollars for per megawatt (MW) of energy consumed per hour and is expressed as AU$/MWh. It was added as a covariate predictor to model *D* in objective two of this study. In certain dispatch events, negative *PR* value was experienced. This happens when the generator stays online, and customers purchase energy at a negative price. Also, when an inter-connector is constrained, energy may not freely flow from one region to another, implying that one's region low-priced generation may not freely flow into another region. This may also push the prices into negatives, depending on the market conditions (AEMO 2020). In this thesis, any *PR* value exceeding $AU500/MW were treated as outliers and replaced with the average value that was computed from the other two years on the same dispatch date.

3.3 Theoretical Framework

3.3.1 Extreme Learning Machine (ELM) model

Extreme learning machine (ELM) is an improved version of ANN and is an algorithm for the single-hidden layer feed-forward neural network which was developed by Huang et al. (2006). Unlike the ANN model, ELM randomly generates input weights, destabilizing the network and creating an overfitting phenomenon, thereby detecting and analysing data features. Therefore, the training or tuning speed of ELM is computationally fast as it does not need iterative modifications of network parameters during training. It also has greater generalization capability and sustainability for non-linear activation and kernel functions and is convenient to use for forecasting applications (Zhang et al. 2013; Deo & Sahin 2015; Deo et al. 2017). Hence, ELM is highly preferred and can be easily updated to track real-time trends in electricity demand while continuously minimizing training error.

In summary, the ELM model operates in three steps where firstly the hidden layer neurons, weights and biases are randomly generated and in this process the system detects and analyses the data features. The inputs are then passed through the hidden layer where it further generates hidden layer features to an output matrix. Finally, the output weights are estimated by inverting the hidden layer output matrix thereby computing its product against the target, resulting in solving the set of linear equations.

The model incorporates several single ELMs. They arbitrarily select a number of hidden nodes within a pre-determined series as well as the input parameters. This results in the generalization of the natural stochastic volatility by reducing large variance and bias, which enhances the ensemble model to generate accurate forecasting results and defeat the volatility issues of single ELMs.

For any training data set with N samples, the output function of the Single Layer Feed-forward Neural Network (SLFN) with y hidden nodes and activation function μ can be expressed as follows (Zhang *et al*. 2013):

$$f_y\ (X_k) = \sum_{i=1}^{y} \varphi_k\ \mu\ (w_k\ x_l + c_k) = t_j\ ,\ j = 1, 2, \ldots N \qquad (1)$$

Equation (1) can be simplified as follows (Huang et al. 2006)

$$H\gamma = T \qquad (2)$$

For a training data set, given the activation function and hidden node number, the ELM learning can be summarized into three steps:

Step 1. Randomly generate the input weights w_k and x_l , $1 \leq i \leq N$;

Step 2. Calculate the hidden layer output matrix **H**;

Step 3. Calculate output weights matrix $\gamma = H^{\dagger} T$;

where $H^{\dagger}$ is the Moore–Penrose (MP) generalized inverse of H.

There are several methods to calculate the MP generalized inverse H†. It is suggested that singular value decomposition is the most appropriate method because of its universality (Huang *et al*. 2006). In marked contrast to traditional ANN learning algorithms, ELM does not require iterative adjustments of network parameters during the training; therefore, its training speed can be thousands of times faster. In the meantime, as proved in Huang *et al*. (2006), it can not only reach the minimized training error $\| H\gamma - T \|$ but also the smallest norm of output weights $\| \gamma \|$. According to ANN theory (HariKumar et al. 2012), for feed-forward neural networks reaching smaller training errors, the smaller the norm of weights are, the better generalization performance the networks tend to have. Another important benefit of ELM is its efficient tuning mechanism given an activation function, where only the hidden neuron nodes number needs to be tuned, which can be efficient.

3.3.2 Copula Theory

This study hybridized the ELM algorithm with copulas. Copulas are rigorous analytical tools used to investigate the joint dependence structure of various time-independent variables (Nelsen 2003). It was first explored as Sklar theorem (Sklar 1959) in a mathematical and statistical framework.

Copula families differ, with each having different model parameters ranging from one to three. The model parameters are indicative of the intensity of correlation of the joint dependence between two or more variables (Balakrishna & Lai 2009). Such correlations are either approximated by way of a theoretical relationship such as Kendall Tau and Linear Spearman coefficients, or suggested from the empirical multivariate probability distribution of the data. Of all the parameter estimation methods, local optimization algorithms which utilizes the Newton-Raphson method (Salvadori & De Michele 2004; Bárdossy 2006; Gräler et al. 2013; Ribatet & Sedki 2013) are the most widely used. There are strengths and weaknesses of local optimization methods. These respectively include the effective search algorithms, as well as their vulnerability to being confined in local optima (Duan et al. 1992). On the other hand, the global optimization parameters based on the Bayesian framework have been extensively used in the recent ages for exploring the best fitted copula parameters (Pitt et al. 2006; Min & Czado 2010; Smith et al. 2012; Parent et al. 2014; Kwon & Lall 2016).

A copula function is basically a mathematical function that is defined from I^2(F , G) to I(H) such that [F(x) , G(y) , H(x,y)] is a point in I^3 with $I \in [0,1]$ and X, Y are continuous random variables with distribution functions $F(x) = P(X \leq x)$ and $G(y) = P(Y \leq y)$, and $H(x,y) = P(X \leq x , Y \leq y)$ is a function that describes their joint distribution.

In this study, we utilize 26 different types of copulas to improve the performance of the ELM-MCMC copula-based model. The primary copulas can be written mathematically as follows:

I. The Gaussian copula (Li *et al*. 2013), expressed as

$$\int_{-\infty}^{\emptyset^{-1}(a)} \int_{-\infty}^{\emptyset^{-1}(b)} \frac{1}{2\pi\sqrt{1 - \phi^2}} \exp\left(\frac{2\phi xy - x^2 - y^2}{2\,(1 - \phi^2}\right) dxdy,\ \phi \in [-1,1] \qquad (3)$$

II. a t-copula (*Li et al.* 2013), formulated as:

$$\int_{-\infty}^{t\phi_2^{-1}(a)} \int_{-\infty}^{t\phi_2^{-1}(b)} \frac{r\left((\phi_2+2)/2\right)}{r(\phi_2/2)\pi\phi_2\sqrt{1-\phi^2}} \left(1 + \frac{x^2-2\phi_1 xy+y^2}{\phi_2}\right)^{(\phi_2+2)/2} dxdy, \phi_1 \epsilon[-1,1], \phi_2 \epsilon(0,\infty) \quad (4)$$

III. a Clayton copula (Clayton 1978), written as:

$$\max^{(a^{-\emptyset} + b^{-\emptyset} - 1,0)-1/\Theta} \phi_1 \epsilon[-1,1], \phi_2 \epsilon[-1,\infty] \setminus 0 \quad (5)$$

IV. a Frank copula (Li *et al*. 2013), defined according to the following mathematical formulation:

$$-\frac{1}{\emptyset} \ln \left(1 + \frac{(\exp(-\phi_a) - 1)(\exp(-\phi_b) - 1)}{\exp(-\emptyset) - 1}\right), \emptyset \epsilon R \setminus 0 \quad (6)$$

V. a Gumbel copula (*Li et al*. 2013), expressed as:

$$exp\left(-(((-ln(a))^{\theta} (-ln(b))^{\theta})^{1/\theta}\right), \theta \epsilon [1,\infty] \quad (7)$$

VI. a Fischer-Hinzmann copula (Fischer & Hinzmann 2006), given as:

$$\left[\phi_1(\min(a,b)))^{\Theta_2} + 1 - \Theta_1(ab)^{\Theta_2}\right]^{1/\Theta_2}, \Theta_1 \epsilon [0,1], \Theta_2 \epsilon R \quad (8)$$

The remaining 20 different types of copulas used in this thesis have been discussed in previous studies (Sadegh et al. 2017). In all types of copula-based models an unknown process κ links observation $\check{Y}$ to parameters Θ^* in the modelling inference analysis (Sadegh *et al*. 2017) and can be given through the following equation.

$$\check{Y} = \kappa(\Theta^*) + \xi \quad (9)$$

Where ξ indicates a vector of measurement errors. The vector e = $\check{Y} - Y$ is called the error residual and e =$[e_1,\ e_2 \dots, e_n]$ where n is the number of observations that include the effects of model structural errors (Sadegh *et al*. 2017). Bayesian analysis was performed for model inference and uncertainty quantification purposes because Bayesian analysis quantifies uncertainty with a probability distribution (Sadegh *et al*. 2017).

Bayes' law attributes all modelling uncertainties to the parameters and estimates the posterior distribution of model parameters by the following equations (Sadegh *et al*. 2017):

$$P\,(\Theta \mid \check{Y}) = \frac{P(\Theta)P\,(\check{Y} \mid \Theta)}{P(\check{Y})} \tag{10}$$

where p(θ) and p (θ | Y̌) defines prior and posterior distribution of parameters, respectively. Further, P (θ | Y̌) ≈ L (Θ *I* Y̌) denotes the likelihood given as:

$$L\,(\Theta \mid \check{Y}) = \frac{n}{2}\,\ln \frac{\sum_{i=1}^{n} [\check{Y}i - \check{Y}i(\Theta)]^2}{n} \tag{11}$$

To solve equation (12) analytically and numerically, a Markov Chain Monte Carlo (MCMC) simulation technique was adopted to sample the posterior distribution. For more details, readers are referred to Sadegh *et al*. (2017).

3.3.3 Markov Chain Monte Carlo (MCMC) copula-based models

The MCMC simulation is performed using the Multivariate Copula Adaptive Toolbox (MvCAT) in MATLAB software. This toolbox is used to analyse the dependence structure of many predictor variables rigorously and comprehensively in solving real-world problems by offering various copula families with varying complexity. The MvCAT applies MCMC simulation within a Bayesian framework to estimate copula parameters and the underlying uncertainties by ranking the best choice of copula for the underlying data. It also undertakes dependence analysis using a stringent and extensive method. The area of interest in probability space is pursued using MCMC simulation whereby many chains are executed in parallel. The chains establish the relationship, matching the characteristics of the posterior area and approximating the global optimum. MvCAT comprises 26 copula families, which are often applied in both local optimization and MCMC methods. It automatically plots posterior parameter distributions of any selected copula(s) and plots their corresponding empirical probability isolines when MCMC is selected and executed. After the execution process, a summary report is automatically generated based on the performance of the best choice of preferred copulas in terms of Max_L, *AIC*, and *BIC*. It also presents the 95% confidence interval, root mean square error (*RMSE*), and Nash Sutcliffe efficiency (N_{SE}) for each of the selected copulas and ranks them in order of the best to the worst based on their respective performance in the probability space (Sadegh et al. 2017).

3.4 Methods

3.4.1 Development of ELM-MCMC copula-based models

The available data of aggregated demand for two years (2017 - 2018) was used to train the ELM model and tested using 2019 *D* data. The analysis of time series plots of *D* data exhibited that *D* was not stationary with time, due to changing consumer demand. Therefore, to forecast future *D*, the Partial Autocorrelation Function (PACF) was used to compute the statistically significant time-lagged *D* from historical *D* which were used as the only predictors. In most timescales, the first lag (t-1) had the highest lagged correlation, therefore, it was used as the only predictor in the development of the ELM model for all timescales despite some timescales (for example; 6-hours and 12-hours) having highest correlation lags other than one. This was done to ensure the achievement of uniform simulation whilst returning an accurate *D*-forecast. **Figure 4** illustrates the sample Pearson correlation coefficient based on PACF of lagged *D* for the daily timescale which was employed for constructing the ELM-MCMC copula-based model.

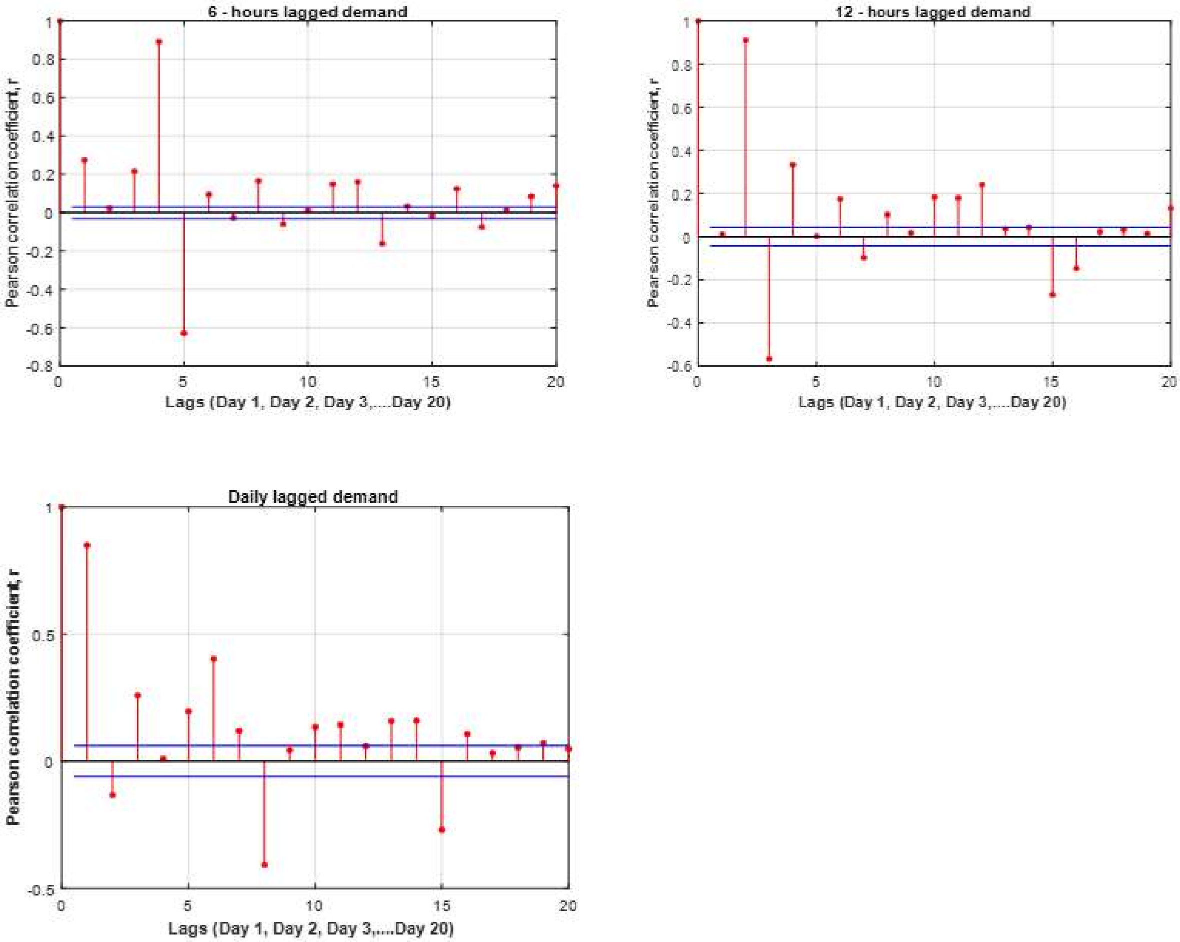


Figure 4: *Statistically significant lags of historical D at the 95% confidence interval as shown in blue lines derived from sample partial autocorrelation function (PACF) from the training sets used as predictors to forecast D at respective timescales (i.e., 6-hours, 12-hours, and daily).*

Table 3 depicts the data points with their corresponding predictor variables for the respective forecast horizons. The only model input for all timescales was the significant lagged *D*, and in this case, it is the first lag (t-1) for all timescales. From **Figure 3**, it is noteworthy that even though 6 and 12-hours had the fourth and second highest significant lags *D* respectively, only the first lagged *D* was used with the intent to achieve uniform forecasting for all timescales.

Table 3: *Details of predictor variables used for constructing the bivariate ELM model for various forecasting horizons*

Forecast Period	Period of data	Number of datum points	Number of input variables	Training data (2017 - 2018)	Testing data (2019)	Number of significant lags
6 hours	01-January 2017 to 31-December 2019	4380	4379	66.66%	33.34%	1
12 hours	01-January 2017 to 31-December 2019	2190	2189	66.66%	33.34%	1
Daily	01-January 2017 to 31-December 2019	1095	1094	66.66%	33.34%	1

The model simulations for both ELM and MCMC were done using MATLAB software on a Pentium 4, 2.93 GHz CPU system. **Table 4** shows the model architecture of the optimal ELM model where the "hit-and-run" approach was employed to determine the optimal activation function for each timescale. The three-year data was sorted into timescales of 6-hours, 12-hours, and daily. Following Ali *et al*. (2018a), the dataset from each timescale was partitioned into subsets of two years (2017 - 2018) for training and one year (2019) for testing. From the training set, 20% were allocated for model validation. Subsequently, the training dataset was used to train the model, then the testing dataset was used to test the model performance for each forecast horizon.

Table 4: *The ELM modelling framework utilized in this study. The transfer functions used are hardlim (hard-limit), tribas (triangular basis), radbas (radial basis), sin, and sig. The optimum functions are boldfaced and were used in the D-forecasting.*

6-hours ELM Model	
Number of layers	3
Neurons	Input: 1 (lagged demand) Hidden: 3…50 Output: 1 (*D*)
Transfer function	Hardlim, tribas, radbas, ***sin,*** sig
Learning rule	ELM for SLFNs
Model architecture (Input-Hidden Neurons-Output)	1-2-1
12-hours ELM Model	
Number of layers	3
Neurons	Input: 1 (lagged demand) Hidden: 3…50 Output: 1 (*D*)
Transfer function	Hardlim, tribas, radbas, ***sin,*** sig
Learning rule	ELM for SLFNs
Model architecture (Input-Hidden Neurons-Output)	1-1-1
Daily ELM Model	
Number of layers	3
Neurons	Input: 1 (lagged demand) Hidden: 3…50 Output: 1 (*D*)
Transfer function	Hardlim, ***tribas***, radbas, sin, sig
Learning rule	ELM for SLFNs
Model architecture (Input-Hidden Neurons-Output)	1-5-1

Table 5 displays the forecasted *D* from the optimal ELM model in the test set of the respective timescales which were then fed into the MCMC copula-based model for final *D*-forecasting. This involved the selection of all 26 families of copula followed by the selection of the local and MCMC (global optimization) parameter estimation methods, which try to match the observed *D* with the best copula estimate in a single execution. This technique assists in

locating the global optimum value, as well as approximating the posterior distribution of the parameters while exploring various copulas in probability space (fitting uncertainties) for matching against observed D (Ali *et al*. 2018b). The copula parameters are estimated using the MCMC simulation and this is achieved inside a Bayesian framework. The testing results were evaluated using stringent statistical metrics and compared against standalone ELM and MCMC copula models to explore the efficiency of the objective model (ELM-MCMC). It is worth noting that both the training and testing data were normalized before feeding the model. This was done to improve computational performance by conditioning them from variations in data patterns, and was achieved by applying the following equation (Hsu & Chen 2003):

$$x_{norm} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (12)$$

Where x denotes any reference point of target and objective variable, x_{max} and x_{min} are the maximum and minimum values of dataset respectively, and x_{norm} is the normalized reference point.

Figure 5 shows the schematics for the methodology applied in exploring *D*-forecasting for objective one of this study as discussed above. It indicates that objective one of the present study utilized the PACF, whereas objective two used a different methodology where the cross-correlation function was applied to forecast *D*.

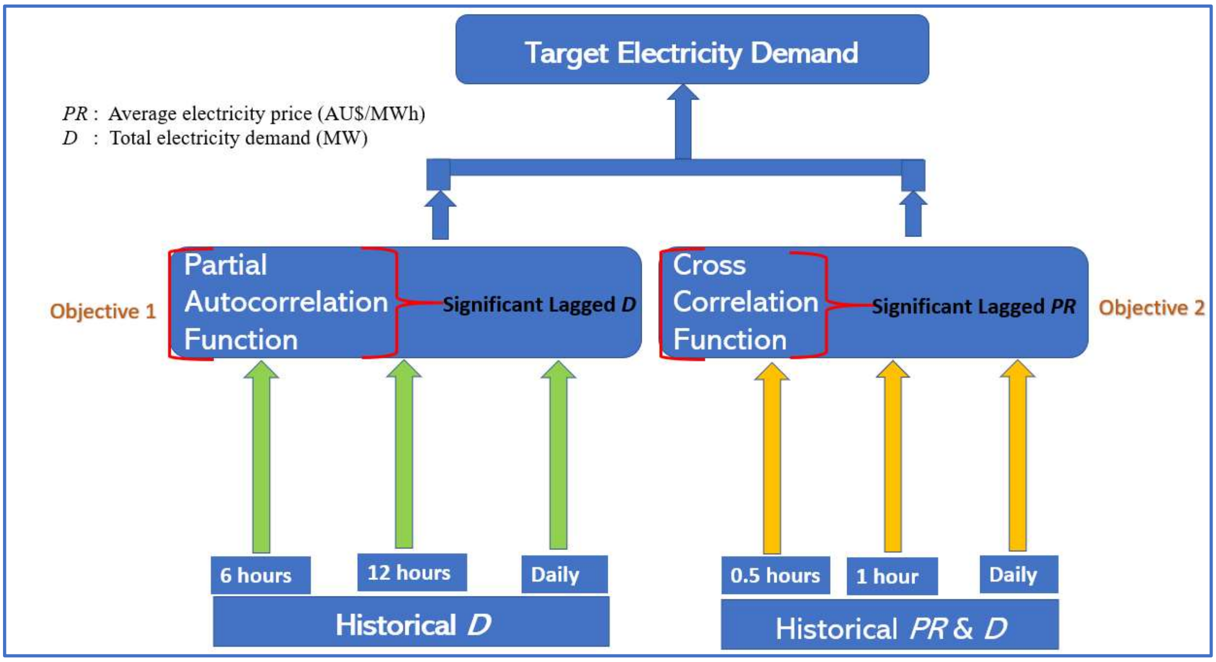


Figure 5: *Schematics of partial autocorrelation function and cross - correlation function of objectives 1 and 2 respectively.*

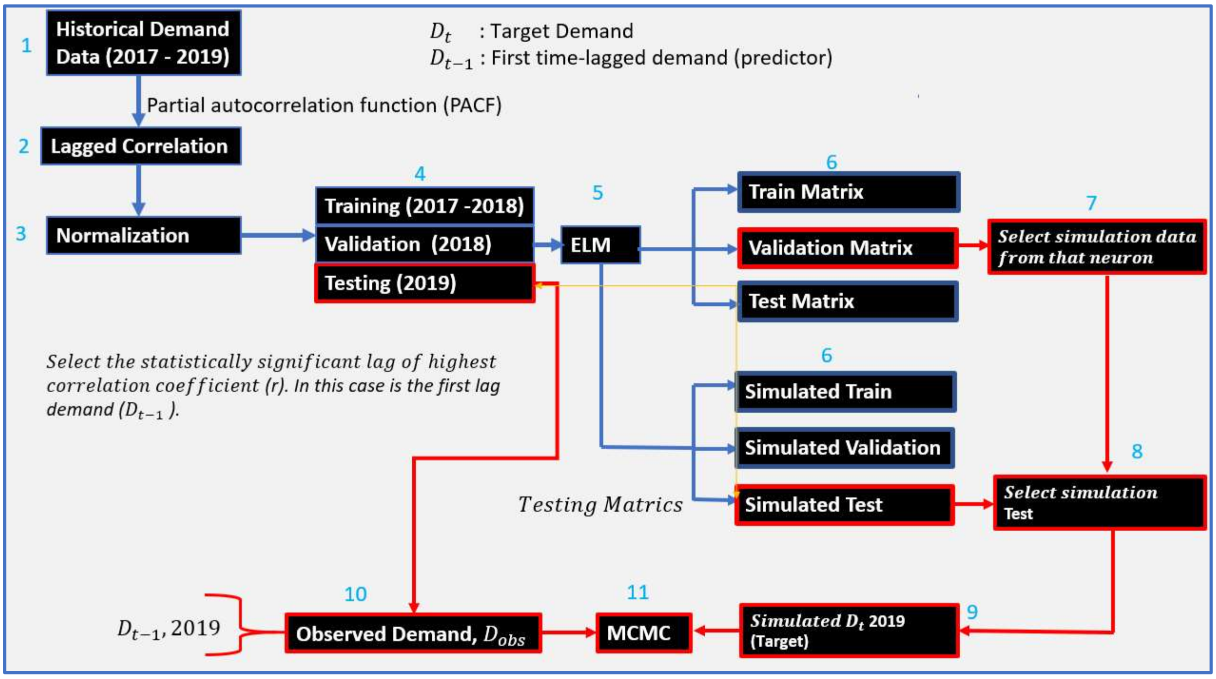


Figure 6: *Flow chart of the learning algorithm of hybrid extreme learning machine and Markov Chain Monte Carlo based copula model for objective one.*

The ELM model as discussed in section 3.2 above, is the ML model of an improved version of the traditional ANN model. It is easy to develop, computationally fast, and generally produces accurate forecast results. Therefore, the ELM model is applicable for real-time *D*-forecasting applications.

The primary role of the ELM algorithm as illustrated in steps 5 to 9 in **Figure 6** above, is to train (2017 - 2018) historical *D* data. This involved the use of a "hit-and-run" approach where several hidden neurons and transfer functions were tested to determine the optimal model architecture (neuronal structure) based on the respective model performance at various forecast horizons. The best model architecture from the training phase was selected and used to test the model using the independent "test" dataset of 2019 data. Following Ali et al. (2018b), the simulated result from the test set was then fed into MCMC copula-based model (steps 10 - 11, **Figure 5**) to further improve forecasting accuracy as well as to rank the 26 types of copulas from best to worst performing copula based on their respective performance matrices. This provides a more robust model in terms of similar results derived from training and testing sets. The best copula, with its corresponding forecasting result, was selected and used for short-term *D*- forecasting.

Three layers were used for the ELM model to design and construct the neuronal structure (see **Table 4**) for short-term *D*-forecasting at 6-hours, 12-hours, and daily timescales using the 2017 - 2018 dataset for training and the 2019 dataset for testing at the respective timescales. The optimal hidden neuron and transfer function was selected based on the "hit-and-run" approach, where the taxonomy of each transfer function (hard-limit, radial basis, sig, sin, and triangular basis) were tested individually. The number of nodes in the hidden layers in each trial was increased by an increment of one. Based on the testing performance, the optimal node of 2 with the sin transfer function was selected for the 6-hours timescale, having a neuronal structure of 1-2-1 for the ELM model. Similarly, a neuronal structure of 1-1-1 with sin transfer function was selected for the 12-hour timescale while a neuronal structure of 1-5-1 with the triangular basis transfer function was selected for the daily timescale. This was done to obtain the optimum forecasted *D* for all timescales.

3.5 Evaluation of machine learning (ML) models

To determine the merits of the hybrid ELM-MCMC model for short-term demand forecasting in Queensland, the results are compared against previous studies. The model also uses various statistical error criteria based on the performance metrics that hold the equations below and depend on the predicted and observed data for D (Willmott 1981, 1982, 1984; Dawson et al. 2007; Willmott et al. 2012; Mohammadi et al. 2015b):

The correlation coefficient (r) is expressed as:

I. $$r = \left(\frac{\sum_{i=1}^{n} \left(\left(D_{obs,i} - \overline{D}_{obs,i}\right) \left(D_{for,i} - \overline{D}_{for,i}\right) \right)}{\sqrt{\sum_{i=}^{n} \left(D_{obs,i} - \overline{D}_{obs,i}\right)^2} \sqrt{\sum_{i=}^{n} \left(D_{for,i} - \overline{D}_{for,i}\right)^2}} \right) \quad (13)$$

Mean square error (*MSE,* MW) is expressed as:

II. $$MSE = \frac{1}{N} \sum_{i=1}^{n} (D_{for,i} - D_{obs,i})^2 \quad (14)$$

Root mean square error (*RMSE,* MW) is expressed as:

III. $$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(D_{for,i} - D_{obs,i}\right)^2} \quad (15)$$

Relative root mean square error (*RRMSE*, %) is expressed as:

IV. $$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (D_{for,i} - D_{obs,i})^2}}{\frac{1}{n} \sum_{i=1}^{n} (D_{obs,i})} \times 100 \quad (16)$$

Mean absolute error (*MAE,* MW) is expressed as:

V. $$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| \left(D_{for,i} - D_{obs,i}\right) \right| \quad (17)$$

Relative mean absolute error (*RMAE*, %) is expressed as:

VI. $$RMAE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{D_{for,i} - D_{obs,i}}{D_{obs,i}} \right| \times 100 \quad (18)$$

Legates and McCabe' index (L_M) is expressed as:

VII. $$L_M = 1 - \left[\frac{\sum_{i=1}^{n} |(D_{for,i} - D_{obs,i})|}{\sum_{i=1}^{n} |(D_{ob\ ,i} - \overline{D}_{obs})|}\right], \infty \leq E_{NS} \leq 1 \quad (19)$$

where $D_{for,i}$ is the data forecasting and $D_{obs,i}$ is the data observation for D in the test period, $D_{for,i}$ and $D_{obs,i}$ are the average of the forecasted ($D_{for,i}$) and observed ($D_{obs,i}$) respectively, and N is the number of data points in the test period.

In particular, r is used to show the relationship between $D_{for,i}$ and $D_{obs,i}$ (r = 1 means there is a strong and positive relationship and when $D_{obs,i}$ increases, $D_{for,i}$ increases too, $r = -1$ means there is a negative relationship and when $D_{obs,i}$ increases, $D_{for,i}$ decreases, and r = 0 means there is no relationship between $D_{obs,i}$ and $D_{for,i}$). The most widely used and important measure to identify model performance is *RMSE,* which calculates the error between $D_{for,i}$, and $D_{obs,i}$, which is always positive and is the best model that produces the smallest *RMSE*. However, the common measure of the forecast error is *MAE*. The ability of the model for data forecasting is measured by L_M and its value is between 0 and 1, but negative values are accepted. The highest value of L_M refers to the best model and the negative value refers to the poorer model performance. The model accuracy is determined according to Li *et al*. (2013) which is excellent when *RRMSE* < 10%; good if 10% < *RRMSE* < 20%; fair if 20% < *RRMSE* < 30%; and poor if *RRMSE* ≥ 30%. The level of overall agreement between $D_{obs,i}$, and $D_{for,i}$ is recorded by *MAPE*.

3.6 Evaluation criteria for copula-based models

The performance of the MCMC-based bivariate copula models are evaluated using statistical metrics different to those of the ML models. These performance matrices are mathematically expressed as follows:

I. Maximum-Likelihood value (Max_L) (Thyer et al. 2009) is computed as:

$$Max_L = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\alpha^2) - \frac{1}{2}\ln(\alpha^{-2})\sum_{i=1}^{n}\left[D_{obs,i} - D_{for,i}\right]^2 \quad (20)$$

II. Akaike Information Criterion (*AIC*) (Akaike 1974) is calculated as:

$$AIC = 2D + n.\ln - \left(\frac{\sum_{i=1}^{n}\left[D_{obs,i} - D_{for,i}\right]^2}{n}\right) - 2CS \quad (21)$$

III. Bayesian Information Criterion (*BIC*) (Schwarz 1978) is expressed as:

$$BIC = D.\ln + n.\ln - \left(\frac{\sum_{i=1}^{n}\left[D_{obs,i} - D_{for,i}\right]^2}{n}\right) - 2CS \quad (22)$$

where $D_{obs,i}$ and $D_{for,i}$ are the observed and forecasted *i*th value of electricity demand and n is the number of datapoints, and D is the number of parameters in the statistical model. The maximum-likelihood (Max_L) reduces the residuals between model forecasts and observations. The *AIC* takes into consideration the model complexity as well as reducing error residuals to provide a more reliable measure of the quality of model predictions. The better and appropriate model fit is denoted by lower *AIC* and *BIC* values. The constants $\alpha = \frac{\sum_{i=1}^{n}[\check{Y}i - \check{Y}i(\Theta)]^2}{n}$, D , and CS, and equations (20) – (22) are used to assess the goodness of fit of the copula models, whereas equations (13) – (19) display the accuracy of the ELM-MCMC model performance.

CHAPTER 4: ELECTRICITY DEMAND FORECASTING USING HISTORICAL DEMAND AND HYBRID EXTREME LEARNING-COPULA MODELS

4.1 Results and Discussion

4.1.1 ELM model performance

The prediction metrics in equations (13) to (19) were used to evaluate the ML model's performance in terms of its ability to forecast *D* based on the observed measurements. In this study, the optimal output of the ELM model was fed as predictors into the MCMC copula-based model for final prediction of *D*. The learning rule utilized for this study for the ELM model is "ELM for Single Layer Feedforward Neural Network" (SLFNs).

Table 5: *Testing of the optimal performance capability of the ELM Model for all timescales (6-hour, 12-hour, and daily)*

	RMSE (MW)	*RRMSE (%)*	*MAE (MW)*	*RMAE (%)*	L_M
6-hours	7066.209	9.476	5848.66	7.877	0.265
12-hours	12523.959	8.398	9720.712	6.417	0.285
Daily	3690.590	1.237	2983.860	0.996	0.851

Table 5 shows the optimal performance of ELM model. It is noteworthy that each of the performance metrics have their strengths and weaknesses depending on the nature and complexity of data in use. Generally, the testing metrics should agree by displaying low *RMSE*, *RRMSE*, *MAE*, *RMAE*, and high L_M. In the case for 6 and 12-hours timescale, the optimal ELM model generated prediction metrics of *RMSE* (7066.209 MW and 12523.959 MW), *RRMSE* (9.476% and 8.398%), *MAE* (5848.66 MW and 9720.712 MW), *RMAE* (7.877 % and 6.417 %), and L_M (0.265 and 0.285) respectively. This indicates that the L_M disagrees with other metrics by exhibiting low value of 0.265 while both *RRMSE* and *RMAE* attained values less than 10%, implying that the model performance is excellent. This disagreement in terms of

percent errors are attributed to high fluctuation of observed data in the testing set (see **Table 2**), which have the skewness and kurtosis of 0.58 and -0.02 respectively, indicating presence of outliers.

Similarly, for the daily timescale, it yielded *RMSE* (3690.590 MW), *RRMSE* (1.237 %), *MAE* (2983.860 MW), *RMAE* (0.996%), and L_M (0.851) for 12-hours and daily timescales respectively. This indicated an agreement in performance metrics whereby the L_M exhibited high value whereas the absolute errors and percent terms attained low values. Hence, it demonstrates excellent model performance.

4.1.2 Performance of ELM-MCMC copula-based model for 6-hours

The forecasting results for the 6-hourly timescale were evaluated using many statistical matrices as individual error terms had their own shortcomings (Willmott 1981, 1982, 1984; Willmott et al. 2009; Willmott et al. 2012). Therefore, it was advisable to evaluate any model using a combination of error terms (as seen in **Table 7**) with a wide range of graphical illustrations to fully define the capability of its predictive performance. In this case, the objective model (*i.e.* ELM-MCMC copula-based model) and comparative MCMC copula-based model were evaluated using both the error terms for ML models and copula models while the standalone ELM model was only evaluated by error terms for ML models. Hence, the performance of the hybrid objective model was evaluated using robust statistical error terms particularly, mean square error (*MSE),* relative mean square error *(RMSE),* relative root mean square error (*RRMSE),* mean absolute error (*MAE*) *,* root mean absolute error (*RMAE*)*,* Legates and McCabe' Index (L_M)*,* Akaike Information Criterion (*AIC*)*,* Bayesian Information Criterion (*BIC*)*,* and Maximum Likelihood (Max_L).

Table 6: *Local and global (MCMC) copula parameters with 95% confidence interval of testing set for hybrid extreme learning machine with Markov Chain Monte Carlo (ELM-MCMC) copula-based and standalone MCMC copula-based model for forecasting electricity demand (D).*

6-hours: ELM-MCMC Model							
Model	Copula	Local para 1	Local para 2	MCMC para 1	MCMC para 2	95% CI Local-MCMC para 1	95% CI Local-MCMC para 2
M1	**Fischer-Hinzmann**	1.00	-34.462	1.00	34.999	[1.00 1.00]	[34.786 34.997]
M2	Linear-Spearman	1.00	[redacted]	1.00	[redacted]	[1.00 1.00]	[redacted]
MCMC Model							
M1	Marshal-Olkin	8.674	[redacted]	0.220	0.909	[0.215 0.225]	[0.845 0.999]
M2	TAWN	0.220	0.916	0.220	0.914	[0.216 0.226]	[0.867 0.982]
12-hours: ELM-MCMC Model							
Model	Copula	Local para 1	Local para 2	MCMC para 1	MCMC para 2	95% CI Local-MCMC para 1	95% CI Local-MCMC para 2
M1	**Fischer-Hinzmann**	1.00	-29.723	1.00	34.995	[1.00 1.00]	[33.272 34.999]
M2	Linear-Spearman	1.00	[redacted]	1.00	[redacted]	[1.00 1.00]	[redacted]
MCMC Model							
M1	Marshal-Olkin	0.050	22.971	0.093	0.802	[0.086 0.102]	[0.631 1.161]
M2	TAWN	0.093	0.804	0.093	0.809	[0.085 0.101]	[0.631 1.161]
Daily: ELM-MCMC Model							
Model	Copula	Local para 1	Local para 2	MCMC para 1	MCMC para 2	95% CI Local-MCMC para 1	95% CI Local-MCMC para 2
M1	**Cuadras-Auge**	0.999	[redacted]	1.00	[redacted]	[1.00 1.00]	[redacted]
M2	Fischer-Hinzmann	0.998	31.775	1.00	34.990	[1.00 1.00]	[34.476 34.995]
MCMC Model							
M1	Burr	0.268	[redacted]	0.268	[redacted]	[0.263 0.274]	[redacted]
M2	Joe	4.535	[redacted]	4.535	[redacted]	[4.459 4.609]	[redacted]

Table 7: *Testing performance for the objective model for 6-hourly timescale: Hybrid extreme learning machine model integrated with Markov Chain Monte Carlo (ELM-MCMC) in comparison with standalone MCMC copula-based model and the ELM model. The results displayed below only shows the best copula in terms of its copula rankings and its associated model performance. The model input for all timescale is the lagged demand* at first lag (t-1). *The boldface (black) text indicates the best model.*

Copula	*RMSE (MW)*	*RRMSE (%)*	*MAE (MW)*	*RMAE (%)*	L_M	*AIC*	*BIC*	*MaxL*
				6-hours				
			ELM-MCMC copula-based model					
Fischer-Hinzmann	285.480	0.345	262.241	0.336	0.980	-63136.10	-63125.53	51570.05
			MCMC copula-based model					
Marshal-Olkin	835.909	1.160	573.119	0.829	0.947	-12109.1	-12098.53	6056.90
			ELM model					
	7066.21	9.476	5848.66	7.878	0.265			

Table 7 shows that the ELM-MCMC-Fischer-Hinzmann copula model attained the highest forecasting accuracy at 6-hours timescale by exhibiting the lowest errors in *RMSE* ≈ 285.480 MW, *RRMSE* ≈ 0.345%, MAE ≈ 262.241, *RMAE* ≈ 0.336%, and highest L_M of 0.98. This model was best fitted by exhibiting the lowest *AIC* and *BIC* values of -63136.10 and -63125.53 respectively, and with much reduction in residual errors in terms of highest Max_L of 51570.05. This was followed by the MCMC-Marshall-Olkin copula with error values of *RMSE* ≈ 835.909 MW, *RRMSE* ≈ 1.160 %, *MAE* ≈ 573.119 MW, *RMAE* ≈ 0.829 %, L_M ≈ 0.947, *AIC* ≈ -12109.10, *BIC* ≈ -12098.53, Max_L ≈ 6056.90, and finally the standalone ELM model with *RMSE* ≈ 7066.21 MW, *RRMSE* ≈ 9.476%, *MAE* ≈ 5848.66 MW, *RMAE* ≈ 7.878 %, and L_M ≈ 0.265.

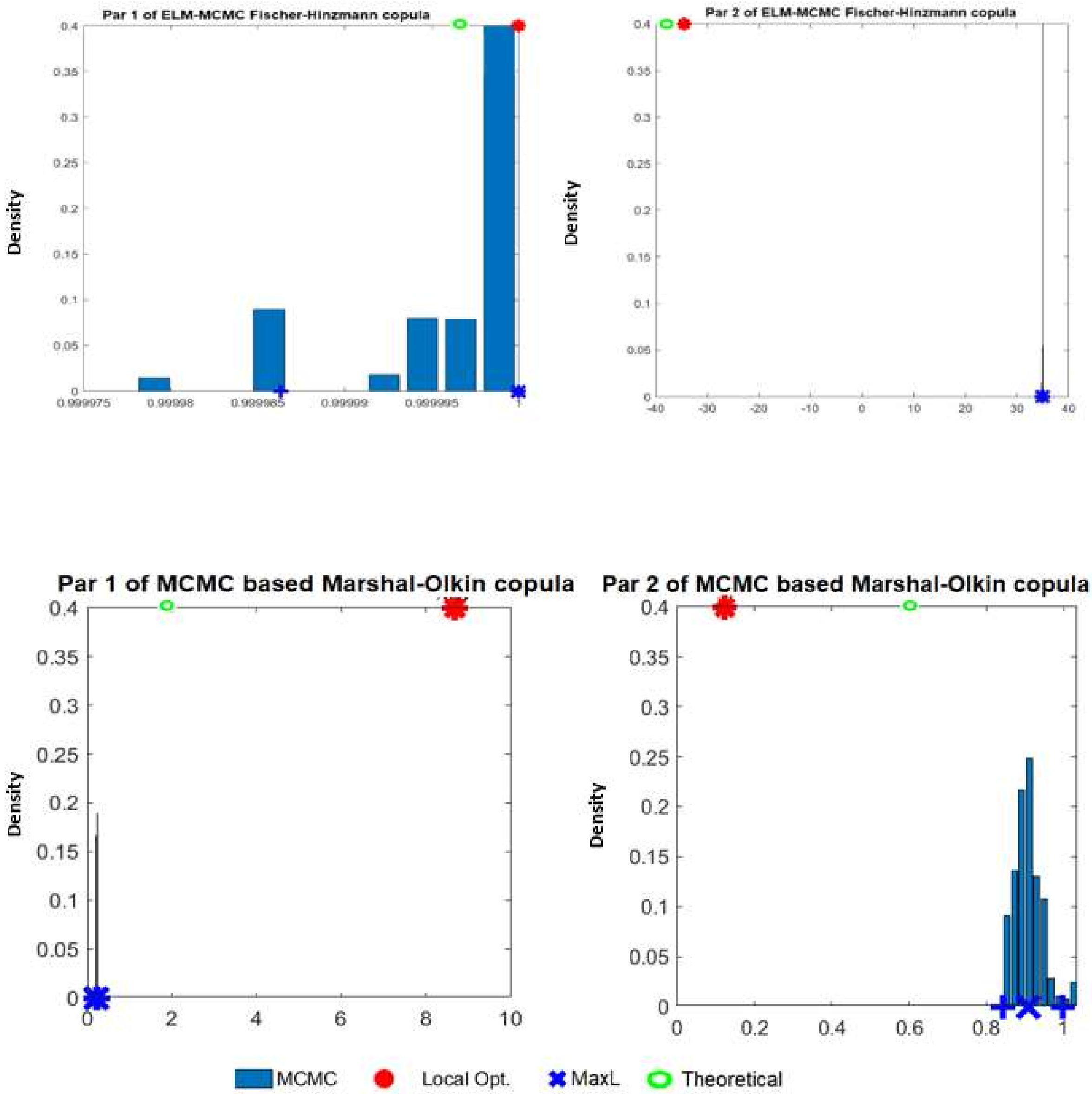


Figure 7: *Posterior distribution of the ELM-MCMC copulas obtained by global MCMC simulation for 6-hours timescale. The blue bins are the MCMC-derived parameters and the crosses (aqua) denote the maximum likelihood parameter of the MCMC. It also shows the local optimization (red circle) method for comparison.*

The results of the ELM-MCMC copula-based models have been compared against MCMC based copula models and the standalone ELM model using the evaluation criteria expressed above (equations 13 – 22). **Figure 7** shows the histogram of posterior parameter distribution (blue bins) of copula uncertainties obtained from MCMC simulation within a Bayesian framework after all 26 families of copula were evaluated by executing the observed and forecasted *D* data in the MvCAT. It shows the posterior parameter distributions of the best ranked ELM-MCMC-Fischer-Hinzmann copula against the MCMC Marshal-Olkin copula using the global MCMC (Bayesian inference) method.

The posterior parameter of the ELM-MCMC Fischer-Hinzmann copula is well constrained compared to the MCMC Marshal-Olkin copula. The first copula parameter (see **Table 6**) obtained by the local optimization method (red circle) coincides with the maximum likelihood parameters (blue cross) of the posterior distributions (blue bins) of the ELM-MCMC-Fischer-Hinzmann copula. The second parameter of both the local and MCMC methods diverge significantly. However, this is not the case for MCMC Marshal-Olkin copula where its suggested local parameters diverge extensively from their counterpart global MCMC parameters. The results also show that the theoretical parameter (green circle) is closest to both the local and global MCMC parameters compared to the local parameters for MCMC based Marshal-Olkin copula.

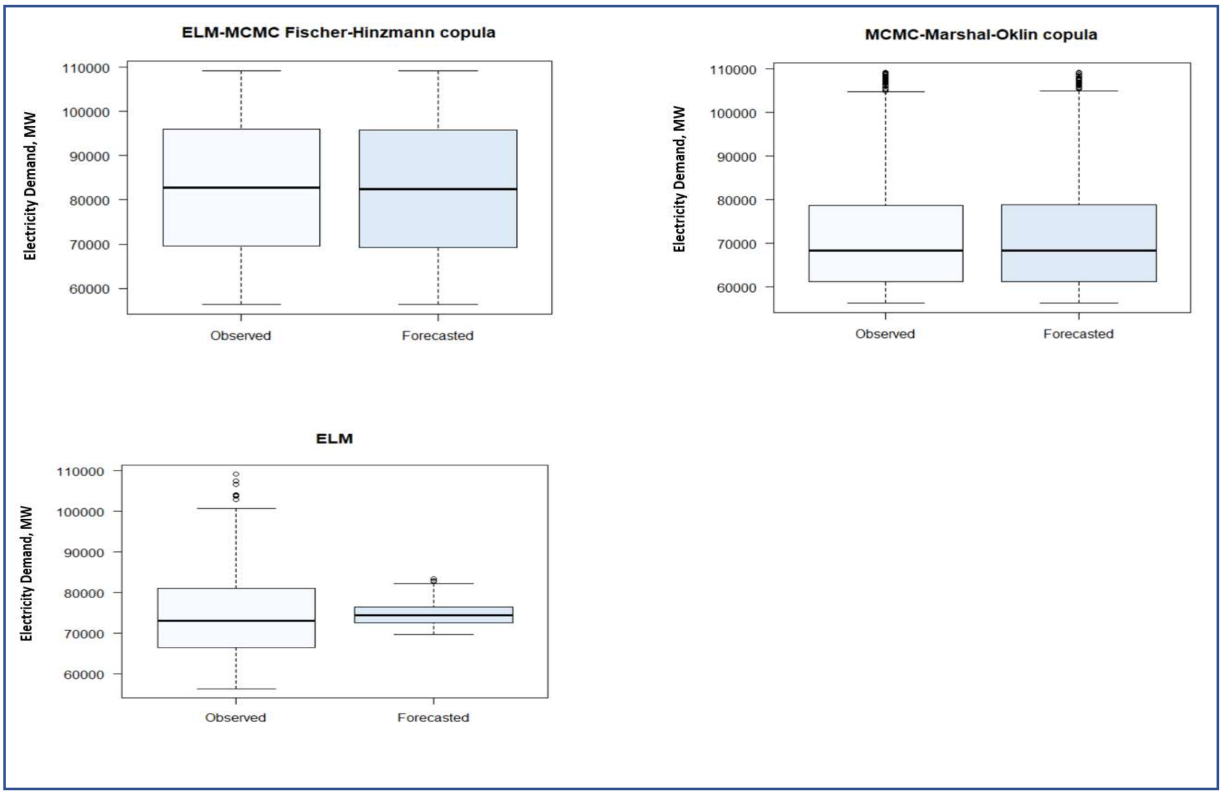


Figure 8: *Boxplots of the observed and forecasted D for ELM-MCMC copula-based models vs MCMC copula-based models and the standalone ELM model for electricity demand for 6-hourly timescale.*

Figure 8 clearly shows that the observed and forecasted D derived from the ELM-MCMC-Fischer-Hinzmann copula-based model is normally distributed and symmetrical, hence more accurate whereas the spread of the *D*-observed and forecast for the other two comparative models was asymmetrical. This shows that forecasting from a standalone AI - based model (*i.e*., an ELM model) is not stable, as confirmed from the test results in **Table 7**. This accords with the view of Cook et al. (2019) that model hybridization improves performance irrespective of the nature and size of data, as each of the amalgamated models contribute to addressing the deficiencies that others.

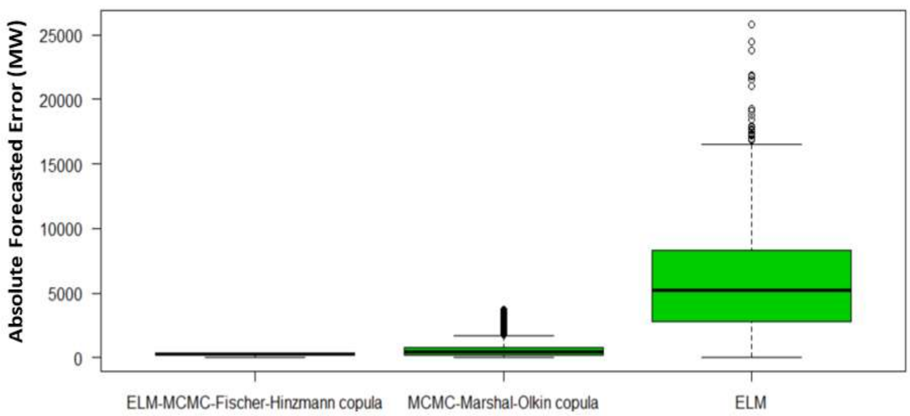


Figure 9: *Boxplots of absolute forecasting error (MW) of ELM-MCMC copula-based models against MCMC copula-based model and the standalone ELM model in forecasted electricity demand (D) for 6-hours timescale.*

Boxplots in **Figure 9** were plotted using the absolute forecasted error, $|FE| = |D_{obs} - D_{for}|$, for 6-hours timescales for the ELM-MCMC copula-based models with the counterpart MCMC copula-based and the standalone ELM model. The outliers are denoted by ° in each boxplot indicating the significant variability of |FE| in the testing period. The first line at the bottom of the boxplot represents the first quartile, the middle line shows the median while the third line at the upper end represents the third quartile of the |FE|. According to those quartiles, the larger spread was evident for both the MCMC copula-based models and standalone ELM models compared to the hybrid ELM-MCMC copula-based model. Accordingly, the ELM-MCMC-

Fischer-Hinzmann copula model exhibited highly accurate performance for *D*-forecasting with less spread of |FE| at 6-hours timescale followed by the MCMC-Marshall-Olkin copula and the standalone ELM model.

4.1.3 Performance of ELM-MCMC copula-based model for 12-hours

The testing performance of the best ELM-MCMC copula-based model against the MCMC copula-based model and standalone ELM model for 12-hours timescale were evaluated using the model performance matrices shown in **Table 8**.

Table 8: *Testing performance for the objective model for 12-hourly timescale: Hybrid extreme learning machine model integrated with Markhov Chain Monte Carlo (ELM-MCMC) in comparison with standalone MCMC copula-based model and the ELM model. The results displayed below only shows the best copula in terms of its copula rankings and its associated model performance. The model input for all timescale is the lagged demand at first lag (t-1). Note that the best model is boldfaced (black).*

	12-hours							
Copula	***RMSE (MW)***	***RRMSE (%)***	***MAE (MW)***	***RMAE (%)***	L_M	***AIC***	***BIC***	***MaxL***
ELM-MCMC copula-based model								
Fischer-Hinzmann	534.09	0.32	490.661	0.309	0.98	-34727.47	-34718.28	17365.73
MCMC copula-based model								
Marshal-Olkin	1635.21	1.140	1255.31	0.902	0.927	-11482.88	-11472.31	5743.60
ELM model								
	12524	8.398	9720.71	6.417	0.285			

For the 12-hour timescale (**Table 8**), again the ELM-MCMC-Fischer-Hinzmann copula appeared to be the best ranked copula, having minimum error terms, maximum $L_{M,}$ lowest *AIC* and *BIC*, and highest Max_L values. Particularly, it attained *RMSE* ≈ 534.09 MW, *RRMSE* ≈ 0.32 %, *MAE* ≈ 490.661 MW, *RMAE* ≈ 0.309 %, L_M ≈ 0.98, *AIC* ≈ -34727.47, *BIC* ≈ -34718.28, Max_L ≈ 17365.73. This was followed by the second best performing model, the MCMC-Marshall-Olkin copula-based model with corresponding error values of *RMSE* 1635.21 MW,

RRMSE ≈ 1.140 %, *MAE* ≈ 1255.31 MW, *RMAE* ≈ 0.902 %, L_M ≈ 0.927, *AIC* ≈ -11482.88, *BIC* ≈ -11472.31, Max_L ≈ 5743.60, and finally the worst performing ELM model with *RMSE* ≈ 12524 MW, *RRMSE* ≈ 8.398 %, *MAE* ≈ 9720.71 MW, *RMAE* ≈ 6.417 %, and L_M ≈ 0.285.

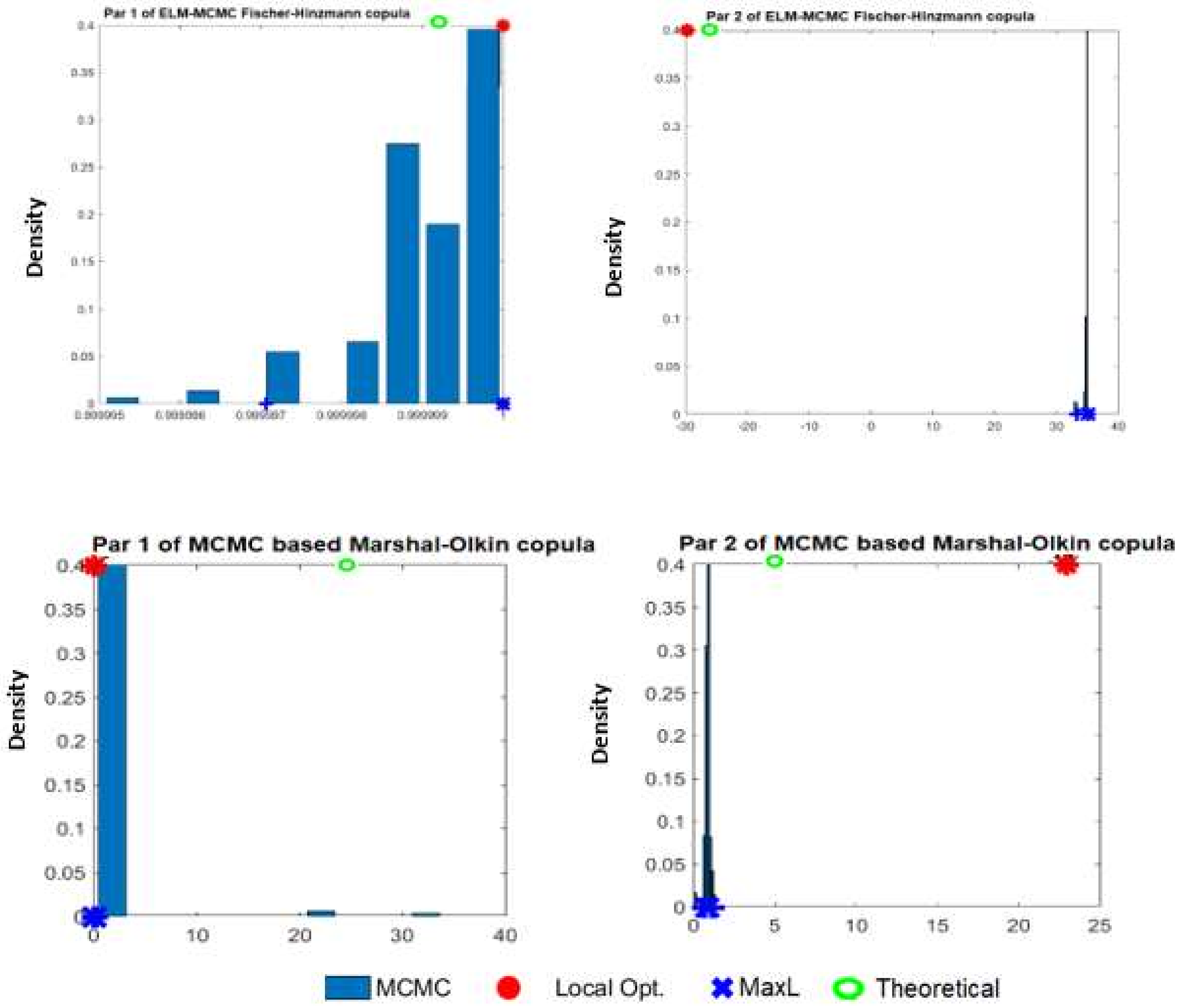


Figure 10: *Posterior distribution of the ELM-MCMC copulas obtained by global MCMC simulation for 12-hours timescale. The blue bins are the MCMC-derived parameters and the crosses (aqua) denotes the maximum likelihood parameter of the MCMC. It also shows the local optimization (red circle) method for comparison.*

Figure 10 shows the posterior distributions (blue bins) of the best ranked ELM-MCMC-Fischer-Hinzmann copula against the MCMC Marshal-Olkin copula-based model once all 26 families of copula were evaluated after executing the observed and forecasted *D* data. The posterior distributions of the ELM-MCMC-Fischer-Hinzmann copula shift to parameter bounds, with the best parameter located on the boundary, suggesting that the optimization algorithm has forcefully gone beyond the bounds to improve model fit. The first parameter of the MCMC Marshal-Olkin copula-based model match with parameter values suggested by both local and global MCMC optimization approaches, with Max_L value of zero corresponding with local optimization method as opposed to the ELM-MCMC copula-based model. The second parameter of the MCMC Marshal-Olkin copula-based model has slightly uneven distribution where the local and global parameters diverge significantly, with best parameter (Max_L) almost being zero. This implies that there is insufficient information to constrain two parameters of the Marshal-Olkin copula. Also, the same applies for the second parameter of the ELM-MCMC Fischer-Hinzmann copula.

The copula parameters obtained by applying the global MCMC algorithm shows that the local optimization (red circle) corresponded with the maximum likelihood parameters (blue cross) of the posterior distributions (blue bins) of the best ELM-MCMC-Fischer-Hinzmann copula-based model and the first parameter of the MCMC-Marshal-Olkin copula-based model. This means both local and global MCMC methods can be used to estimate the best copula parameters for the 12-hours timescale. However, the best parameter (Max_L) value of ELM-MCMC-Fischer-Hinzmann copula registered a higher value compared to the MCMC Marshal-Olkin copula which was almost zero. This implies that the ELM-MCMC-Fischer-Hinzmann copula had a lower residual error and has been therefore accurate compared to its counterpart MCMC Marshal-Olkin copula. The results show that ELM-MCMC-Fischer-Hinzmann copula-based model is the optimal model for *D*-forecasting as its local parameters (red circle) equally corresponded with the global MCMC parameters, thereby achieving high projection accuracy.

The results also show that the theoretical parameter (green circle) is closest to the global MCMC parameters compared to the local parameters for MCMC based Marshal-Olkin copula. This means that for 12-hours timescale, the global MCMC copula parameter estimation is accurate compared to its counterpart local optimization method.

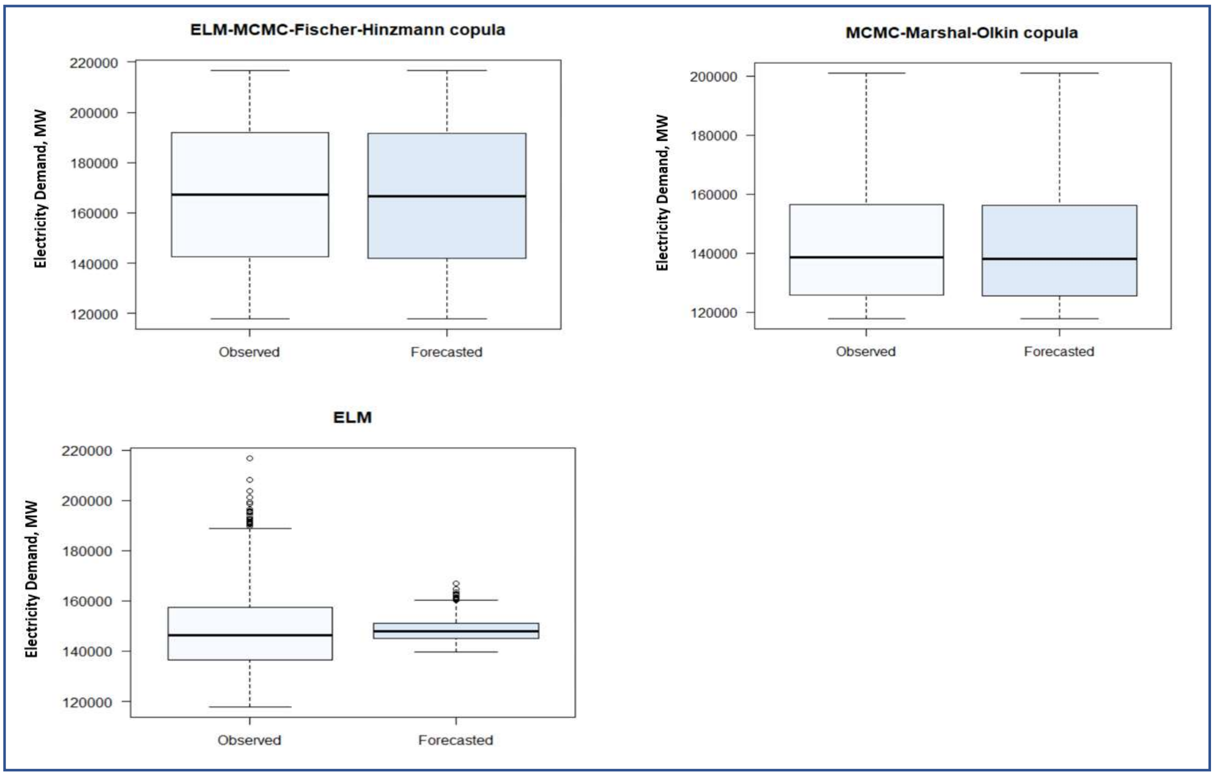


Figure 11: *Boxplots of the observed and forecasted D for ELM-MCMC copula-based models vs MCMC copula-based models and the standalone ELM model for electricity demand for 12-hourly timescale.*

Figure 11 displays the observed and forecasted *D* derived from the best ELM-MCMC-Fischer-Hinzmann copula-based model against the best MCMC-copula-based model and the standalone ELM model for 12-hours timescale. It is evident that the observed and forecasted *D* from the ELM-MCMC-Fischer-Hinzmann copula-based model is normally distributed and symmetrical as it demonstrates equal spread between the first and third inter-quartile range (IQR) with the mean of *D*-observed equating to *D*-forecast and vice versa. Hence, the model is optimal, more stable, and accurate whereas the IQR of both *D*-observed and forecast for the MCMC-Marshall-Olkin copula model are equal, except that the whisker between the third quartile and maximum values is bigger than that between the first quartile and minimum values. For the standalone ELM model, there is significant mismatch between the observed and forecasted *D*, hence the model is not accurate.

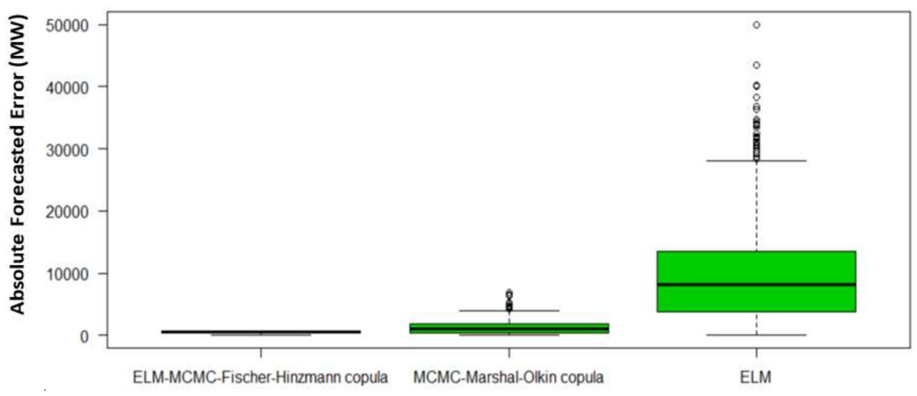


Figure 12: *Boxplots of absolute forecasting error (MW) of ELM-MCMC copula-based models against MCMC copula-based model and the standalone ELM model in forecasted electricity demand (D) for 12-hours timescale.*

Boxplots in **Figure 12** were plotted using the absolute forecasted error, |FE| = $|D_{\text{obs}} - D_{\text{for}}|$, for 12-hours timescales for the ELM-MCMC copula-based models with the counterpart MCMC copula-based and the standalone ELM model. The outliers are denoted by ° in each boxplot indicating the significant variability of |FE| in the testing period. The first line at the bottom of the boxplot represents the first quartile, the middle line shows the median while the third line at the upper end represents the third quartile of the |FE|. According to those quartiles, the larger spread was evident for both the MCMC copula-based models and standalone ELM models compared to the hybrid ELM-MCMC copula-based model. Accordingly, the ELM-MCMC-Fischer-Hinzmann copula model exhibited highly accurate performance for *D*-forecasting with less spread of |FE| at 12-hours timescale followed by the MCMC-Marshall-Olkin copula and the standalone ELM model.

4.1.4 Performance of ELM-MCMC copula-based model for 24-hours

The testing performance of the best ELM-MCMC copula-based model against the MCMC copula-based model and standalone ELM model for 24-hours (daily) timescale was evaluated using the model performance matrices shown in **Table 10**.

Table 9: *Testing performance for the objective model for daily timescale: Hybrid extreme learning machine model integrated with Markov Chain Monte Carlo (ELM-MCMC) in comparison with standalone MCMC copula-based model and the ELM model. The results displayed below only show the best copula in terms of its copula rankings and its associated model performance. The model input for all timescales is the lagged demand at first lag (t-1). The boldface (black) text indicates the best model.*

Copula	*RMSE (MW)*	*RRMSE (%)*	*MAE (MW)*	**Daily** *RMAE (%)*	L_M	*AIC*	*BIC*	*MaxL*
				ELM-MCMC copula-based model				
Cuadras-Auge	694.769	0.22	638.365	0.208	0.98	-14514.31	-14510.41	7258.16
				MCMC copula-based model				
Burr	835.91	1.16	573.119	0.829	0.947	-3641.12	-3637.22	1821.60
				ELM model				
	3690.59	1.234	2983.86	0.996	0.851			

Table 9 presents the model performance in terms of the error terms of the ELM-MCMC copula-based model against two comparative models, the MCMC copula-based model and the standalone ELM model for the daily timescale. It is evident that the ELM-MCMC Cuadras-Auge copula outperformed both of its comparative models with corresponding error matrices of *RMSE* ≈ 694.769 WM, *RRMSE* ≈ 0.22 %, *MAE* ≈ 638.365 WM, *RMAE* ≈ 0.208 %, L_M ≈ 0.98, *AIC* ≈ -14514.31, *BIC* ≈ -14510.41, and Max_L ≈ 7258.16. These performance metric were compared against the counterpart MCMC-based Burr copula with *RMSE* ≈ 835.91 MW, *RRMSE* ≈ 1.16 %, *MAE* ≈ 573.119 WM, *RMAE* ≈ 0.829 %, L_M ≈ 0.947, *AIC* ≈ -3641.12, *BIC* ≈ -3637.22, and Max_L ≈ 1821.60, followed by the standalone ELM model with *RMSE* ≈ 3690.59 MW, *RRMSE* ≈ 1.234 %, *MAE* ≈ 2983.86 MW, *RMAE* ≈ 0.996 %, and L_M ≈ 0.851.

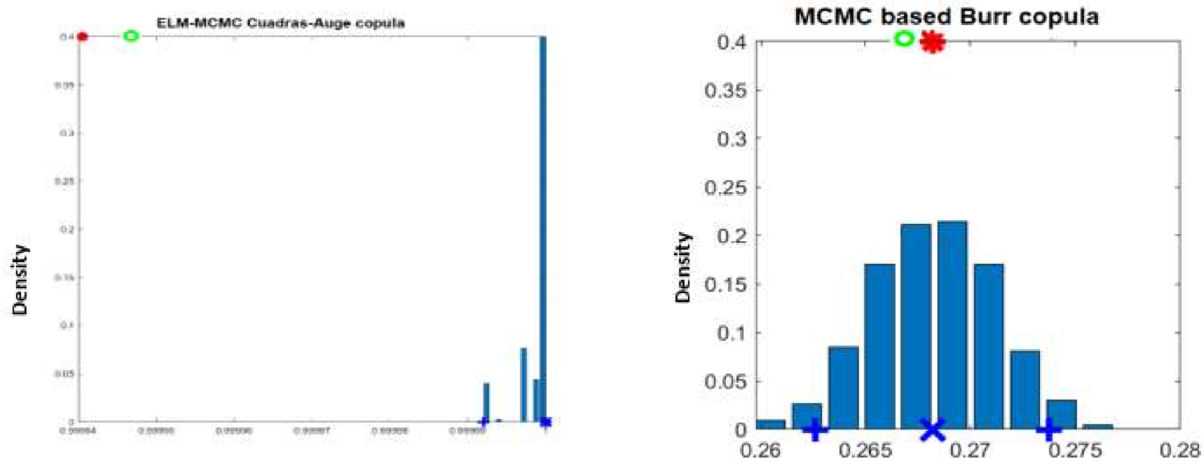


Figure 13: *Posterior distribution of the ELM-MCMC copulas obtained by global MCMC simulation daily timescale. The blue bins are the MCMC-derived parameters and the crosses (aqua) denotes the maximum likelihood parameter of the MCMC. It also shows the local optimization (red circle) method for comparison.*

Figure 13 shows the posterior distributions (blue bins) of the best ranked ELM-MCMC-Cuadras-Auge copula and the MCMC-Burr copula once all 26 families of copula were ranked after running the observed and forecasted *D* data for the daily timescale. The copula parameters obtained by applying the global MCMC algorithm shows that the local parameter (red circle) which corresponded with maximum likelihood parameters (blue cross) of the posterior distributions (blue bins) of the best ELM-MCMC-Cuardas-Auge copula model and the MCMC-Burr copula model. The posterior distributions of the ELM-MCMC-Cuadras-Auge copula shift to parameter bounds, with the best parameter located on the boundary. This suggests that the optimization algorithm has forcefully gone beyond the bound to improve model fit. The result also shows that the local parameter (red circle) of the ELM-MCMC-Caudras-Auge copula model is closest to theoretical parameter (green circle) compared to its counterpart MCMC parameter. The posterior distributions (blue bins) and maximum likelihood (blue cross) resulting from the global MCMC algorithm for MCMC Burr copula coincided with the copula parameters obtained by the local optimization method (red circle). The MCMC-Burr copula has uniform distribution where the parameters obtained by local optimization technique coincided with the global MCMC method. However, its best parameter (Max_L) is less compared to the counterpart ELM-MCMC-Cuadras-Auge model.

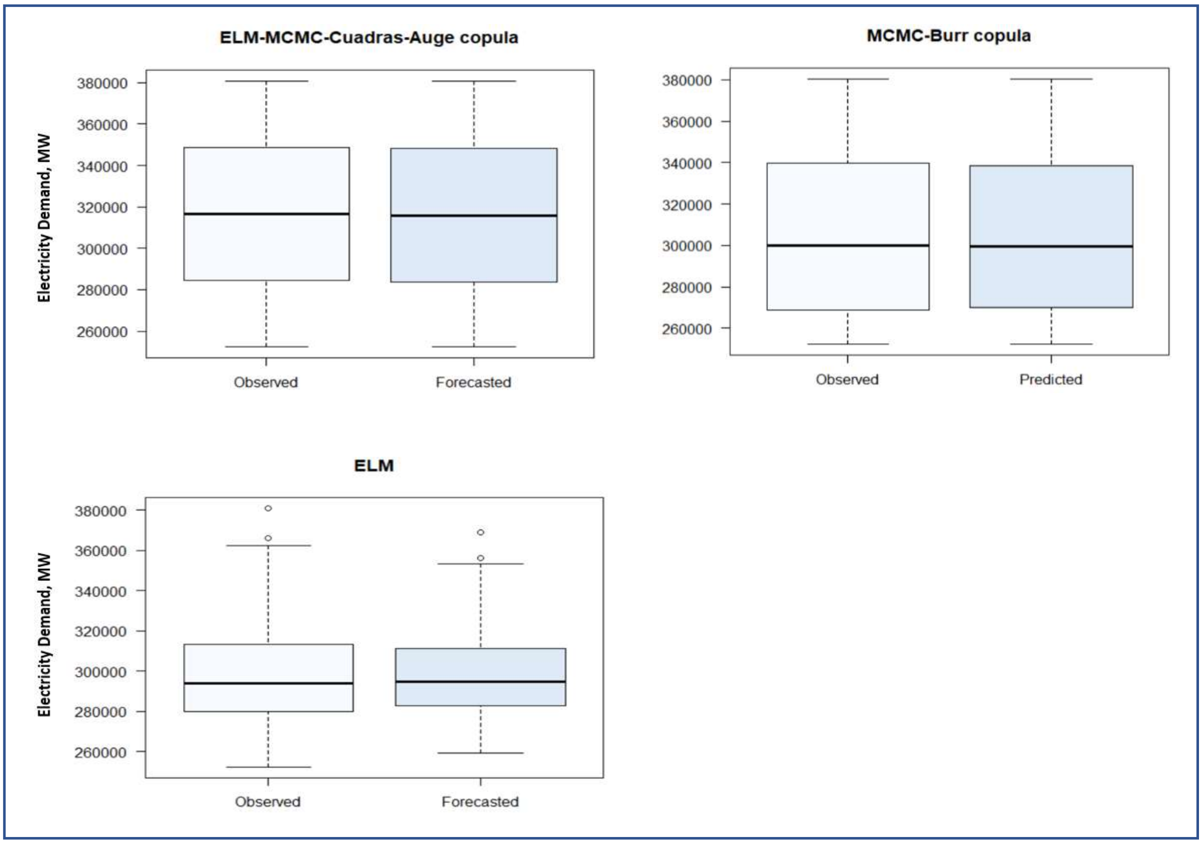


Figure 14: *Boxplots of ELM-MCMC copula-based models vs MCMC copula-based models and the standalone ELM model for electricity demand for daily timescale.*

Figure 14 shows the observed and forecasted *D* derived from the best ELM-MCMC-Cuadras-Auge copula-based model against the best MCMC-Burr copula model and the standalone ELM model for the daily timescale. It is evident that the observed and forecasted *D* from the ELM-MCMC-Cuadras-Auge copula-based model is normally distributed and symmetrical as it demonstrates equal spread between the first and third inter quartile range (IQR) with the mean of *D*-observed equating to *D*-forecast and vice versa. Hence, the model is optimal, more stable, and accurate whereas the IQR of both *D*-observed and forecast for the MCMC-Marshall-Burr copula model are equal, except that the whisker between the third quartile and maximum values is bigger than that between the first quartile and minimum values. For the standalone ELM model, there is significant mismatch between the observed and forecasted *D*, denoting high residual error, hence the model is not accurate.

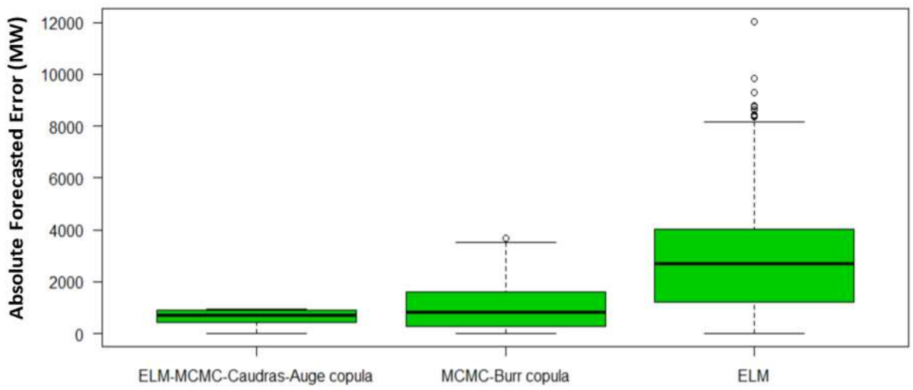


Figure 15: *Boxplots of absolute forecasting error (MW) of the best ELM-MCMC-Cuadras-Auge copula model against the best MCMC-Burr copula model and the standalone ELM model for daily timescale.*

Boxplots in **Figure 15** were plotted using the absolute forecasted error, FE = $|D_{obs} - D_{for}|$, for the daily timescale for the best ELM-MCMC-Cuardas-Auge copula model in comparison with the MCMC-Burr copula model and the standalone ELM model. The outliers are denoted by ° in each boxplot indicating the significant variability of FE in the testing period. The first line at the bottom of the boxplot represents the first quartile, the middle line shows the median while the third line at the upper end represents the third quartile of the FE. According to those quartiles, the standalone ELM model had the largest spread followed by the MCMC-Burr copula model, while the ELM-MCMC-Cuadras-Auge copula had the least spread of FE. Therefore, the ELM-MCMC-Cuadras-Auge copula model exhibited highly accurate performance for *D*-forecasting exhibiting least FE at daily timescale followed by the MCMC-Burr copula and finally the standalone ELM model.

4.2 Comparisons of all timescales

The performance of the objective model was compared against the comparative models for all timescales (6-hours, 12-hours, and daily) in terms of the relative root mean square error (*RRMSE*) and relative mean absolute error (*RMAE*).

Table 10: *Periodic comparison of the predictive performance of the best ELM-MCMC vs. the counterpart MCMC copula-based models and standalone ELM model which are evaluated in terms of relative root mean square error (RRMSE), relative mean absolute error (RMAE), and the Legates and McCabe' Index (L_M). Note that the robust model is coloured black and bolded.*

6-hours ELM-MCMC			
Copula	***RRMSE (%)***	***RMAE (%)***	***L_M***
Fischer-Hinzmann	0.345	0.336	0.980
MCMC			
Marshal-Olkin	1.160	0.829	0.947
ELM			
	9.476	7.878	0.265

12-hours ELM-MCMC			
Copula	***RRMSE (%)***	***RMAE (%)***	***L_M***
Fischer-Hinzmann	0.32	0.309	0.98
MCMC			
Marshal-Olkin	1.14015	0.902	0.927
ELM			
	8.398	6.417	0.285

Daily ELM-MCMC			
Copula	***RRMSE (%)***	***RMAE (%)***	***L_M***
Cuadras-Auge	0.22	0.208	0.98
MCMC			
Burr	1.16	0.829	0.947
ELM			
	1.234	0.996	0.851

Table 10 shows the periodic comparisons of the ELM-MCMC copula-based models against the MCMC-copula-based models and the standalone ELM model based on the *RRMSE*, *RMAE*

and L_M for different timescales (6-hours, 12-hours, and daily). In terms of periodic comparisons, the daily timescale employing the ELM-MCMC-Cuadras-Auge copula performed accurately by displaying the smallest percentage errors for both *RRMSE* (0.22 %) and *RMAE* (0.208 %) with maximum L_M (0.980). It was followed by the 12-hourly (*RRMSE* ≈ *0.320 %, RMAE* ≈ *0.309 %,* L_M ≈ 0.980) and finally the 6-hourly (*RRMSE* ≈ *0.345 %, RMAE* ≈ *0.336 %,* L_M ≈ 0.980) respectively, with both timescales employing the ELM-MCMC-Fischer-Hinzmann copula model. The overall forecast generated at each timescale by the respective models displayed excellent performance as their relative errors were less than the 10% threshold as stated in Ertekin and Yaldiz (2000) and Mohammadi et al. (2015a).

4.3 Concluding Remarks

This chapter investigated the performance of a novel hybrid ELM-MCMC copula-based model for forecasting *D* at various timescales (6-hours, 12-hours, and daily). Based on the results in sections 4.6.2–4.6.4, the hybrid ELM-MCMC copula-based model displayed highly accurate results for all timescales. For both 6-hours and 12-hours, the ELM-MCMC-Fischer-Hinzmann copula outclassed its counterparts by means of displaying consistent results in these timescales. Hence, the results of the study in this chapter accords with the view of Cook *et al.* (2019) that hybridization of models improves forecasting accuracy. Moreover, the performance ranking of both objective and comparative models for all timescales in descending order is the ELM-MCMC copula-based models followed by the MCMC-copula-based models and closing with the ELM model.

CHAPTER 5: ELECTRICITY DEMAND FORECASTING USING HISTORICAL PRICE AND PROBABALISTIC-COPULA MODELS

5.1 Development of conditional bivariate copula

A copula is a joint distribution with two or more random variables in play in a particular probability problem where each of the single-dimensional margins is marginally uniformly distributed over the distribution scale (0,1). The copula models can attract random time-independent variables regardless of the type of their marginal distributions. According to Sklar's theorem (Sklar 1959), presuming that each continuous variable X_i (*i.e.* D and PR in this study) with its independent marginal cumulative distribution function (CDF) is expressed as $F_k(X_k)$ while the probability distribution function (PDF) is given as $f_k(X_k)$. Given the S-dimensional variables $(X_1 \ldots . X_s)$, the CDF can be expressed as:

$$F(x_1, x_2 \ldots, x_s) = C[F_1(x_1), F_2(x_2), \ldots, F_s(x_s)] \qquad 23$$

And the equivalent marginal PDF can be expressed as:

$$f(x_1, \ldots, x_s) = [\coprod_{i=1}^{s} f_i(x_i)] c[F_1(x_1), \ldots, F_s(x_s)] \qquad 24$$

The unit hypercube $C: [0,1]^s \rightarrow [0,1]$ relates to s-variate distribution with a copula (c) of unique uniform marginal distribution with its associated copula density $c = \frac{\partial^s}{\partial_1 \ldots \partial_s} C(u_1, \ldots, u_s)$. In this function, u_1 represents $F_1(x_1)$ while $u_1 \epsilon (0,1)$ is defined as the probability integral transform (PIT), (Killiches et al. 2017; Kraus & Czado 2017). This *c* can further be disintegrated to create a conditional bivariate copula densities s(s-1)/2. This result is consistent with equation (29) which clearly shows that both the dependence structure and marginal distributions can be modeled independently which gives rise to copula model approach.

5.1.1 Type of copula families

There is a wide range of copula families such as elliptical, Archimedean, vine, extreme value, among others (Czado et al. 2013; Tosunoglu & Singh 2018; Tosunoglu et al. 2020). The Archimedean copula family is used extensively in different applications and can either have symmetric or asymmetric distributions. Also, the use of multivariate elliptical copula is prevalent in real-world applications but does have its shortfalls so as the Archimedean copula. For example, the performance capability in higher dimensions is limited when elliptical copulas are applied as they portray symmetric tail dependence while Archimedean copulas possess numerous dependence parameters. However, such limitations can be subdued by adoption of vine copulas in disintegrating the multivariate copulas into many bivariate copulas (Aas et al. 2009; Fischer et al. 2017) to identify the dependence between complex variables.

5.1.2 Vine Copula

The three basic types of vine copulas are the Regular vine, Drawable vine, and Canonical vine. The concepts of vine copulas were introduced by Joe (1997). Bedford and Cooke (2001) then assessed their performance in depth by way of graphical reliance models in utilizing Markov trees for developing bivariate copulas as well as defining multivariate variables. Studies by AghaKouchak et al. (2010) and Nguyen-Huy *et al*. (2017) for precipitation forecasting indicated that vine copulas suitably address tail dependence and asymmetries issues.

5.2 Conditional probability of electricity demand given lagged value of price

Copulas are statistical models having computational capability for solving non-linear dependence. This is efficiently achieved by ranking the dependence structure using Kendall's tau correlations, which the linear correlation coefficient is unable to solve (Manner et al. 2019). Copulas employ advanced statistical techniques that can model multivariate data for investigating dependence. The method also includes the determination of bivariate/multivariate distributions having non-linear dependence. Moreover, the marginal distributions sometimes exhibit asymmetric distributions.

This method develops a copula-statistical model (Sklar 1959; Sklar 1996) for probabilistic forecasting of *D* using significant lagged correlation of *PR*. It involves the use of conditional bivariate copula models to examine the joint behaviour of *D* and the significant lagged *PR* for probabilistic forecasting of *D*. Thus, it is necessary to investigate the lagged correlation between *PR* and *D* first. In this study, the correlation between *PR* and *D* was computed using the cross-correlation function (CCF). Based on the CCF plots, the first lag (t-1) had the highest correlation (refer to **Figure 16**). Therefore, it was used as the covariate predictor in the bivariate copula model for probabilistic forecasting of *D*. **Figure 17** summarizes the main steps for developing a probabilistic forecast model. This approach provides a more robust fit than single-equation models and computes accurate forecasts for spike probabilities (Manner et al. 2016).

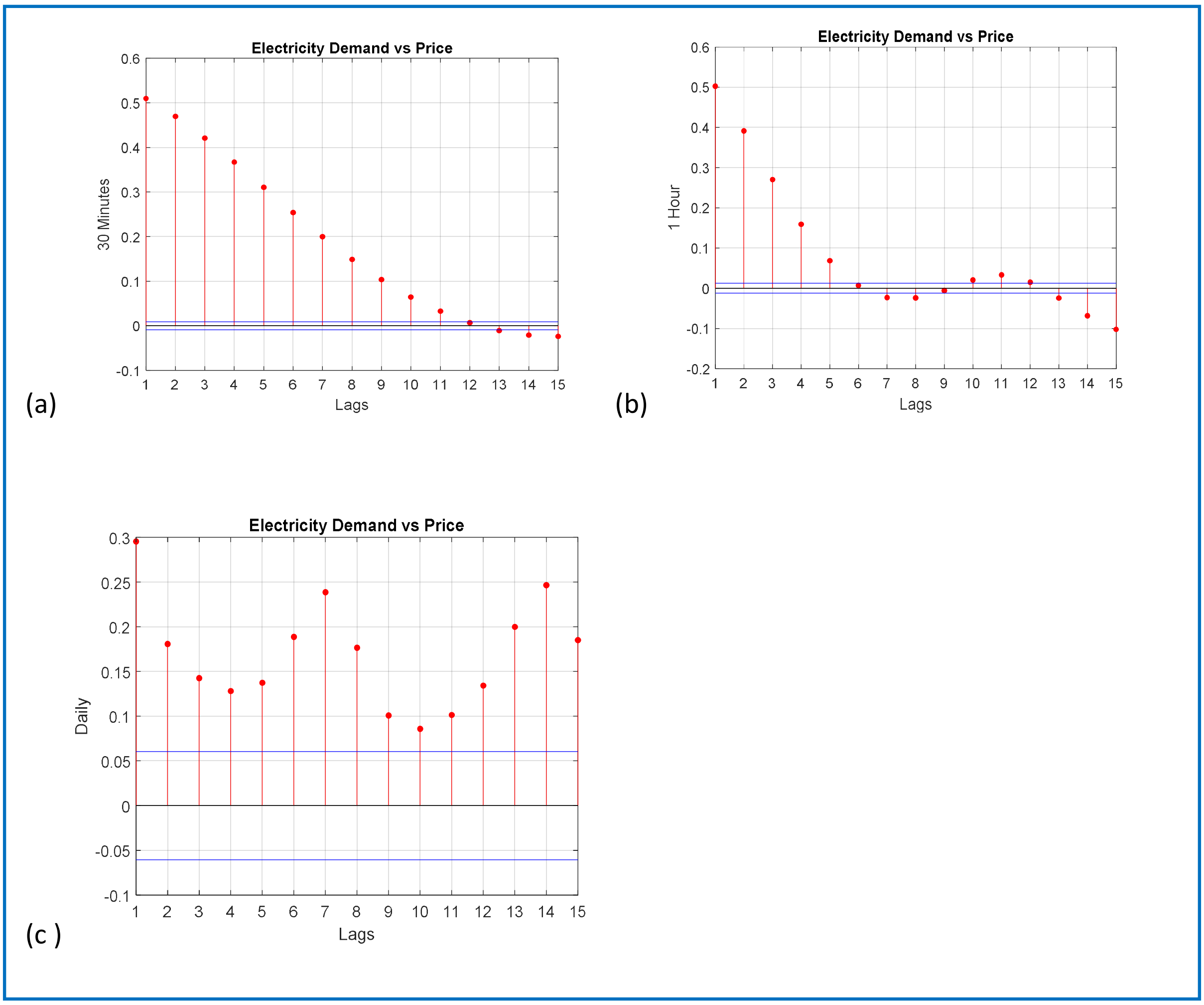


Figure 16: *Crosscorrelation plots of electricity demand and price for (a) 30-minutes, (b) 1-hour, (c) and daily timescales from 2017 to 2019.*

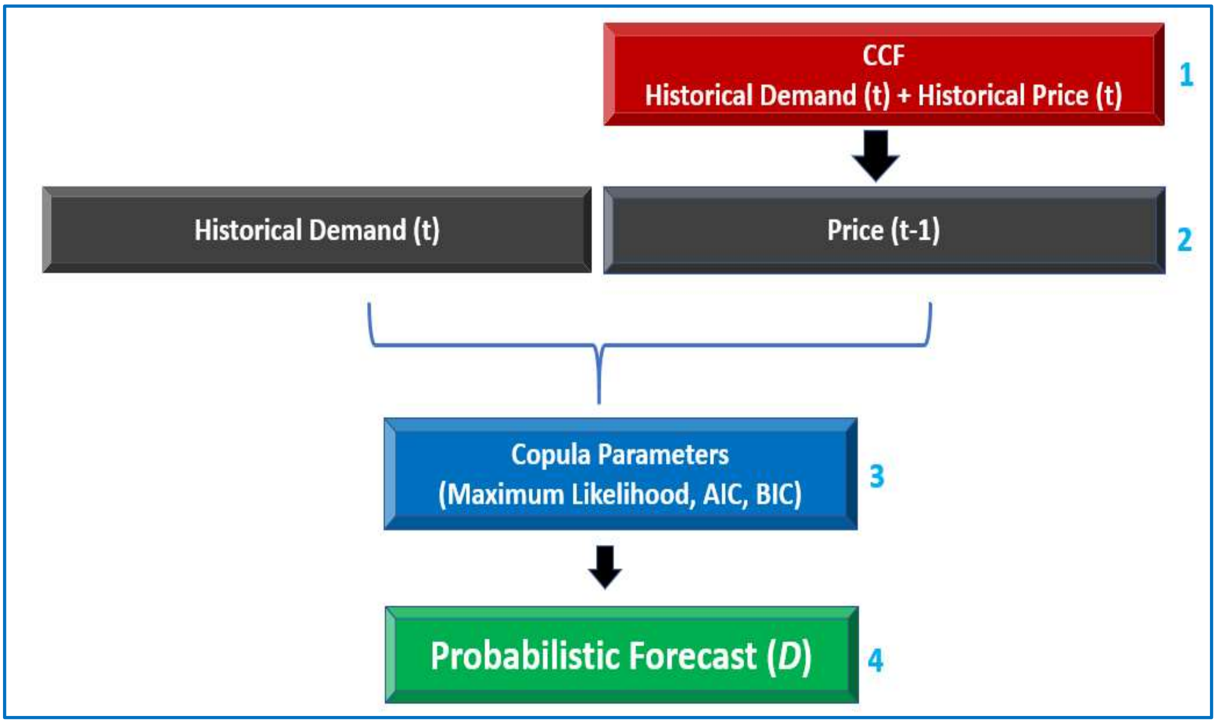


Figure 17: *Steps for developing probabilistic forecast model for objective two.*

In the next step, a copula function (C) was used to model the joint probabilistic behaviour of the joint cumulative distribution function (CDF) $F_{PRD}(pr,d)$ between *PR* and *D, i.e.* $F_{PRD}(pr,d) = C[F_{PR}(pr), F_D(d)]$. The most appropriate copula function used to describe the dependence structure between *PR* and *D*, among many candidates, was selected based on different criteria including *AIC, BIC,* and logarithmic likelihood (*logLik*). It is noted that the copula parameters were estimated using the maximum likelihood estimate (*i.e.* local method), which is different to the one in the MCMC model mentioned in Chapters 3 and 4. Also, it is worth repeating that the ranked Kendall tau correlation coefficient was used to model non-linear dependence between *D* and *PR*. Finally, the probabilistic forecast of *D* based on the first lag of *PR* (*i.e.*, *PR*(t-1)) was derived via the conditional probability P (X<x | Y =y). Three different timescales (30-minutes, 1-hour, and daily) were selected to examine the model performance.

This chapter studies two scenarios of *D*-forecast based on the first lag of *PR* and the copula model. First, we explore the nonexceedance probabilities of *D,* given that the *PR* peak exceeds certain thresholds of *pr.* According to Madadgar and Moradkhani (2013), the conditional probability based on bivariate copulas is expressed as follows:

$$P(D \leq d | PR \geq pr) = \frac{P(D \leq d, PR \geq pr)}{P(PR \geq pr)} = \frac{F_D(d) - F_{PRD}(pr,d)}{1 - F_{PR}(pr)}$$

$$= \frac{F_D(d) - C[F_{PR}(pr), F_D(d)]}{1 - F_{PR}(pr)} \quad (25)$$

Conversely, the exceedance probabilities of *D,* given that the *PR* peak exceeds certain thresholds of *pr,* can be expressed as (Nguyen-Huy et al. 2018):

$$P(D \geq d | PR \geq pr) = \frac{P(D \geq d, \; PR \geq pr)}{P(PR \geq pr)}$$

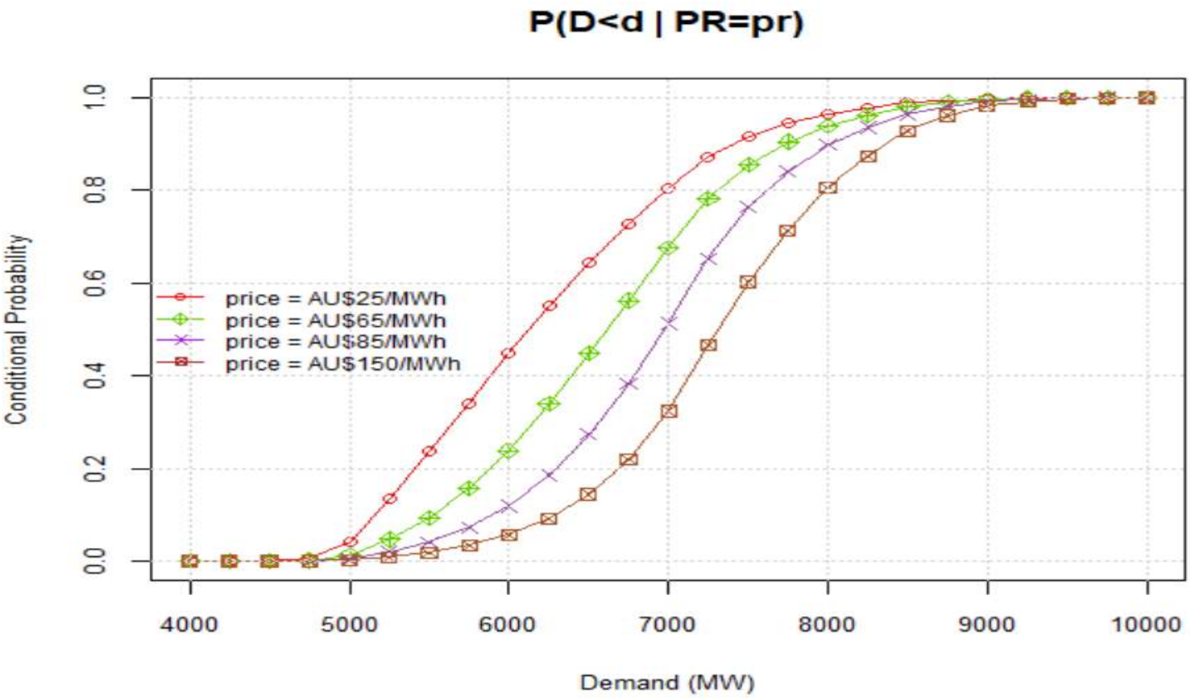
$$= \frac{1 - F_{PR}(pr) - F_D(d) - C[F_{PR}(pr), F_D(d)]}{1 - F_{PR}(pr)} \quad (26)$$

5.3 Results and Discussion

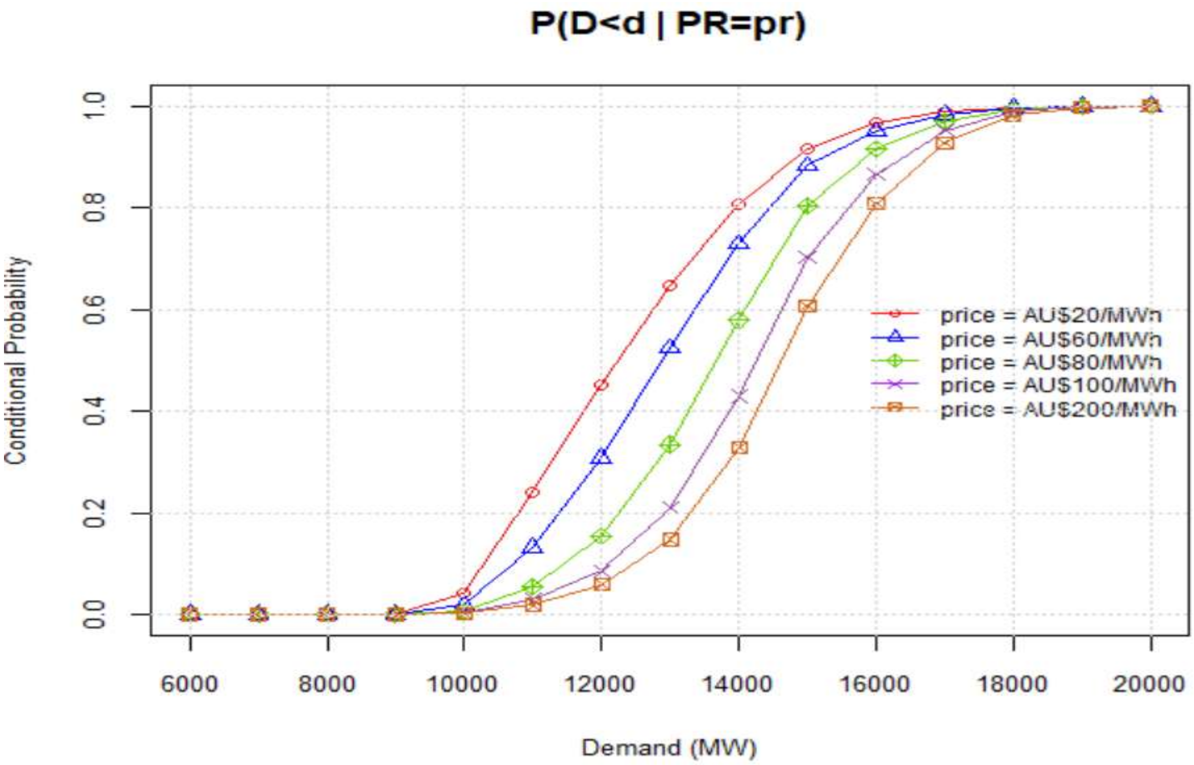
5.3.1 Conditional nonexceedance probabilistic forecasting of electricity demand (D).

The results for conditional nonexceedance probabilistic forecasting of *D* for 30-minutes, 1-hour, and daily timescales, given lagged price *PR* (t-1) greater than or equal to certain threshold price (AU$/MW) by applying conditional bivariate copula models are provided in **Figure 18 (a - c)**.

(a)


P(D<d | PR=pr)
Conditional Probability
Demand (MW)
price = AU$25/MWh
price = AU$65/MWh
price = AU$85/MWh
price = AU$150/MWh
0.0
0.2
0.4
0.6
0.8
1.0
4000
5000
6000
7000
8000
9000
10000

(b)


P(D<d | PR=pr)
Conditional Probability
Demand (MW)
price = AU$20/MWh
price = AU$60/MWh
price = AU$80/MWh
price = AU$100/MWh
price = AU$200/MWh
0.0
0.2
0.4
0.6
0.8
1.0
6000
8000
10000
12000
14000
16000
18000
20000

(c)

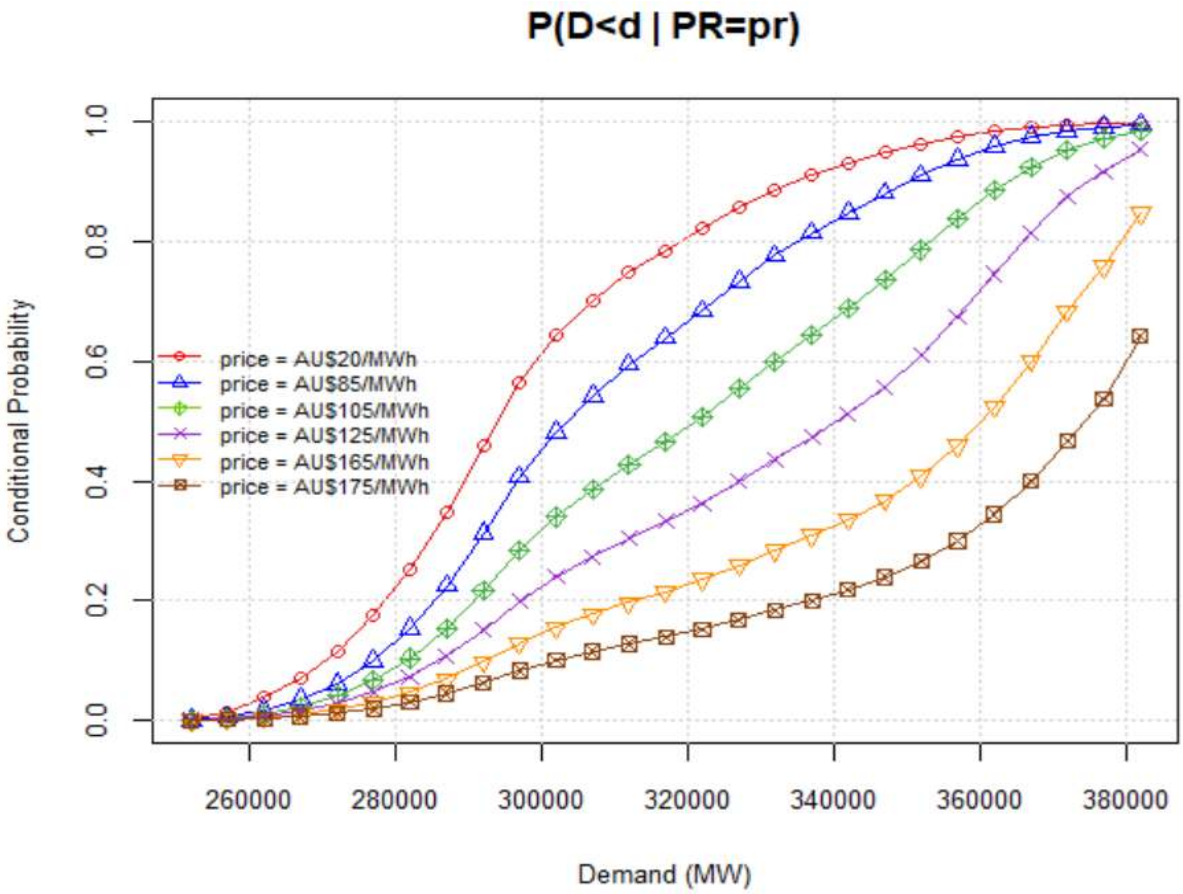


Figure 18: *Conditional nonexceedance probability forecasting of electricity demand (D) for various timescales (a) 30-minutes, (b) 1-hour, and (c) daily, given lagged price PR (t-1) greater than or equal to a certain threshold price (AU$/MW) by applying conditional bivariate copula models*

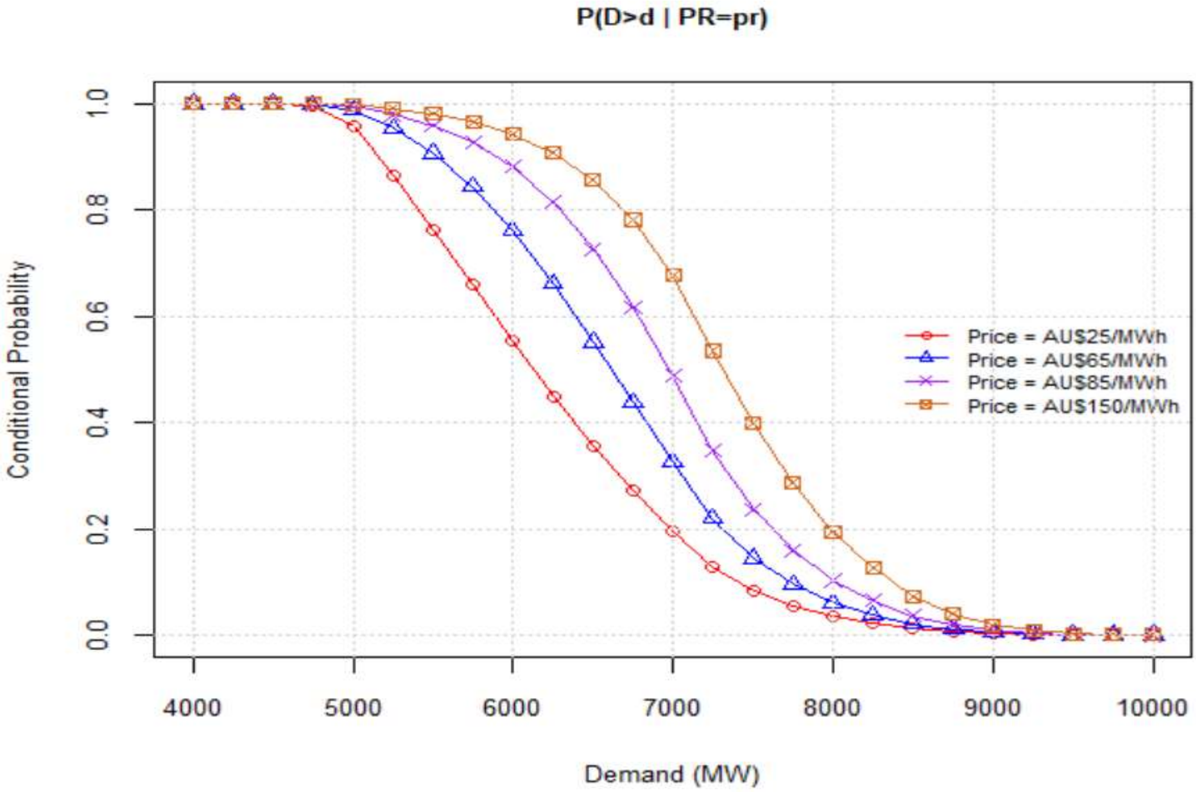
Figure 18 (a - c) portrays the joint probabilistic forecasting model derived from conditional bivariate copulas for predicting *D* at 30-minutes, 1-hour, and daily timescales. In Particular, it illustrates the demand of the conditional nonexceedance probability given the price value exceeding a certain threshold when the bivariate BB8 copula is utilized in accordance with *AIC* criteria for 30-minutes and 1-hour while the bivariate BB7 copula is utilized for the daily timescale. As shown in **Figure 18 (a - c)**, if the given price exceeds AU$150/MW, AU$200/MW, and AU$175/MW for 30-minutes, 1-hour, and daily timescales respectively, a shortage for electricity demand is not likely to happen at any time, meaning the electricity supply will be sufficient to meet consumer-demand. **Figure 18** presents a decision-making tool which energy forecasters and planners could use to appropriately estimate the provision of electricity demand for users at a certain time and locality.

In **Figure 18 (a),** for the 30-minutes timescale the conditional probability of electricity demand less than 7000MW given a price greater than AU$25/MWh (*i.e*., normal price) and AU$150/MWh (*i.e*., maximum price) could be approximately 80% and 32% respectively. This indicates that consumers practice cost saving measures in consuming less electricity when the price is high and vice versa. Also, for Figure **18 (b),** for the 1-hour timescale, the conditional probability of electricity demand less than 14000 MW given the price greater than AU$60/MWh and AU$100/MWh could be approximately 72% and 32% respectively. Likewise for the daily timescale (**Figure 18-c**), the conditional probability of electricity demand less than 360,000MW, given a price greater than AU$20/MWh, AU$85/MWh, AU$105/MWh, AU$165/MWh, and AU$175/MWh could be approximately 98%, 97%, 89%, 70%, 50% and 32% respectively.

5.3.2 Conditional exceedance probabilistic forecasting of electricity demand (*D*).

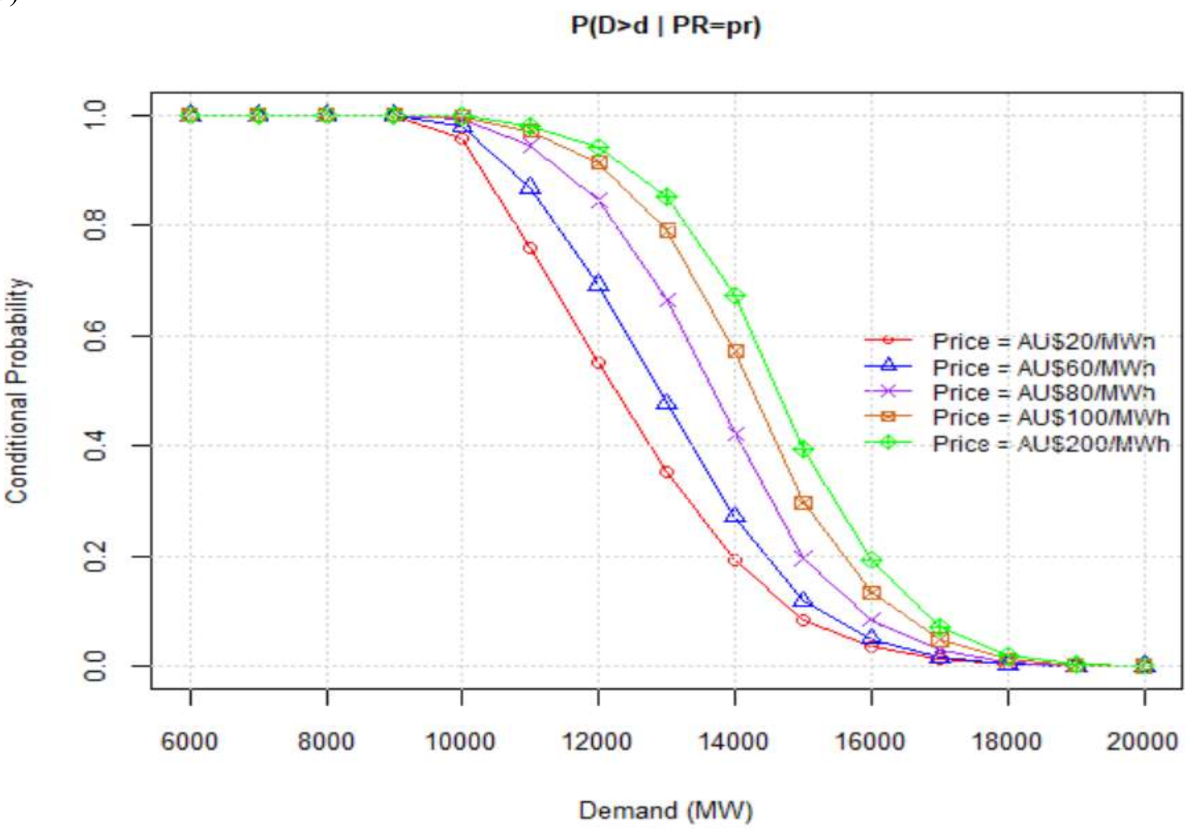
The results for conditional exceedance probabilistic forecasting of *D* for 30-minutes, 1-hour, and daily timescales, given lagged price *PR* (t-1) greater than or equal to a certain threshold price (AU$/MW) by applying conditional bivariate copula models are provided below in **Figure 19 (a - c)**.

(a)



P(D>d | PR=pr)
Conditional Probability
Demand (MW)
Price = AU$25/MWh
Price = AU$65/MWh
Price = AU$85/MWh
Price = AU$150/MWh
0.0
0.2
0.4
0.6
0.8
1.0
4000
5000
6000
7000
8000
9000
10000

(b)



P(D>d | PR=pr)
Conditional Probability
Demand (MW)
Price = AU$20/MWh
Price = AU$60/MWh
Price = AU$80/MWh
Price = AU$100/MWh
Price = AU$200/MWh
0.0
0.2
0.4
0.6
0.8
1.0
6000
8000
10000
12000
14000
16000
18000
20000

(c)

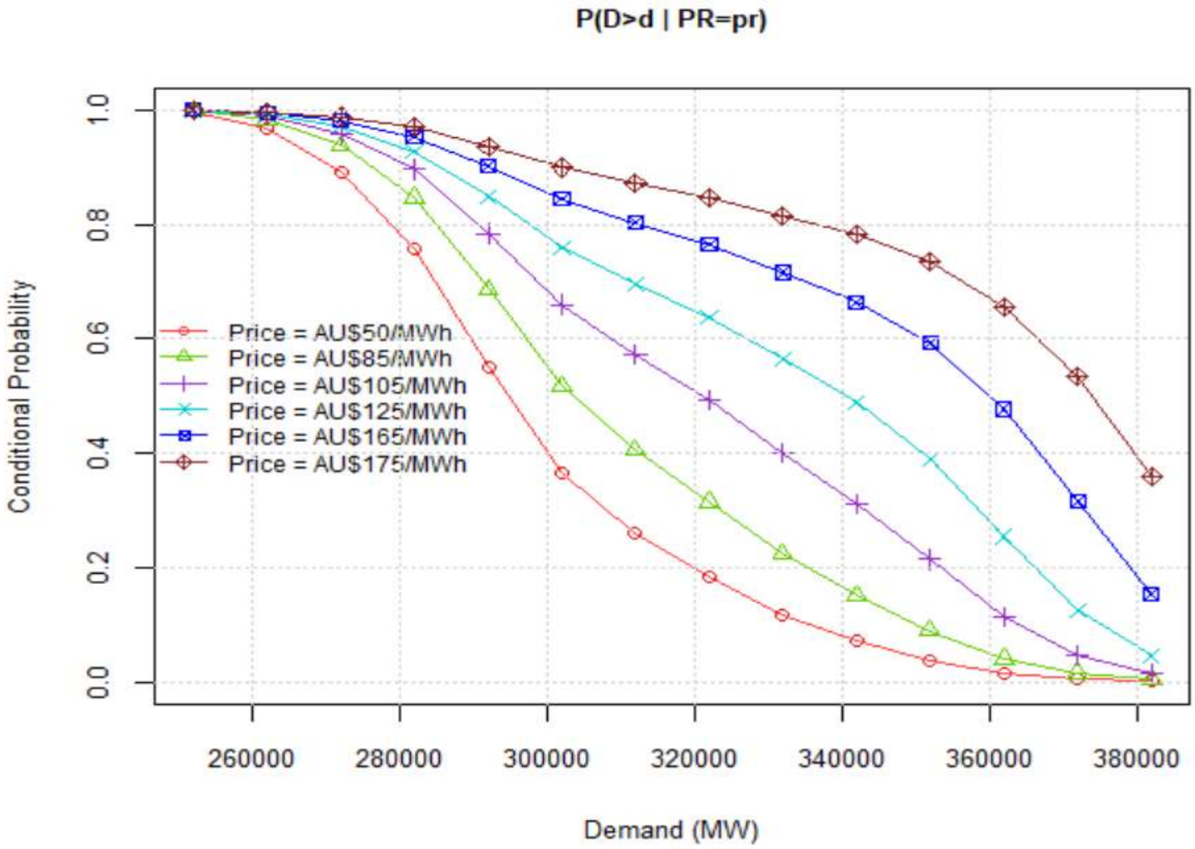


Figure 19: *Conditional exceedance probability forecasting of electricity demand (D) for (a) 30-minutes, (b) 1-hour, and (c) daily, given lagged price PR (t-1) exceeding certain threshold price (AU$/MW) by applying conditional bivariate copula models*

Conversely, the interpretation of **Figure 19 (a - c)** is straightforward. For the 30-minutes timescale (**Figure 19-a**), the conditional exceedance probability of electricity demand greater than 7000 MW given a price greater than AU$25/MWh (*i.e.*, normal price) and AU$150/MWh (*i.e.*, maximum price) could be approximately 20% and 68% respectively. For the 1-hour timescale (**Figure 19-b**), the conditional probability of electricity demand greater than 14,000 MW given a price greater than AU$60/MWh and AU$100/MWh could be approximately 30% and 48% respectively. Likewise for the daily timescale **(Figure 19-c**), the conditional probability of electricity demand greater than 360,000MW given a price greater than

AU$50/MWh, AU$85/MWh, AU$105/MWh, AU$125/MWh, AU$165/MWh, and AU$175/MWh could be approximately 2%, 3%, 11%, 28%, 50% and 68% respectively.

5.3.3 Local method for copula parameters estimate

The local method for copula parameter estimates is derived by maximum likelihood estimation. **Table 11** shows the details of the tail dependence measures and model fit statistics of the respective conditional bivariate copulas for each timescale (30-minutes, 1-hour, and daily).

Table 11: *Copula parameters, tail dependence measures, and model fit statistics derived from the local (maximum likelihood) method*

	30-minutes	1-hour	Daily
Best copula	BB8	BB8	BB7
Parameter 1	6	6	1.27
Parameter 2	0.57	0.57	0.16
Kendall Tau	0.43 (empirical = 0.44, p value < 0.01)	0.42 (empirical = 0.42, p value < 0. 01)	0.19 empirical = 0.16, p value < 0.01
Upper tail dependence	0	0	0.28
Lower tail dependence	0	0	0.01
logLik	12277.73	5800.93	64.23
AIC	-24551.45	-11597.86	-124.45
BIC	-24533.71	-11581.51	-114.45

Table 11 shows that BB8 copula appeared to be the best copula for 30-minutes probabilistic forecasting of *D* in the State of Queensland. It attained the first and second local copula parameters of 6 and 0.57 with logarithmic likelihood, *AIC*, and *BIC* of 12277.73, -24551.45, and -24533.71, respectively. For 1-hour timescale, also the BB8 copula was ranked highly with

first and second local copula parameters values of 6 and 0.57, respectively. It also exhibited the logarithmic likelihood, *AIC*, and *BIC* of 5800.93, -11597.86, and -11581.51, respectively. Similarly, for the daily timescale, BB7 copulas was ranked highly with its corresponding local copula parameter 1 and 2 values of 1.27 and 0.16, respectively, and logarithmic likelihood, *AIC*, and *BIC* of 64.23, -124.45, and -114.45, respectively.

CHAPTER 6: CONCLUSION

6.1 Summary of the research findings

Electricity demand forecasting was undertaken for the state of Queensland, which is the second largest state in Australia and where electricity demand is ever increasing. The hybrid extreme learning-copula models generated accurate *D*-forecast for all timescales (6-hours, 12-hours, and daily). A copula approach was also utilized to evaluate the joint behaviour of *D* and *PR* based on their historical lagged relationship. The results illustrated that BB7 copula was suitable for the probabilistic forecasting of daily *D* while BB8 was best suited for 30-minutes and 1-hour timescales. The analytical results from the conditional probability indicated that prices greater than or equal to AU$50/MW, AU$75/MW, AU$100/MW and, AU$150/MW show the lowest requirements for *D,;* whereas prices greater than or equal to AU$150/MW displayed the highest *D* requirements by consumers. This indicated that consumers were cost conscious where they practiced cost saving measures during price surge in electricity supply.

This study also explored both the local and global (Bayesian inference) optimization methods to estimate the best predictive uncertainties of copula parameters for developing a model for *D*-forecasting in the State of Queensland. The limitation of the local optimization for approximating copula parameters is that it frequently gets confined in the local minima (optima) and is not able to deliver any approximation of fundamental uncertainties. Hence, it may lead to providing biased results. In contrast, the global MCMC optimization method also estimates copula parameters by way of posterior distributions as well as presenting the uncertainty range. It generates and explores a good estimate of the global optimum as well as presenting an estimation of the primary uncertainty in a global MCMC simulation within a Bayesian framework. Therefore, in this study, the global MCMC approach is the best method for approximating the predictive uncertainties of copula parameters for developing models for 6 and 12-hours while local method is best for daily *D*-forecasting model in the State of Queensland.

6.2 Synthesis

The results presented in this study confirm that hybrid models deliver accurate more *D*-forecasting results than standalone models. The integration of ELM with MCMC copula-based models further improved forecasting efficiency by ranking the best performing copulas based on their respective performance metrics in achieving robust prediction for short-term *D*-forecasting in Queensland. The results generated by the hybrid ELM-MCMC copula-based models, namely the ELM-MCMC-Fischer-Hinzmann copula model for both 6-hours and 12-hours timescales, and the ELM-MCMC-Cuadras-Auge copula model for daily timescale, are highly accurate than that presented in previous studies by Al-Musaylh et al. (2018a) for short-term *D*-forecasting in Queensland. Also, the *D*-forecasting results in terms of the evaluation metrics for this study are highly accurate when compared with a similar study by Ali et al. (2018b) for predicting cotton yield in Pakistan. Therefore, the novel hybrid ELM-MCMC copula-based models developed in this study are highly accurate and reliable for short-term *D*-forecasting in Queensland as well as other regions.

This study also developed a copula-statistical model for the probabilistic forecasting of *D* using significant lagged correlation of *PR* as a covariate. It explored a joint probabilistic forecasting model from conditional bivariate copulas for predicting *D* at 30-minutes, 1-hour, and daily timescales under two conditions. Firstly, the probabilistic forecasting of *D* was achieved when the conditional nonexceedance probability given price (*PR*) exceeding a certain threshold when the bivariate BB8 copula model was executed in accordance with the *AIC* criteria. Secondly, *D* was probabilistically forecasted when the conditional exceedance probability given PR exceeding a certain threshold when the bivariate BB8 copula model was used in respect of the *AIC* criteria. This method attained a novel forecasting model by utilizing copula-statistical models, which had never been applied for short-term probabilistic forecasting of *D* anywhere previously. It therefore presents a vital decision-making tool which energy forecasters and planners can use to appropriately estimate the provision of electricity demand for users at certain times and localities.

6.3 Limitations and opportunity for future research

Since the scope of this Master's degree project was restricted to only the specified objectives set - out for this study over a duration of 1.5 years, the following recommendations are highlighted for future researchers to pursue.

I. This study only used the historical data for average electricity price (*PR*) and demand (*D*) to forecast future *D*. Since *D*-forecasting is a complex problem that involves various interconnected variables, future research may be targeted at investigating the lagged relationship between climate variables, population density, and GDP data, and incorporate these data to further enhance the forecasting of *D*.

II. This study used aggregated *D* and *PR* data for the entire Queensland region for model training and testing. For site-specific *D*-forecasting, it may be valuable to source data for *D* and *PR* for substations to build forecasting models to determine respective *D* requirements as each site may experience varying climatic patterns that may influence *D*. These results may be useful for planners and forecasters in the energy sectors to plan and execute appropriate electricity generation systems and distribution to specific sites where required.

REFERENCES

Aas, K, Czado, C, Frigessi, A, et al. 2009, 'Pair-copula constructions of multiple dependence', *Insurance: Mathematics and economics*, vol. 44, no. 2, pp. 182-98.

ABS 2020, *Population growth, Queensland, December quarter 2019,* ABo Statistics, Queensland, Australia, https://www.qgso.qld.gov.au/statistics/theme/population/population-estimates/state-territories>.

AEMO 2020, *Aggregated price and demand data - Historical*, Australian Energy Market Operator (AEMO), The Australian Energy Market Operator, https://aemo.com.au/en/energy-systems/electricity/national-electricity-market-nem/data-nem/aggregated-data>.

AghaKouchak, A, Bárdossy, A & Habib, E 2010, 'Copula-based uncertainty modelling: application to multisensor precipitation estimates', *Hydrological Processes*, vol. 24, no. 15, pp. 2111-24.

Akaike, H 1974, 'A new look at the statistical model identification', *IEEE transactions on automatic control*, vol. 19, no. 6, pp. 716-23.

Akay, D & Atak, M 2007, 'Grey prediction with rolling mechanism for electricity demand forecasting of Turkey', *Energy*, vol. 32, no. 9, pp. 1670-5.

Al-Alawi, SM & Islam, SM 1996, 'Principles of electricity demand forecasting. I. Methodologies', *Power Engineering Journal*, vol. 10, no. 3, pp. 139-43.

Al-Musaylh, MS, Deo, RC, Adamowski, JF, et al. 2018a, 'Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia', *Advanced Engineering Informatics*, vol. 35, no. 1, pp. 1-16.

Al-Musaylh, MS, Deo, RC, Li, Y, et al. 2018b, 'Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting', *Applied Energy*, vol. 217, pp. 422-39.

Al-Musaylh, MS, Deo, RC, Adamowski, JF, et al. 2019, 'Short-term electricity demand forecasting using machine learning methods enriched with ground-based climate and ECMWF Reanalysis atmospheric predictors in southeast Queensland, Australia', *Renewable and Sustainable Energy Reviews*, vol. 113, pp. 2-22.

Alexander, C 2004, *Managing Energy Price Risk: The New Challenges and Solutions*, London: Risk Books.

Ali, M, Deo, RC, Downs, NJ, et al. 2018a, 'Multi-stage hybridized online sequential extreme learning machine integrated with Markov Chain Monte Carlo copula-Bat algorithm for rainfall forecasting', *Atmospheric Research*, vol. 213, pp. 450-64.

Ali, M, Deo, RC, Downs, NJ, et al. 2018b, 'Cotton yield prediction with Markov Chain Monte Carlo-based simulation model integrated with genetic programing algorithm: A new hybrid copula-driven approach', *Agricultural and Forest Meteorology*, vol. 263, pp. 428-48.

Ali, M, Deo, RC, Downs, NJ, et al. 2020, 'Monthly rainfall forecasting with Markov Chain Monte Carlo simulations integrated with statistical bivariate copulas', in *Handbook of Probabilistic Models*, pp. 89-105.

Altinay, G & Karagol, E 2005, 'Electricity consumption and economic growth: Evidence from Turkey', *Energy Economics*, vol. 27, no. 6, pp. 849-56.

Assembly, G 2015, *Resolution adopted by the General Assembly on 19 September 2016*, A/RES/71/1, 3 October 2016 (The New York Declaration).

Balakrishna, N & Lai, CD 2009, 'Distributions expressed as copulas', in *Continuous Bivariate Distributions*, Springer, pp. 67-103.

Ball, A, Ahmad, S, McCluskey, C, et al. 2016, *Australian Energy Update*, Department of Industry, Innovation and Science, Canberra, http://www. industry. gov. au/Office-of-the-Chief-Economist/Publicat ions/Documents/aes/2015-australian-energy-statistics>.

Bárdossy, A 2006, 'Copula-based geostatistical models for groundwater quality parameters', *Water Resources Research*, vol. 42, no. 11.

Bedford, T & Cooke, RM 2001, 'Probability density decomposition for conditionally dependent random variables modeled by vines', *Annals of Mathematics and Artificial intelligence*, vol. 32, no. 1-4, pp. 245-68.

Bello, A, Reneses, J & Muñoz, A 2016, 'Medium-term probabilistic forecasting of extremely low prices in electricity markets: Application to the Spanish case', *Energies*, vol. 9, no. 3, pp. 1-27.

Benaouda, D, Murtagh, F, Starck, JL, et al. 2006, 'Wavelet-based nonlinear multiscale decomposition model for electricity load forecasting', *Neurocomputing*, vol. 70, no. 1, pp. 139-54.

Brinsmead, TS, Hayward, J & Graham, P 2014, *Australian electricity market analysis report to 2020 and 2030*, CSIRO Technical Report No. EP141067.

Bunn, D & Farmer, ED 1985, *Comparative models for electrical load forecasting*, United States.

Campbell, A 2018, 'Price and income elasticities of electricity demand: Evidence from Jamaica', *Energy Economics*, vol. 69, pp. 19-32.

Cao, Z, Liu, L, Hu, B, et al. 2019, 'Short-Term Load Forecasting Based on Variational Modal Decomposition and optimization Model', *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, pp. 121-6.

Chen, H, Canizares, CA & Singh, A 2001, 'ANN-based short-term load forecasting in electricity markets', *2001 IEEE power engineering society winter meeting. Conference proceedings (Cat. No. 01CH37194)*, IEEE, pp. 411-5.

Chui, F, Elkamel, A, Surit, R, et al. 2009, 'Long-term electricity demand forecasting for power system planning using economic, demographic and climatic variables', *European Journal of Industrial Engineering*, vol. 3, no. 3, pp. 277-304.

Clayton, DG 1978, 'A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence', *Biometrika*, vol. 65, no. 1, pp. 141-51.

Clements, AE, Hurn, AS & Li, Z 2016, 'Forecasting day-ahead electricity load using a multiple equation time series approach', *European Journal of Operational Research*, vol. 251, no. 2, pp. 522-30.

Contreras, J, Espinola, R, Nogales, FJ, et al. 2003, 'ARIMA models to predict next-day electricity prices', *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014-20.

Cook, R, Lapeyre, J, Ma, H, et al. 2019, 'Prediction of Compressive Strength of Concrete: Critical Comparison of Performance of a Hybrid Machine Learning Model with Standalone Models', *Journal of Materials in Civil Engineering*, vol. 31, no. 11, p. 04019255.

Czado, C, Brechmann, EC & Gruber, L 2013, 'Selection of vine copulas', in *Copulae in mathematical and quantitative finance*, Springer, pp. 17-37.

da Silva, FLC, Cyrino Oliveira, FL & Souza, RC 2019, 'A bottom-up bayesian extension for long term electricity consumption forecasting', *Energy*, vol. 167, pp. 198-210.

Dalkey, N & Helmer, O 1963, 'An experimental application of the Delphi method to the use of experts', *Management science*, vol. 9, no. 3, pp. 458-67.

Dawson, CW, Abrahart, RJ & See, LM 2007, 'HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts', *Environmental Modelling & Software*, vol. 22, no. 7, pp. 1034-52.

De Michele, C, Salvadori, G, Canossi, M, et al. 2005, 'Bivariate statistical approach to check adequacy of dam spillway', *Journal of hydrologic engineering*, vol. 10, no. 1, pp. 50-7.

Deo, RC & Şahin, M 2015, 'Application of the Artificial Neural Network model for prediction of monthly Standardized Precipitation and Evapotranspiration Index using hydrometeorological parameters and climate indices in eastern Australia', *Atmospheric Research*, vol. 161-162, pp. 65-81.

Deo, RC & Sahin, M 2015, 'Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia', *Atmospheric Research*, vol. 153, pp. 512-25.

Deo, RC & Sahin, M 2017, 'Forecasting long-term global solar radiation with an ANN algorithm coupled with satellite-derived (MODIS) land surface temperature (LST) for regional locations in Queensland', *Renewable & Sustainable Energy Reviews*, vol. 72, pp. 828-48.

Deo, RC, Samui, P & Kim, D 2016, 'Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models', *Stochastic environmental research and risk assessment*, vol. 30, no. 6, pp. 1769-84.

Deo, RC, Tiwari, MK, Adamowski, JF, et al. 2017, 'Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model', *Stochastic environmental research and risk assessment*, vol. 31, no. 5, pp. 1211-40.

Duan, Q, Sorooshian, S & Gupta, V 1992, 'Effective and efficient global optimization for conceptual rainfall-runoff models', *Water Resources Research*, vol. 28, no. 4, pp. 1015-31.

Erdogdu, E 2007, 'Electricity demand analysis using cointegration and ARIMA modelling: A case study of Turkey', *Energy Policy*, vol. 35, no. 2, pp. 1129-46.

Ertekin, C & Yaldiz, O 2000, 'Comparison of some existing models for estimating global solar radiation for Antalya (Turkey)', *Energy Conversion and Management*, vol. 41, no. 4, pp. 311-30.

Fan, S & Chen, L 2006, 'Short-term load forecasting based on an adaptive hybrid method', *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 392-401.

Fan, S & Chen, LN 2006, 'Short-term load forecasting based on an adaptive hybrid method', *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 392-401.

Fischer, M, Kraus, D, Pfeuffer, M, et al. 2017, 'Stress testing German industry sectors: Results from a vine copula based quantile regression', *Risks*, vol. 5, no. 3, p. 38.

Fischer, MJ & Hinzmann, G 2006, *A new class of copulas with tail dependence and a generalized tail dependence estimator*, Diskussionspapier.

Florens, MSL, Cairns, IH, Knock, SA, et al. 2007, 'Data-driven solar wind model and prediction of type II bursts', *Geophysical Research Letters*, vol. 34, no. 4, pp. 1-5.

Fousekis, P & Grigoriadis, V 2017, 'Joint price dynamics of quality differentiated commodities: copula evidence from coffee varieties', *European Review of Agricultural Economics*, vol. 44, no. 2, pp. 337-58.

Gräler, B, van den Berg, M, Vandenberghe, S, et al. 2013, 'Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation', *Hydrology and Earth System Sciences*, vol. 17, no. 4, pp. 1281-96.

Grégoire, V, Genest, C & Gendron, M 2008, 'Using copulas to model price dependence in energy markets', *Energy risk*, vol. 5, no. 5, pp. 58-64.

Haida, T & Muto, S 1994, 'Regression-Based Peak Load Forecasting Using a Transformation Technique', *IEEE Transactions on Power Systems*, vol. 9, no. 4, pp. 1788-94.

Hao, Z & AghaKouchak, A 2013, 'Multivariate standardized drought index: a parametric multi-index model', *Advances in Water Resources*, vol. 57, pp. 12-8.

HariKumar, R, Vasanthi, N & Balasubramani, M 2012, 'Performance analysis of artificial neural networks and statistical methods in classification of oral and breast cancer stages', *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 2, no. 3, pp. 263-9.

He, Y, Liu, R, Li, H, et al. 2017, 'Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory', *Applied Energy*, vol. 185, pp. 254-66.

Hsu, C-C & Chen, C-Y 2003, 'Regional load forecasting in Taiwan—applications of artificial neural networks', *Energy Conversion and Management*, vol. 44, no. 12, pp. 1941-9.

Hu, ZY, Bao, YK & Xiong, T 2013, 'Electricity Load Forecasting Using Support Vector Regression with Memetic Algorithms', *Scientific World Journal*, vol. 2013, pp. 2-10.

Huang, GB, Zhu, QY & Siew, CK 2006, 'Extreme learning machine: Theory and applications', *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501.

Hyndman, RJ & Fan, S 2010, 'Density Forecasting for Long-Term Peak Electricity Demand', *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 1142-53.

Joe, H 1997, *Multivariate models and multivariate dependence concepts*, CRC Press, Edinburgh, England.

Jongman, B, Hochrainer-Stigler, S, Feyen, L, et al. 2014, 'Increasing stress on disaster-risk finance due to large floods', *Nature Climate Change*, vol. 4, no. 4, pp. 264-8.

Khan, S, Javaid, N, Chand, A, et al. 2019, 'Forecasting day, week and month ahead electricity load consumption of a building using empirical mode decomposition and extreme learning machine', *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 1600-5.

Khotanzad, A, Afkhami-Rohani, R, Lu, T-L, et al. 1997, 'ANNSTLF-a neural-network-based electric load forecasting system', *IEEE Transactions on Neural networks*, vol. 8, no. 4, pp. 835-46.

Killiches, M, Kraus, D & Czado, C 2017, 'Examination and visualisation of the simplifying assumption for vine copulas in three dimensions', *Australian & New Zealand Journal of Statistics*, vol. 59, no. 1, pp. 95-117.

Kraus, D & Czado, C 2017, 'D-vine copula based quantile regression', *Computational Statistics & Data Analysis*, vol. 110, pp. 1-18.

Kwon, HH & Lall, U 2016, 'A copula-based nonstationary frequency analysis for the 2012–2015 drought in California', *Water Resources Research*, vol. 52, no. 7, pp. 5662-75.

Lebotsa, ME, Sigauke, C, Bere, A, et al. 2018, 'Short term electricity demand forecasting using partially linear additive quantile regression with an application to the unit commitment problem', *Applied Energy*, vol. 222, pp. 104-18.

Li, C, Singh, VP & Mishra, AK 2013, 'A bivariate mixed distribution with a heavy-tailed component and its application to single-site daily rainfall simulation', *Water Resources Research*, vol. 49, no. 2, pp. 767-89.

Madadgar, S & Moradkhani, H 2013, 'Drought Analysis under Climate Change Using Copula', *Journal of hydrologic engineering*, vol. 18, no. 7, pp. 746-59.

Mamun, MA & Nagasaka, K 2004, 'Artificial neural networks applied to long-term electricity demand forecasting', *Fourth International Conference on Hybrid Intelligent Systems (HIS'04)*, IEEE, pp. 204-9.

Manner, H, Türk, D & Eichler, M 2016, 'Modeling and forecasting multivariate electricity price spikes', *Energy Economics*, vol. 60, pp. 255-65.

Manner, H, Fard, FA, Pourkhanali, A, et al. 2019, 'Forecasting the joint distribution of Australian electricity prices using dynamic vine copulae', *Energy Economics*, vol. 78, pp. 143-64.

Marwala, L & Twala, B 2017, 'Electricity Load Forecasting Using an Ensemble of Optimally-Pruned and Basic Extreme Learning Machines', *2017 9th IEEE-GCC Conference and Exhibition (GCCCE)*, pp. 1-6.

Marzband, M, Azarinejadian, F, Savaghebi, M, et al. 2015, 'An optimal energy management system for islanded microgrids based on multiperiod artificial bee colony combined with Markov chain', *IEEE Systems Journal*, vol. 11, no. 3, pp. 1712-22.

Marzband, M, Ardeshiri, RR, Moafi, M, et al. 2017, 'Distributed generation for economic benefit maximization through coalition formation–based game theory concept', *International Transactions on Electrical Energy Systems*, vol. 27, no. 6, p. e2313.

Min, A & Czado, C 2010, 'Bayesian inference for multivariate copulas using pair-copula constructions', *Journal of Financial Econometrics*, vol. 8, no. 4, pp. 511-46.

Mirasgedis, S, Sarafidis, Y, Georgopoulou, E, et al. 2006, 'Models for mid-term electricity demand forecasting incorporating weather influences', *Energy*, vol. 31, no. 2-3, pp. 208-27.

Mohammadi, K, Shamshirband, S, Anisi, MH, et al. 2015a, 'Support vector regression based prediction of global solar radiation on a horizontal surface', *Energy Conversion and Management*, vol. 91, pp. 433-41.

Mohammadi, K, Shamshirband, S, Tong, CW, et al. 2015b, 'A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation', *Energy Conversion and Management*, vol. 92, pp. 162-71.

Morimoto, R & Hope, C 2004, 'The impact of electricity supply on economic growth in Sri Lanka', *Energy Economics*, vol. 26, no. 1, pp. 77-85.

Nasr, GE, Badr, EA & Younes, MR 2002, 'Neural networks in forecasting electrical energy consumption: univariate and multivariate approaches', *International Journal of Energy Research*, vol. 26, no. 1, pp. 67-78.

Nelsen, RB 2003, 'Properties and applications of copulas: A brief survey', *Proceedings of the first brazilian conference on statistical modeling in insurance and finance*, Citeseer, pp. 10-28.

Nguyen-Huy, T, Deo, RC, An-Vo, D-A, et al. 2017, 'Copula-statistical precipitation forecasting model in Australia's agro-ecological zones', *Agricultural Water Management*, vol. 191, pp. 153-72.

Nguyen-Huy, T, Deo, RC, Mushtaq, S, et al. 2018, 'Modeling the joint influence of multiple synoptic-scale, climate mode indices on Australian wheat yield using a vine copula-based approach', *European Journal of Agronomy*, vol. 98, pp. 65-81.

Nti, IK, Teimeh, M, Nyarko-Boateng, O, et al. 2020, 'Electricity load forecasting: a systematic review', *Journal of Electrical Systems and Information Technology*, vol. 7, no. 1, pp. 1-19.

Nwankwo, O & Njogo, B 2013, 'The effect of electricity supply on industrial production within the Nigerian economy (1970–2010)', *Journal of Energy Technologies and Policy*, vol. 3, no. 4, pp. 34-42.

Odhiambo, NM 2009, 'Electricity consumption and economic growth in South Africa: A trivariate causality test', *Energy Economics*, vol. 31, no. 5, pp. 635-40.

Pai, P-F & Hong, W-C 2005, 'Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms', *Electric Power Systems Research*, vol. 74, no. 3, pp. 417-25.

Papadopoulos, S & Karakatsanis, I 2015, 'Short-term electricity load forecasting using time series and ensemble learning methods', *2015 IEEE Power and Energy Conference at Illinois (PECI)*, pp. 1-6.

Papalexopoulos, AD, Hao, S & Peng, T-M 1994, 'An implementation of a neural network based load forecasting model for the EMS', *IEEE Transactions on Power Systems*, vol. 9, no. 4, pp. 1956-62.

Parent, E, Favre, A-C, Bernier, J, et al. 2014, 'Copula models for frequency analysis what can be learned from a Bayesian perspective?', *Advances in Water Resources*, vol. 63, pp. 91-103.

Pitt, M, Chan, D & Kohn, R 2006, 'Efficient Bayesian inference for Gaussian copula regression models', *Biometrika*, vol. 93, no. 3, pp. 537-54.

Ribatet, M & Sedki, M 2013, 'Extreme value copulas and max-stable processes', *Journal de la Société Française de Statistique*, vol. 154, no. 1, pp. 138-50.

Sadegh, M, Ragno, E & AghaKouchak, A 2017, 'Multivariate C opula A nalysis T oolbox (MvCAT): describing dependence and underlying uncertainty using a B ayesian framework', *Water Resources Research*, vol. 53, no. 6, pp. 5166-83.

Salcedo-Sanz, S, Rojo-Alvarez, JL, Martinez-Ramon, M, et al. 2014, 'Support vector machines in engineering: an overview', *Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery*, vol. 4, no. 3, pp. 234-67.

Salvadori, G 2003, 'A generalized pareto intensity duration model of storm rainfall exploiting 2-copulas', *J Geophys Res*, vol. 108, no. 2, p. 4067.

Salvadori, G & De Michele, C 2004, 'Frequency analysis via copulas: Theoretical aspects and applications to hydrological events', *Water Resources Research*, vol. 40, no. 12.

Schwarz, G 1978, 'Estimating the dimension of a model', *The annals of statistics*, vol. 6, no. 2, pp. 461-4.

Sharda, V, Prasher, S, Patel, R, et al. 2008, 'Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in mid-Himalayan micro-watersheds with limited data/Performances de régressions par splines multiples et adaptives (MARS) pour la prévision d'écoulement au sein de micro-bassins versants Himalayens d'altitudes intermédiaires avec peu de données', *Hydrological sciences journal*, vol. 53, no. 6, pp. 1165-75.

Sheikh, SK & Unde, M 2012, 'Short term load forecasting using ann technique', *International Journal of Engineering Sciences & Emerging Technologies*, vol. 1, no. 2, pp. 97-107.

Sigauke, C & Chikobvu, D 2010, 'Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach', *ORiON*, vol. 26, no. 2.

Singh, AK, Ibraheem, SK, Muazzam, M, et al. 2013, 'An overview of electricity demand forecasting techniques', *Network and Complex Systems*, vol. 3, no. 3, pp. 38-48.

Sivapragasam, C & Liong, S-Y 2005, 'Flow categorization model for improving forecasting', *Hydrology Research*, vol. 36, no. 1, pp. 37-48.

Sklar, A 1996, 'Random variables, distribution functions, and copulas: a personal look backward and forward', *Lecture notes-monograph series*, pp. 1-14.

Sklar, M 1959, 'Fonctions de repartition an dimensions et leurs marges', *Publ. inst. statist. univ. Paris*, vol. 8, pp. 229-31.

Smith, MS, Gan, Q & Kohn, RJ 2012, 'Modelling dependence using skew t copulas: Bayesian inference and applications', *Journal of Applied Econometrics*, vol. 27, no. 3, pp. 500-22.

Smola, AJ & Schölkopf, B 1998, *Learning with kernels*, vol. 4, Citeseer.

Sozen, A & Akcayol, MA 2004, 'Modelling (using artificial neural-networks) the performance parameters of a solar-driven ejector-absorption cycle', *Applied Energy*, vol. 79, no. 3, pp. 309-25.

Stern, DI, Burke, PJ & Bruns, SB 2019, 'The impact of electricity on economic development: A macroeconomic perspective'.

Suganthi, L & Jagadeesan, T 1992, 'Energy substitution methodology for optimum demand variation using Delphi technique', *International Journal of Energy Research*, vol. 16, no. 9, pp. 917-28.

Suganthi, L & Samuel, AA 2012, 'Energy models for demand forecasting—A review', *Renewable and Sustainable Energy Reviews*, vol. 16, no. 2, pp. 1223-40.

Taylor, JW & Buizza, R 2002, 'Neural network load forecasting with weather ensemble predictions', *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 626-32.

Thyer, M, Renard, B, Kavetski, D, et al. 2009, 'Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis', *Water Resources Research*, vol. 45, no. 12, pp. 1-22.

Toman, MA & Jemelkova, B 2003, 'Energy and Economic Development: An Assessment of the State of Knowledge', *Energy Journal*, vol. 24, no. 4, pp. 93-112.

Tosunoglu, F & Singh, VP 2018, 'Multivariate modeling of annual instantaneous maximum flows using copulas', *Journal of hydrologic engineering*, vol. 23, no. 3, p. 04018003.

Tosunoglu, F, Gürbüz, F & İspirli, MN 2020, 'Multivariate modeling of flood characteristics using Vine copulas', *Environmental Earth Sciences*, vol. 79, no. 19, pp. 1-21.

Türkay, BE & Demren, D 2011, 'Electrical load forecasting using support vector machines', *2011 7th International Conference on Electrical and Electronics Engineering (ELECO)*, IEEE, pp. I-49-I-53.

Vernieuwe, H, Vandenberghe, S, De Baets, B, et al. 2015, 'A continuous rainfall model based on vine copulas', *Hydrology and Earth System Sciences*, vol. 19, no. 6, pp. 2685-99.

Wang, AJ & Ramsay, B 1998, 'A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays', *Neurocomputing*, vol. 23, no. 1-3, pp. 47-57.

Willmott, CJ 1981, 'On the validation of models', *Physical Geography*, vol. 2, no. 2, pp. 184-94.

Willmott, CJ 1982, 'Some comments on the evaluation of model performance', *Bulletin of the American Meteorological Society*, vol. 63, no. 11, pp. 1309-13.

Willmott, CJ 1984, 'On the evaluation of model performance in physical geography', in *Spatial statistics and models*, Springer, pp. 443-60.

Willmott, CJ, Matsuura, K & Robeson, SM 2009, 'Ambiguities inherent in sums-of-squares-based error statistics', *Atmospheric Environment*, vol. 43, no. 3, pp. 749-52.

Willmott, CJ, Robeson, SM & Matsuura, K 2012, 'A refined index of model performance', *International Journal of Climatology*, vol. 32, no. 13, pp. 2088-94.

Wu, J, Cui, Z, Chen, Y, et al. 2019, 'A new hybrid model to predict the electrical load in five states of Australia', *Energy*, vol. 166, pp. 598-609.

Xydas, E, Marmaras, C, Cipcigan, LM, et al. 2016, 'A data-driven approach for characterising the charging demand of electric vehicles: A UK case study', *Applied Energy*, vol. 162, pp. 763-71.

Yang, J, Rivard, H & Zmeureanu, R 2005a, 'On-line building energy prediction using adaptive artificial neural networks'.

Yang, J, Rivard, H & Zmeureanu, R 2005b, 'Building energy prediction with adaptive artificial neural networks', *Ninth International IBPSA Conference Montréal*, pp. 15-8.

Yu, P-S, Chen, S-T & Chang, IF 2006, 'Support vector regression for real-time flood stage forecasting', *Journal of Hydrology*, vol. 328, no. 3, pp. 704-16.

Zareipour, H, Bhattacharya, K & Canizares, CA 2006a, 'Forecasting the hourly Ontario energy price by multivariate adaptive regression splines', *2006 IEEE Power Engineering Society General Meeting*: *Proceedings of the 2006 IEEE Power Engineering Society General Meeting* IEEE, pp. 1-7.

Zareipour, H, Bhattacharya, K & Canizares, CA 2006b, 'Forecasting the hourly Ontario energy price by multivariate adaptive regression splines', *2006 IEEE Power Engineering Society General Meeting*, IEEE, p. 7 pp.

Zhang, R, Dong, ZY, Xu, Y, et al. 2013, 'Short-term load forecasting of Australian National Electricity Market by an ensemble model of extreme learning machine', *Iet Generation Transmission & Distribution*, vol. 7, no. 4, pp. 391-7.

Ziel, F & Steinert, R 2018, 'Probabilistic mid- and long-term electricity price forecasting', *Renewable & Sustainable Energy Reviews*, vol. 94, pp. 251-66.