Design of Smart Energy Generation and Demand Response System in Saudi Arabia

A thesis submitted in partial fulfilment of the requirements for the degree

of Doctor of Philosophy

by

Fahad Aljahdali

Department of Electronic and Computer Engineering

College of Engineering, Design and Physical Sciences

Brunel University London

United Kingdom

August 2019

Publications Based on this Research

[1]. Aljahdali, F., & Abbod, M., 2017. Design of smart energy generation and demand response system in Saudi Arabia. In *2017 52nd International Universities Power Engineering Conference (UPEC),* pp. 1-5, IEEE.

[2]. Aljahdali, F., & Abbod, M., 2018. Design of Smart Generation by Integrating Renewable Energy into Western Power Grid of Saudi Arabia. In *2018 53rd International Universities Power Engineering Conference (UPEC)*, pp. 1-6, IEEE.

[3]. Aljahdali, F., Abbod, M., 2018. Integration of RE into 380 kV transmission grid in Western part of Kingdom of Saudi Arabia Smart Grid Conference 2018- Jeddah, Saudi Arabia. (in press)

Declaration

It is hereby declared that the thesis in focus is the author's own work and is submitted for the first time to the Post Graduate Research Office. The study was originated, composed and reviewed by the mentioned author in the Department of Electronic and Computer Engineering, College of Engineering, Design and Physical Sciences, Brunel University London, UK. All the information derived from other works has been properly referenced and acknowledged.

Fahad Aljahdali

August 2019

London, UK

Acknowledgements

Abstract

Though the promising benefits of renewable sources have already pushed many countries into implementing RE units, Saudi Arabia is stll highly dependent on fossil fuel. However, the decrease in value of oil reserves has enforced Saudi Arabia to prioritize renewable energy sources in the next decades. Such energy sources are highly dependable on accurate forecasts, due to their intermittence and operability. The present research has the objective to develop models that can accurately forecast energy load for implementation in a decision-making system. The case investigated is the western region of Saudi Arabia. Two modelling approaches were evaluated, linear regression (LR) and artificial neural network (ANN). This last one was chosen because it is a mathematical model able to deal with non-linear relationship among input(s) and output(s) in the data. Time series (past load data) and multivariate data from 2010 until 2016 were investigated A hybrid model structure (combiner) was implemented to analyse the effects of combining outputs of two models in a single one. This hybrid model consisted of a regular average and weighted average of the time series and multivariate model, with calibration through Fuzzy and Particle Swarm Optimisation. These two were selected because, while Particle Swarm Optimization is an optimization algorithm, Fuzzy consists in a complete structured model. The forecasted load and the available input were used in the last chapter for power generation planning and decision-making support. The software used for the modelling and simulation is ETAP®. Different scenarios for replacement of fossil fuel power plants by renewable units were tested considering the network of western Saudi Arabia. The results show that Artificial Neural Network with time series input and 15 neurons in hidden layer shows superior performance (MSE $3.7*10^5$ and R^2 equals 99%) compared to other neural networks and linear regression. Though the application of combiner models did not significantly improve model performance, the Fuzzy Combiner shows the best one (MSE $5.8*10^5$ and R^2 equals 93%) since it incorporates information from time series and multivariate data. It is important to mention that all the modelling approaches evaluated have some limitations, such as the necessity of accurate input data and they are limited in capability of extrapolating over the training range. In the last section, it was observed that renewable energy sources can be integrated in the grid network without excessive risk regarding demand. This occurs because the current energy management policy of western Saudi Arabia enables the use of energy units with fast compensation (using gas units) in the case of demand increase or decrease in solar or wind power.

Thesis Contents

List of Figures

ix

List of Tables

List of Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
ARIMA	Auto Regressive Integrated Moving Average
BP	Backpropagation
BPNN	Backpropagation Neural Network
CC	Combined Cycle
DER	Distribution Energy Resources
DHI	Diffuse Horizon Irradiance
DNI	Direct Normal Irradiance
DR	Demand Response
EIA	Energy Information Administration
EMC	Energy Market Company
ES	Early Stopping
FFN	Feed Forward Neuron
FIS	Fuzzy Inference System
FL	Fuzzy Logic
GCC	Gulf Cooperation Council
GHG	Greenhouse Gas
GHI	Global Horizontal Irradiance
GSA	Gravitational Search Algorithm
LR	Linear Regression

MAD	Mean Absolute Distribution
MAE	Mean Absolute Error
MF	Membership Functions
MIMO	Multi Input Multi Output
MLP	Multi- Layer Perceptron
MPLSR	Modified Partial Least Squares Regression
MTLBO	Modified Teaching Learning- Based Optimisation
MTLF	Medium-Term Load Forecasting
MVA	Mega Voltage Ampere
MW	Mega Watt
MZEC	Ma Zoon Electricity Company
NI	National Instrument
OLSR	Ordinary Least Squared Regression
PCR	Principal Component Regression
PE	Processing Elements
PLSR	Partial Least Squared Regression
PSF	Pattern Sequence- based Forecasting
PSO	Particle Swarm Optimisation
PV	Photovoltaic
RES	Renewable Energy Sources
RMSE	Root Mean Squared Error
SA	Sensitivity Analysis
SARIMA	Seasonal Auto Regressive Integrated Moving Average

SEC   Saudi Electricity Company

SSE   Sum of Squared Errors

STF   Short Term Forecasting

STLF   Short Term Load Forecasting

SVM   Support Vector Machine

SVR   Support Vector Regression

TAF   Terminal Aerodrome Forecast

TEC   Teheran Electricity Company

TS   Time Series

UN   United Nations

VSTLF   Very Short-Term Load Forecasting

WOA   Western Operational Network

Chapter 1.
Introduction

1.1. Introduction

The current energy scenario in Saudi Arabia is highly dependent on fossil fuel sources, two viewpoints being clearly divisible. Saudi oil reserves have been decreasing in volume, with a recent decrease of 185×10^9 barrels/year, as the daily production is reaching 12 million barrels/day. Increase in the demand can even worsen such scenario, with sharper reduction of reserves (Energy Information Administration (EIA), 2014). Global economy forecasts suggest that there is a high probability of increase in oil demand all over the world. Moreover, the forecast for demand of electricity against growth in power production in Saudi Arabia is expected to increase as a result of increasing population, commercial plans, rural areas electrification and industrial projects. Fossil fuel-based generating plants can meet the existing and planned power demand; however, the CO_2 emission will increase and will be affected by reduction in oil production and increased prices.

As a participant to the United Nations (UN) agenda principle on Climate Change, Saudi Arabia obviously recognizes its accountability and obligation to mitigate environmental effects caused by its economy. Hence, the focus is diverted in investing in renewable energy generation to reduce impact on the environment, especially as the kingdom is rich in both solar and wind resources. However, storage of renewable energy is not an easy or cheap task. Hence, grid-scale short and long-term generation/demand management is a key process for the success of optimum utilisation of renewable resources and reduction in fossil fuel usage. Saudi Arabia's location is the middle of hot countries in the Middle East, where the sun shines for an average of 14 hours per day, with 2,550 kWh of solar radiation per square metre; making Saudi Arabia weather an ideal place for photovoltaic (PV) investment. Moreover, the long coastal side of the Red Sea and Arabian Gulf, with high wind speed also motivates investment in wind energy. Hence, for the next

decade, renewable energy sources available in Saudi Arabia should be the priority for private companies and government (EIA, 2014).

On the other hand, there are other causes why developing a practical renewable energy market is significant, even for a country that is home to rich hydrocarbon reserves. These reasons have raised a great motivation on Saudi Arabia to invest on RE. Electric power demands will increase year on year, whereby renewable energy resources can comply with this through providing clean and reliable supply from different sources other than fossil fuels, while maintaining low CO_2 emission. This thesis seeks to approach intelligent generation, expansion and demand response systems for such kind of application and compare the possible schemes to achieve the most effective renewable energy integration and operation in the Saudi Arabian grid (Taher and Al-Hajjar, 2013)

1.2. Power Demand Prediction

Electric power systems have become more and more essential for our life. Therefore, achieving a reliable power system is a very important trend for electric energy production facilities. A more reliable power system is a more economical and less costly one. A good system planning will lead to a certain compromise between its reliability degree and its economics, taking into account the load demand profile over time. Electrical power systems consist of different group of elements, such as generators, distribution networks, transformers and loads.

The main issue of a power system is to supply the exact demand to the customers at all times. As customer numbers are increasing day after day, good power system planning must be done to face this increment. On the other hand, during the daily operation, the load itself changes up and down, so a good control strategy will be very beneficial to ensure power supply continuity without load shedding (Yan et al, 2019). The primary aim of this research work is to investigate the fundamentals of power system planning and control (load flow, economic dispatch, the power frequency control). A power distribution system is a crucial part of electric networks between generation and consumer service locations. Distribution networks represent a huge capital investment.

The planning and design of electric distribution systems are divided into long-term strategic planning for future expansions and short-term operation reliability for demand

2

management and minimising losses. The study of power distribution system covers many areas, such as planning and design of the distribution systems, load studies, system protection, reliability and maintenance (Yan et al, 2019).

1.3. Motivation of the Research

The Saudi Arabia government has developed a plan for replacement of energy sources, from fossil to renewable sources (RES). However, such source is dependent on various factors, among them weather conditions, solar irradiation, wind speed, etc.

In this regard, the current research investigates different modelling techniques for accurate load forecasting, since such step is vital for the safe maintenance of the electricity grid in a healthy state with complete supply of all demands. It is expected that the presented models will be potentially implemented in production in Saudi Arabia, assisting the decision-making systems which operate with RES or RES/fossil fuel systems.

1.4. Aim and Objectives

The main aim of this thesis is to design an intelligent prediction system to be implemented in decision-making systems for the power grid in Saudi Arabia. This will enable management to meet expansion and demand response requirements of the power grid, based on utilisation of distributed generation. Prediction horizon is one-day-ahead load forecasting, since this time window meets the requirements of management in Saudi Arabian energy companies.

In order to achieve this aim, the following objectives are set:

1. Investigate the available opportunities to improve the electric energy generation in Saudi Arabia through utilisation of renewable energy resources (solar and wind).

2. Investigate modern intelligent techniques used for load forecasting and generation management, such as regression models, time series prediction and other optimisation/prediction methods in Saudi Arabian grid using MATLAB® software.

3. Investigate the impacts of existing/future load demand on energy generation on the Saudi Arabian grid, including possible renewable energy resources.

4. Use the models developed in MATLAB® to develop a generation management system for recommendations that is suitable for the Saudi Arabian grid when renewable energy resources are integrated using ETAP® software.

1.5. Contributions to Knowledge

The work will investigate the opportunities to use the renewable and distributed energy resources in the Kingdom of Saudi Arabia in the next 11 years, which will help in building an image which will be used as a strong base for decision-making in relation to integrating renewable and distributed energy in the country. Moreover, it will estimate the impacts of existing and future increase in the energy consumptions on the Saudi national grid, whereby it will provide optimal prospective for renewable energy and distributed generation integration and management in Saudi Arabian grid to achieve quick demand response based on the historical environmental and grid operation information available.

1.6. Research Methodology

Quantitative research methods are concerned with investigating natural scientific phenomena, whereas qualitative research methods are related to perceiving human experience and knowledge associated with social and cultural phenomena. Qualitative research uses a systematic process to enable researchers to understand various aspects associated with people and cultural problems. Any research method could be influenced by the philosophical assumptions of the researcher and explains the way by which information will be collected and processed to achieve the objectives of the subject being examined. In general, there is no common philosophical research paradigm, while the three main principles are thought to be positivism, interpretivist and critical theory positivism. Positivism is related to qualitative research while interpretivist interprets reality as a social construction by observations that differ from one person to another as a result of different social perspectives. Hence, it is more likely linked to qualitative research. Critical theory research is defined as interpreting the assumption that 'people can consciously act to change their social and economic conditions' (Creswell, 2013).

This research is a quantitative research project in its nature, which requires solid data collection in relation to the topic under investigation in the Kingdom of Saudi Arabia. This work is concerned with how to improve the current situation of power generation and

demand management into smart power generation and demand management systems in the Kingdom of Saudi Arabian power grid when renewable resources and distributed generation are integrated in the power network. Consequently, solid and exact interpretation of the current power grid infrastructure in the kingdom, in addition to historical data and future plans, have to be collected in an accurate way so as to enable trusted investigations and final results. Hence, the work will need to investigate different parameters and relevant data that affect production of electric power from different resources to minimise CO_2 emission. It will need also to study generation and demand as well as modern management techniques, including time series prediction to minimise generation/demand mismatch almost instantaneously (within half an hour or more) using long and short-term load forecasting techniques.

Through applying time series prediction, optimisation techniques, demand analysis and load forecasting, renewable and distributed generation energy resources available in the kingdom will be integrated in the grid and will be optimally operated. Hence, reduction in the cost of generating and CO_2 emission will be achieved in the kingdom's power network. The research stream tends to depend on previously identified parameters in related works and matching with the case of the Saudi Arabian grid. In general, the methodology will be based on the following steps:

1. Review state-of-the-art in this area, including intelligent prediction techniques.

2. Obtain related data records of the Saudi Arabian grid.

3. Obtain information related to promising solar stations, wind turbines locations as well as forecasted and planned renewable energy resources generation output. throughout the government.

4. Obtain information related to current and future load demand in the power grid.

5. Simulate the Saudi Arabian electricity network with potential renewable and distributed generation energy resources integrated.

6. Design an intelligent management system that will optimally operate the generating resources to achieve quick demand response while reducing CO_2 emission.

1.7. Thesis Outline

Based on the findings of the research and the objectives achieved, there are seven chapters in this thesis:

Chapter 2 shows the current state-of-the-art investigations on energy load forecast, usage of renewable energy sources, smart energy development and maintenance as well as models currently used for prediction and optimisation when analysing energy load and the factors involved on it.

The background theory and methodology involved in the mathematical approach used in the current research are discussed in Chapter 3 and the structure of the models implemented is explained, namely Linear Regression and Artificial Neural Network. The structure, concept and theory behind the Combiner model, a superior model which couples two models, are also discussed. The metrics used to evaluate performance goodness are also presented in this chapter, along with its mathematical structure.

Chapter 4 presents the case description and the results obtained for the two models used (Linear Regression and Artificial Neural Network) with two types of model input (Time series and Multivariate), for energy load forecasting. Time series input refers to past available load, while multivariate consists of exogenous inputs, such as temperature, humidity, etc.

The results obtained for the combiner models, a superior structure which couples multiple models to generate a single prediction, are shown in Chapter 5. The results provide evidence for discussion of the pros and cons of each combiner, and a sensitivity analysis of each factor used a predictor for energy load forecast.

Chapter 6 displays the analysis of the network grid of western Saudi Arabia, for smart energy management and operation. The focus is the incorporation of renewable energy units with minimum effect on the load pattern, as well as guaranteeing the complete supply of energy as per demand. Different scenarios were investigated to confirm the requirements for the implementation of solar and wind power.

Chapter 7 summarises the research's main conclusions and significant contributions. It also highlights future work in this field and provides suggestions of further areas for study.

Chapter 2.

Literature Review

2.1. Introduction

Availability of data has changed the landscape of the process of production and consumption patterns and paved the way for strengthening the operational efficiency of the renewable energy (Zhou et al., 2016). Furthermore, large data availability results in providing opportunities for smart energy management, specifically assisting the public and private sector electricity producers and distributors in achieving their goals of clean power generation, efficient power transmission and dynamic distribution (Diamantoulakis et al., 2015). All these benefits of energy data could not be achieved without using the data in some sort of modelling techniques needed for controlling information flow. Smart energy management is about collecting a large amount of data related to the electricity consumption and production and then using modelling techniques to predict the energy demand in real time (Yang and Shen, 2013).

This chapter, through literature review, presents the current state-of-the-art of energy load modelling, forecasting and optimisation using various modelling techniques available.

2.2. Importance of Producing Electricity via Renewable Energy Sources

As mentioned by Mitchell et al. (2012), the modern world in which we live today is dependent on energy supplies for its economic and social development. Usage of traditional sources of energy, such as oil, gas and coal, leaves devastating environmental damage causing catastrophe to stratospheric ozone and increasing air acidity (Aloughani, 2015). Kranzberg (2012) conducted sustainability analysis of producing electricity from renewable and fossil fuels energy sources on the basis of various factors, such as cost, raw materials, water consumption and greenhouse gas (GHG) emission. On the cost element, it was found that the cost of solar energy sources is estimated to be six times more than energy from hydro source, as shown in Figure 2-1 below. Similarly, wind energy has been found to be

costly comparable to other energy sources. Key reasons behind the high cost of producing electricity from RES lies in the fact that such production requires a large area of land for installing wind turbines, materials needed for electricity processing and manufacturing.
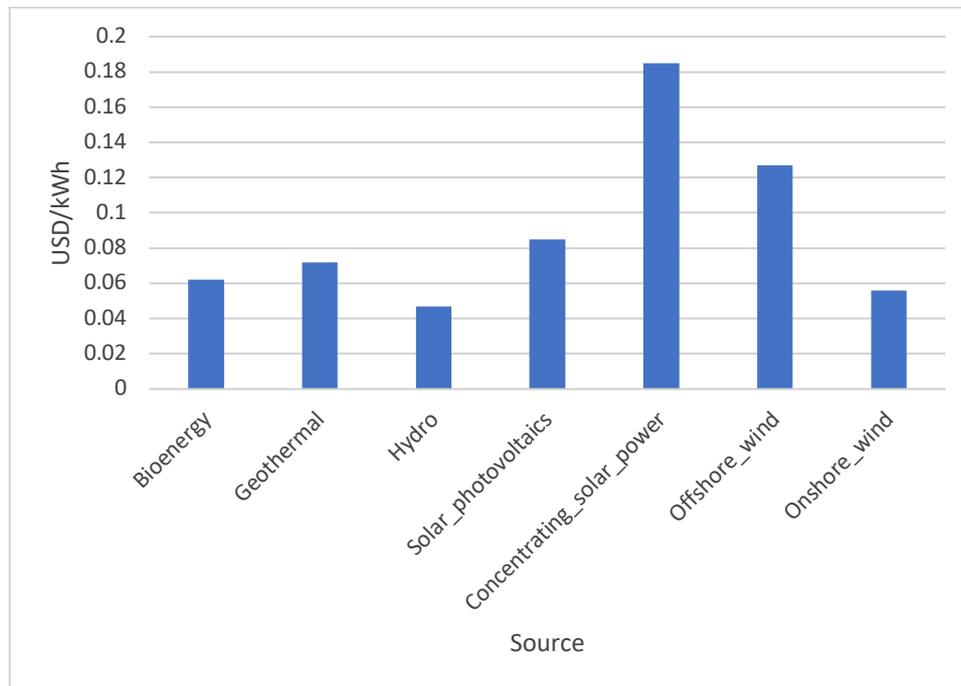
Figure 2-1 - Cost of electricity production from different sources. Source: IRENA., 2018

In contrast to cost, RES requires some definite elements of raw materials, such as wind turbines, converters, etc. However, extraction and manufacturing of wind turbines and other materials releases GHG (Kranzberg, 2012). Yet, production of electricity from RES requires fewer raw materials since they produce electricity at the source, whereas fuels need to be shipped to processing facilities to produce electricity.

On the other hand, when it comes to water consumption required for producing electricity, RES appears to be winning the race. According to Kranzberg (2012), sources such as nuclear and fossil fuels have been found to consume 60,000 and 50,000 gallons per megawatt per hour (MWh), respectively, as compared to merely one gallon water used per megawatt hour by renewable energy sources such as wind. However, solar source such as

sunlight has been found to consume 30 gal/MWh. Thus, in terms of water consumption, renewable energy appears to be more sustainable in relation to fossil fuels.

Moving towards assessing GHG emission remitted in generating electricity, Kranzberg (2012) found that, globally, electricity production produces 10 billion tons of CO_2 annually, understandably being classified as a health risk. While there is no direct CO_2 emission in using wind power to produce electricity, fossil fuels such as coal and gas emit 1017g of CO_2/kWh and 790g of CO_2/kWh of GHG into the environment when burned for energy production (Kranzberg, 2012, p.16). Thus, the carbon footprint of RES is massively low for producing electricity as compared to that of fossil fuels.

2.2.1. Usage of RES to Produce Electricity for the Kingdom of Saudi Arabia

Saudi Arabia is one of the largest countries in the Middle Eastern region comprising total area of 2,149,690 km^2 (Rehman et al., 2007). Saudi Arabia is a relatively rich country as it sits on some of the largest reserves of oil in the globe, amounting to an approximately 268.3 billion barrels. Such number represents 18% of global oil reserves (OPEC, 2017). Having abundance of oil means higher consumption of oil at domestic level.

According to Estimo (2016), domestically, Saudi Arabia consumes around 40 barrels per capita. If current rate of consumption continues, then it is predicted that, by 2038, Saudi Arabia will have to import oil to meet its domestic energy needs.

Electricity is produced in the country through oil. Electricity consumption in the country is five times greater than that of the US due to extreme temperature (Sughai, 2014). Saudi Electricity Company (SEC) is the key supplier of electricity throughout the country. Due to rise in the consumption of power transmission rate, SEC has grown by 50-60% and so has the number of customers by 70% during 2000-2010 (Taha, 2012). Looking at past trends, it is obvious that future demand for electricity is going to rise in the country.

According to an estimate, electricity generation consumes 25% of Saudi Arabia's oil production (Ratikainen, 2017). Fossil fuels used for the generation of electricity in the country are a cause for concern, not only for depleting oil reserves in the country, but also for the adverse environmental impact of fossil fuels (Ratikainen, 2017). According to the report by Climate Transparency, Saudi Arabia is the second largest emitter of greenhouse gases (GHG) in all of the G20 countries with per capital emission rising by 75% during

1990-2013. One solution to reduce such high rates is to increase current supply of electricity through addition of electricity produced by renewable sources (Ratikainen, 2017, p.19).

The Saudi Arabia Government has already started to take strategic initiatives to produce electricity using RES. Among the different possibilities, solar and wind power have been identified to be the most important players (Reed, 2018). According Ratikainen (2017, p.23) solar and wind power provide clear advantages to Saudi Arabia since there is enough area in the country to accommodate solar parks and wind farms.

Vast inhabited desert land in Saudi Arabia with sunshine throughout the year affords favourable conditions for solar and wind power sources. Feasibility studies have already been conducted to use wind and solar power to generate electricity and all of them have come up with positive final answers (Shalwala and Bleijs, 2009; Al-Ammar and Yousef, 2010; Shalwala and Bleijs, 2010).

It may be noticed that past studies have already presented positive feasibility for producing electricity in Saudi Arabia through wind and solar energy sources. However, the complete integration of such system into Saudi Arabia's energy grid is only feasible after assessing load forecasting. To this end, the western region of Saudi Arabia is ideal for producing energy through solar and wind power since it covers area with deserts and extremely hot climate during the day. Very high temperatures and humidity mean people in the area are compelled to use air conditioning systems in their daily routine lives, hence, causing hike in electricity demand, specifically during hot summer times when temperature goes up 54^0C (Weather Online, 2019). Therefore, it becomes necessary to cover a wide geographical area with PV and wind sources, since high temperature may cause the PV generation efficiency o to drop.

Apart from weather conditions, another reason to advocate better power management in the western region of Saudi Arabia lies in the fact that its cities, such as Mecca and Medina, are global hubs for religious tourism (Henderson, 2011). According to Arab News (2018), during 2018 more than two million people performed *Hajj,* hence, requiring more electricity to meet the needs of pilgrims. According to an estimate during *Hajj* peak time (the 10th *Eid al-Adha*), Mecca registered extremely high electricity loads of 284 megawatts

during 2018. That was 11% higher than the 256 megawatts load registered during 2017 (Al-Arabia News, 2018).

To ensure reliable transmission of power, the Saudi Electricity Company has to take pre-emptive measures, such as maintaining facilities, grid stations and electrical equipment costing millions of dollars. High electricity consumption during *Hajj* times directly contributes towards high carbon emission, which, in turn, poses colossal environmental challenges (Bhattay, 2017). To this end, the Saudi government has taken initiatives to promote solar and wind energy usage to produce electricity with an investment of 50 billion dollars (Bhattay, 2017). However, having electricity production facility based on RES is one thing, but actually incorporating them into the system is another, hence, the requirement for robust load forecasting system (Gajowniczek and Ząbkowski, 2016).

2.3. Smart Energy Development

It has been found that electricity from renewable energy tends to be sporadic, resulting in unreliable network to meet the demand (Suganthi et al., 2015). The need for a reliable system for forecasting electricity production of renewable energy lies in the fact that such an approach will result in providing reliable and environmentally friendly energy, hence, will not only aid in effective usage of energy sources, but also will leave behind a healthier environment for the generation to come (Yaïci and Entchev, 2016).

The need for the development of smart ways to develop energy stems from the necessity to reduce its impact on the environment (Kanchev et al., 2011). While various methods for producing renewable energy such as photovoltaic (PV), wind power, etc., have been developed there is, however, a lack of system that could control and synergise the demand and supply in real time (Byun et al., 2011). It is this need for controlling demand and supply of power in real time that has led to the development of the various electricity modelling techniques (Li et al., 2011). More practical reasons for the development of smart energy development modelling techniques come from the facts related to demographic reasons, such as rise in population, economic growth, etc., demand for electricity increases thus requires effective power planning at the state level (Deka et al., 2015). Given the abundance of possible sites for developing renewable energy for solar and wind farms in Saudi Arabia, rising energy needs and the government's initiative to decrease its reliance on

the fossil fuels to run the economy, yet its failure to carry forward any project (Strait Times, 2018), establishing compelling reasons to identify the energy modelling techniques to produce diversify economy and reinvigorate growth.

2.4. Energy Management Models

Management of energy in an optimal manner is an utmost priority since decisions made today will determine the future of the next generation (Fan et al., 2014). To this end, having data for renewable energy for the sake of it would not bring any advantage unless used efficiently so that quality, availability and environmental effects of renewable energy are understood in their totality. The key aim of the predictive models is to anticipate the demand to ensure effective use of renewable energy resources and establish reliability in the supply chain. Predictive energy models have been classified in various ways, such as static versus dynamic and univariate versus multivariate. Regardless of which model is used, the key aim is to develop scenarios and project the energy requirements for both residential and industrial purposes (Kaytez et al., 2015). In their review of the energy models for demand forecasting, Suganthu and Samuel (2012) identified various energy models, including predictive models, artificial neural models, power network and planning and modelling and simulations.

2.5. Power System Forecasting Models

According to Ertugrul (2016), forecasting long-term electricity demand is vital for identifying planning needs. In the event that overestimations are made while forecasting electricity demand or supply, it may lead to superfluous capacity resulting in wasted resources. Similarly, underestimation could result in shortage of power, leading to higher operational cost, inconvenience for the general public and could become a cause of energy outage.

The report by the World Data on energy consumption shows that Saudi Arabia is completely self-sufficient in fulfilling its electricity needs, which, as of 2015, stood at 292.80 billion Kilowatts Per Hour (kWh) (World Data, 2015). In contrast, during 2015, Saudi Arabia produced 318 billion kWh energy, that is, 109% of its own requirement. This means a surplus of 9%, which ended up being wasted electricity. A key aspect that worsens this

fact is that the production of electricity in Saudi Arabia also results in the carbon emission (CO_2) left into the air as a result of the overproduction.

According to the World Data (2015), during 2015, electricity production accounted for 601.05 metric ton (mt) of CO_2 in Saudi Arabia. In line with the estimations provided above, it can be argued that 9% of the CO_2 emission, that is 54.09 mt of CO_2, could have been avoided through applying proper prediction technique. The situation becomes even worse since 99.9% of the electricity was produced from sources related to fossil fuels and only 0.1% was produced through using renewable energy. Thus, it can be argued here that, through using appropriate estimation technique and integrating it with the renewable sources of generating electricity, Saudi Arabia can reduce its CO_2 emission by a big margin. Regression and time series methods are the two most traditional methods used for forecasting electricity demand and consumption. To this end, over the years, various predictive models have emerged, which are as follows:

- Regression models;

- Time series;

- Artificial neural network (ANN) models;

- Fuzzy logic;

- Hybrid prediction models;

- Deep Learning (RNN models);

According to Warrior et al. (2016), in the field of electricity forecasting there has been found several applications ranging from predicting power demand and prices to the strengths of solar and wind power output. The significance of using a good prediction model for forecasting purposes lies in the fact that power demand and supply relies on various environmental factors. On the demand side, factors such as weather conditions, human needs, time of the day and year could cause problems in accurately predicting demand. In contrast, on the supply side, weather conditions and wind power directly contribute to complexity of generating power from these sources at a given period of time. Behavioural and environmental factors make things worse when it comes to short-term power forecasting (STF) since they cause nonlinear problems (Mares and Mercado, 2017). To this end, Chen

et al. (2010) proposed various algorithm-modelling techniques, such as time series, neural networks and autoregressive integrated moving averages. All these models are nonlinear and suited for short-term load forecasting.

Bendary et al. (2013) investigated the impact of long-term electric load forecast on the management of load in demand side management (DSM) in the context of various customer segments. Customer segments investigated by Bendary et al. (2013) included households, public and commercial organisations in Egypt. The authors aimed to investigate the load of the above-mentioned sectors on the national network of power systems in Egypt for the period of 2009-10 to forecast load through using an adaptive network called Fuzzy interference system (ANFIS) for the period of 2010/11 and 2020-2021. To this end, daily load curve for the load demand in Egypt for the month of August during 2009 and 2010 was taken into consideration as expressed on the country's power system. Daily load forecast was predicted using three network systems, namely, ANFIS, artificial neural network (ANN) and regression programming. The aim of the applications of three systems was to estimate loads for the period of 2010-2011 and 2020-2021. Research concluded that ANFIS system was more reliable than the other two methods because ANFIS produced lowest system error.

In contrast, in their study Cheah et al. (2011) found ANN more accurate when it comes to short-term load forecasting (STLF) in the context of Singapore. To this end, Cheah et al. (2011) used ANN to estimate quarter-hourly ahead load. The ANN was programmed using national instrument (NI) LabView software that then aligned with three-layers feed-forward ANN for which backpropagation (BP) learning method was used to train neurons. To improve performance of the learning method, early stopping (ES) algorithm was applied to ensure that the system avoided danger of over-fitting while training of the data. Historical load data were gathered from Energy Market Company (EMC) in Singapore for the period of January 2010 to March 2010. Findings of Cheah et al.'s (2011) study indicated satisfactory results for ANN. It was found that application of ANN could be used to accurately forecast daily unit commitment, that would not only minimise cost of scheduling, but also maximise revenue of the power grid in Singapore.

On the other hand, Tan et al. (2013) tested the ability of renewable energy sources for reducing power loss, increased system reliability and low investment requirement. To

14

this end, they adopted a multi-objective approach to test impact of optimal size determination and allocation of multi distribution generation systems that also linked power system increase. They (2013) tested the application of a new hybrid population-based algorithm in combination to practical swarm optimisation (PSO) and gravitational search algorithm (GSA) models. To ensure robustness of the system, various factors were taken into account, including real power losses, voltage profiles and mega voltage ampere (MVA). Findings revealed that the expected combined algorithm conformed to effective, robust and handling of nonlinear optimisation problems.

Xing et al. (2015) tested applicability of integrating energy from a wind power generation system. It was assumed that increased penetration of renewable energy would push the grid sufficiently to create a balance between power supply and demand. Xing et al. (2015) concluded that combining a demand response (DR) system with a wind farm economic dispatch model could compensate situations that demand higher peak times load. The objective of the function of a combination of DR and wind farm economic dispatch model was to reduce operational cost, which includes fuel cost, initiation cost, reduction in CO_2 emission and reduce demand response cost. In order to achieve the research goals, Xing et al. (2015) further added DR maximum participation time and DR ramp rate constraints into the model. The overall model was tested on the basis of ten traditional units that included one large-scale wind farm and three types of demand response, such as interruptible load, direct load control and load capacity. The reason behind including traditional units into the model was to validate the effectiveness of the proposed model. Furthermore, to examine the model's effectiveness, four DR penetrations rates were used, included 0%, 10%, 20% and 30%. Findings indicated that DR penetration rate tends to reduce total system operations costs, indicating that lower priced DR are given priority. Thus, the study asserted that it is important to integrate DR and wind farm integration model to gain better economic results.

Yammani et al. (2014) used modified teaching learning-based optimisation (MTLBO) algorithm to examine the optimal placement and size of distribution energy resources (DERs) units in a power distribution system. Yammani et al. (2014) connected DERs near load in the electric distribution systems and integrated various RES into an electric distribution network, which included wind, solar and fossil fuel-based generation, such as micro turbines, along with other distributed energy elements. Every combination of

DERs was monitored for seek impact indices, active and reactive losses and voltage profiles. For each case of test, uniform data were collected from eighty-four businesses in the Taiwan Power Company distribution system. Findings revealed that integration of DERs at different places and sizes reduces overall network losses.

2.5.1. Regression

Recognition of the need for preserving energy across the globe has resulted in both public and private sector organisations striving to search for methods for improving energy performance (Fumo and Biswas, 2015). To this end, investments have been made towards identifying ways of reducing grid energy consumption through investigating various energy alternative sources, such as solar energy and wind power. The aim of the research has been to boost efficiency of the grid through integrating renewable energy sources (RES) (Norrga et al., 2012). However, integration of RES into a grid system requires building alternative modelling designs. Efficiency of new designs can be assessed through simulations that would lead to determination of the most efficient and less costly methods. Fumo (2014) identified three approaches to building new energy consumption modelling techniques, such as statistical, hybrid and engineering. Among these three techniques, the statistical approach has been recognised as the most effective since it relies on measured data. Other benefits of using regression analysis for energy consumption modelling include comparative ease of implementation, less need for computation power relative to other statistical approaches, good predictive capability and ease of availability of data from grid stations or metering (Foucquier et al., 2013).

Wang et al. (2012) attempted to use a regression model to eliminate the impacts of residual disturbance from traditional semi-parametric model of electricity consumption. To this end, a case study of electricity consumption in China was used to examine the effectiveness of an improved semi-parametric regression model. China presents a unique case study to evaluate the effectiveness of regression modelling technique for electricity consumption forecasting because of its unprecedented economic growth that led to increase in electricity consumption by more than 70% during 1980 to 2010. Noting that 1978 was turning point for China's economic growth, the authors used data between 1980-2005 for the purpose of building the model and data for the period of 2006-2010 were used to test the proposed regression model. In order to improve efficiency of the semi-parametric model,

three statistical measures were used to evaluate accuracy of the prediction model: mean absolute error (MAE), mean absolute distribution (MAD) and mean squared error (MSE). Results revealed that this modelling approach can assist in capturing the complicated cooperative relationship between electricity consumption and economic drivers. Forecast results presented satisfactory performance of the model in predicting future electricity consumption and, thus, demonstrated reliability and effectiveness of the regression modelling technique for forecasting electricity consumption.

Another study by Satiawan et al. (2009) tested the use of the support vector regression (SVR) model to predict load demand in Australia for every five minutes, characterising the very short-term load forecasting (VSTLF). An Australian national energy market operator known as NEMMCO uses VSTLF as an approach to issue every five-minute production schedule for generating electricity. Five-minute regional demand methodology is aimed to forecast demand at the beginning of dispatch interval until the consumption. Electricity generators then use this forecast to adjust their supply of electricity. To test the viability of the SVR model to forecast five-minute load, VSTLF data from 2006, 2007 and 2008 were gathered from NEMMCO. The key objective of the study was to check the applicability of VSTLF in the Australian power system. To this end, performance of SVR was tested against three traditional methods: linear regression (LR), least mean square (LMS) and backpropagation neural network (BPNN) algorithm. Four sets of time period on a specific day were used to evaluate the performance of each modelling technique. Findings revealed that regression models such as SVR, LR and LMS obtained uniformed mean absolute error of 41% as compared to the 53% of BPNN, thus, the regression modelling technique appears to perform better than ANN modelling technique when it comes to VSTLF. Moreover, it has been reported that, although SVR showed promise as a prediction model in outperforming BPNN, LR and LMS also produced similar results and were even faster to design comparative to SVR.

Azad et al. (2018) recognised challenges facing the Gulf countries in efficient electricity load prediction, particularly in the case of the United Arab Emirates. According to the report by the UAE State of Energy, as of 2015, residents in the UAE were using about 20-30 kilowatt/hour of electricity per day, which was expected to grow by 9% annually (Government.ae, 2019). To this end, Azad et al. (2018) tested use of the Support Vector

Regression (SVR) model for predicting peak load demand through using historical data on peak load consumption records. The authors supported use of the SVR model because of its ability of minimising structural risk instead of minimising training errors, as happens in ANN modelling.

For the purpose of SVR modelling, Hamdan et al. (2017) collected data for the period of 2014-2016 on daily records of electricity peak load consumption and corresponding temperature and humidity. Data contained peak load consumption and corresponding temperature and humidity of Sharjah City, UAE. Azad et al. (2018) implemented SVR using R studio software. Data for the period of 2014-15 were used for training the Sector Vector Machine (SVM), the results of which were then trained in SVR. In contrast, data for the year 2016 were used to test ability of SVR to predict peak load time of that same year. Evaluation of findings concluded that the SVR model showed excellent prediction capability for both weekends, weekdays and seasonal data.

Another prominent study was conducted by Yang et al. (2016) to analyse and predict Beijing's (Capital city of China) electricity consumption using various multivariate regression models. Multivariate regression models used included ordinary least squared regression (OSLR), principal component regression (PCR), partial least squares regression (PLSR) and modified partial least squares regression (MPLSR). The rationale for using multivariate methods to predict electricity consumption was presented through arguing that multivariate modelling assists in building a fitted model to accurately predict electricity demand. For the purpose of analysis, a dataset from Chinese Statistical Yearbook was gathered consisting of information on electricity consumption and corresponding economic variables. To this end, economic growth factors, such as GDP, primary industry, secondary industry, tertiary industry and total population, were used as independent variables and electricity consumption was used as dependent variable. Due to the multicollinearity distribution of data, OLSR modelling technique could not be used because OLSR is inherently designed to solve the problem of collineation. Based on theoretical constructs of PCR, PLSR and MPLSR, an electricity consumption model was conducted using identified dependent and independent variables. Findings revealed that, compared to other multivariate models, PLSR presented the best predictive capability as, when used in the electric grid

18

system, it resulted in saving cost on manpower and financial resources required to distribute electricity in Beijing.

2.5.2. Neural Networks

While the linear regression-modelling techniques discussed above are suitable for modelling of data presenting linear relationships between different variables, they are unable to provide accurate load prediction with dataset presenting nonlinear relationship (Zhao and Magoules, 2012). That is particularly important when forecasting load, since there are nonlinear relationships between load and various underlying factors, namely, weather conditions such as temperature, humidity and air pressure, and environmental factors such as time of day, seasons, etc. To this end, the artificial neural network (ANN) approach provides a unique solution to identify relationship between load and underlying factors (Quan et al., 2014). ANN is defined as a data-driven nonlinear mapping tool that can assist in resolving nonlinear relationships between various factors through data training.

One of the key problems associated with ANN is the time period required to train the network and attain data with minimum error (Ding and Wu, 2012). Training of data is a process that involves determination of optimal weights of neural in ANN conducted through evaluating a cost function (i.e. mean squared error (MSE), sum of squared errors (SSE)) between the output of network and desired target and minimising the error with respect to the weight.

Backpropagation (BP) is a supervised training method that has been employed consistently for ANN training. Supervised method of training data involves interactive updating of actual numerical weights (W) to elements of inputs (X) such as time, day, temperature, humidity, etc. Numerical weights and inputs are then matched with historical data to attain desired outputs. However, such process could take a lot of time, as identified by Ding and Wu (2012). To overcome this limitation of ANN modelling technique, various methods for training algorithms have been proposed, each of which possess their own advantages and disadvantages.

Pan et al. (2013) compared six methods for training data under BP technique, such as Additive Momentum, Self-Adaptive Learning Rate, Conjugate Gradient Backpropagation, Quasi-Newton, Resilient Backpropagation and Bayesian Regulation

19

Backpropagation. To test the efficiency of each of the six methods, four-phase framework was applied under which the dataset contained two parts: actual records and labelled data. The first part of data included actual data of hourly load and weather records of the city of Toronto, Canada, for the period of July 2009 to June 2011. In contrast, labelled dataset included logical information aimed at identifying load difference at particular day, such as weekend, weekday and holiday. Nine weather conditions were used as input to the network: mean, lowest and highest temperature, mean sea level pressure, mean humidity, precipitation amount, mean visibility, mean wind speed, maximum sustained wind speed, logical adjustment of weekday or weekend and load records k. The model was designed to test the accuracy of the algorithm at 5% mean squared error training goal. Each of the models was compared on five capability factors: parameter setting complexity, training speed, convergence, accuracy of prediction and stability of results. Analysis of results revealed that, when it comes to processing speed and prediction accuracy, Resilient Backpropagation and Conjugate Gradient Backpropagation are better than other types of training methods. However, others, such as Quasi-Newton and Self-Adaptive Learning methods, lack accurate predictability yet they are fast to model. Bayesian Regulation Backpropagation achieved good accuracy and was able to process large data.

Despite the complexity of the ANN training process, various analysts and academics have used this model in the electricity load forecasting field. Badri et al. (2012) compared the capability of ANN and fuzzy logic (FL) forecasting methods to predict the short-term (hourly) electricity load in the capital of Iran, Tehran. Like ANN, FL has also been identified as an efficient tool to identify nonlinear map from input space to output space. FL technique tends to apply IF-THEN rule to evaluate mapping behaviour. Twenty-four hour data were attained from Tehran Electricity Company (TEC), which included information on each hour load amount and corresponding temperature. Analysis of the data in MATLAB ® software revealed that ANN showed significantly better accuracy than FL modelling technique. Badri et al. (2012) attributed such difference to the inaccurate regression curve determined by FL, while ANN reproduced the historical data through an accurate regression.

Feilat and Bouzguenda (2011), in their quest to forecast electricity load at short-term level in Oman, compared linear regression and ANN tools. To this end, data for the period of 2007-2010 were collected from the Mazoon Electricity Company (MZEC). Exploratory

20

data analysis revealed U-shaped relationship between electricity consumption during seasons, with high usage at summer and lower in the winter. Data for the period of 2006-2009 were used to evaluate the models using MATLAB® computer simulation program. Results of the simulation showed that neural networks outperformed linear regression technique and ANN training enabled it to capture the monthly peaks in historical data.

Similarly, Deng (2010) attempted to identify electricity demand of China until 2050 using ANN and compared results with linear models. To this end, feed-forward multilayer perceptron with BP training algorithm was used as a typology of neural network. Data for the study were gathered from National Bureau of Statistics of China's information service in four independent variables: GDP, population, import and export amounts. Data from 1990 to 2008 were used to train data and then test the linear and ANN models for their capability to accurately identify future demand of electricity in China till 2050. Findings revealed that energy demand in China would peak between 2026-2035 and would decrease gradually. Analysis of the data in MATLAB® software revealed that Feed-Forward ANN model would be able to better estimate energy demand comparative to linear regression model optimised through least squares method (LSM) and using as cost function the root mean squared error (RMSE).

Saravanan et al. (2012), while forecasting electricity load in India, also agreed with Deng (2010) as they also found that ANN presents more encouraging results to predict the future demand of electricity in India when compared to linear regression analysis.

According to Jain et al. (1996), neurons' ability to learn is a fundamental trait of ANN. In the context of ANN, learning is viewed as resolving the problem of updating connection weights to allow the network to perform a specified task efficiently (Jain et al., 1996). Although there are different training algorithms, they do not follow any preconceived rule, rather, the neurons tend to learn as a result of establishment of input/output relationships.

Jain et al. (1996, p.35) identified three artificial intelligence algorithms, as supervised, unsupervised and hybrid. Table 2.1 provides differences between the three methods of learning.

21

Table 2.1: Difference between supervised, unsupervised and hybrid learning. Source: Adapted from Jain et al., 1996, p.35.

	Supervised	**Unsupervised**	**Hybrid**
Provision of output	Network is provided with correct answer for every input pattern	Explores underlying structure in the data to look for correlations, dispersion to identify correlation	Combines supervised and unsupervised learning methods to identify patterns
Weights	Updated according to a cost function to minimise the error between predicted and observed output.	Weights are determined by the system itself	Parts identified by system feeder and others by system itself

2.5.3. Fuzzy Modelling

Fuzzy logic modelling technique has been lauded for conceptualising the fuzziness in the system for forecasting electricity from a renewable energy source into a quantifiable parameter, hence, it can be adopted for effective energy planning. It is the accuracy of the fuzzy modelling technique that has led to its increasing use in the renewable energy systems. For instance, Groscurth and Kress (1998) adopted fuzzy modelling technique to compress the data for minimising the primary energy inputs and emission of pollutant matters. Similarly, in their investigation of effective energy management system for micro-grid based, they lauded the use of fuzzy modelling technique motivated by the possibility of identifying both nonlinear dynamic behaviour and uncertainty in the data emerging from the renewable energy source. Such approach allows the establishment of more accurate forecasting.

2.5.4. Hybrid Systems

The hybrid approach allows usage of different models in combination for electricity forecasting (Teixeria et al., 2017). The underlying assumption of the hybrid modelling approach is that it promotes usage of multiple modelling techniques for the purpose of electricity forecasting, overcoming the weakness of the single method through a combination of models. Yamin (2004), in his review of the methods of generating schedules in the electricity power system, lauded the use of a hybrid approach to electricity forecasting, arguing that it results in gaining a more accurate result for forecasting single unit commitment.

Koprinska et al. (2013) proposed a hybrid approach to forecast electricity demand and even prices for STLF. A hybrid approach is defined as a modelling technique under which ANN is combined with other statistical models to conduct forward analysis (Li et al., 2013). Koprinska et al.(2013) supported using a hybrid approach to electricity load forecasting arguing that using single model for forecasting purpose merely results in predicting value for single future point, such as single modelling merely allows to use values at time t to predict at time $t+h$, where h is called as forecasting horizon. However, when it comes to predicting value for more than one step ahead, such as forecasting electricity load for a forecast horizon using hourly, daily or monthly time period, represented as $t+h$, $t+h+1$, $t+h+(n-1)$, a hybrid approach would be suitable.

Koprinska et al. (2013) proposed the combination of pattern sequence-based forecasting (PSF) algorithm with ANN to forecast electricity demand and prices for next twenty-four hours. The underlying concept of the PSF algorithm stems from time series statistical analysis that has been recognised for its ability to process numbers in terms of time, thus developing a series of changes in chronological order (Lee and Hong, 2015). Similar, to past researchers, Koprinska et al. (2013) used BP neural network structure to investigate its predictive capabilities when applied to data from China's power company in Beijing. Data span the period from February 1st to 30th April. Findings revealed that complementing ANN with time series prediction model provided accurate predictions.

The performance of a linear-nonlinear model for energy forecast was studied by Barak and Sandegh (2016). The linear model consists of Auto-Regressive Input with

Moving Average model (ARIMA) which was used as an input to the ANFIS (nonlinear) model, with other inputs for energy consumption forecast. By testing different ANFIS patterns, the authors found that using AdaBoost diversification schema for dealing with insufficient data possessed superior performance over traditional modelling technique.

2.5.5. Particle Swarm Optimisation

Particle Swarm Optimisation (PSO) approach is an algorithm-based model proposed by Kennedy and Eberhart in 1995 (Huang et al., 2005). The key advantage of using PSO is that it can be applied to situations where linear correlation or identifiable pattern does not exist. The PSO approach to electricity modelling uses various economic factors for the purpose of electricity forecasting, which includes Gross Domestic Product (GDP), population growth rates and imports and exports of a country as key indicators for the future electricity demand. The key benefit of using PSO methodology for electricity forecasting is that it can be used for both short and long-term electricity forecasting (Suganthi and Samuel, 2012).

2.5.6. Model Sensitivity Analysis

Sensitivity analysis is defined as a process of investigating potential changes or errors and their impact on the outcome of implementing a model or framework to resolve a problem (Pannell, 1997). The benefit of using sensitivity analysis for predicting the future lies in the fact that it is not only easy to understand, but also it is easy to use in a variety of matters, be it a social or scientific phenomenon (Borgonovo and Plischke, 2016).

Bauucells and Borgonovo (2013) argued that sensitivity analysis has been widely used in scientific and engineering fields to predict the behaviour of certain systems, such as forecasting the energy production from renewable energy source. Specifically, in the context of commercial building load forecasting, the assistance of sensitivity analysis has been highly lauded. The process of forecasting electricity for commercial building involves a high number of variables, management of which for forecasting could become a burdensome task. To this end, sensitivity analysis can assist through identifying the possible influence of most important parameters that have negligible effects to the outcome. This results in narrowing down the possible outcomes and strengthening the final conclusion (Borgonovo and Plischke, 2016).

2.6. Summary

Discussion in this chapter aimed to evaluate past studies regarding usage of various modelling techniques for electricity load forecasting. Discussion on predictive models revealed various modelling techniques that can be used to study the load forecasting, including time series, neural networks and regression models. All of which can be used for both short, medium and long-term load forecasting.

Overall, past researchers presented conflicting findings such as while some found that regression and FL models appear to be better than ANN (Satiawan et al., 2009; Azad et al., 2018; Wang et al., 2012; Bendery et al., 2013; Yang et al., 2016), others found the opposite (Cheah et al., 2011; Feilat and Bouzguenda, 2011; Badri et al., 2012; Ding and Wu, 2012; Saravanan et al., 2012; Pan et al., 2013). Additionally, other authors supported the use of hybrid approach for forecasting future electricity demands (Tan et al., 2013; Yammani et al., 2014). Discussion on hybrid approach to load forecasting revealed that, when ANN is combined with time series, it provides accurate results compared to the combination of ANN and regression analysis (Koprinska et al., 2013). The conflicting nature of past studies shows that there is a gap in the research about evaluating the capability of different models to predict future electricity load.

Past researches also explored the limitations of each modelling approach. For both ANN and neural structures (such as neuro fuzzy) in general, its black-box structure hinders its interpretability. Also, as mentioned by Ding and Wu (2012), it requires data as every data-driven approach and are limited in capability of extrapolating over the training range.

Review of the literature indicates that having robust energy prediction model results in establishing an understanding of the quality, availability and environmental effects of using renewable energy. Furthermore, it was found that, in the context of Saudi Arabia, it is important to use an effective energy modelling technique to eliminate the possibility of energy waste. The next chapter moves towards analysing the usage of renewable energy sources for the production of electricity in both public and private sectors.

Chapter 3.

Theory and Methodology

3.1. Regression Model

Drawing on Draper and Smith (2014), regression analysis is a statistical tool used to find a functional relationship among different responses or independent and dependent variables and predictors. In the event that regression analysis deals with one variable, then regression is called univariate regression and when multiple variables are used in the analysis then it is called as multivariate regression. The discussion below provides descriptive analysis of both linear and multiple regression modelling techniques.

3.2. Linear Regression – Single Input, Single Output

The simple linear regression consists of the following base formula:

$$Y = \beta_0 + \beta_1 X + \varepsilon \qquad (3.1)$$

where,

Y = Response variable

X = Predictor variable

β_0 and β_1 = Regression coefficient or regression parameters

ε = Error for the discrepancy between predicated data and observable data

For the purpose of using a linear regression model for predicting a scenario such as electricity consumption, an exponent (^) sign is added on top of parameters, as shown in the equation below:

$$\hat{Y} = \widehat{\beta_0} + \widehat{\beta_1} X \qquad (3.2)$$

where,

$\hat{Y}$ = Fitted or predicted value

$\widehat{\beta}$ = Estimates of the regression coefficient

The difference between the observed and fitted values is used to calculate the actual error of the model. This error can be minimized by using an optimization algorithm so as to build a model that is able to accurately estimate the electricity supply and consumption.

3.3. Multiple Regressions – Multiple Input, Multiple Output

Multivariate modelling technique is used in the scenario when there are many variables that also include control variables (Murtagh and Heck, 2012). To predict electricity demand and consumption, multiple regression uses weighted least squares estimation (Fumo and Biswas, 2015). The key aim is to identify the statistical relationship between total load and weather conditions along with influence of the day type. A defined amount of historical load data is used to compute the regression coefficients through an equally or exponentially weighted least squares estimation. To this end, the following equation is used:

$$\widehat{Y_\iota} = \hat{\beta}_{i,0} + \hat{\beta}_{i,1}X_1 + \hat{\beta}_{i,2}X_2 + \cdots \hat{\beta}_{i,p}X_p \tag{3.3}$$

There are two approaches to apply regression analysis for the purpose of predicting electricity: top-down and bottom-up approach (Braun et al., 2014). The top-down approach tends to use information on higher level of energy consumption and means considering variables defining overall energy consumption instead of use of energy by the individual end-users. Thus, the top-down approach to estimating energy consumption is usually used for the long-term purposes. In contrast, the bottom-up approach aims to estimate the expected electricity consumption at the houses in a particular region through using models that define the use of electricity at the individual house level. To this end, a segmentation tool can be used to characterise the residential segment of interest. Since, the aim of this study will be to use regression technique to predict the consumption and potential supply of electricity at the short-term level, bottom-up approach to regression analysis method will, therefore, be used.

3.4. Types of Model Input

According to the type of data being input to the model, it may be classified in two classes:

- Time series input: In the context of energy load forecasting, time series consists of using past energy load data to predict future ones.

- Multivariate input: data representing another phenomena or information, such as temperature, humidity, day of week, etc.

Drawing on Palma (2016), time series is a statistical tool that includes a collection of observations derived sequentially in time. Time at which data were taken can be either discrete or continuous. It is the inclusion of time period in the statistical analysis of time series that has resulted in it being likened to a stochastic process corresponding to random variables noted across the time. There are various fields in which time series is used, such as finance, economic, sociology, energy, medicine, climatology and transport. Various modelling techniques can be used under time series analysis, such as discrete time models, serial dependence, non-stationary, whiteness testing, parametric and non-parametric modelling.

The process of conducting time series analysis starts with drawing data at equal time intervals, which could be every minute, every hour or every day (Box et al., 2015). The underlying assumption behind using equal time intervals in time series analysis is that there exists some pattern in historical observations that is expected to be repeated in the future. Thus, effective use of time series requires no structural breaks or routine switching in the process of data collection.

It is expected that usage of historical data as a variable in prediction model would result in providing accurate forecasting (Aryaputera et al., 2015). Time series (TS) data are defined as large data that include a series of numerical values obtained over a period of time. Values in data are measured in identical pattern of time intervals, such as every hour, day, week and month. Furthermore, TS technique is recognised for its effectiveness in situations where there is high need for accuracy (Wan et al., 2015).

28

The feasibility of accurately predicting future load through using TS modelling could assist decision-makers at SEC to plan and manage the demand response. Moreover, TS modelling provides potential to design a smart grid station system connected with RES, such as solar and wind energy in the context of the present research; however, for this to happen it is important to relate TS with electricity load forecasting. Another advantage of TS in short-term load forecasting (STLF) is to re-regulate integration and promotion of RES into a grid system (Martin et al., 2010; Yang et al., 2013).

3.5. Artificial Neural Network Model

Artificial Neural Networks (ANNs) are mathematical models that, similar to linear regression, are capable of performing prediction of future or unknown states of a system given some known input data. However, differently from the linear regression, ANN are far more complex mathematical structures which are able to deal with highly non-linear relationship between input(s) and output(s). Therefore, it is appropriate to compare this type of predictive modelling with linear regression.

The underlying concept of Artificial Neural Networks (ANN) has emerged from the theory of artificial intelligence (AI) (Agatonovic-Kustrin and Beresford, 2000). AI is an area of computer science that uses software that is capable of utilising sophisticated intelligence and computations similar to those used by the human brain. The key tools used in AI include logical and inductive knowledge acquisition and reasoning capability for solving problems. ANN is a type of AI that models the way the human brain works through using an expert knowledge-based system which allows the way rules stimulate the process of thinking and derives conclusions and solves problems through using pre-set rules. Therefore, it would not be wrong to suggest that ANN is a digitalised model of the human brain that learns through experience. In the ANN, neural networks are used to gather knowledge through detecting patterns and relationships within the data. To form neural structure in an ANN, artificial neurons are usually connected with the coefficients, also known as processing elements (PE). Each of the PE in the ANN tends to have weighted inputs, transfer function and one output. The key aim of the PE is usually to balance the inputs and outputs.

When it comes to predicting future energy needs, various forms of ANN models have been proposed that include: Feed-forward; Recurrent; Hopfield; Elman, Bi-directional, Self-Organising maps, Stochastic and Physical ANN (Yuan et al., 2018).

ANNs are mathematical tools originally developed for the purpose of identifying the way the human brain processes information (Schalkoff, 1997). They are composed by simple elements that operate in parallel pattern. The topology of the network is usually determined by the connection between elements. ANNs can be trained to perform specific function. The training method is conducted through adjusting values of connections known as weights among elements and is usually conducted to allow inputs to gain particular output, thus, the process of adjustment of network tends to continue until output matches the target. In reality, the process of adjustment requires adjustment of many input/output pairs to train a network.

Processing of neurons is conducted internally in two stages. In the first stage, input values are combined internally whose result triggers (non)linear activation function. Secondly, a combination of neurons uses weights of each connection with a consistent bias term. The neuron output comprises the following function:

$$Y = f\left(\sum W_t X_t + \theta\right) \qquad (3.4)$$

In Equation 3.4 above, X_t represents neuron input multiplied by weight link W_t and θ represents the characteristics of neuron offset (bias) and finally f represents the activation function.

Mostly, ANN is structured in multilayer perceptron (MLP) type that organises the layers. Neurons in each layer could share similar inputs that may not be connected to each other, in which case the architecture is feed forward neuron (FFN). The output of one layer is considered as an input of the next layer. The layers between input and output layer are called hidden layers. Figure 3-1 presents an example of three-layer FFN perceptron network with three input neurons, four neurons as hidden layer and three output layers. Each layer possesses specific number of nodes, the interconnection between which is only with adjacent layers. Each neuron belongs to a layer and is connected to an adjacent layer.
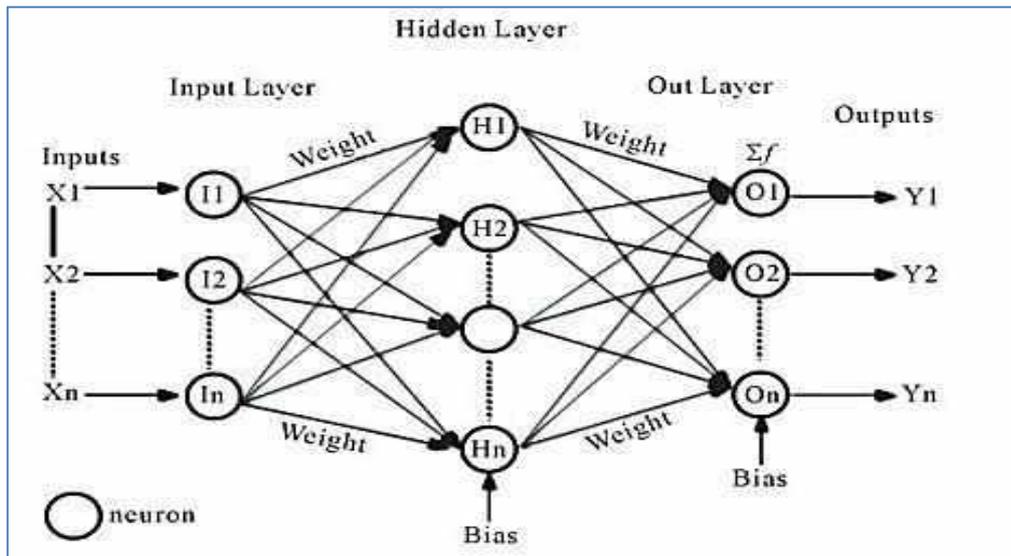
Figure 3-1 - Architecture of FFNN. Source: Alshareef and Abbod, 2005.

Parameters of the FFN presented in Figure 3-1 are weight matrix and the bias. The process of prediction of parameters is called training of the network. The method used for training algorithm in load forecasting and is called a backpropagation algorithm. The default in MATLAB® is Levenberg-Marquadt. It sets the problem of optimising the weights of the neural network in a similar structure as the least squares problem. It can be stated that it tries to minimise the parameter values by finding the parameters which finds the minimum of square deviation. Mathematically:

$$\beta_{opt} = \min_{\beta} \sum (y_i - f(x,\beta))^2 \qquad (3.5)$$

where:

β_{opt} – optimum set of parameters;

y_i – observed values;

$f(x,\beta)$ – nonlinear function A(NN) to be optimised.

First step of the solution approach is to linearise the ANN using Taylor expansion:

$$f(x,\beta + \delta) = f(x_0,\beta_0) + J_i \delta_i \qquad (3.6)$$

where:

31

$$J = \frac{\delta f(x, \beta)}{\delta B} \tag{3.7}$$

J consists in the gradient vector of the cost function.

Applying the algorithm, performing rearrangements and other operations as necessary, the final equation implemented by Levenberg-Marquadt, with a damping factor, is:

$$(J^T J + \lambda I)\delta = J^T [y - f(\beta)] \tag{3.8}$$

Using backpropagation algorithms, one can train neurons, get network weights and biases in the direction in which performance function creates least error.

The activation function used for hidden layer is the default one from MATLAB®, which is hyperbolic tangent sigmoid function:

$$tansig = \frac{2}{1 + e^{-2n}} - 1 \tag{3.9}$$

Figure 3-2 illustrates the output obtained from this activation function.
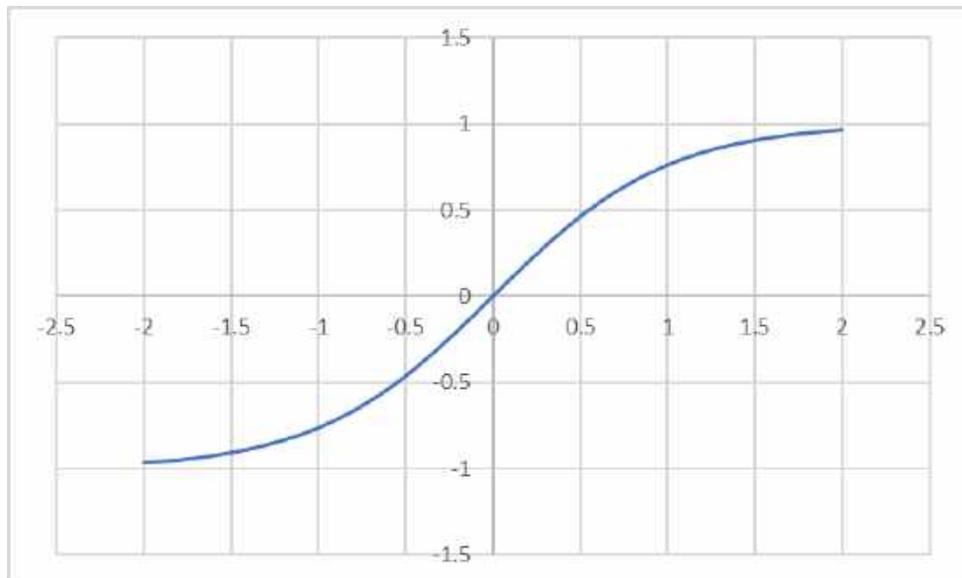
Figure 3-2 – Hyperbolic tangent sigmoid function – Activation function in ANN.

The basic form of multilayer FFNN would be most suitable for the present study because of the involvement of a large amount of data yet prior little prior knowledge

regarding the regulatory system that manages the generation of data (Haykin and Network, 2004).

3.6. Hybrid Systems

Different types of models can be combined in a single superior stage structure, generating a hybrid system. Technologies such as ANN, POS, ANFIS, LR and other modelling techniques are combined in such a way that one can take profit of their advantages while minimising the deficiencies of each one. There is a variety of ways of performing such integration and many authors have showed the efficacy and usefulness of such methodology when handling real life problems (Sarabia et al., 2005; Wdaa, 2018).

3.6.1. Model Combiner

A model combiner consists of a superior modelling stage where the outputs of different models are combined to generate a third output, which is intended to be superior to the former ones, through a balance of the model output(s). This balance may be linear or nonlinear.

In the research, two superior modelling stages were evaluated: Fuzzy and POS. In both cases, the combined structure consists of a linear balance (weighted average) of the two internal models. Considering y_{pred} as the output of the LR model and z_{pred} as the output of the ANN model, the combined one generates the output w_{pred} which comprises the linear equation:

$$w = ay_{pred} + bz_{pred} \qquad (3.10)$$

Ideally, the accuracy of w would be higher than y_{pred} or z_{pred}. The two parameters of the superior model (*a* and *b*) are optimised through the superior stage model.

For comparison, a third superior model was evaluated which fixes the parameters *a* and *b* to 70% and 30% of the best and worst model, respectively.

The following subsections give a short theory background on the Fuzzy Combiner concept and the PSO concept.

3.6.2. Fuzzy Model Combiner

Fuzzy modelling consists of inferring information from data using a set of rules which are diffuse by concept. A rule can be understood as a conditional statement, as an if-else expression.
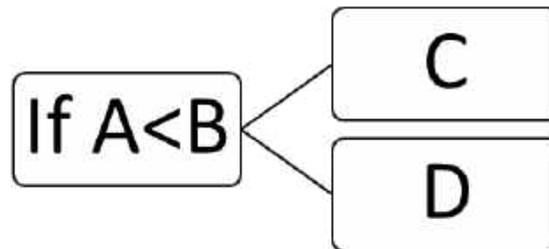
Figure 3-3 – Example of simple if-else fuzzy rule.

Figure 3-3 shows a simple if-else fuzzy rule as example, where A and B are descriptors of pieces of knowledge and the rule is the block statement which gives a body to the relationship between variables involved in the problem. The descriptors, or information granules can be sets, rough rules or fuzzy sets.

The rules can be categorised as:

- multi input multi output (MIMO) fuzzy rules:

  e.g. : If (A) and (B) and (C) then:

  Y1 and Y2 and Y3

- Certainty-qualified rules:

  e.g.: If (A) and (B) and (C) then:

  Y1 = C with certainty P

where P is the degree of certainty. Usually $0<P<1$. $P = 1$ means that the rule is 100% certain.

- Gradual rules.

  e.g.: If A is higher, then B is lower.

This type of rule expresses relationship in the behaviour of variables. It also captures tendencies between pairs of collections of information granules.

 - Functional fuzzy rules:

e.g.: If (A) then y =f(x)

The rules establish that the function y is restricted to the information A. This brings modularity to the model.

Figure 3-4 illustrates the structure of a general FIS (Fuzzy Inference System).
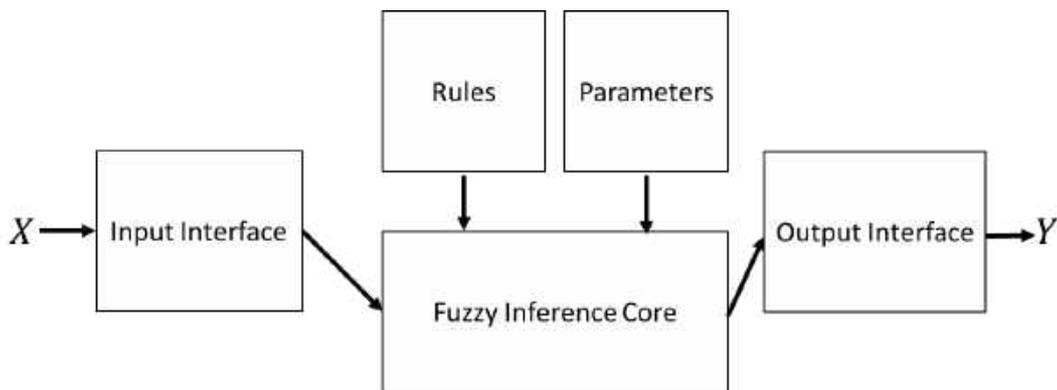
Figure 3-4 – Fuzzy Inference System Architecture.

A fuzzy inference system consists of the following five steps (MATLAB®):

- Fuzzification of the input variables;

- Application of the fuzzy operator (AND or OR) in the antecedent;

- Implication from the antecedent to the consequent;

- Aggregation of the consequents across the rules;

- Defuzzification.

In MATLAB®, the fuzzification of the inputs is done by determining the degree of each one through membership functions. This consists of an always crisp numerical value limited to the discourse range of the variable. The output of this transformation is a fuzzy degree of membership qualifying the set in a linguistic sense. In summary, this process generates a table lookup or a function evaluation.

Fuzzified inputs are clear in the sense to which degree they satisfy the rules of its antecedent. In the case that an antecedent has multiple parts, the fuzzy operator obtains one number indicating the result of the rule antecedent. This value is then applied to the output function, generating a single truth value.

The transformed data are almost ready to receive implication method, but, before that, one must determine the rule weight. The weights are applied to the number given by the antecedent. In most cases, the weights are 1, thus not altering the implication process. Still, one can decrease the effect of a certain rule in relation to others through modification of the weight value to less or more than 1.

Weighted rules and transformed input can be given as input to the implication methods. The consequent consists of a fuzzy set. This is represented by a membership function, where the linguistic characteristics are properly weighted and attributed to it. Then, the consequent is reshaped using a function associated with the antecedent. Although the input of this process is a single number, the output is a fuzzy set.

The combination of outputs generated from the previous step is performed through aggregation. In this step, the fuzzy sets of each rule are combined into a single fuzzy set. This is done once for each output variable, before applying defuzzification. A fuzzy set is generated as output of the aggregation step, one for each output variable.

The final step consists of performing defuzzification of the aggregated output, generating a single number for each output variable. There are different ways of performing this process, calculating the centroid of the set being the most common. Other methods are bisector, middle of maximum, largest of maximum, etc.

3.6.3. Particle Swarm Optimisation (PSO) Combiner

PSO derives its name from the emulation of groups of birds randomly searching food in a region. Mathematically, the so-called food is the minima of the function of interest. The birds are the particles used to optimise parameters. As in nature, the birds (particles) do not know where exactly the food (minima) is in the area (function space). But, at each iteration, they are aware of how far food is. The strategy used by the flock is to follow the bird nearest to the food, thus moving in the function space as a group.

Each bird (particle) consists of a possible solution in the function space. As such, it has a cost function associate, which is related to the distance to the food (minima). Each particle moves in the function space with a different velocity pointing towards the direction and speed of it. The direction is indicated by the current optimum particles.

The algorithm is initialised by setting the particles randomly in the space and updating its position at each iteration. Such update considers two best values: the best solution achieved so far (or local best) and the best one achieved so far by any particle in the population (the global best).

At each iteration, the speed and position of the particles are updated according to the following equation:

$$v_p(i + 1) = v_p(i) + \left[c_i \times rand(0,1) \times \left(p_p^{best} - p_p(i) \right) \right] \tag{3.11}$$
$$+ \left[c_2 \times rand(0,1) \times \left(p^{gbest} - p_p(i) \right) \right]$$

where

$v_p(i + 1)$ – speed of particle p at iteration i+1

c_i – weighting coefficient of local (1) and global best (2)

$p_p(i)$ – position of particle p at iteration i

p_p^{best} – best particle position at iteration (local best)

p^{gbest} – global best particle position

$rand(0,1)$ - uniformly random generated value between 0 and 1

The position of each particle is updated as follows:

$$p(i + 1) = p(i) + v(i) \tag{3.12}$$

The algorithm for PSO is as follows.

- Initialise particles

- While error tolerance is not achieved:

- For each particle:

- Calculate Cost Function value C

- If C<pBest:

- set current value C as pBest

- Choose particles best positioned as the hBest.

- For each particle:

- Calculate particle velocity

- Calculate particle position

In this work, this optimisation algorithm is used to calibrate the parameters of the combiner model by minimising the accuracy metrics, i.e MSE. It is incorporated as shown in Figure 3-5.
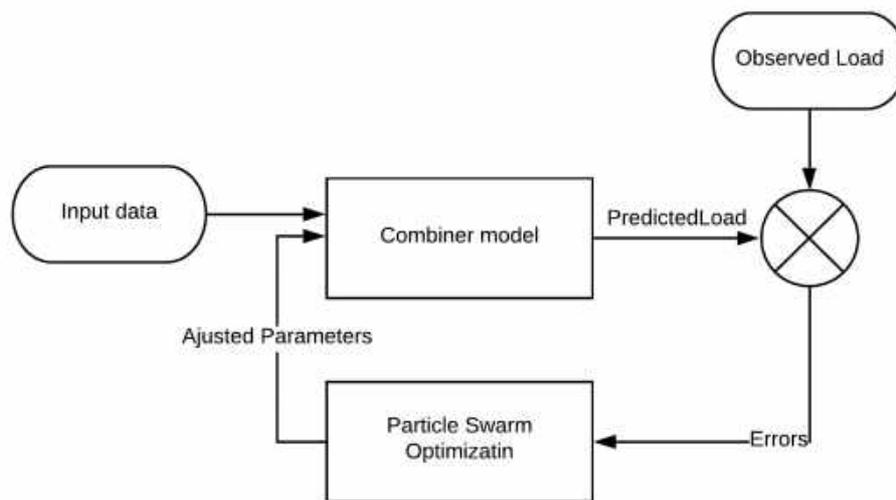
Figure 3-5 Incorporation of PSO to calibration of Combiner model.

3.7. Model Performance Analysis

A variety of metrics can be used to investigate model performance, that is, how accurate are the predictions issued by a certain modelling approach when compared with observed data. In this regard, two metrics were considered in the current research: Mean Squared Error (MSE) and correlation coefficient (R^2).

To obtain the MSE, the difference between each observed and predicted data point is calculated and squared. The final value is obtained by averaging the errors, i.e. dividing by the total number of data points. In summary:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_{obs} - y_{pred}\right)^2 \tag{3.13}$$

where:

n – number of observations (data points.);

y_{obs} – observed energy load;

y_{pred} – forecasted energy load;

The second metrics, correlation coefficient (R^2), is useful to investigate the level of correlation between the observed and the predicted energy load. It is obtained by first calculating the SSE (sum of squared errors) as follows:

$$SSE = \sum_{1}^{n}\left(y_{obs}^n - y_{pred}^n\right)^2 \tag{3.14}$$

Another relevant variable is the SST (sum of squared totals), which is calculated as:

$$SST = \sum_{1}^{n}(y_{obs}^n - \frac{1}{n}\sum_{1}^{j}y_{obs}^j) \tag{3.15}$$

with SSE and SST, the coefficient of determination R^2 can be obtained as follows.

$$R^2 = 1 - \frac{SSE}{SST} \tag{3.16}$$

The value of R^2 ranges from $-\infty$ to 1. An ideal model that exactly matches the observations with no error has R^2=1.0. Accurate models are usually reported to have $0.8 < R^2 < 1.0$. R^2=0.0 indicates that the regression line between observed and predicted is a flat, horizontal line. This means that the model is not able to capture the observed patterns, having the same accuracy as the use of average as a predictor.

3.8. Model Sensitivity Analysis

It is imperative to understand the effect that each predictor (input) to the model has on the output (energy load). Different analysis may be performed in this direction, such as Principal Component Analysis or Performance Error. In this investigation, Performance Error metrics are used to retrieve the input, which causes more deviations on the load forecasting. The metrics used are MSE applied to each isolate input. Higher values indicate higher sensitivity of the model to such input, while lower values indicate more robustness or less value variation, causing smaller deviations.

3.9. Power Network and Planning

In literal terms, power network and planning are about conducting analysis as to when, where, and what type of new transmission lines are needed to be built to supply the energy to consumers (Qui et al., 2017). Electricity network lines are then connected to the grid that controls the balance between supply and demand. Over the years, the increasing number of renewable energy distribution requires strategies for the operation and management of the electricity grid so that the grid system can be switched between different sources of energy systems. This means that new grid management systems would be needed. To this end, two innovative factors are driving the energy industry, the development of fast semi-conductor switches and real-time computer controllers (Mwasilu et al., 2014). While fast semi-conductor switches allow quick switching and handling of high power, real-time computer controllers can use advanced and complex control algorithms, hence, leading to application of cost-effective and grid friendly converters. The purpose of power network and planning in the context of this research is twofold. Firstly, it is about the ability of the grid system to convert the source of energy from fossil fuels, such as gas generation units, to renewable energy. Secondly, power and planning infrastructure will be needed to create a balance between demand of electricity and creating a balance with the forecasted electricity and renewable energy source.

3.10. Summary

This chapter discussed the theory background and structure of the modelling techniques used for energy load forecasting and investigation. Two types of model were

presented: Linear Regression and Artificial Neural Network. The first has the advantage of possessing a simple structure with few parameters and model linearity, which makes for easier analysis and understanding of its behaviour. ANN, on the other hand, consists of a more advanced model structure, with nonlinear behaviour but with reportedly higher performance.

In energy load forecasting, there are two basic types of input data that may be used, previous load data (time series) or other variables, such as temperature, humidity, day of week, etc (multivariate data). The two models (linear regression and ANN) can be evaluated using two types of input. For analysis of potential improvement, model combiner technique is implemented. This consists of a higher stage model which can couple the output of multiple models, generating a third one, potentially better than the two given as input. Specifically, it combines the outputs of one model using timeseries as input and another model using multivariate data as input into one output. The parameters that weights each model output is calibrated using PSO and Fuzzy, in an effort to minimize the error between prediction and observation.

Lastly, the theory and methodology used for analysis of power network and planning was presented. This investigated the possibility of integrating renewable energy sources in Saudi Arabia's network, through the switch of other power plants, reducing greenhouse gas emission but still meeting demand.

Chapter 4.

Modelling and Simulations

4.1. Introduction

The promising benefits of the renewable sources based on distributed generation are pushing the future energy markets to invest more into the available renewable systems. This research focuses on integrating available renewable energy resources in the Kingdom of Saudi Arabia in the electric grid to minimise the energy production from fossil fuels through continuous prediction and forecast of demand requirements. The work will concentrate on linking modern forecasting techniques with load history, weather information and the grid, with the objective to minimise CO_2 emission over both short and long-term forecasting periods. A simulation model will be developed to validate the proposed solutions using MATLAB® software environment. With the aim of testing the linear module performance, historical load data from January 2010 to August 2016 were collected from the Saudi Arabia Electrical Company for the western part of the country.

This chapter will start by reviewing significant literature related to the integration of electricity produced by renewable energy sources (RES). Afterwards, discussion will move to evaluate the ways electricity produced by RES can be tapped into the electricity grid system based on load forecasting. To do this, discussion will highlight the significance of load forecasting and present electricity load data gathered from Saudi Electricity Company (SEC). This data will then be subjected to multivariate and time series statistical analyses that can be used to make decision regarding integrating RES into the electricity network. Discussion will conclude by evaluating the accuracy of time series and multivariate modelling techniques to forecast electricity load.

4.1.1. Load Forecasting

According to Phuangpornpitak and Tia (2013), integration of electricity produced through renewable energy into existing electricity transmission network is one of the key challenges. For the electricity grid system to accommodate electricity produced through

RES would mean alteration in backup and energy storage capability. Such alteration would be necessary to meet challenges of variations in the amount of power being generated through RES.

Additionally, energy from RES tends to be highly volatile due to its dependability on various exogenous factors, such as season, weather, ecological and environmental factors, sunlight and wind. It is due to inherent variability of the electricity production system through RES that makes integration of RES into electric grid system a challenging task. That is why the need for accurate load forecasting models, able to predict the power accurately, is emphasised.

The concept of load forecasting possesses a central importance in effective planning and operations of electricity utilisation across the globe (Alfares and Nazeeruddin, 2002). There are three types of time periods for which electricity load is forecasted: short-term load forecasting (STLF), medium-term load forecasting (MTLF) and long-term load forecasting (LTLF) (Srinivasan and Lee, 1995). Areas with challenging nature (such as the cities of Jeddah, Medina and Makkah) mean high variation in electricity demand, requiring the use of STLF modelling technique. Singh et al. (2014) highlighted the use of STLF technique for providing micro level information on electricity demand that can be used in the existing grid system to devise efficient cost and pricing strategies. From the cost perspective, STLF could assist in controlling operations and decisions of dispatching, unit commitment and resources allocation needed to produce electricity, be it renewable or other sources. Hobbs et al. (1999) investigated the dollar value of accurate load forecasting through STFL modelling technique and concluded that 1% reduction in the error of electricity forecast could result in savings of hundreds of thousands or even millions of dollars.

4.2. Case Description and Data Collection

The need for accurate forecasting of load for the western region of Saudi Arabia lies in the fact that this area holds the economic, social and religious hubs of the country. Saudi Electricity Company (SEC) is in charge of providing reliable and consistent electricity to the region. Figure 4-1 presents the western operational network (WOA) of the SEC covering several cities such as Makkah, Medina, Jeddah, Taif and Yanbu, as shown in Figure 4-1.
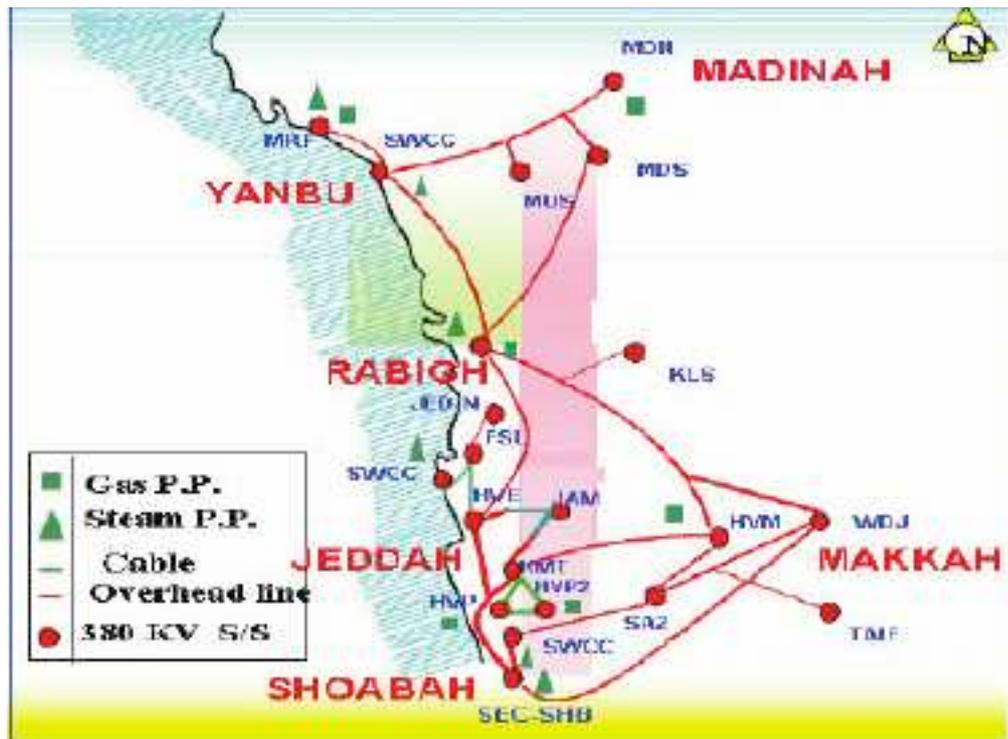
Figure 4-1 - Network of western operational area of SEC. Source: Alshareef et al., 2008.

The data used for the current project consist of historical records from January 2010 to August 2016 received from the SEC. Data contained various parameters such as date, month, year, time (hourly), day of week (Monday, Tuesday, etc.) temperature, humidity and load. Data are hourly-based in which peak period, starting from 13:00 till 15:00 is presented in half hourly periods. There are two half hourly slots. The first one lies in the 13:30, 14:30 and 15:30. The next half hourly slot starts from 18:00 until 20:00, hence presented as 18:30, 19:00, 19:30, 20:00 and 20:30. The reason for presenting evening data in half hourly slot lies in the fact that evening time is the time when everyone is at home and usage of air conditioning is possibly at maximum level. Initial analysis of the data resulted in identifying various patterns.

Figure 4-2 presents load pattern in MW for the first day of the months January, June and December during 2015. It can be seen that, while load is higher for the month of June (compared to January and December 2015), in every month load pattern is similar. For instance, temperatures at early morning and late evening have similar values, rising along the days. Maximum values may be observed during the day (from 10:00 to 18:00), after

which it decreases, reaching a steady value. Hourly difference in the load can be linked to the domestic and commercial activities. For instance, usage of air conditioning tends to be low during early hour time before children and adults go to school and office, respectively. Moreover, at the time of noon prayers (between 12:00 to 13:00) load tends to reach its maximum level until 18:00, starting to decline as the sun starts to set in the evening.
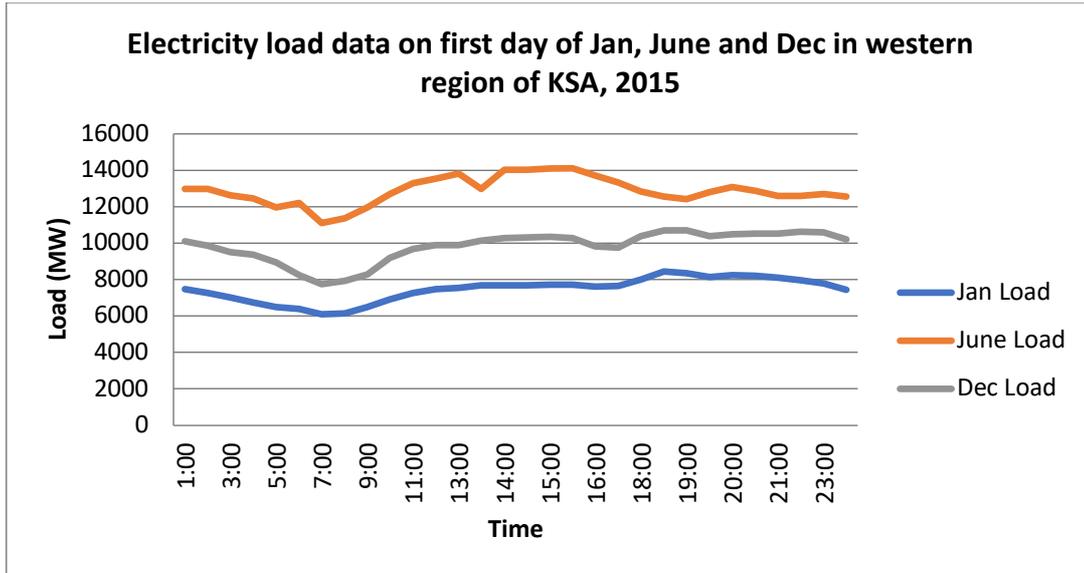
Figure 4-2 - Electricity load in the first week of the month

Figure 4-3 presents the peak load in MW for the year 2015. It is visible that there is a difference between electrical load consumption during summer and winter period. Key source of the difference lies in heavy usage of home appliances, such as air conditioning, at both domestic and commercial level during summertime. The energy consumption varies, according to the day and time of the day, and is strongly correlated with aspects such as: work and school hours, prayer times, weekends, special event and holidays. Hence, the load is influenced by people's lifestyles, their behaviour and other factors, such as the social life.
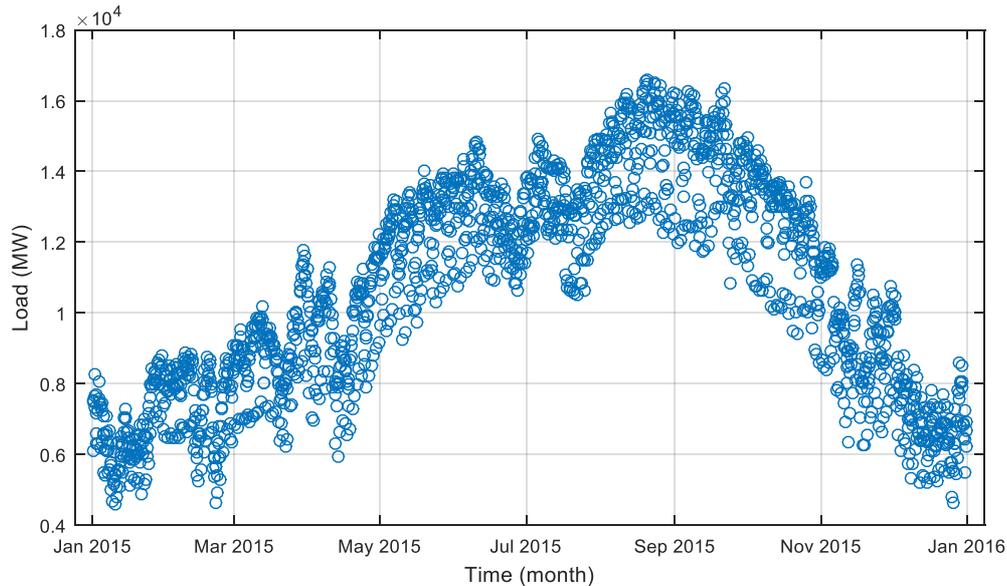
Figure 4-3 - Load profile for the year 2015.

Figure 4-4 presents average temperature for the Western region of Saudi Arabia in degrees Celsius (^{0}C) for the period of January to December 2015. It can be seen that highest levels during the summertime may reach more than 45^{0}C and less than 14^{0}C during wintertime. This behaviour can be justified by the influence of maritime effects and the presence of mountains surrounding, which minimises the desert effect.

As already mentioned, it is important to notice that higher temperatures (above 25° C) cause a drop in the efficiency of the PV modules, even under full sunshine. When the PV energy production diminishes, then more is required from wind power.

Another observation is that, though the highest temperatures are observable in August, highest load demand appears in September. This may be attributed to the effect of holidays, in which people use to travel during summer time and come back to work/ school by September, thus consuming more energy even though the highest peak of temperature has passed.

46
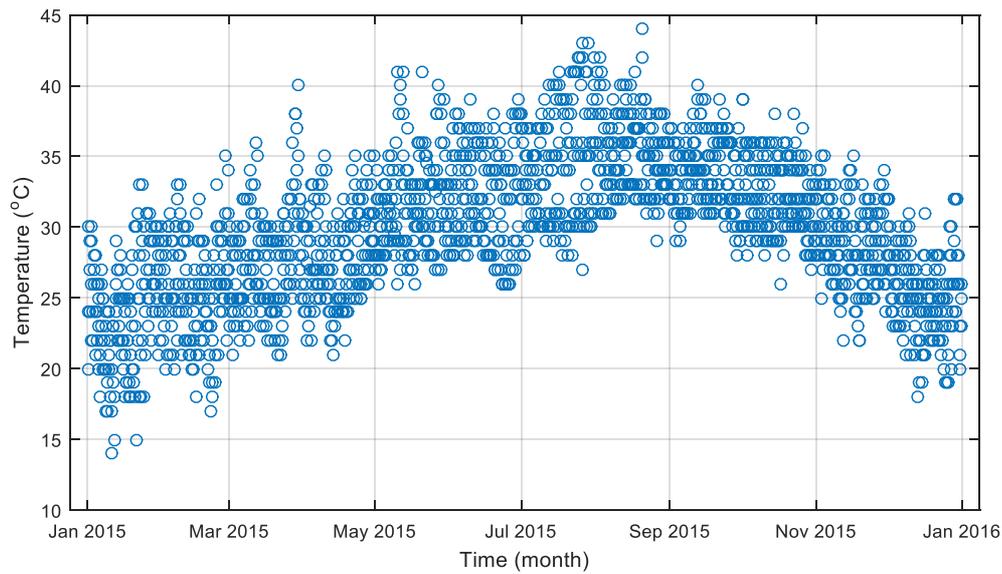
Figure 4-4 - Temperature profile training data for the year 2015 in Saudi Arabia.

Humidity levels are illustrated in Figure 4-5 in percentage levels (%) from January to December 2015. It can be seen that humidity levels span wide ranges along the year, without a clear pattern for seasons. There is just an almost imperceptible drop in humidity during the summer time, which can be inferred from the lack of high humidity (above 80%) for July and August. This is in accordance with the temperature, which is highest during this time of the year. Again, the proximity to the sea plays a key role on the humidity of the region, making it completely different from a desert climate, which would be more representative in the Kingdom of Saudi Arabia.

47
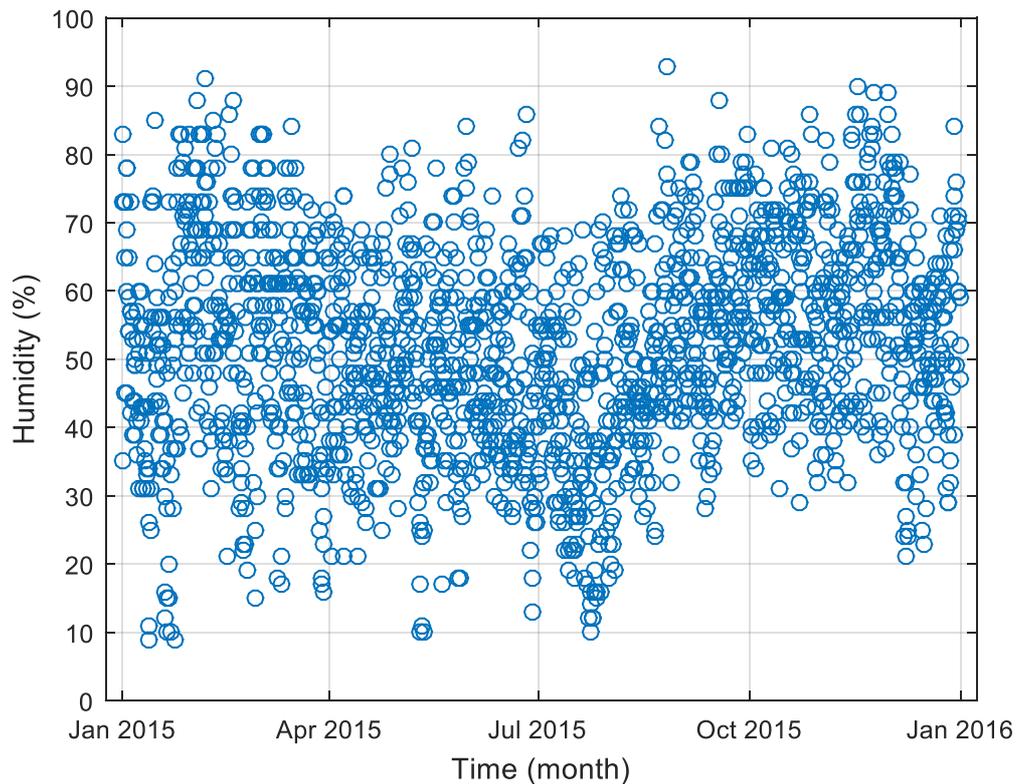
Figure 4-5 - Humidity profile during 2015 in Saudi Arabia.

It is clearly visible from the above figures that both the temperature and humidity level are the highest during the summertime, namely, from April to August. This pattern maintains unchangeable in other years, which have been taken under consideration in the obtained data.

4.2.1. Data Split: Training and Testing.

Since load, temperature and humidity follow similar daily patterns, data were divided yearly. The data available from 2010 to 2015 were used for training while data available for 2016 (until August) were used for testing.

4.3. Modelling methodology

The developed models for forecasting the load conditions are based on two approaches: multivariate (MV) and time series (TS) prediction. The modelling is done using artificial neural networks (ANN) compared to linear regression (LR) models. The time series model is based on the historical load conditions, while the multivariate model is based on

generating a model as a function of the weather conditions, month, day type and time. The multivariate model consists of eight inputs and one output, while the time series model is based on six-day history, and one-day-ahead prediction. The regression models are fitted using MATLAB® toolbox, while the neural network model multivariate model is based on a topology of one hidden layer with 15 hidden neurons, eight inputs and a single output neuron. The time series model is based on one hidden layer with 15 neurons, six inputs and one output.

4.3.1. Multivariate Modelling

This technique is one of the most commonly used predictive modelling techniques (Faul et al., 2009). Its aim is to model relationship between one dependent variable and multiple independent ones through fitting a linear equation into experimental data. Such approach, called multivariate or multivariate modelling, differs from the univariate, where one variable tends to be explanatory while the other is the dependent variable. The key problem with univariate linear regression is that it uses only one independent variable, hence, would not fit into the context of the present study considering the fact that electric load forecasting requires integration of multiple independent variables, and, therefore, requires a multiple regression approach also known as multivariate regression Menge et al. (1977) and Panagiotelis and Smith (2008) lauded the use of multivariate modelling technique for electricity load forecasting.

Multivariate linear regression model is a load forecasting approach that serves scenarios where there is more than one input variable (Lang et al., 2018). The model is able to analyse complex data through its ability of identifying relationships between the single dependent variable and multi independent variables. For the purpose of energy regression modelling, Fumo and Biswas (2015) differentiated between two types of multivariate modelling techniques: top-down and bottom-up approach. In this paper top-down approach will be used because it identifies factors that define energy consumptions, such as energy demand, environmental factors, such as weather and wind condition, and behavioural factors, such as day, time and special occasions that include religious holidays, Jumma prayers time, etc. in the context of Saudi Arabia. Bottom-up approach does not suit the present study because such approach is suitable for microanalysis of energy consumption,

such as house type and archetype at particular geographic regions (Amasayli and El-Gohary, 2018).

4.3.2. Time Series Prediction Model

Drawing on Taylor and McSharry (2007), the time series model naturally fits into trajectory of load forecasting. In literal terms, time series is a sequence of vectors that relies on historical data to predict possible future outcome. The efficacy of using the time series model for load forecasting relies on the fact that electricity load tends to be strongly linked to seasonal trends, thus making it possible to use supporting modelling for nonlinear condition (Khashei and Bijari, 2010). Furthermore, the time series model has enjoyed recognition in literature for its ability to accurately predict future electricity load. The model possesses capability of dealing with large sequences of numeric data gathered through repeated number of times based on identical intervals, such as every hour, day, week or month. The ability of the time series model to accurately predict load allows decision-makers to efficiently deal with the complexity of integrating renewable energy sources, such as solar and wind power, into the grid system (Barta et al., 2016; Panapakidis and Dagoumas, 2016).

4.3.3. ANN Configuration of Input Layer

The input vector in this research follows a similar pattern to that used by Al-Shareef and Al-Judaibi (2008) who investigated the STLF predictability using ANN in the western region of Saudi Arabia. Typically, input vector is configured from month, year, hour, type of date, temperature, relative humidity and historical load for the previous one hour, two hours or day. Table 4.1 provides numerical indexes representing inputs for the forecasted hour under both ANN and multivariate modelling.

Table 4.1: Configuration of input vector modelling in multivariate model.

Input	Description
Date	1-31 depending on Gregorian calendar month
Month	1-12 (Gregorian calendar)
Year	January 2010 to August 2016
Day type	Weekdays or weekend

Time of the day	1-24 hours with half hourly data configuration to accommodate high load variations. Half hourly configuration was inserted for the hours 13:30, 14:30, 15:30, 18:30, 19:30 and 20:30. Half hourly configuration was presented in fractions such as 13.4, 14.5, 15.5, 18.5, 19.5 and 20.5.
Humidity	Humidity for the period specified above
Temperature	Temperature for the period specified above
Load	Hourly and half hourly reading

For one-day-ahead prediction, up to one previous day energy load timesteps were used to structure the input layer of the models.

4.3.4. ANN Hidden Layer

For the purpose of setting up the simulation model in this study, one hidden layer is proposed. The reason lies in the fact that ANN with more than one hidden layer tends to become complex and time-consuming (Sietsma and Dow, 1991). Furthermore, the number of neurons in hidden layers tends to have greatest effect on the performance of network since it possesses direct implications for training of neurons in the network (Murata et al., 1994). For instance, increase in the number of hidden neurons can lead to overparametrisation of the model, which can cause overfitting, i.e., the model captures white-noise in the training data, thus, not performing well on the testing set. On the other hand, ANN with no hidden layer may lack capability of capturing nonlinearities on the data, resembling linear models. The number of neurons chosen at the hidden layer was chosen by investigating the one that produced smallest errors. Specifically, 15 neurons were chosen in a hidden layer.

4.4. ANN Training Method

Supervised learning method is used since it matches with the criteria of backpropagation learning algorithm and multi-layer perceptron. The default MATLAB® algorithm was used, namely Levenberg-Marquadt. Using the approach mentioned, error correction rules are needed to be applied because, during the training process, actual output generated by the network may not be equal to the desired output. The main principle of error

correction in the network is to use error signal that modifies connection weights to reduce the error. Error signs are determined by the difference between actual output and desired output. Setting up of error correction rate is important to avoid the overfitting problem associated with training (Rani and Dalal, 2016).

A technique used by MATLAB® Neural Network Toolbox to avoid overfitting is cross-validation. It consists of constantly checking a separate dataset (validation set) for the errors during training. Overfitting is identified by an increase of the error on the validation set. That is when the training halts to avoid such problem.

Data pre-processing is required before feeding it into the neural network. For that purpose, all inputs are rescaled in between -1 and +1. Additionally, outputs are rescaled from -1 and +1 to the original range before being analysed.

In multi-hidden layer ANNs, connecting neurons with adjacent layers by weight could be problematic due to the involvement of nonlinear function between different variables. For that reason, the ANNs developed throughout the present work were limited to a single hidden layer, followed by a linear output layer.

4.5. Results and Comparison

The following section presents the results obtained using LR and ANN for energy load forecast. Both models used multivariate and time series input data. The performance was evaluated using mean squared error (MSE). Identification of data dispersion was investigated through population distribution analysis. One indicator of population spreading is standard deviation, a value that assists in identifying the extent to which data values of different variables (identified in Table 1 above) are spread around the MSE range. In the event that standard deviation is close to the MSE, then it would mean high probability of predicting with the same accuracy as average and higher volatility in the data in the opposite scenario (Leys et al., 2013). Using weekly and daily data, standard deviation of load forecasting error has been calculated.

Different ANN architectures were evaluated for investigation:

- 1 hidden layer, 10 neurons;

- 1 hidden layer, 15 neurons;

- 1 hidden layer, 20 neurons;

- 2 hidden layers, 20 and 10 neurons;

- 2 hidden layers, 20 and 15 neurons.

Table 4.2 presents results of the analysis of regression and ANN for energy load forecasting from 2010-2016. Data have been chosen as 90% for training and 10% for testing for ANN. Linear regression model showed higher MSE values for the entire 2010-2016 period, indicating that ANN is a better modelling approach to integrate RES into the existing grid system in the western region of Saudi Arabia.

Table 4.2 also shows that combined analysis of all loads results during 2010-2016 presented lowest errors on single hidden layer such as Hidden Layer 10 and Hidden layer 15 compared with the architectures containing two hidden layers (20-10 and 20-15). Thus, single hidden layer results would be preferable. Finally, when comparing data between ANN with single hidden layer, such as 10, 15 and 20, it can be seen that the one with 15 neurons in the hidden layer presents the smallest error on both average and standard deviation, thus, it will be chosen for time series modelling for forecasting the load for one day ahead to integrate RES into the 380 kV network grid system in the western region of Saudi Arabia.

Table 4.2 - Summary of performance results and standard deviation (STD) for linear regression and ANNs with different architectures – one hidden layer (10, 15 and 20 neurons) and two hidden layers (20:10 and 20:15 neurons).

Year	Linear Regression		Neural network																				
	Perf (MSE) MW^2.10^5		Hidden layer 10				Hidden layer 15				Hidden layer 20				Hidden layer 20:10				Hidden layer 20:15			
			Perf (MSE) MW^2.10^5		STD		Perf (MSE) MW^2.10^5		STD		Perf (MSE) MW^2.10^5		STD		Perf (MSE) MW^2.10^5		STD		Perf (MSE) MW^2.10^5		STD	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
2010	5.6	2.4	2.2	5.4	8.1	6.5	1.8	8.4	5.7	6.0	1.6	9.6	4.8	1.5	8.4	1.2	1.4	8.6	1.1	1.3	2.1	4.2
2011	1.2	4.2	2.4	2.2	3.9	1.1	1.9	8.6	2.0	3.0	1.8	1.3	6.5	5.3	1.3	3.7	2.2	1.8	1.0	3.5	1.0	1.0
2011	1.4	1.1	2.9	3.2	2.8	4.6	2.7	3.3	7.0	6.2	2.4	2.6	1.1	1.0	3.1	4.7	1.0	1.6	1.1	3.5	8.2	2.5
2013	1.8	3.8	3.1	2.9	9.0	3.9	3.2	1.9	5.1	4.4	2.1	2.9	5.5	4.1	1.4	2.1	1.8	7.5	1.4	2.5	8.2	2.4
2014	1.4	4.6	2.9	5.2	5.5	1.1	2.5	5.2	1.6	8.2	2.2	4.2	1.1	8.7	1.3	7.1	2.3	2.7	1.1	3.6	8.5	8.9
2015	1.9	5.3	3.4	6.6	1.4	1.7	2.6	5.9	7.6	8.6	2.3	1.4	1.2	5.2	1.3	1.4	1.3	3.0	2.3	1.4	1.2	5.2
2016	1.0	2.0	3.2	--	2.0	--	2.4	--	1.8	---	1.3	--	1.1	--	2.1	--	1.5	--	1.3	--	1.1	-
ALL	4.4	7.0	2.2	3.9	2.0	6.3	**3.5**	**3.7**	**4.6**	**1.2**	3.2	3.9	8.5	1.1	2.5	5.7	1.6	3.2	2.1	6.0	2.3	5.2

The best ANN architecture was found to be one hidden layer with 15 neurons. In general, it can be seen that ANN outperformed linear regression, an expected result since ANN has many more degrees of freedom.

4.5.1. Results for LR with Multivariate Input

Figure 4-6 shows the energy load observed and predicted using LR with multivariate input for data available in 2016.
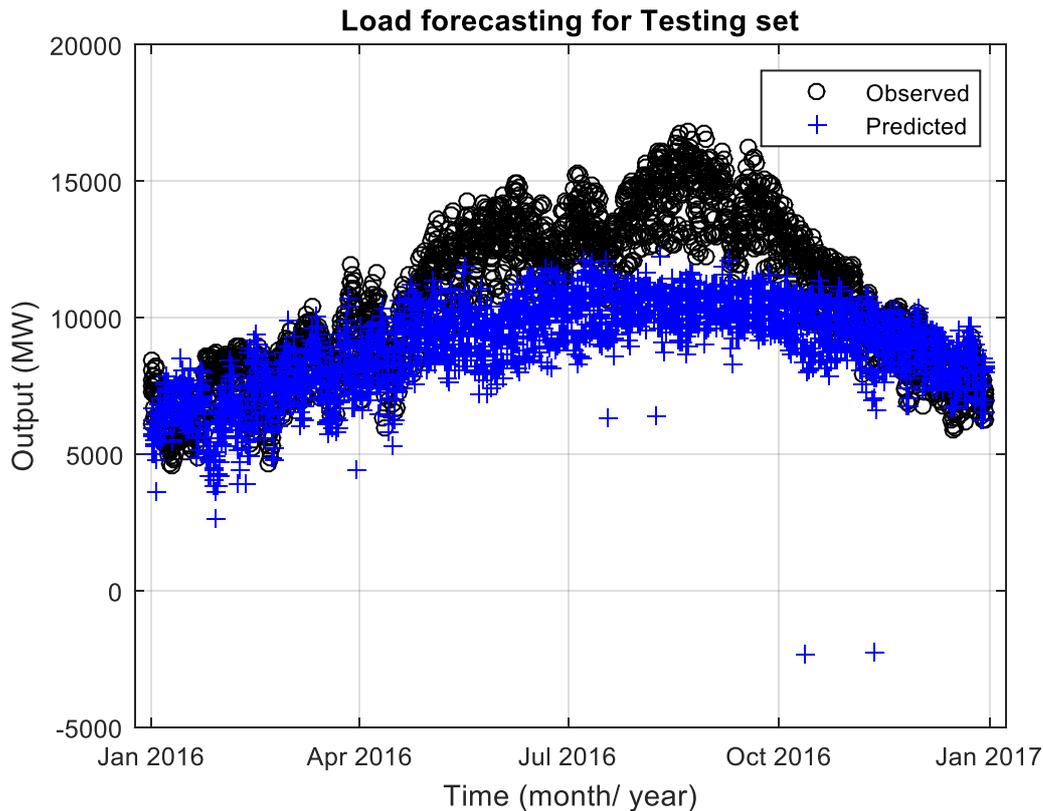
Figure 4-6 – Energy load forecasting for Testing set. Data are resampled for better visualisation.

The LR model shows reasonable agreement of energy load prediction on periods of low demand. These periods are winter season (beginning of the year at the left side of the plot) and when high usage of energy decreases by the end of summer (right side of the plot). However, there is an underestimation of energy peaks, especially at high demand periods, starting from the end of springtime to the middle of summertime. Such underestimation of the load can be observed in the middle of the plot, when LR predicts average loads of approximately 10,000 MW while the observed load is up to 15,000 MW.

Additionally, there is the presence of outliers which adds to the error on the model accuracy. These outliers can be seen around datapoints 6,500 and 8,500.

To complement the analysis of such model, Figure 4-7 shows the Observed x Predicted plot for this same model, in training and testing periods.
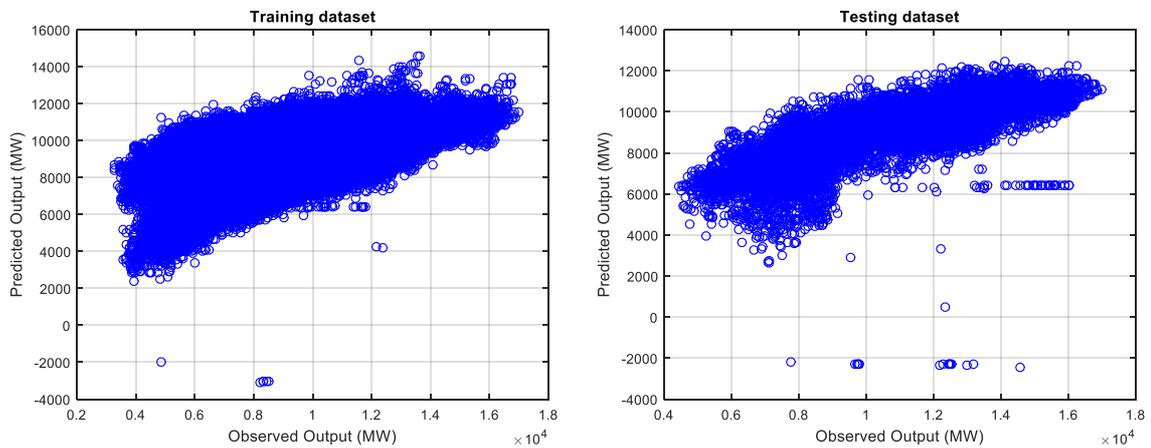
Figure 4-7 – LR with MV input – Observed x Predicted plot.

The Observed x Predicted plot of the Testing data makes even clearer the presence of outliers, which causes the underestimation of load. Some observed outputs in the range of 0.8, 1, 1.2, 1.4 and 1.6 MW have a predicted negative value. This may have been caused by the presence of values in the multivariate data at testing set which were not captured in the training set. A second possibility is that this pattern exists in the training pattern, and such possibility is reasonable since some outliers can also be observed in the Observed x Predicted plot of the training set (figure to the left).

4.5.2. Results for LR with Time Series Input

Figure 4-8 shows the observed and predicted energy load for the year 2016 (testing set) using LR with time series data. The results consist on one-day-ahead prediction, without recursion.
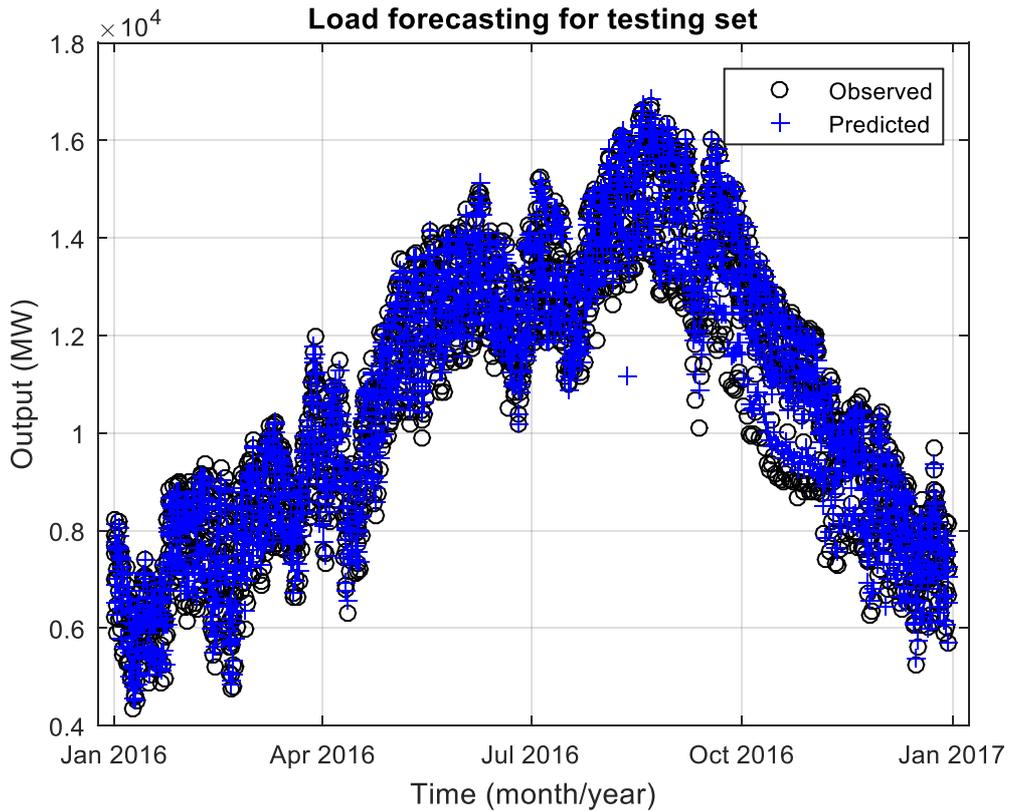
Figure 4-8 – Energy load forecasting for Testing set.

Different from the LR using multivariate data as input, LR with time series shows high accuracy and could successfully capture the patterns of energy load in one-day-ahead step. The accuracy seen is high for both wintertime and season time, indicating that this model can be used in any time of the year for such prediction. There is almost no presence of outliers. Such inference can be confirmed by Figure 4-9, which shows the Observed versus Predicted plot for training and testing sets. This indicates that the previous values of load are better correlated with the current load than this one with the exogenous variables (time of week, temperature, humidity). By looking at the previous 7 values of load, the model was able to predict the next load assuming a linear relationship.
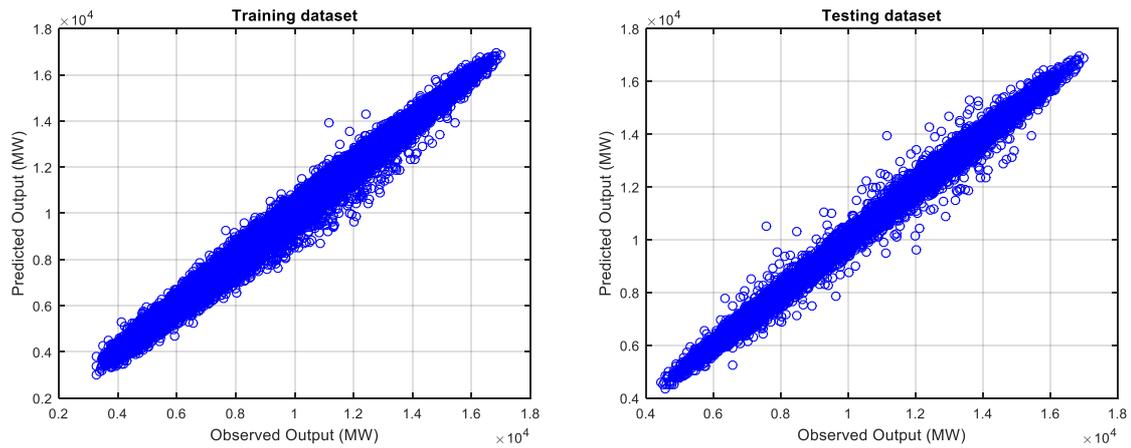
Figure 4-9 – LR with TS data - Observed x Predicted plot.

High correlation can be observed in both plots above, due to the high proximity of the points to the diagonal. It becomes even more difficult to notice the difference in performance between the training and testing datasets just by observation of the plot above.

Such results for LR show that, between multivariate input and time series input, the latter has shown to be the most appropriate when performing one-day-ahead energy load forecasting for the current scenario (western region of Saudi Arabia).

4.5.3. Results for ANN with Multivariate Input

ANN requires iterative training before being investigated. Figure 4.10 shows the evolution of training and validation performance for the ANN using multivariate input.
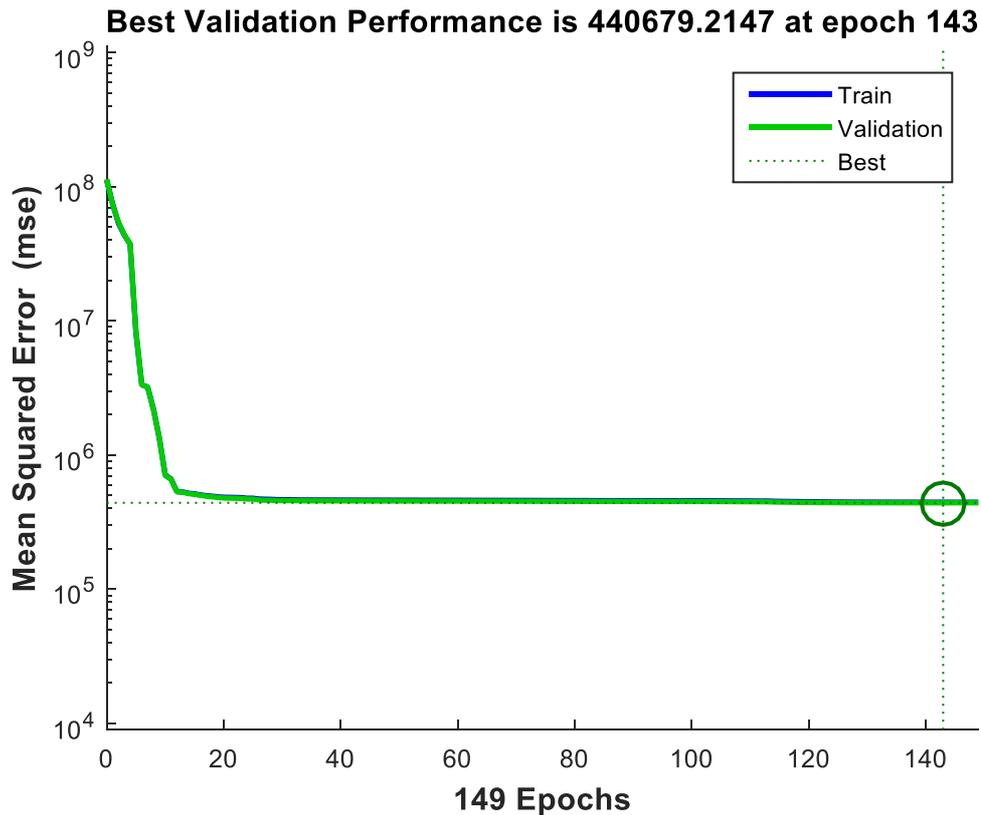
Figure 4-10 - Training evolution for ANN with multivariate input.

The error falls drastically at the initial iterations, due to the random nature of the initialisation of the weights. Also, there is an overlapping of the training and validation curves, which is only perceptible as a "shadow" behind the curve. After a certain point it reaches a plateau (flat line), indicating that the model has converged to a potential minimum of the cost function, so it stops reducing. Confirmation of this is acquired by letting the process iterate multiple times without leaving the plateau. Therefore, the convergence of this training indicates that it has successfully found a minimum value and that the neural network is ready to be evaluated on a testing set.

A comparison between the observed and predicted loads for data available in 2016 is shown in Figure 4-11.
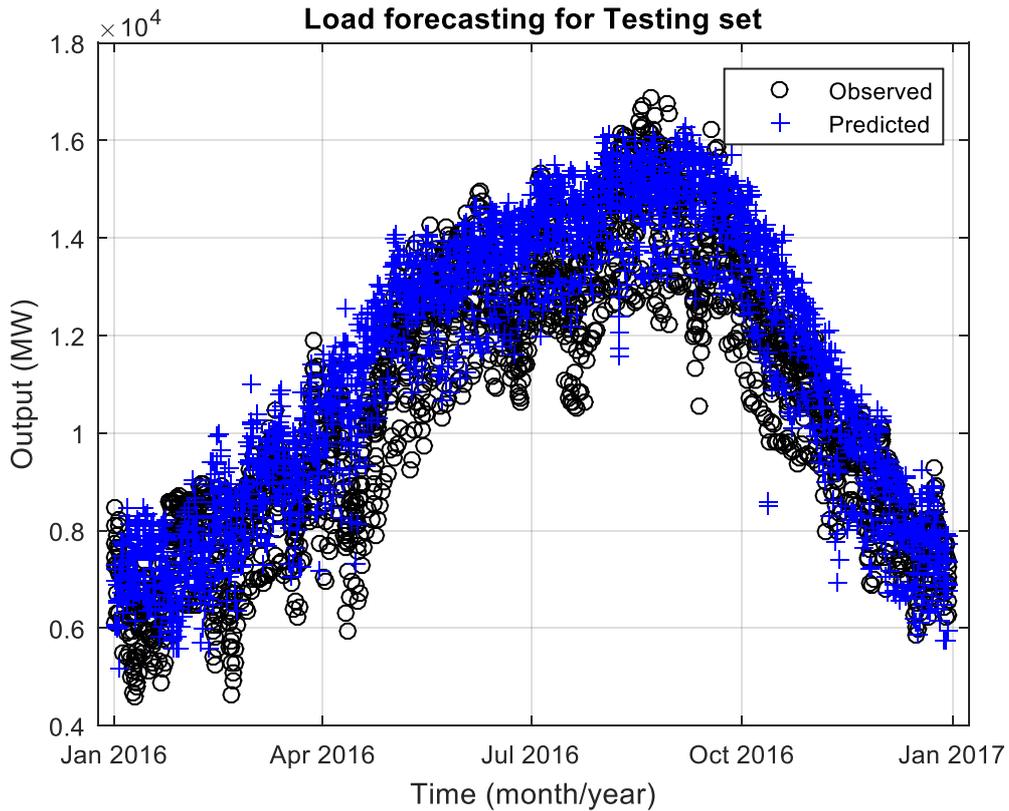
Figure 4-11 – ANN with MV input. Energy load forecasting for Testing set.

The model captures the expected energy load with reasonable agreement all over the year. However, there is a slight overestimation of the lower peaks lying on the lower bound of the observed load window. Specifically, there are clear moments where the load drastically falls during the whole year. The model is not able to see such pattern, still it mimics the general tendency of the data, i.e. low values at winter time and higher values at summer, slightly increasing in the interstice period. It is noticeable the presence of a few predictions, specially in the region of higher peaks, with some underestimation of predicted, load. This becomes more present after the peak, with a few points representing underestimation of load values.

This indicates that the neural network model requires additional information or data to be able to represent the complete behaviour of the load rather than the one present only on the multivariate data.

The reasonable agreement between model and data can be noted in Figure 4-12, which shows the Observed x Predicted plot for training and testing set.
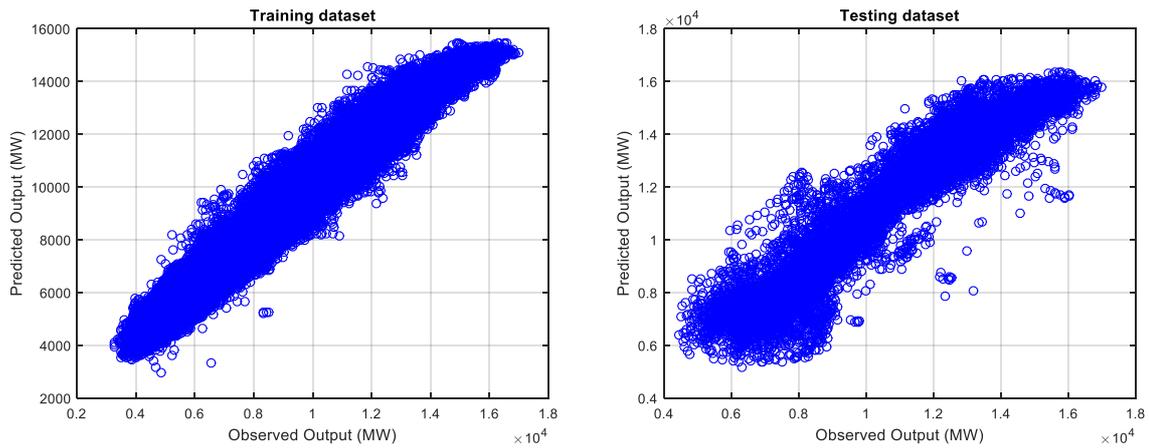
Figure 4-12 – ANN with MV input: Observed x Predicted plot.

The Observed x Predicted plot for the testing set enables the identification of a few outliers, which drastically reduces the expected performance of such modelling technique. Also, it is noticeable the fall in performance from the training to the testing set, since the training points on the left are clearly aligned with the diagonal, while the testing points on the right are much less aligned, with a small curvature on higher values of load.

4.5.4. Results for ANN with Time Series Input

Model performance evolution during training is illustrated in Figure 4-13, for ANN using time series data as input.

Best Validation Performance is 47057.4662 at epoch 102
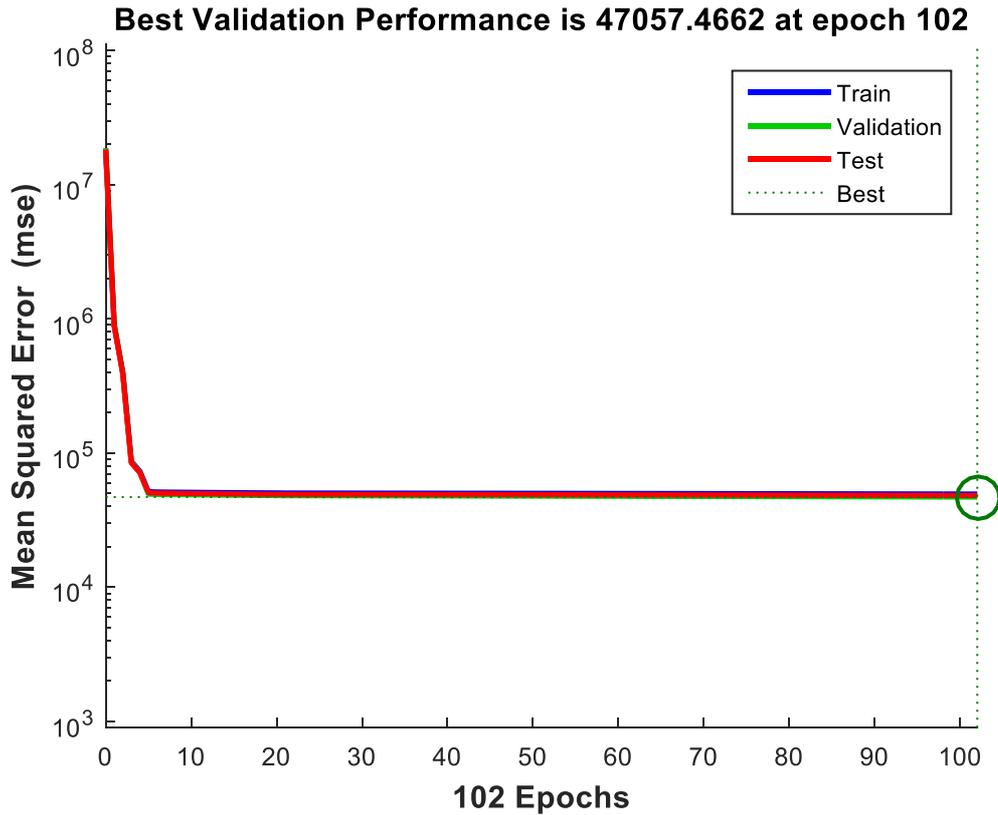
Figure 4-13 - ANN with TS input: Training performance per iteration.

As observed in the previous section, the neural network performance converges to a minimum, confirmed by the plateau clearly seen in the training performance results at each iteration after the first 10 epochs. Again, there is an overlapping of the curves, creating a "shadow" behind the curve on the top (Test).

The energy load predicted for year 2016, compared with observed data for the same period, is seen in Figure 4-14.
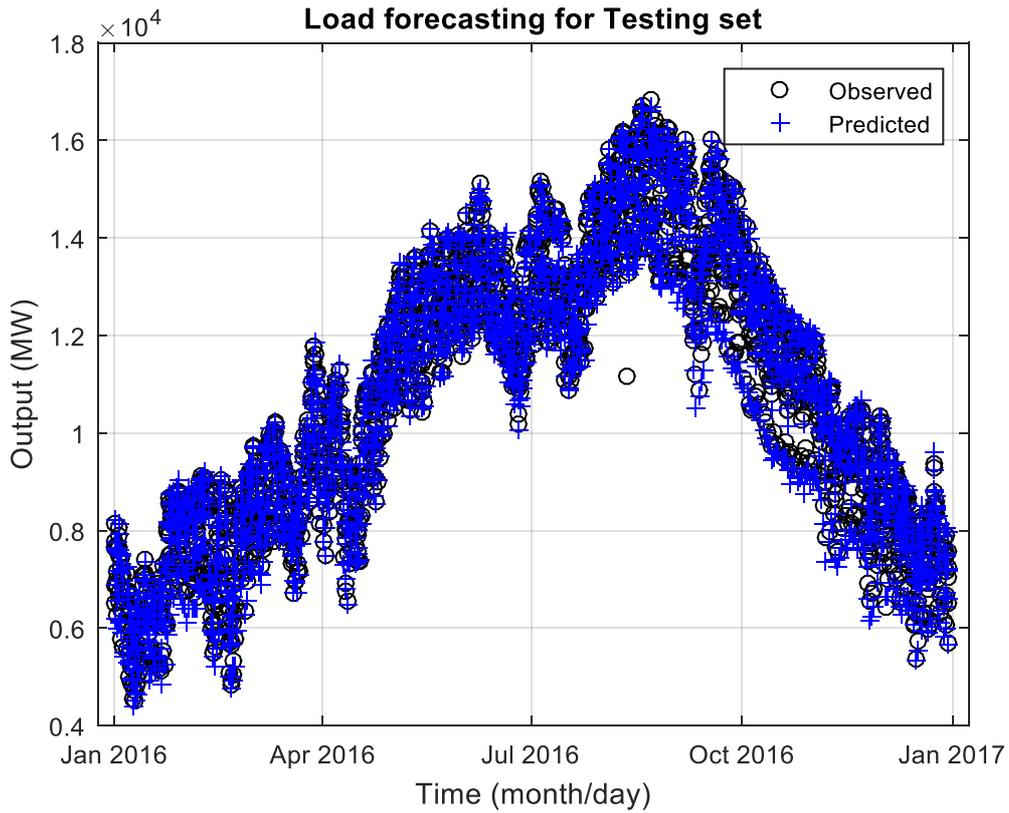
Figure 4-14 - ANN with TS input – Energy load forecasting for Testing set.

The model showed relatively high accuracy, being able to capture patterns of energy load in the long and short-term. In the long-term, the model mimics the decrease of energy load in the winter season as well as the constant increase during mid-seasons, reaching superior values during the mid of summertime. In the short-term, it successfully mimics the short rise and fall in-between data points. The model even captures the well-behaved variation of energy load just after mid-summer, with high usage on one day and lower usage on another.

The high performance of this modelling approach is confirmed in the Observed versus Predicted plot shown in Figure 4-15.
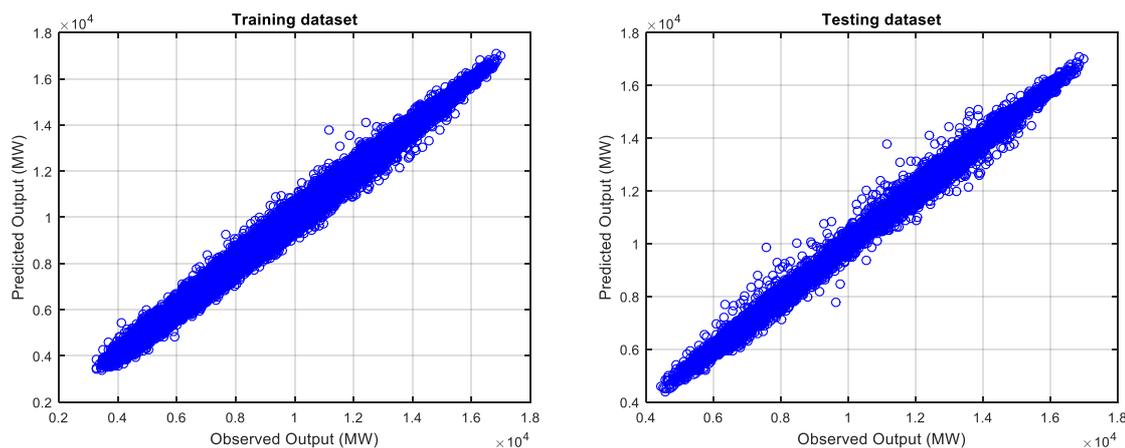
Figure 4-15 - ANN with TS input – Observed x Predicted plot.

As already expected, most points are well-aligned in the neighbouring area of the plot diagonal, with very low spreading. This can be inferred from the training set, as well as from the testing set.

4.5.5. Comparison Between Results

The following table summarises model performance for all the approaches mentioned in the previous sections. The main performance metrics used is MSE (mean squared error), with R² being an auxiliary metrics, which indicates the correlation between observed and predicted data. MSE gives a general sense of the magnitude of the errors, while the R^2 indicates the correlation (or perfectness) of the model, being 100% a perfect model (no errors) while 0% indicates that the average value is a predictor as good as the model.

Table 4.3 - Performance (MSE) for time series and multivariate modelling approach. Values shown in parenthesis indicate R² (%).

Model	Time series		Multivariate	
	Training MSE ($MW^2 \times 10^4$)	Testing MSE ($MW^2 \times 10^4$)	Training MSE ($MW^2 \times 10^4$)	Testing MSE ($MW^2 \times 10^4$)
Regression	6.19 (99%)	7.51 (99%)	442 (40%)	670 (14%)
FFNN	**4.27** **(99%)**	**5.55** (99%)	37.4 (95%)	142 (82%)

Table 4.3 clearly shows the superior performance of the time series approach for one-day-ahead load forecasting. Correlation for both LR and ANN was above 99%, with MSE as low as 4.27×10^4 for ANN with 15 neurons in hidden layer. This error indicates an average deviation of approximately 5% on the energy load estimation. Using LR with time series, the deviation raises to approximately 7%, which is still relatively low.

Though lower, the performance of ANN using multivariate input is still useful, with approximately 82% correlation on the testing set. The errors average to approximately 20 times greater than the time series input. The multivariate approach is not usable for LR, since the performance is much lower than 50%, with errors higher than 100% on the estimation of energy load.

4.6. Results for One-Week Ahead Load Forecasting

The results obtained for one day ahead forecasting motivated the investigation of model performance for prediction of one-week-ahead load. Figure 4-16 shows the results obtained for training and testing datasets using artificial neural network with 15 neurons in one hidden layer and time series input. Error minimisation is accomplished via Levenberg-Marquadt algorithm.
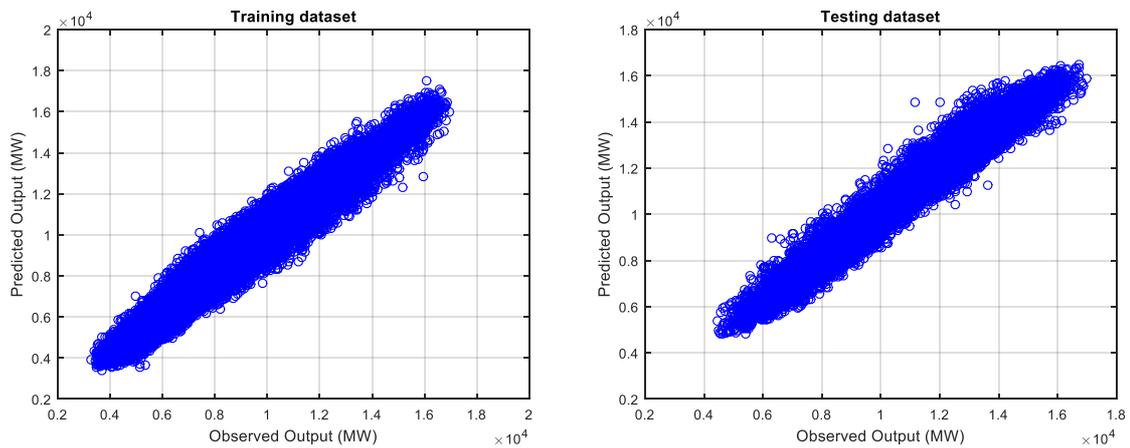
Figure 4-16 – Predicted versus Observed plot in training and testing dataset for linear regression using time series input

Figure 4.16 shows that the neural network is able to model the load forecast with high accuracy. In fact, the coefficient of determination (R^2) obtained for the training set is 0.997, with similar value obtained for the testing dataset. This level of correlation can be also observed from the relatively low spreading of the data.

From the results, it can also be inferred that such high performance may be attributed to high cross-correlation in the load forecasting data. That is probably because doing one-week-ahead forecast means using the same weekday as the last predictor to the next weekday in the time series.

In summary, the performance metrics for one-week-ahead forecast using ANN with 15 neurons with one hidden layer are:

Mean Squared Error (Training) = $1.7 \times 10^5 \, \text{MW}^2$

MSE Standard Deviation (Training) = $707.2 \, \text{MW}^2$

Mean Squared Error (Testing) = $6.5 \times 10^5 \, \text{MW}^2$

MSE Standard Deviation (Testing) = $9.6 \times 10^3 \, \text{MW}^2$

From the results above, it can be stated that the errors obtained were relatively small since the time range is considerably greater when compared with one-day-ahead forecast. However, a relevant observation is that the standard deviation of the performance on the testing set is reasonably greater than on the training set, indicating that the performance varies more on data sets which have not been used for training. Still, that is expected from an artificial intelligence model.

4.7. Summary

Discussion in this chapter evaluated applicability of multivariate and time series data as input. The two models evaluated were linear regression and artificial neural network (ANN). Data were collected from Saudi Electricity Company (SEC). For multivariate data, factors included date, month, year, day type, time of the day, humidity, temperature and load. For time series data, the last 7 days of load were used a predictor of the 8th day. Data range was set at annually from 2010 to 2016. The data available from 2010 to 2015 were used for training while data available for 2016 (until August) were used for testing.

The number of hidden layers in the neural network was fixed to one because in comparison with double hidden layer, single hidden layer has been identified to provide better performance in load forecasting. Moreover, supervised method was used to train data conducted through using backpropagation algorithm. The optimization algorithm used was Levenberg-Marquadt (MATLAB® default) with cross-validation checking. The maximum number of iterations were fixed to 1000.

Results revealed that, when comparing the four combinations (LR with TS, LR with MV, ANN with TS and ANN with MV), ANN with TS input showed superior performance (MSE $5.55*10^5$ or R^2 99%), followed by LR with TS (MSE $7.51*10^5$ or R^2 99%). Using multivariate data as input showed to issue non-reliable results, with low performance for both ANN (MSE $142*10^5$ or R^2 82%) and LR (MSE $670*10^5$ or R^2 14%). On the other hand, the use of time series input showed high accuracy and correlation, indicating that such approach has potential to be used as a prediction tool in Saudi Arabia.

Chapter 5.

Hybrid Prediction Modelling and Simulation

5.1. Introduction

Consistent supply of electricity is not only vital for living beings, but also for the economic development of a country (Wang et al., 2017). Particularly, in the wake of factors such as growing population, electricity companies are required to not only manage current supply of electricity, but also to ensure means by which transmission power of the electric grid system can be improved to meet future demands (Du et al., 2017). Therefore, it is imperative that, in the short-term, the electricity load forecasting model is reliable so as to ensure undisrupted supply of electricity to the population in the western region of Saudi Arabia.

Electricity companies use various time periods, such as short-term and long-term, to project electricity so that the grid can be injected into transmission lines. This type of energy is base load (Ozerdem et al., 2017). However, considering the fact that electricity demand fluctuates, electricity companies tends to store little more than base load energy to meet fluctuating demand. Therefore, injection of renewable energy sources in the electricity transmission lines would mean Saudi Electricity Company (SEC) would end up generating more energy than demand required. This extra energy is called spinning reserve and it can be made available in the event of additional energy demand. From the economic perspective, having spinning reserve is not suitable because it costs a significant amount of money. Thus, to avoid the cost of spinning reserve, it is important to conduct accurate short-term load forecasting (Chen et al., 2013). To this end, efficient grid planning is necessary. However, due to impact of environmental factors, such as rain and dry periods, etc., efficient grid planning becomes a challenging task. A variety of models have been developed to forecast electricity; however, each of these models possesses limited scope and is not feasible to accurately forecast electricity (Lee and Hong, 2015).

The aim of this chapter is to propose a hybrid modelling technique to accurately forecast electricity. Reported advantages of a hybrid model are superior performance when compared with LR (linear regression) and ANN, as well as the ability of incorporating forms of reducing the degrees of freedom.

5.2. Combination Methods

Current models proposed to forecast electricity can be categorised into three major categories, such as artificial intelligence and time series models (Zhang et al., 2018). There is a variety of time series models that have been proposed under artificial intelligence, which includes expert systems, support vector machines, fuzzy logic, artificial neural network (ANN), echo state networks, fuzzy logic and support vector regression (Pai and Hong, 2005; Pandian et al., 2006; Yang et al., 2016). Moreover, particle swarm optimisation (PSO) has also emerged as an evolutionary feed-forward artificial intelligence model that combines architectural evolution with weighted learning (Zhang et al., 2000). Unlike, traditional artificial intelligence models, PSO modelling provides strategy for adding nodes and partial training algorithm to maintain behavioural link between electricity generation sources and demand. Hybrid models have been reported as superior to purely machine learning algorithms, also assisting in decision-making under uncertain electricity generation capabilities linked with renewable energy sources (Raza and Khosravi, 2015), they have been criticised for complex application for the neural training needed to train the overall network (Zhang et al., 2018, p.774).

In contrast, time series models, such as linear regression, seasonal autoregressive, autoregressive integrated moving average (ARIMA), threshold autoregressive, Kalman filtering and seasonal autoregressive integrated moving averages (SARIMA) have been used for forecasting electricity load (Zhang et al., 2018, p.774). However, linear regression models have been criticised for the inability to efficiently detect behaviour of supply and demand of electricity. In summary, time series models have been reported as not being able to produce high performance electricity load forecasting results (Che et al., 2012).

In the lack of both time series and artificial intelligence models to predict electricity in their entirety, hybrid modelling approach has emerged. Hybrid modelling technique consists of the integration of different models, being one with artificial intelligence

characteristics. The objective of this approach is to improve accuracy of electricity forecasting (Cho et al., 2013). Integration of time series and artificial intelligence models has been identified as a good choice for improving electricity forecasting accuracy for a variety of reasons. Firstly, hybrid modelling possesses the ability to anticipate different features of electricity load. Secondly, combining one model with another in an integrated form can work in complementary form, such as overcoming the weaknesses of one another (Ghiasi et al., 2019). This chapter proposes usage of two artificial intelligence models, PSO and fuzzy logic, to predict short-term load in the western region of Saudi Arabia. Various past researchers have used fuzzy logic and PSO methods individually for forecasting load (Khamis et al., 2011; Yu et al., 2012; Minhas et al., 2017). However, the application of PSO and Fuzzy Logic in a hybrid approach is not mature and most researches are still in very early development phase. The uniqueness of the present work consists of the implementation and comparison of both modelling techniques, which will bring solid information about the performance enhancement of such approach over pure machine learning or pure algebraic models.

For a better understanding of the combination method, Figure 5-1 illustrates the architecture of the superior combination model used in the current research.
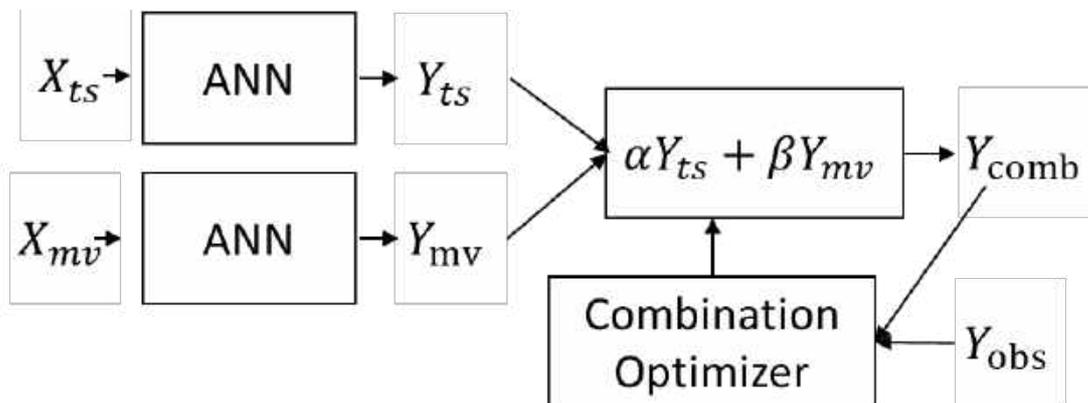
Figure 5-1 – Superior combination model.

In Figure 5.1, the input vectors X_{ts} (time series) and X_{mv} are given as input to both the time series and multivariate model. Each model produces an output, Y_{ts} and $Y_{mv,}$

respectively. These outputs are combined in a single linear model which contains two parameters, α and β, each one performing product with one model output. This combination generates a third output Y_{comb}, which can be understood as the weighted average of both models. The parameters α and β are then calibrated to match the observed output on the best way possible. This calibration is done using a combination optimiser. In the current research, the two options investigated were the PSO algorithm and the Fuzzy Model.

To have a baseline, α and β are manually set and compared to the performance of this model, referred to as Average Model, with the one using PSO (PSO model) and Fuzzy Model. For each case, it was evaluated both LR and ANN models.

5.2.1. Average Combiner with LR

To generate a baseline model for comparison with other combiners, the multivariate and time series LR models were combined using fixed, manually-set values for the parameters of the linear combination, α and β.

Because of the superior performance of the time series model relative to the multivariate model, the values set were:

$\alpha = 0.7$;

$\beta = 0.3$;

with specific attention to the fact that $\alpha + \beta = 1.0$, meaning that the energy load forecast is not over nor underestimated, but consists in a weighted average between the two model outputs.

Figure 5-2 shows the Observed versus Predicted plot for such approach, at both training and testing sets. The MSE obtained for the training set (there is no training in this case) is 1.51×10^6, while, for testing set, is 2.17×10^6. The correlation coefficient is 0.79 for training set and 0.72 for testing set.
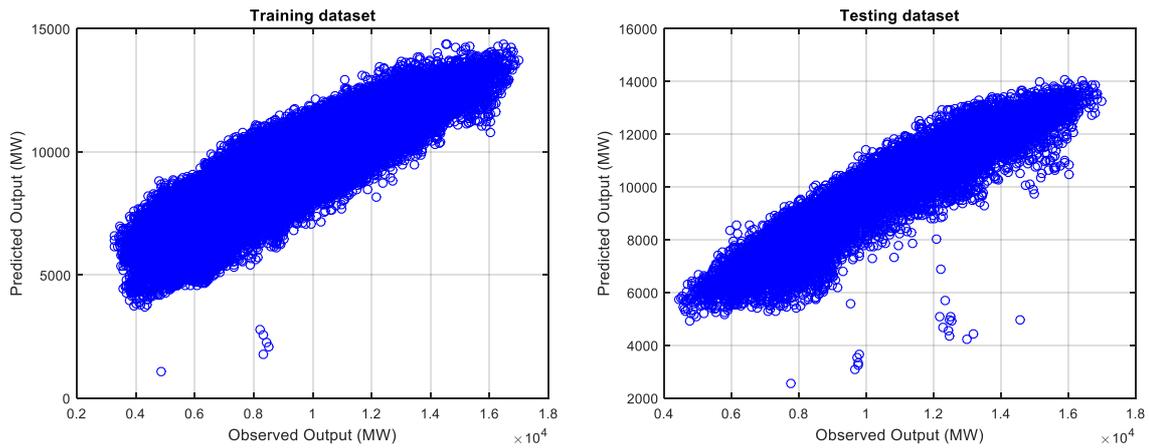
Figure 5-2- Average combiner with LR – Observed x Predicted plot.

From the results expressed above, it can be stated that, as expected, there is an increase in performance when compared with multivariate data, achieved through the compensation of the time series model performance. However, the values obtained are still higher than the ones obtained for ANN (purely time series or purely multivariate), which is expected since linear models are not able to capture nonlinearities of the dataset.

5.2.2. Average Combiner with ANN

The same approach expressed above for the LR model(s) was used for the ANN models. Figure 5-3 shows the Observed x Predicted plot for both training and testing datasets. The MSE obtained for the training set is 6.51×10^5, while, for testing, the value obtained was 7.57×10^5. Correlation coefficient for the training set is 0.91, with approximately the same value being obtained for the testing set.
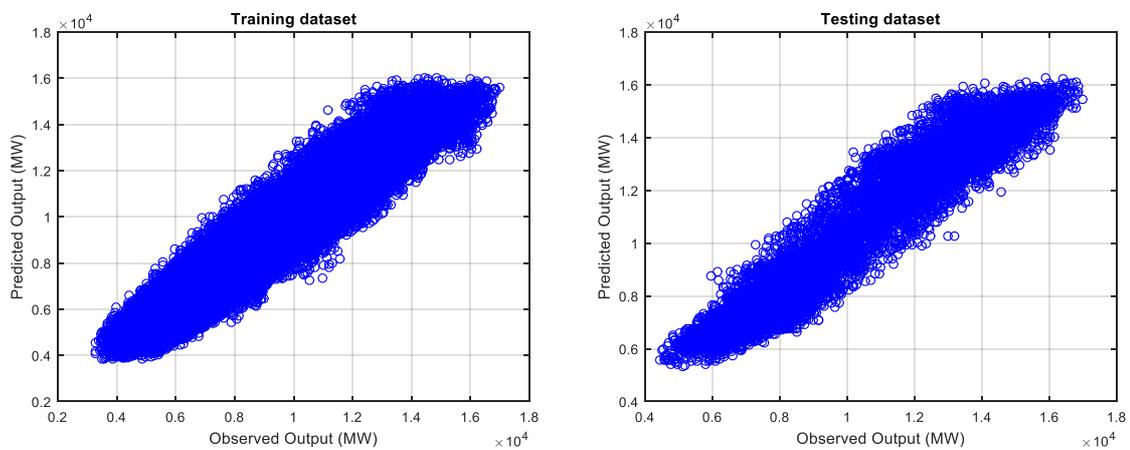
Figure 5-3 - Average combiner with ANN - Observed x Predicted plot

As expected, the ANN Average model outperformed the LR one, with drastic reduction on the MSE error when compared with the multivariate ANN model. The correlation coefficient above 90% indicates that this approach gives reasonable performance and may be relatively robust to be used as an energy load predictor.

5.2.3. PSO Combiner with LR

PSO optimisation algorithm was combined with linear regression using time series and multivariate data, which is expected to enhance model performance by finding global minima. Figure 5-4 shows the particles position at the initialisation of the optimisation process, with random positioning of the particles in the optimisation space. The optimisation started by selecting random weight values on the feature space between 0 and 1. The x-axis on the Figure indicates the value of the combiner model weight for the timeseries model while the y-axis indicates the weight for the multivariate model.
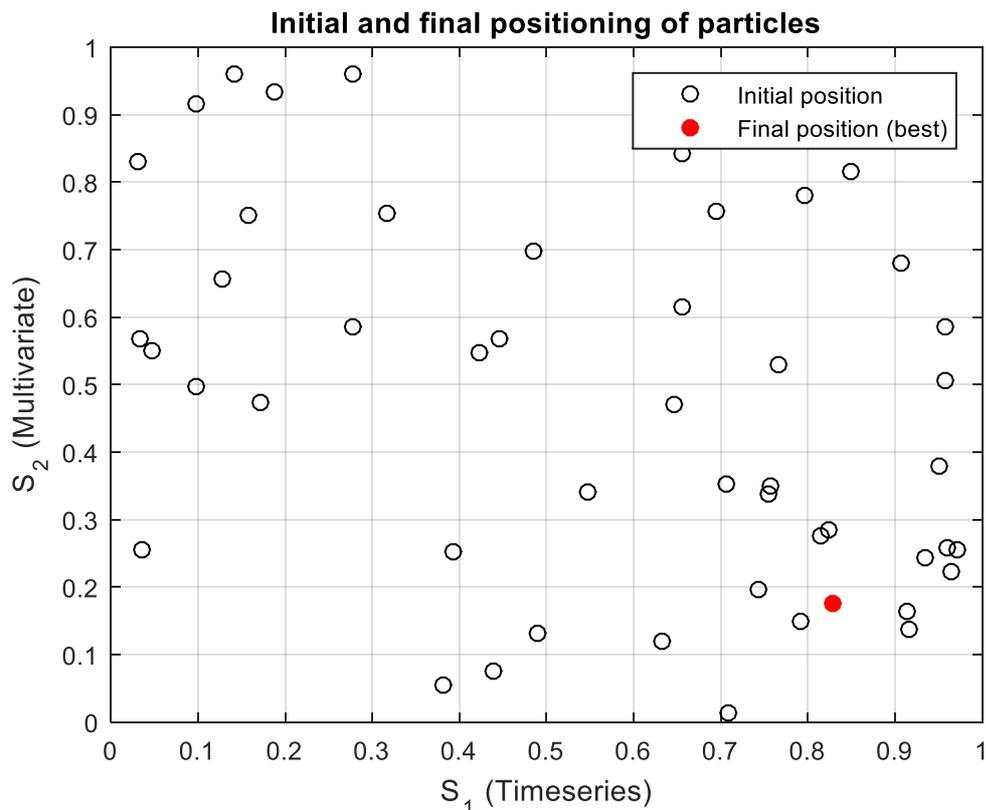
Figure 5-4 – PSO combiner with LR – random initialisation of particles and final position.

After optimisation, particles arrive in a point referred to as global optimum position. The filled point in Figure 5-4 shows the particles position at the global optimal weights of the LR models determined by PSO combiner algorithm.

It is noticeable that the particles converged to a single optimum point, indicating that PSO has successfully managed to determine the best combination of the time series input model with the multivariate input one. This optimum values is used in the combiner model to evaluate the performance of the hybrid structure.

Figure 5-5 illustrates the progression of error at each iteration during model optimisation.
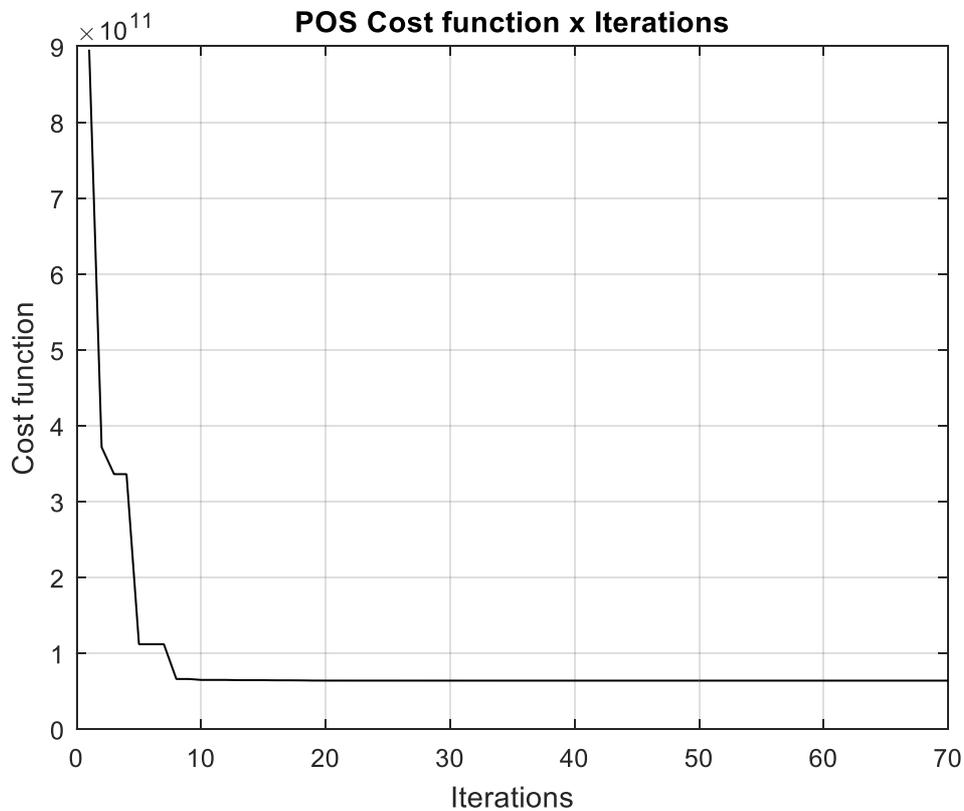
Figure 5-5: Combined LR model error per iteration.

The cost function here is the SSE (sum of squared errors). This value harshly drops after the first iteration, reaching a plateau for some iterations. This plateau can be identified as a local minimum in the cost function. The PSO algorithm is able to leave the local minima and find global minima by reaching a new plateau, lower than the former, and staying on it, which can be seen by the flat line on Figure 5-5.

After optimisation, the combined LR model performance was evaluated on both training and testing dataset. Figure 5-6 illustrates the Observed x Predicted plot for both datasets. The correlation coefficient found for the training was 0.86, with approximately the same value being obtained for the testing set.
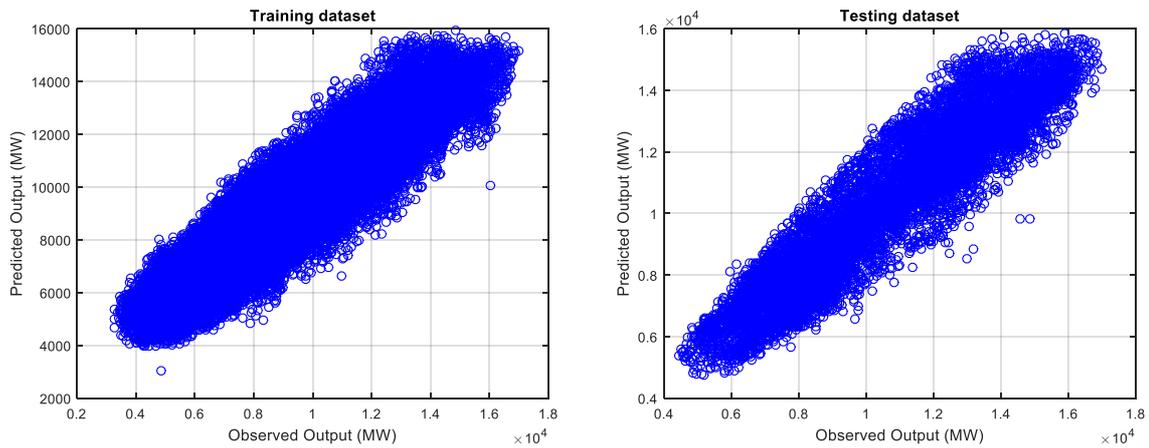
Figure 5-6: Observed versus Predicted plot for PSO combiner with LR. Training (left) and Testing (right).

Training error was noted to be 9.73×10^5 and testing error was found to be 1.09×10^6. For better understanding, correlation coefficient was also calculated. For the training set, it is 0.86, with the same value being obtained for the testing set.

The results indicate that there is a performance improvement in relation to the multivariate approach, but such improvement is compensated on the time series, which decreases in performance. Therefore, it can be a reasonable replacement for the purely multivariate model, though it cannot over perform the purely time series model.

5.2.4. PSO Combiner with ANN

Potential enhancements of the ANN performance were investigated through the combination of time series and multivariate data into a single model, using PSO algorithm for optimisation. Similarly to the procedure followed in obtaining PSO combiner with LR, the particles were initialised in the subspace [0,1]. Figure 5-7 shows the positioning of the particles at random points before the optimisation starts and at the optimum point after the

optimisation process. Here the x-axis indicates the weight for the multivariate model, while x-axis indicates the weight for the timeseries model.
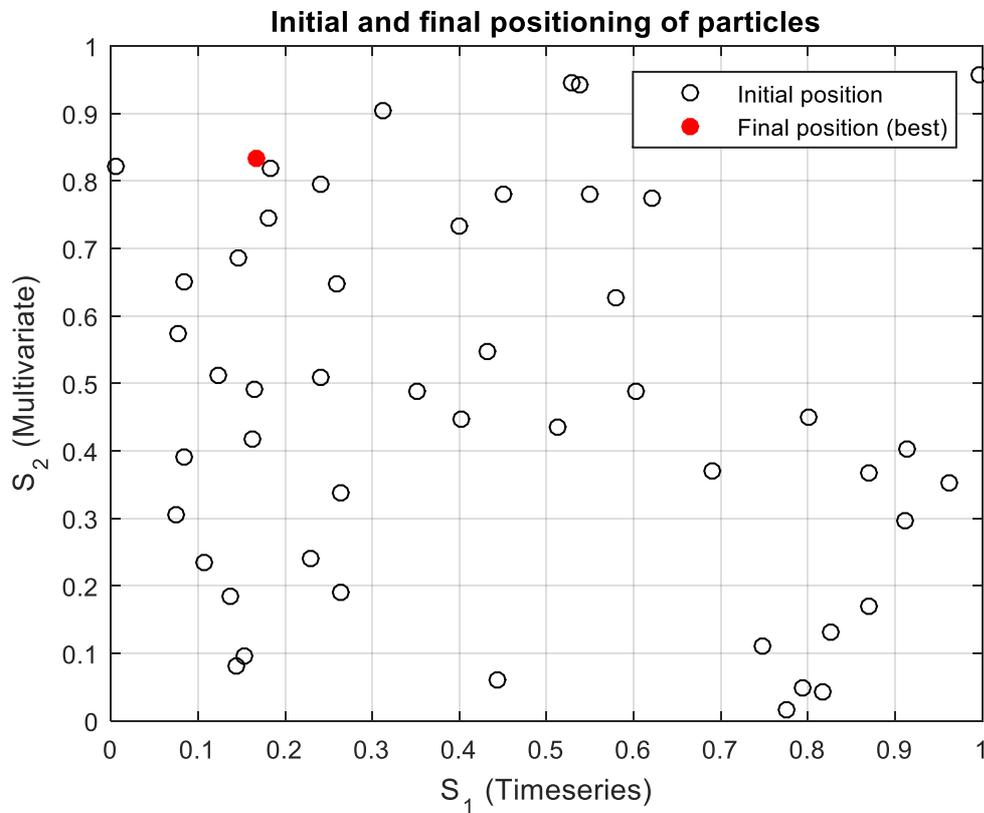
Figure 5-7 – Random initialisation and final position of particles in PSO optimisation – ANN.

As in LR, it can be observed that the particles converged to a position, indicating the presence of a global minima at that point (filled mark). The slight separation between weight values for each layer is also expected due to the nonlinear transformation carried out in the hidden layer. Additionally, one can see that the values are slightly different to the LR, which is attributed to the differences in the topology of both models, which produces different cost functions, the first one strictly convex (LR) and the second one non-strictly convex (ANN).

The evolution of error minimisation using PSO algorithm in ANN is shown in Figure 5-8, at each optimisation iteration.
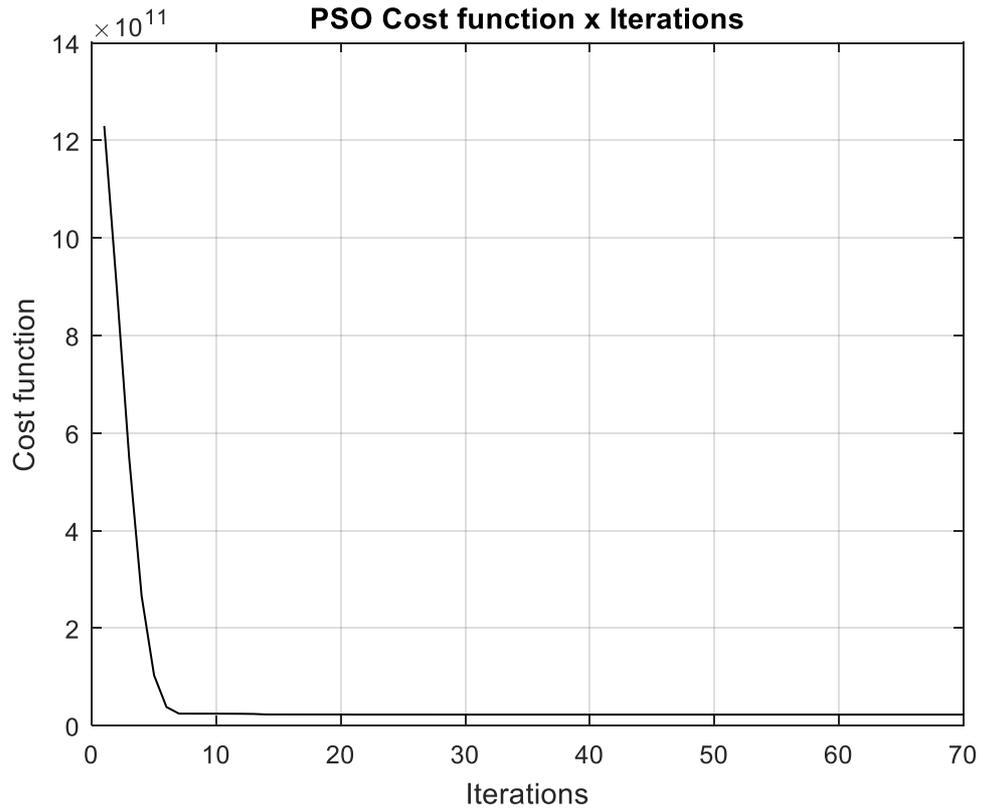
Figure 5-8 – Evolution of PSO optimisation per iteration - ANN.

As in the PSO Combiner with LR, the Cost function here measures the SSE (sum of squared errors). Even though the ANN cost function is highly non-linear, the PSO algorithm is able to determine the global minima of the function, observable in the final plateau obtained. It can be seen a steep drop on the cost function at the beginning with relatively fast convergence to a global minima. This indicates the potentiality of such approach when used to optimise nonlinear models and to determine the global minima.

PSO combiner with ANN was evaluated into both training and testing dataset. Figure 5-9 illustrates the Observed x Predicted plot for such model. The training correlation coefficient is 0.95 (95%), while the testing one is 0.86 (86%).

77
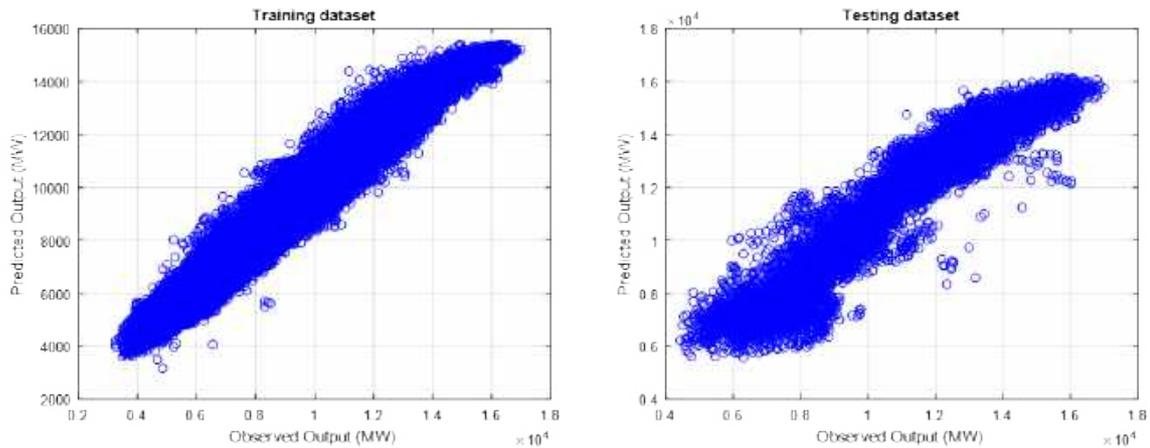
Figure 5-9 – PSO combiner with ANN – Observed x Predicted plot for Training (left) and Testing (right).

The MSE obtained for training is 3.44×10^5, while for testing is 1.06×10^6. These values are lower than the one obtained for the multivariate, but higher than obtained for the purely time series model. This is attributed to the weighted average characteristic of the model, which compensates the low performance of one model through the decrease on the performance of the other.

Still, it can be clearly observed that the use of PSO combiner with ANN can improve the performance through the maintenance of the same topology of the model, but replacing the optimisation algorithm to determine optimum values for the weights.

5.2.5. Fuzzy Combiner with LR

Concept of fuzzy logic modelling was presented by Tanaka et al. (1988) to minimise errors in load modelling based on linear programming via utilising systematic triangular fuzzy parameters. In a nutshell, fuzzy modelling is a non-statistical method for systems that represents probabilistic situation, for instance high rate of variations in electricity load data presented in previous chapters for the western region of Saudi Arabia. Ying and Pan (2008) also supported fuzzy modelling method to capture nonlinear relationship between input components of electricity load forecasting, such as previous day, week, years, day and time, and output components, such as forecasting future load, that cannot be possible through using intuitive understanding via using mathematical methods.

78

Fuzzy logic modelling tends to be based on two methodologies, fuzzification and fuzzy rule base (Ganguly et al., 2017). Fuzzification is referred to as a process in which numerical load values are converted into degree of membership linked to the correspondence set of fuzzy models. Fuzzy modelling creates sets of membership functions (MF). The MF set takes numerical values of load as their argument and returns the value representing a specific set of MF. In contrast, fuzzy rule method stores heuristic knowledge in an if-then rule. The processed information in if-then analysis is then sent to the fuzzy interference system (FIS). Such system processes the information from inference mechanism to quantitative examine information embedded in the fuzzy rule to create output. The process of fuzzy inference is defined as the one in which maps are formulated from a given input to an output through fuzzy logic. Mapping then forms the foundation for the decision-making process.

After the combination of the model using fuzzy at superior stage, the correlation between the predicted and observed values for energy load were compared using Observed x Predicted Plot. The graphs for the training and testing set are show in Figure 5-10. The correlation coefficient for the training dataset is 0.84, while for testing is 0.73.
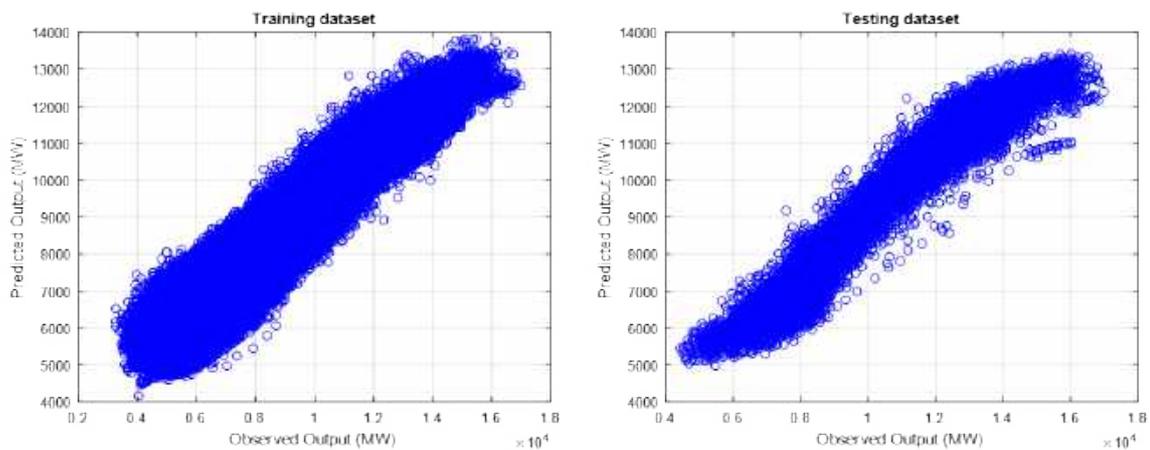
Figure 5-10 - Observed x Predicted plot for PSO fuzzy combiner using time series and multivariate data. Training (left) and Testing (right) sets.

In contrast to PSO with LR and with ANN, fuzzy combiner provided training MSE of 1.16×10^6 and testing MSE of 2.13×10^6. This error is higher than the one obtained for the PSO combiner, but still lower than the purely multivariate data. However, the compensation produced by the performance decline in the time series model may not justify the use of such combiner.

5.2.6. Fuzzy Combiner with ANN

Similar to the LR with fuzzy combiner, time series and multivariate ANN models were combined in a single superior structure with the aim of improving model performance and increasing robustness of each structure.

After combination, the output produced by the combined model was plotted against the observed energy load, generating the Observed versus Predicted plot shown in Figure 5-11. The MSE for training set is 4.22×10^5 and for testing is 5.82×10^5. Correlation coefficient obtained is 0.94 for training and 0.93 for testing.
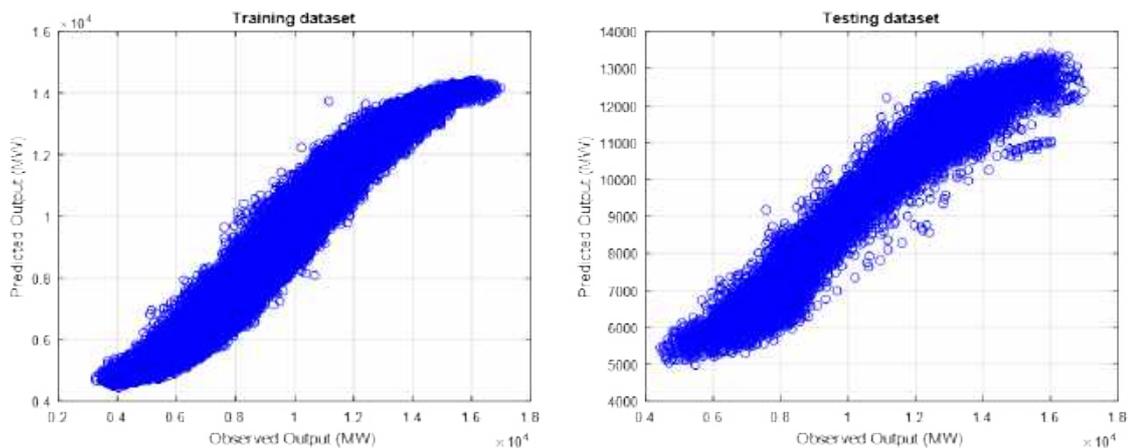
Figure 5-11 – Fuzzy Combiner with ANN – Observed x Predicted plot for Training (left) and Testing (right).

Fuzzy combiner with ANN showed severely higher performance when compared to fuzzy combiner with LR. As already expected, the compensation characteristics of the model caused a decrease of performance when compared with time series, but with an increase on the multivariate one.

5.2.7. Summary of Combiner Models' Performance

Table 5.1 provides a comprehensive summary of the performance metrics obtained for all three combiner models used (Average, PSO and Fuzzy) at each modelling technique (LR and ANN). The two metrics shown here are the MSE (in MW^2) and the R^2 (in percentage, being 100% perfect model and 0% a model as good as using the average value of load).

Table 5.1 – Summary of model performance for combiners applied to LR and ANN models. Values are MSE ($MW^2 * 10^5$) and inside parentheses are R^2 (%).

Model	FIS LR	FIS ANN	PSO LR	PSO ANN	AVG LR	AVG ANN
Train x 10^5	12 (84%)	**4.2** **(94%)**	9.7 (86%)	3.4 (95%)	15.1 (79%)	6.5 (91%)
Test x 10^5	21.3 (73%)	**5.8** **(93%)**	10.9 (86%)	10.6 (86%)	21.7 (72%)	7.6 (90%)

From the results above, it can be clearly stated that, considering both training and testing, the best result was obtained through the combination of fuzzy superior model with ANN. Such result is clearly distinguishable from the others, though in the training set ANN with PSO combiner also demonstrated a high performance. However, in this case, performance severely dropped on the testing set, what prejudiced the judgement of such result.

5.3. Sensitivity Analysis and Results

In this research, sensitivity analysis (SA) has been conducted to test the effect of each environmental factor (input) on the load forecasting. Pianosi et al. (2016) defined SA as a method to investigate the way variations in output of numerical modelling are affected by the inputs. In the broad context, the type of approach and level of complexity of SA will

depend upon the domain of modelling and its specific aims and objectives. In this work, the aim of sensitivity analysis is to assess the impact of various environmental factors, such as month, year, day type, time of the day, temperature, day sensitivity such as first day of month, and, finally, humidity.

Suitability of conducting SA in this research lies in the fact that it allows dealing with complex simulation scenarios such as changing climate and atmosphere on electricity production using renewable energy sources (RES) through using 'what if' analysis (Devenish et al., 2012). What if analysis assists in making decisions while facing complex situations such as where inputs of a model vary significantly (Hong and Fan, 2016).

Sangrody et al. (2017) investigated the issues surrounding weather forecasting for solar energy prediction. The authors suggested that, when integrating RES in the electric grid, electricity forecasting itself become an SA exercise. That can be stated, particularly for RES, due to the high-level intermittence in the availability. Under these scenarios, SA can help by identifying the input factors that could cause high level of variation in the forecasted load.

SA can be conducted under four basic conditions, namely, factor prioritisation, factor fixation, cutting variance and factor mapping (Saltelli et al., 2004). Factor prioritisation assists in ranking the input factors according to their ability to reduce the variability in the output. On the other hand, factor fixation assists management in identifying factors that do not have any influence on the output. In contrast, variance cutting method assists management in fixing sub-set factors to achieve desirable output. Finally, factor mapping assists the identification of the most responsible factors for causing variance in the output. In this research, factor mapping tool of SA will be used since it aims to identify the most responsible environmental factors impacting the load forecasting of different hybrid models.

Sensitivity analysis was performed using input data previously mentioned for the year 2010 until 2016. The analysis was carried out in MATLAB® to assess performance of neural networks based on training and testing data. Results of this analysis are presented in Table 5.2 and Table 5.3. This performance was measures in terms of MSE, which indicates the amount of error that each feature affects.

Table 5.2 – Sensitivity Analysis for Training set.

	PERFORMANCE (MSE) (MW^2*10^6)	STANDARD DEVIATION (MW^2*10^3)
Month	**0.847**	**7.19**
Year	1.225	**9.007**
Time (Hourly and Half hourly)	**0.497**	**3.804**
Weekday	0.363	4.328
Temperature	0.363	4.328
Month day	**0.413**	2.68
Humidity	0.432	6.918

Table 5.3 – Sensitivity Analysis for Test Set.

	PERFORMANCE (MSE) (MW^2*10^6)	STANDARD DEVIATION (MW^2*10^3)
Month	**3.546**	**0.0805**
Year	2.755	**0.0089**
Time (Hourly and Half hourly)	**52.4**	**79.38**
Weekday	3.51	0.32
Temperature	3.51	0.32
Month day	**4.811**	0.875
Humidity	2.494	0.442

From Table 5.2 and Table 5.3, one can notice that time, month and month day are the most important factors since they show highest performance error in both testing and training sets. In contrast, error standard deviation of month and year appears to be the most impacting factors causing variations in the output of load forecasting.

Figure 5-12 illustrates the Observed x Predicted plot for the sensitivity analysis of month factor as an input to the neural network with 15 neurons in one hidden layer.
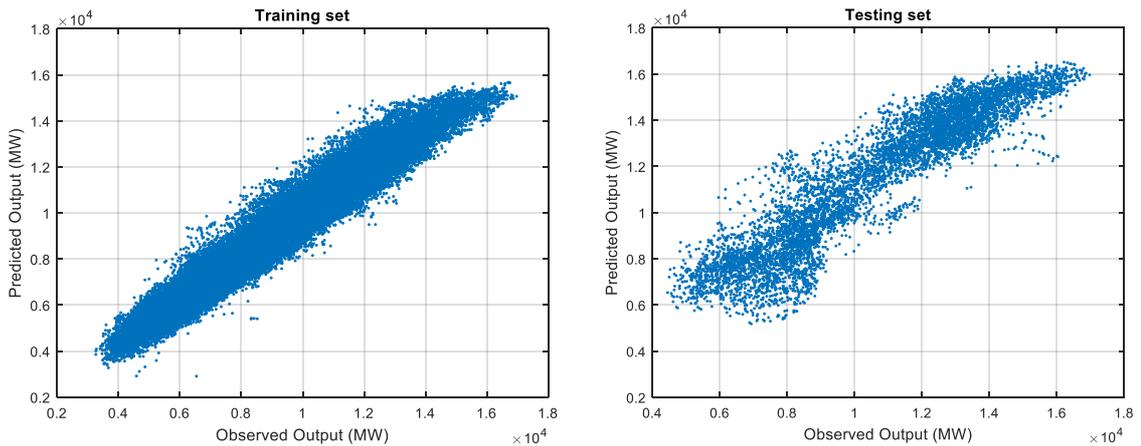
Figure 5-12 - Sensitivity analysis of month factor (Observed x Predicted plot). Training (left) and Testing (right).

The scattered data observed in the testing data show how the change of month deeply affects the load forecast, confirming the high sensitivity that the model presents towards this factor. In the training set, lower spreading of data is observed, which indicates that patterns on this factor changed from the time period used in the training set to the time period presented in the testing set.

Figure 5-13 shows Observed versus Predicted plot for the sensitivity analysis of the load forecasting with respect to the year input data. The model used here is the neural network with 15 neurons in one hidden layer.
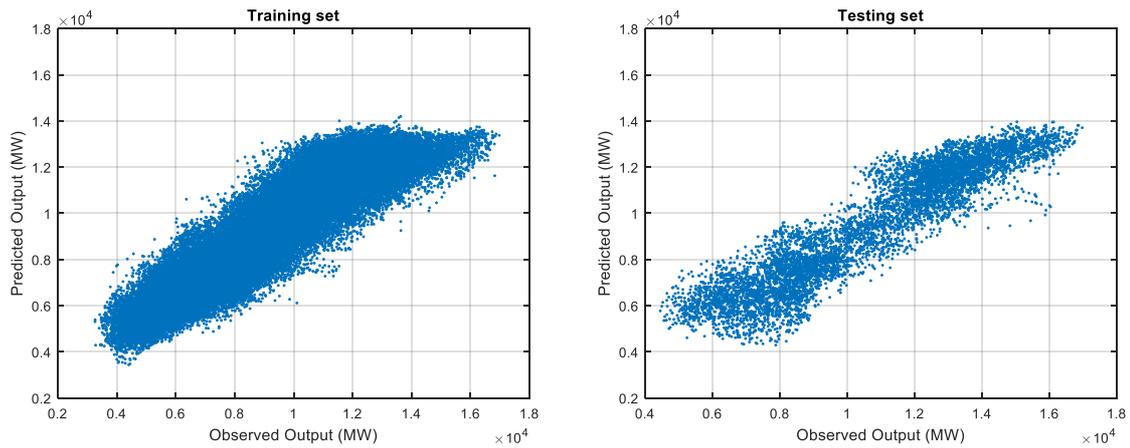
Figure 5-13 - Sensitivity analysis of year factor (Observed x Predicted plot). Training (left) and Testing (right)

It can be seen that, though severely affected by this input, the neural network was able to capture the relation between year and load forecast with relatively reasonable performance. Slightly higher scattering of data is observed in the testing set in relation to the training set. Such scattering may indicate a small change on the energy load pattern for years after the training set.

Figure 5-14 shows Observed versus Predicted plot for the time factor using neural network with 15 neurons in one hidden layer.
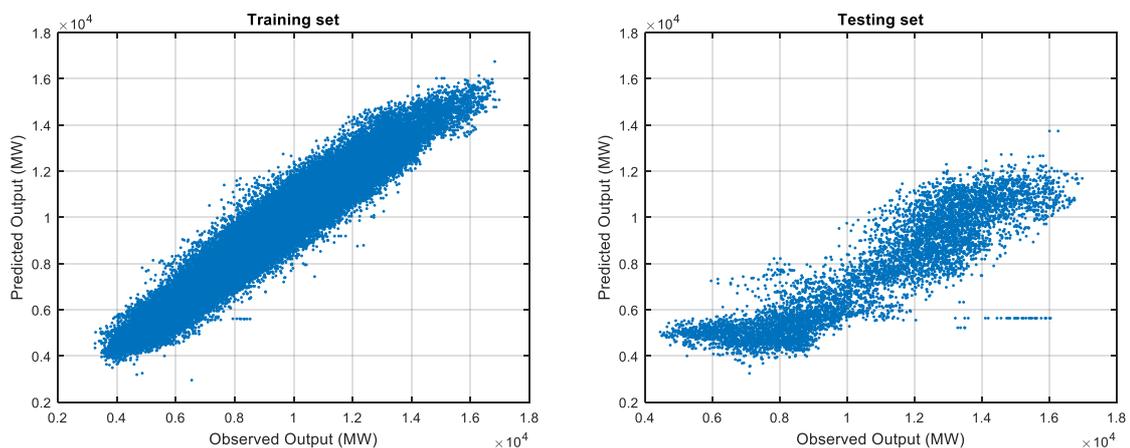
Figure 5-14 - Sensitivity analysis of time factor (Observed x Predicted plot). Training (left) and Testing (right)

It can be seen that, though in the training set the model is able to capture the relation between energy load and time, in the testing set such performance decreases drastically. Such statement can be confirmed by the scattering of data in the figure above.

Observed versus Predicted plot for temperature is shown in Figure 5-15. The model used, as in the previous analyses shown, is the neural network with 15 neurons in one hidden layer.
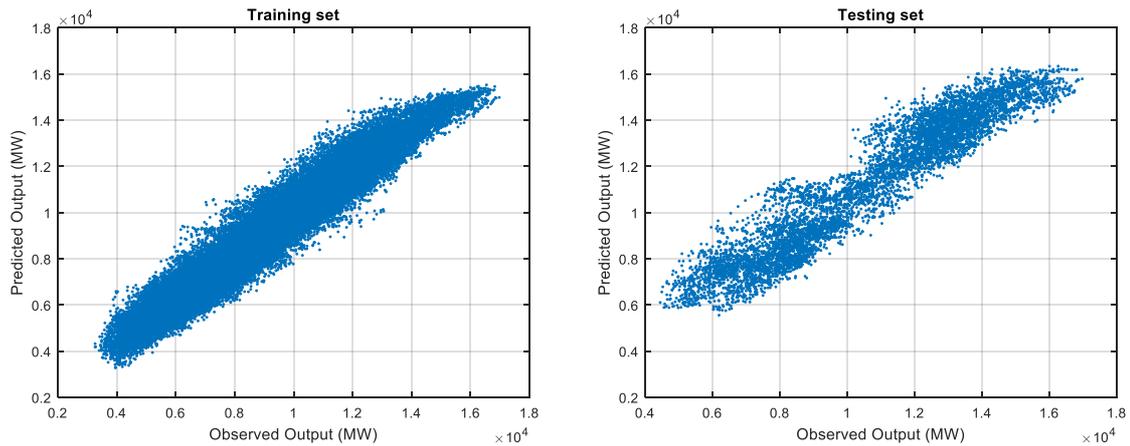
Figure 5-15 - Sensitivity analysis of temperature factor (Observed x Predicted plot). Training (left) and Testing (right).

In both training and testing set, correlation coefficient indicates that temperature had little impact on model sensitivity. This also indicates that the temperature pattern is maintained for the whole period of training and testing set.

Day of month sensitivity results are illustrated in Figure 5-16, through Observed versus Predicted plot. The model used is the same as in previous cases, i.e. neural network with 15 neurons in one hidden layer.
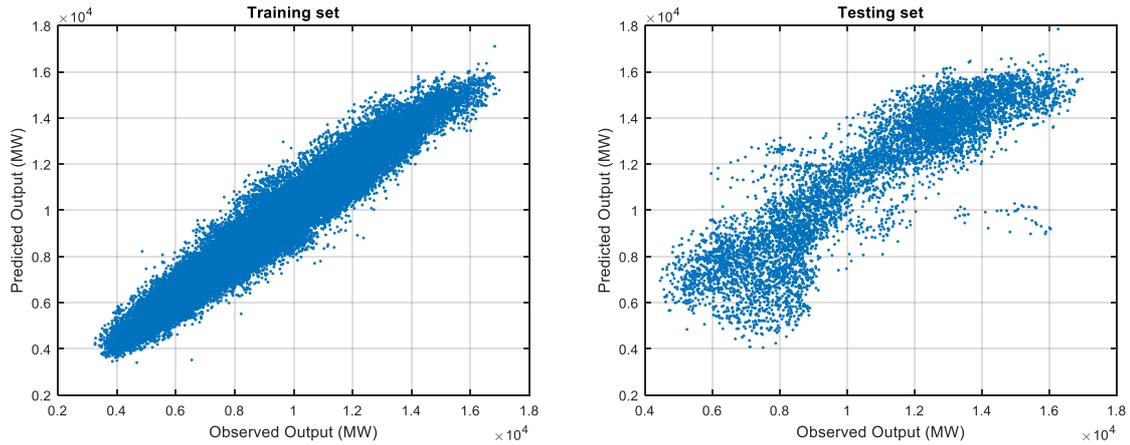
Figure 5-16 - Sensitivity analysis of month day factor (Observed versus Predicted plot). Training (left) and Testing (right).

Relatively high scattering of data in the testing set shows the high effect that month has on the sensitivity of the model. Interestingly, the spreading is drastically less on the training set, indicating that the model was able to capture the relation of this variable with the load for the period used in training (2010-2015). However, the high decrease in performance may indicate a change of pattern on the relation between such variables.

Figure 5-17 presents humidity sensitivity results using neural network with 15 neurons in one hidden layer.
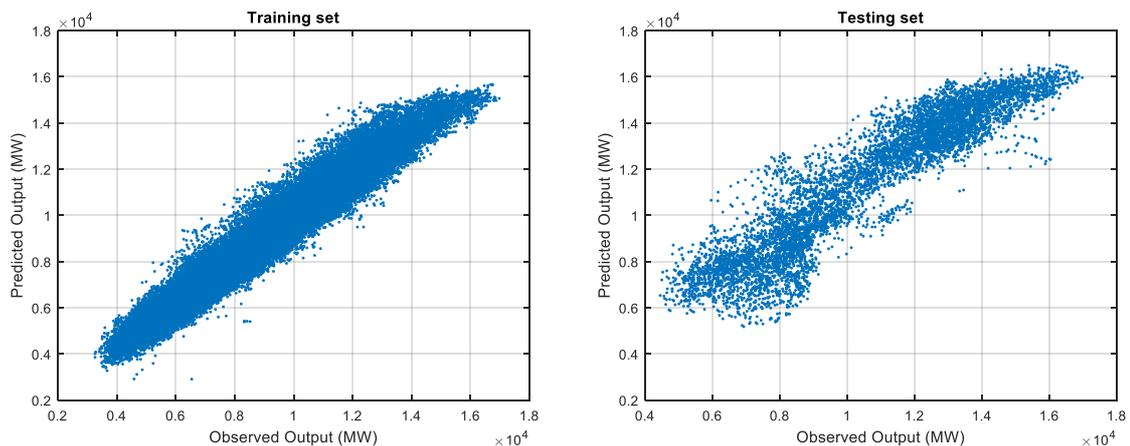
Figure 5-17 - Sensitivity analysis of humidity factor (Observed x Predicted plot). Training (left) and Testing (right).

The relatively low scattering of data in both training and testing sets confirms the relatively low impact of such factor in model sensitivity for energy load prediction.

5.4. Summary

In the current chapter, PSO and fuzzy techniques were used as higher level model combiners to incorporate the advantages of multiple model embedding. Additionally, the combiner model should provide a balance between the timeseries input model and the multivariate one, with the intent of providing robustness and increase the accuracy. The final output consists in a weighted average between the two models. This approach was evaluated for both the linear regression and neural network model.

The PSO and fuzzy techniques uses different approaches to determine the parameters that weights each model. The results of these techniques were compared with hard fixing the weights in an approach referred to as "Average Combiner". The best output was obtained using the Fuzzy Combiner with Artificial neural network models. The correlation on this case was 93% with MSE $5.8*10^5$ MW^2 in the test set. On the other hand, the combiner with the least performance was Average Combiner with linear regression, with correlation of 72% with MSE $21.7*10^5$ MW^2 at the test set. Fuzzy combiner used on the same model (LR) slightly outperformed this one (R^2 73% with MSE $21.3*10^5$ MW^2), with the least performance among the PSO and fuzzy combiners.

Though the MSE obtained for the pure ANN with TS was lower than the one obtained for the combiner models, the FIS with ANN would be evaluated as a better model, since it is able to perform predictions based on more information, i.e not only the timeseries but also the multivariate data.

Regarding sensitivity analysis, findings show that input factors such as month and day are key factors that impact the load forecasting since these factors presented highest error. It also confirms that neural networks are able of capturing the relations between these variables with high accuracy.

Though the modelling in this chapter was done with data from west part of Saudi Arabia, the methodology employed could be generalised to different parts of the country. This will enable the development of a general power system for planning and optimisation

of the grid with the inclusion of RES. For that, the models can be generalised for each region, according its data to predict power demand. Integrating the model with weather conditions can enable the decision making system to plan ahead and satisfy the demand using Res without compromising the safety of the network.

Chapter 6.

Power Network and Planning

6.1. Introduction

After generating reliable and accurate models (as done in Chapter 4 and 5), it is necessary to develop a robust decision making system that these models can be integrated to fulfil the purposes of the management team. Such purposes may be to maintain the production in accordance to the demand level, or more advanced, to enable the incorporation of RES as a replacement of traditional power plants, without compromising the demand.

This second purpose is important, since application of linearization and other forecasting methods has revealed that global fossil fuels such as oil and gas production have already peaked and could well have entered declining mode (Brown and Foucher, 2008). Expected decline in fossil fuels could result in massive discrepancy in demand and supply of energy produced through fossil fuels. Such fact creates a need for integrating renewable energy sources in the electric grid system (Al-Garni et al., 2016). However, another reason for introducing RE sources in electricity generation lies in that fact that carbon emission released by burning fossil fuels releases greenhouse gases (GHG) into the environment. This raises environmental concerns (Ramli et al., 2017). Although Saudi Arabia sits on some of the largest reserves of fossil fuels in the world, it also possesses potential to produce electricity through renewable energy sources due to the country's favourable climate conditions.

Abundance of RE sources such as solar photovoltaic (PV) and wind energy in Saudi Arabia can be used as alternative methods for energy production. This policy would reduce the country's reliance on fossil fuels. While undoubtedly the country has potential to produce electricity from RE sources, such move would require alteration in how to manage, distribute and maintain the grid (Mahmud et al., 2018). To this end, various approaches to simulation and modelling in power system have been developed with the Electrical Transient Analyser Program (ETAP). This tool has been recognised as an effective one to deal with the complex task of distributed energy system (Kulkarni and Sontakke, 2017).

In this study, four offshore wind farms in Rabigh and Shoiba with the capacity of 600 MW were used. Two test cases were evaluated. The first one consists of switching off 17 gas units using solar and wind sources. In the second case, 33 gas units are switched off. In both cases, 380kV network RES electricity was injected to the transmission grid with estimated green power generation capacity of 2,490 MW. This value represents almost 12% to total fuel generation.

This chapter will evaluate five capacity scenarios of integrating RES into the electric grid system in the western region of Saudi Arabia. These scenarios consist of investigating the impact of switching off 17 gas units and 33 gas units, respectively, on the peak and off-peak load under five capacity scenarios: 100%, 95%, 75%, 50% and 25%. To this end, discussion will start by providing a brief description of Electrical Power System Analysis (ETAP®) modelling technique followed by simulation results, analysis of quality of network generation and implications for planning and renewable energy utilisation. The main idea is that, using the knowledge from previous chapters, available data can be used to perform prediction on the future load, and this generated data can be integrated in a robust, centred decision-support system.

Having identified capacity of power stations without inclusion of RES, discussion presents data on solar and wind power availability in the western region of Saudi Arabia. To this end, it provides the concept of smart grid and its role in transmitting power followed by availability. Moreover, the theoretical background of the modelling technique will be presented followed by presenting data on power generation capacity of grid stations in different cities and data on wind and solar power.

ETAP® provides a comprehensive analysis platform for design, simulation, transmission and optimisation automation of energy generation in a distributed energy production grid system (Kulkarni and Sontakke, 2017). ETAP® has been recognised as a comprehensively integrated software solution to integrate RE into the electricity transmission system since it provides software solutions to integrate a variety of aspects, such as arc flash, load flow, short circuit, relay coordination, cable capacity, transient stability, optimal power flow, etc. (Kulkarni and Sontakke, 2017).

In the previous chapters, models were developed to perform one-day-ahead load forecasting. This required either multivariate data (humidity, temperature, etc.) or time series data. In both cases, data were generated for one-day-ahead load. In the current chapter, this same data are used to simulate the grid. Finally, the recommendation system can provide results regarding the operation of the power generation.

Load forecast of the western region of Saudi Arabia will be used to assess the possibility of integrating RE via using ETAP® software. It is expected that ETAP software will assist in providing error report and highlight the gap to diffuse the ability of RES in fulfilling the need of energy requirement of the Western region of Saudi Arabia.

6.2. Electrical Power System Analysis (ETAP) Modelling

Figure 6-1 presents the 380 kV electricity transmission network of the western region of Saudi Arabia on ETAP® software. A schematic diagram provided by SEC of the same network is shown in Figure 6-2.
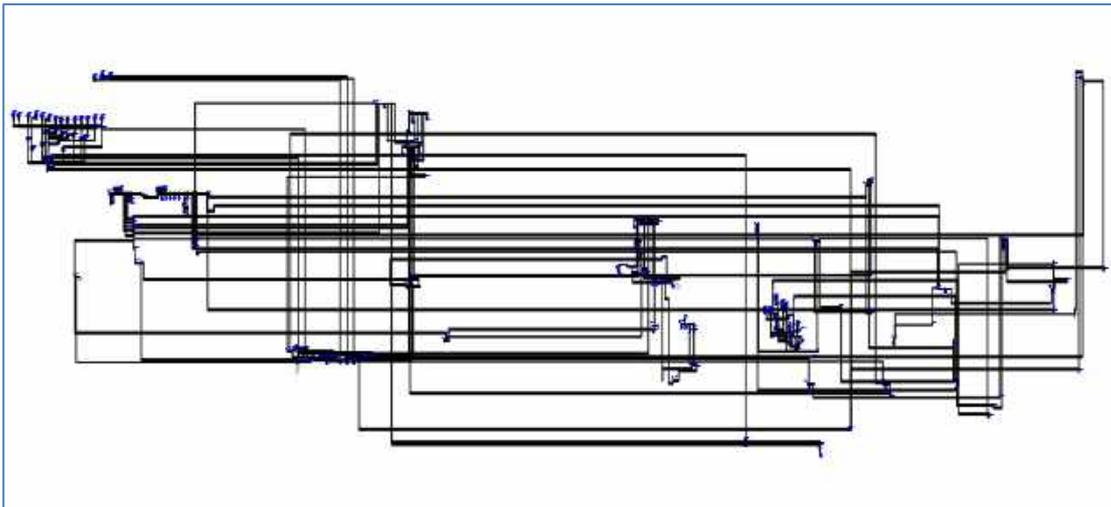
Figure 6-1 - 380kV network of the western region of Saudi Arabia.
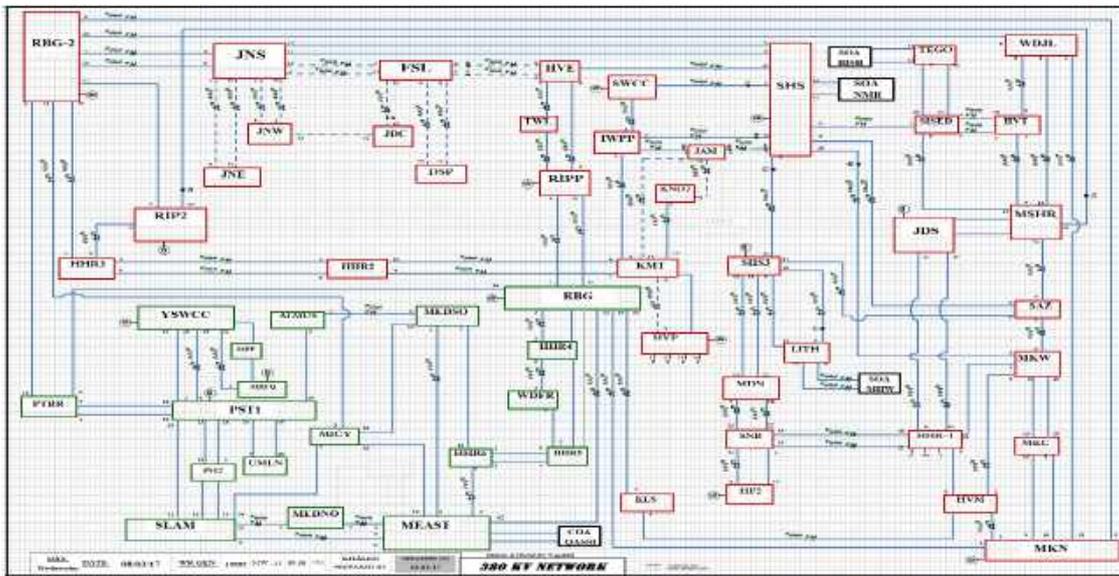
Figure 6-2 - Schematic diagram of the 380 kV network for the western region of Saudi Arabia.

Figure 6-3 presents the two wind farms; one of them is off-service and one in service. The numbers of gas units are switched off while not on service. Figure 6-3 is taken at the time of ETAP modelling.
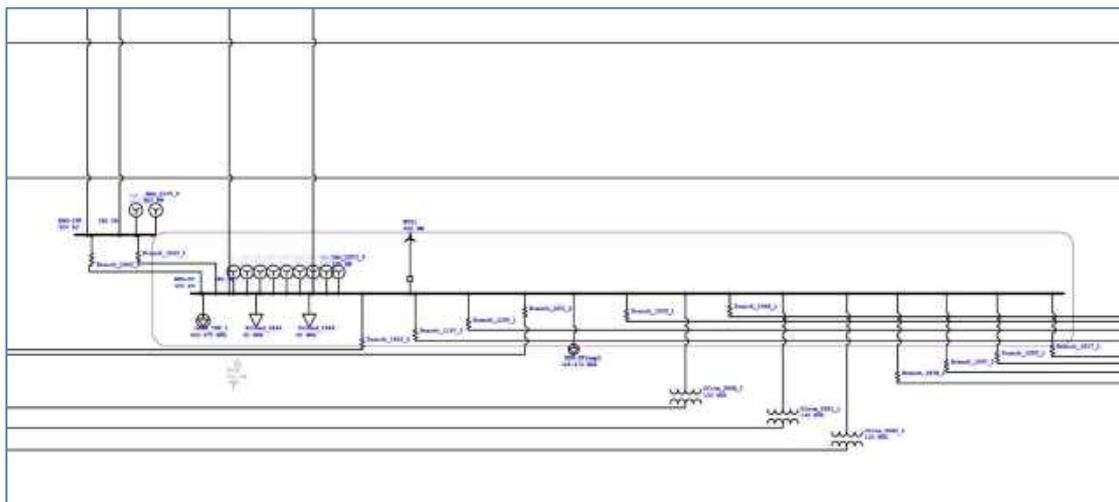
Figure 6-3 - Two wind farms. Off-service (light grey) and on-service (wind turbine symbol).

Figure 6-4 presents two solar panels integrated in Medina. Figure was generated during ETAP modelling.
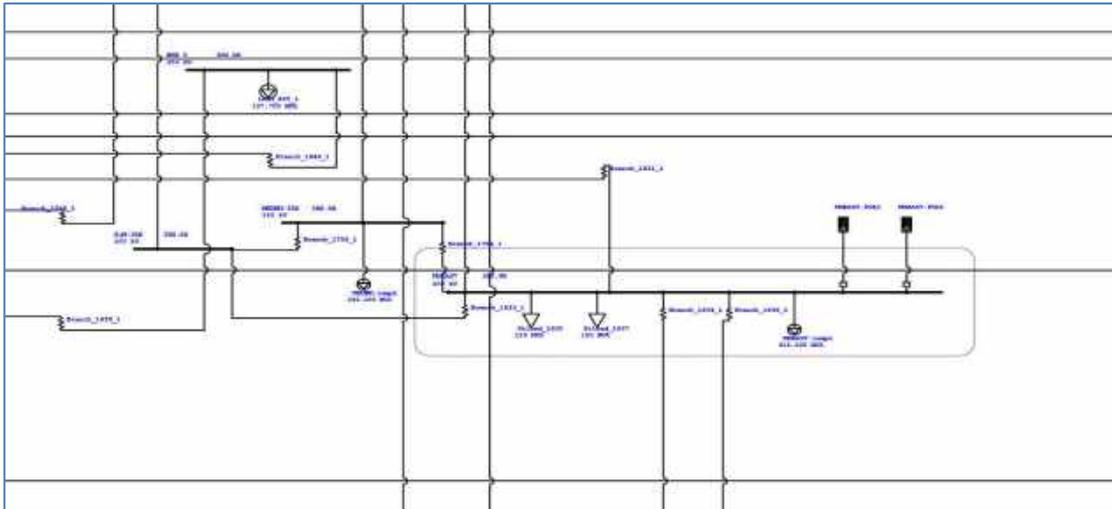
Figure 6-4 - Two solar panels in Medina.

6.3. Case Study

Table 6.1 presents the power plants' capacities in the main five cities in the western region of Saudi Arabia.

Table 6.1 - Generation capacity of power grids in the western region of Saudi Arabia

Stations	Number of units	Number of Power Plants	Capacity based on primary fuel (MW)
Jeddah	39	4	4026
Rabigh	54	4	8056
Yanbu	3	1	45
Shoiba	29	3	6738
Makkah	14	1	666
Medina	8	1	293
Taif	6	1	100

The total generation capacity of all the grids is 19,933 MW, without inclusion of RES. Generation, transmission and distribution data were provided by SEC. Data on distribution network of Jeddah, Rabigh, Yanbu, Shoiba, Makkah, Medina and Taif were removed. This was done because modelling in this study did not include network distribution. Therefore, only data on generation and transmission were used for integration of RES and assessment of different scenarios of power flow.

Chapter 4 has already provided discussion on load forecasting and data for the year of 2010-2016 gathered from Saudi Electricity Company (SEC). Data contained information on various factors, such as weather conditions (temperature and humidity), date (time, day, month, year) and, finally, the load. For the purpose of ETAP modelling, time series (TS) technique will be used.

6.4. Smart Grid Role

The power grid operates a system that transmits power from various central power generating stations to population for both domestic and commercial use (Liserre et al., 2010). In contrast, the smart grid system tends to possess ability to connect the grid system with different sources of energy and is also capable of rerouting power in more efficient manner. Smart grids tend to possess control systems that analyses performance through using autonomous learning controller, thus possessing capability to adjust supply and demand of energy. Moreover, it is easier to connect RES, such as micro turbines, solar power and wind power, at not only local, but also regional and international levels. However, connection to RES results in bulk increase in the capacity of energy generation, thus requires improvement in transmission management. Through using a smart grid, management can optimise power flow at both peak and off-peak times. Furthermore, the smart grid can also adjust them through switching between power generation sources, thus overcoming the issue of intermittent nature of power generated through RES (Liserre et al., 2010). Two RES have been selected for the purpose of simulation exercise, namely, solar and wind power.

6.5. Solar Energy in Saudi Arabia

Past studies have established Saudi Arabia's capability for solar energy due to the country's geographical location (Almasoud and Gandayh, 2015: Al-Sharafi et al., 2017).

Figure 6-5 shows that the Middle Eastern region, including Saudi Arabia, is located in the geographically best locations to take advantage of solar energy. According to estimate by Gulf Cooperation Council (GCC), annual solar radiation in Middle Eastern countries reaches to about $250W/m^2$ where, the same for countries in North America, Latin America and Western Asian countries, it is between 100-200 W/m^2 (Pazheri et al., 2012). Furthermore, producing electricity from solar energy has been noted to cost far less in areas with strong solar radiation (Pazheri et al., 2012). Moreover, Saudi Arabia can also avail economic benefit of using its solar energy. According Matar et al. (2017), the strength of solar energy in Saudi Arabia means that the country could go beyond producing enough energy to harvest its domestic demand for electricity. It could also export electricity to other countries of East Asia and Latin America which need such resource.
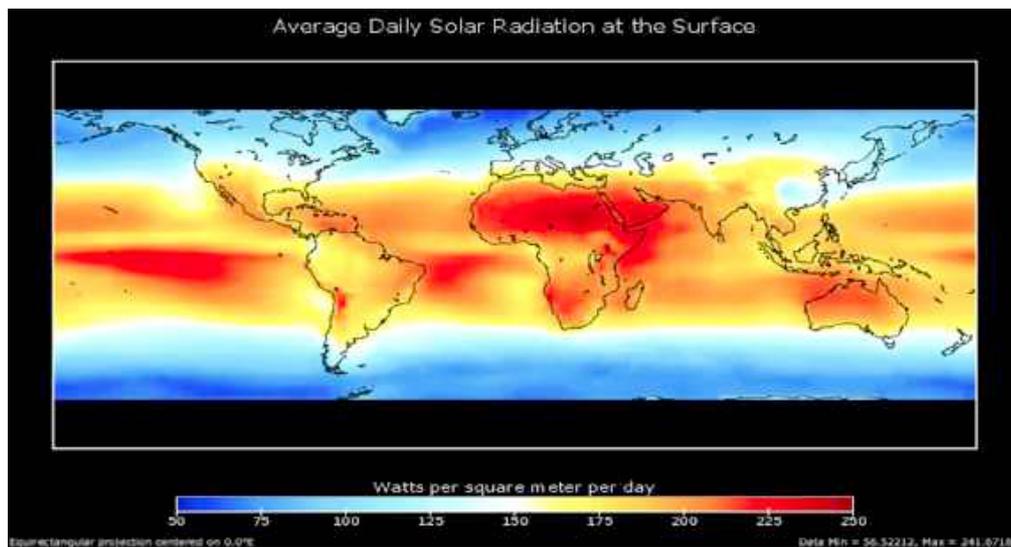
Figure 6-5 - Solar radiation distribution in Saudi Arabia. Source: Pazheri et al., 2012.

6.6. Wind Energy in Saudi Arabia

Six regions in Saudi Arabia, Dhahran, Jouf, Turaif, Yanbu and Alwajh, lie along the coastline of the Arabian Gulf and Red Sea and thus are characterised as a natural source of wind power (Tagle et al., 2017). From Figure 6-6, it can be seen that there are many regions in Saudi Arabia with high wind speed. Mean annual wind speed of the above-mentioned four windy regions of Saudi Arabia has been estimated to exceed 16.7 km/h with maximum of 16-19 km/h to 22 km/h. However, the region on the coastal area of the Red Sea can

produce density of at least 250 to 500 kWh/m², whereas same in inland area drops to about 50 kWh/m². Energy produced through wind power has been recognised to bring a variety of benefits. For a start, wind energy has been recognised to be pollution free, unlike other sources of energy such as coal, gas, petroleum and other fossil fuels (Saidur et al., 2011). Therefore, electricity produced through wind energy can directly contribute towards reduction of carbon dioxide, nitrogen oxide and sulphur dioxide and, thus, would directly contribute towards preserving the environment and reduce the impact of climate change. Economically, energy produced through wind power has been recognised to be relatively low-cost and could bring increased security for the supply of electricity from non-grid connected locations, thus providing protection against hike in the prices of electricity (Saidur et al., 2011).
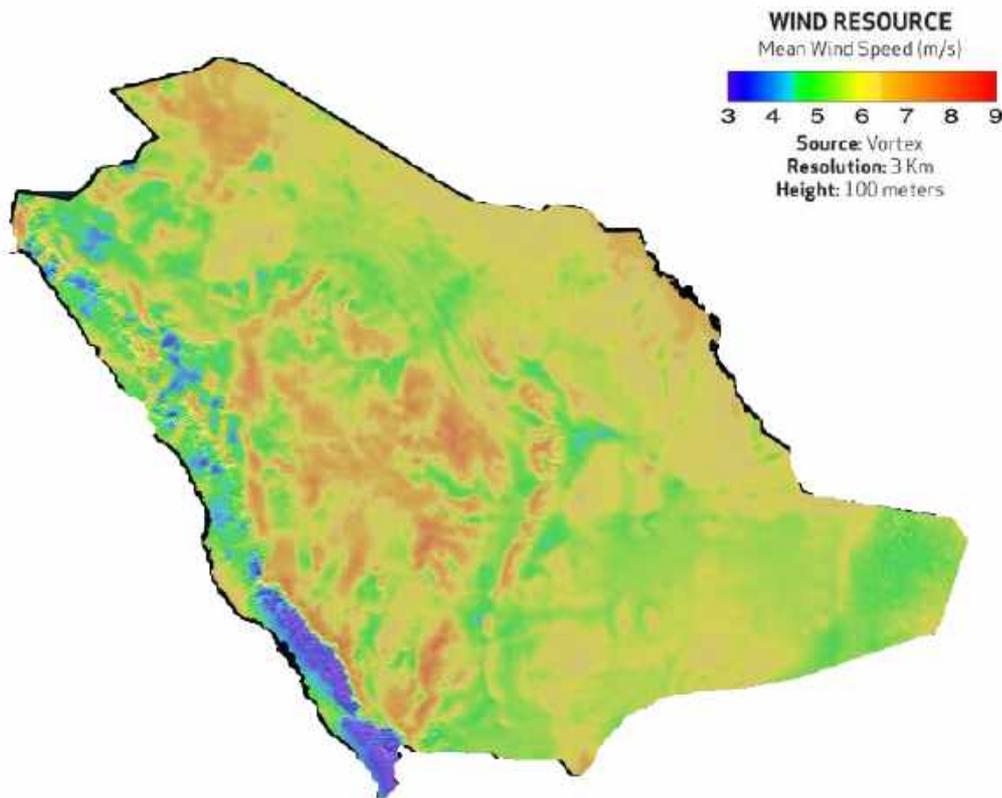
Figure 6-6 - Wind map of Saudi Arabia. Adapted from: Vortex, 2014

6.7. Challenges Using RES for Power Generation

While solar and wind power has been recognised as both economically and environmentally feasible solutions to produce electricity, questions have been raised on the reliability of RES for consistent electricity supply (Sovacool, 2009). A report by the US Energy Information Administration (EIA) suggested that energy produced from RES tends to be too scarce, too diffused and often located at distant locations. This limits the use of such resource as a consistent supply of electricity. Base-load power needs to be always on to meet minimum requirement of electricity demand. There is consensus among grid operators that the intermittency of RES is the main obstacle in the usage of wind and solar energy in the supply of electricity.

The western region of Saudi Arabia is located in geographically favourable location to obtain solar and wind energy. Nonetheless, weather condition cannot be guaranteed. This factor impacts plant performance. Furthermore, favourable locations for the installation of solar and wind power units are mostly located in very arid lands and sands. This situation is illustrated in Figure 6-7.

Figure 6-7 - Geographical locations for solar and wind power. Source: Google Earth.

The regions illustrated above indicate that locations for solar and wind power are characterised by dusty conditions, high level of humidity and lack of rain. Dust storms are regular in these locations. Such storms may potentially damage photovoltaic (PV) module, blocking irradiation to reach the cells via the glass cover and consequently impacting the electricity generation (Rahman et al., 2012; Aloughani, 2015). To bypass such situation, PV would require constant cleaning so that it may operate at the correct and planned conditions.

The situation mentioned above could become worse during the season of dust storms in which brings inhuman conditions that would make it difficult for the workers to clean the plants. According to data issued by the Climatology Department in Saudi Arabia, during the past decade, on average thirty-six dust storms were recorded in the western region of Saudi Arabia annually. It is expected that, in the event that solar plants are not maintained, they could wear out and even lose their efficiency over time (Al-Dowsari et al., 2014; Alharthi et al., 2017).

Hermann et al. (2015), in their study on the effect of dust on PV modules, concluded that, under the prevailing conditions that are embroiled by dust, it results in dust particles adhering to the surface of PV modules thus causing soiling that, in effect, causes substantial losses in PV yield. Variations in availability of solar and wind power energy sources in the Saudi Arabia means that there is a need to identify fit between RE and electricity demand, thus emphasising importance of load forecasting and modelling technique to forecast off-peak and peak load.

6.8. Weather Data

In this section it is presented collected data from Yanbu Airport and Jeddah Airport (wind data for 1 week), 14 years of wind data from the Authority of Meteorology and Environmental Protection. The other data (temperature, humidity) was obtained from Saudi Electrical Company. From the data collected, it was used 1 day wind data from Yanbu and Jeddah, 1 year of wind data (minimum, average and maximum speed average) from the Authority of Meteorology and Environmental Protection. From the Saudi Electrical Company it was used the data from 2010 to 2016.

6.8.1. Solar Energy Data

Data for the period of January 2010 to August 2016 have been collected from Saudi Electricity Company (SEC) and configured into eight indices, such as time, date (day, month and year), type of the day (weekday or weekend), temperature, humidity and, finally, load on grid station. While data range has been collected in hourly form, peak load reading has, however, been collected in half hour slots. Peak load time includes one o'clock in afternoon till four o'clock and then six o'clock in evening till nine o'clock. Therefore, during peak load time, data range is made in half-hourly intervals such as 13; 13.5, 14; 14.5, 15; 15.5 for the afternoon period and then 18; 18.5, 17; 17.5, 18; 18.5, 19; 19.5, 20; 20.5. Therefore, for the period from 21:00 in the night till the next afternoon 13:00 hourly data have been selected. The reasons for half hourly and hourly segregation of data have been provided in the previous chapter. Figure 6-8 below presents data on temperature for the year 2016 in Saudi Arabia. It can be seen from the data that temperature starts to rise from February, reaches to its maximum in June and starts to decline from September, thus providing eight months of high temperature. During summertime, temperature could reach to highs of 45^0C. According to estimates issued by the Saudi Meteorological Department, sunshine in Saudi Arabia can last for a maximum of 9.4 hours to a minimum 7.4 hours and, on average, it accounts for 8.90 hours per day and thus could prove to be as an asset for producing electricity from solar energy (Almasoud and Gandayh, 2015).
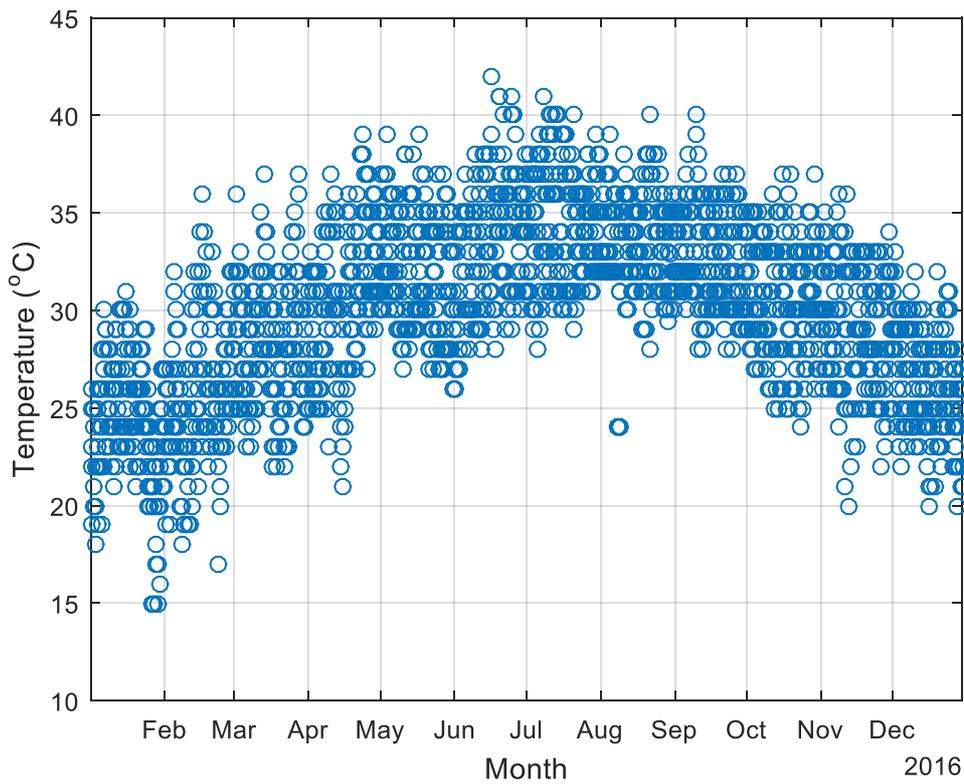
Figure 6-8 - Temperature during 2016 in the western region of Saudi Arabia.

Due to high temperature, Saudi Arabia also observes high level of humidity throughout the year, as shown in Figure 6-9 below. This fact causes increased demand, since people tends to make more use of air conditioning. Additionally, it is known that high levels of humidity can cause the formation of droplets in the PV panel, reducing its efficiency. Combined with temperature, these factors can cause accelerated degradation of the module. (Panjwani & Narejo, 2014). On the other hand, consistent solar radiation, long sunshine hours and relative humidity in the western region of Saudi Arabia makes the region suitable for producing electricity through using solar energy (Said et al., 2008).
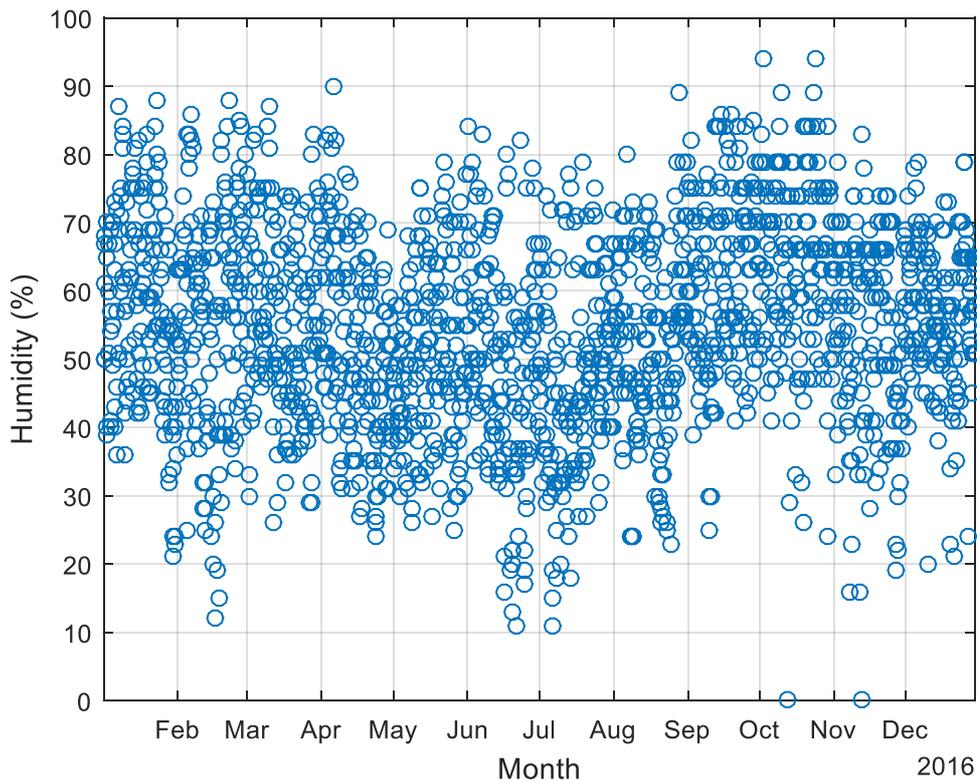
Figure 6-9 - Humidity level in western region of Saudi Arabia during 2016.

Identification of availability of solar energy is only part of the solution since the main determinant for integrating RES into the grid station is to meet the future demand for electricity loads. It is due to this reason that the first aim of the smart grid station designer is to assess the electrical load and load profile that RES should meet (Baras et al., 2012). Therefore, the aim of one-day-ahead forecasting electricity load in this chapter is to formulate, evaluate and analyse alternative plans for adding to the capacity of grid stations to serve the future electricity load.

There tends to be various factors that influence electricity load to go up and down during typical day and time of year. Variation in electric load is more significant for the western region of Saudi Arabia as it hosts two religious cities, Mecca and Medina, both of which attract millions of religious tourists every year, as explained in the previous chapter. Impact of environmental factors on electricity load can be seen in Figure 6-10 below that, starting from April till September, annual demand for electricity tends to be high. It is also

because of the high use of air conditioning in the summertime at both domestic and commercial level.
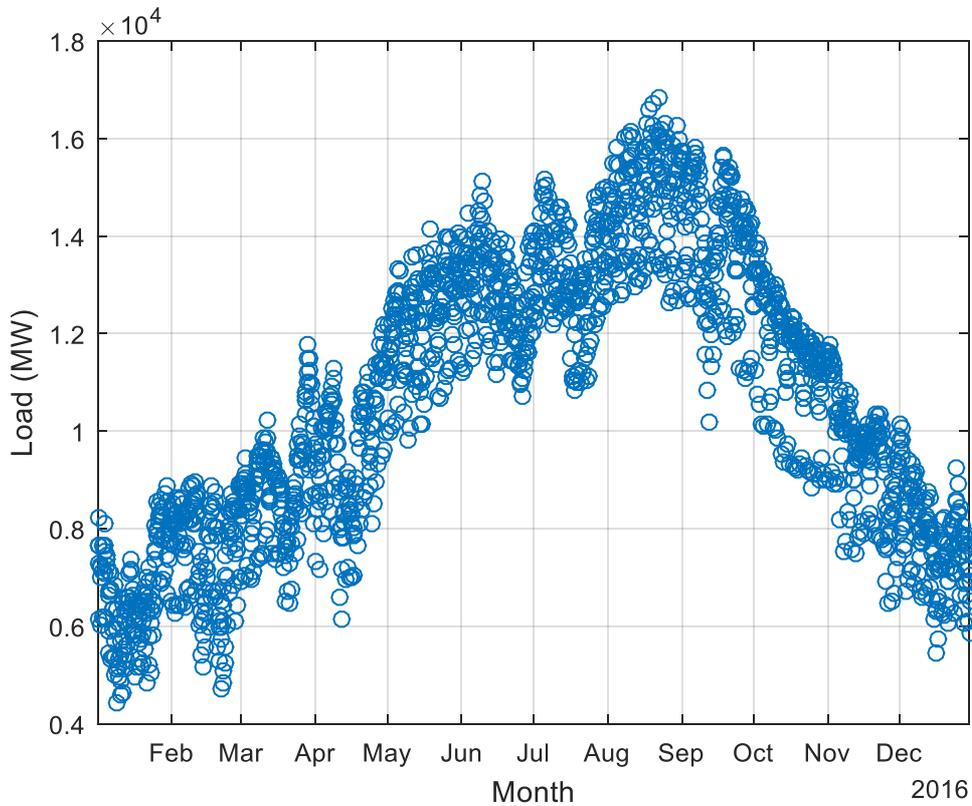
Figure 6-10 - Electricity load in western region of Saudi Arabia during 2016.

6.8.2. Wind Data

Among all the RES, wind energy has been identified to possess an edge over all others due to the requirement of simple infrastructure, maturity of technologies and low-cost (Rafique et al., 2018). Moreover, long-life of wind turbine, effectiveness in installation and operation has resulted in fast growing interest in wind power for producing electricity. However, like solar energy, electricity output produced through wind power cannot be relied upon for meeting load demand (Al-Busaidi et al., 2016). It means that there remains uncertainty and randomness in producing electricity through using wind power, both of which could impact normal generation and distribution of power.

Previous studies on using wind power to produce electricity in Saudi Arabia have largely focused on assessing the feasibility of wind power to produce consistent electricity, which includes implementation of wind power system, optimisation of wind system layout and analysis of wind power distribution and estimation across different regions of Saudi Arabia (Nematollahi et al., 2016; Al-Sharafi et al., 2017). Most past studies concluded that the effectiveness of utilisation of wind energy in electricity system can only be achieved through accurate forecast of wind data (Simao et al., 2017).

For the purpose of this study, wind data have been collected from Aviation Dispatch in Jeddah City, Saudi Arabia. Data consists in speed measurements (knots, later converted to km/h) covering one-day period such as between 1st May 2018 to 9th May 2018. METAR and Terminal Aerodrome Forecast (TAF) data have been gathered, also in unit knots (converted to km/h). METAR data include actual data on hourly weather data, which include hourly data on wind speed, time, gust and direction. In contrast, TAF is a format in which data on the weather condition are provided and include information on wind speed, gust and expected time for weather change and usually available for the period of the next 30 hours (Cunningham et al., 2014). Data of METAR's and TAF cover the area at sea and are reported in Zulu time. Time period of wind data includes the period between 00:00 and 23:00 local time in Saudi Arabia. All the data was obtained as plain data from reports. Figure 8 below shows actual data on wind speed for 20 hours on 7th May 2018. It can be seen from Figure 6-11 below that wind speed starts to pick up from seven o'clock in the morning before reaching to its peak at 1 o'clock and then starts to reduce.

104
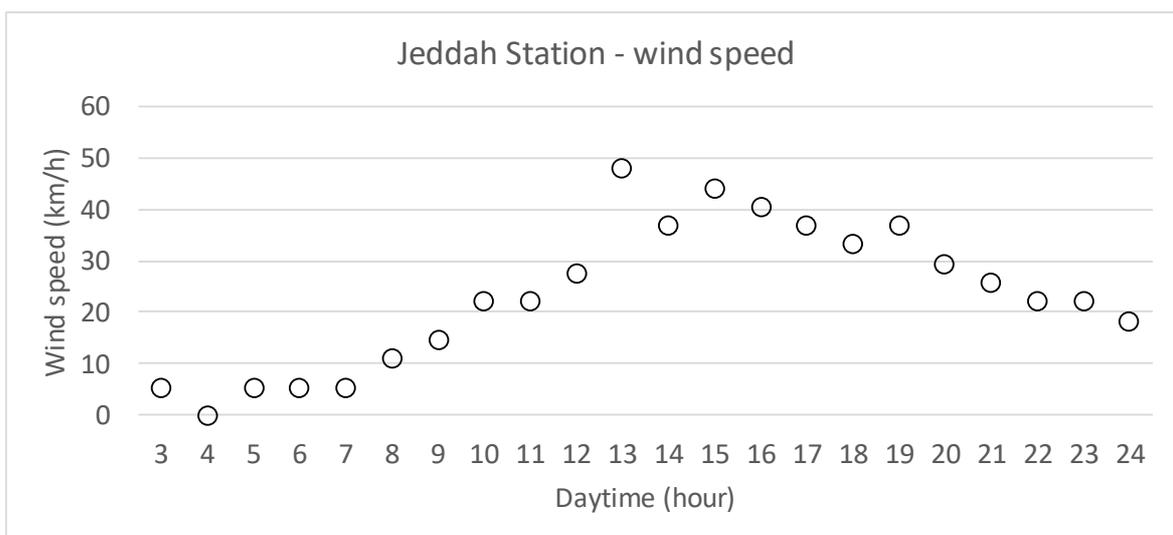
Figure 6-11 – 20-hour wind speed on Jeddah Station.
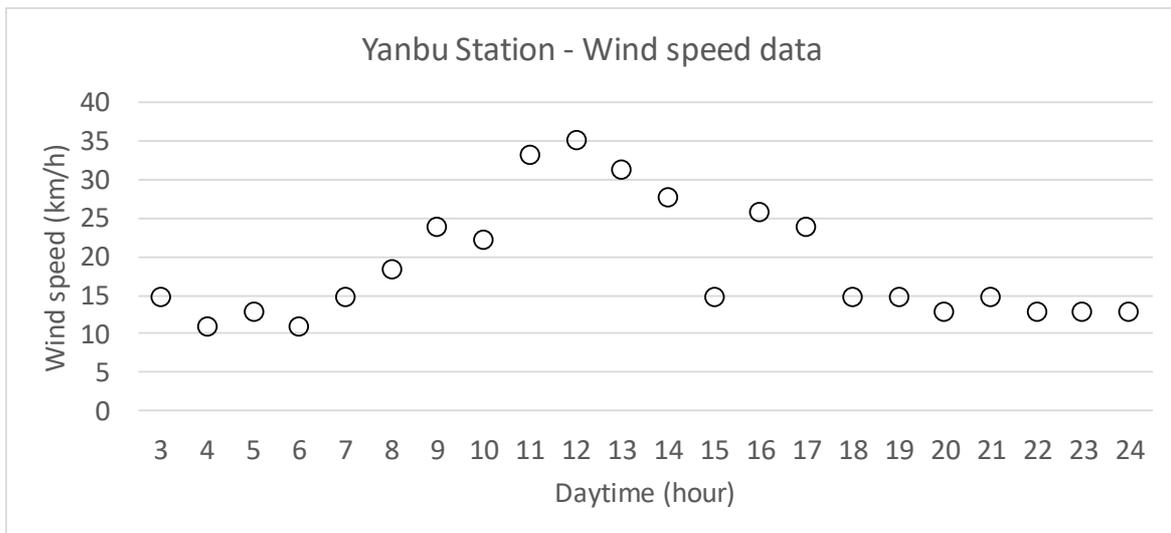
Figure 6-12 – 21 hour wind speed on Yanbu Station (7th May 2018).

Table 6.2 presents forecast for TAF data for one day prior to 6th May 2018. One may note that, from 9:00 to 15:00, wind speed is reported at 33.33 km/h with gusting wind. In contrast, from 00:15 to 9:00, while wind speed stays at 14.4 km/h, gusting wind runs at the speed of 33.33 km/h.

Table 6.2 - Wind forecast TAF one day before the actual

Time (24 Hour)	Wind speed (km/h)	Gusting wind (km/h)
9:00 to 15:00	33.33	0
15:00 to 9:00	14.4	33.33

Table 6.3 presents TAF wind forecast for one day after the one above (7th May 2018). From 9:00 to 15:00, wind speed was 11.11km/h with gusting speed of 29.63km/h. A one-hour window between 14:00 to 15:00 shows that, while wind speed stayed the same at 11.11 km/h, gusting speed, however, increased to 33.33 km/hr. Finally, for the rest of the time from 16:00 to 9:00, wind speed increased to 14.8km/h; however, gusting wind stayed same at 33.33 km/h.

105

Table 6.3: Wind forecast TAF one day after the actual (7th May 2018)

Time (24-Hour)	Wind speed (km/h)	Gusting wind (km/h)
9:00 to 15:00	11.11	29.63
14:00 to 15:00	11.11	33.33
16:00 to 9:00	14.8	33.33

Connecting RES into electricity grid station depends upon finding a match between ability of wind power to produce electricity at a given time and demand for electricity (Baras et al., 2012). Therefore, it is relevant to identify demand for electricity load in the same period for which wind speed is identified. Figure 6-13 identifies the hourly load demand on 7th of May 2018.
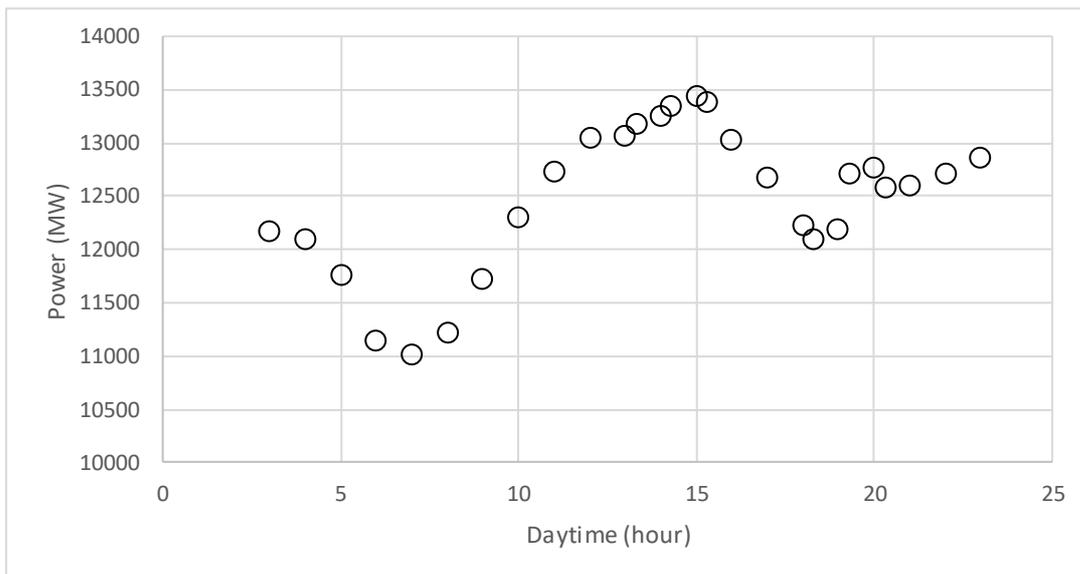
Figure 6-13 – 20-hour electricity load demand as on 7th May 2016

From the data shown above, it can be seen that both wind speed and electricity load demand increase at the same time, typically from 12:00 to 20:00. This pattern has remained

the same while considering other years, thus providing potential for integrating wind from 380 kV network and providing possibility to decrease carbon emission.

In relation to the western region of Saudi Arabia, both solar and wind energy could face problems in providing reliable electricity because the region is characterised by harsh weather, yet with good potential for wind energy. Specifically, the western region's advantage comes from its geographical location. The Red Sea is located on the west side, while the east is surrounded by mountain. That makes the region suitable for accumulating wind which is both stable and possesses enough strength to generate electricity.

Figure 6-14 presents average minimum wind speed during 2017. From the figure it can be seen that wind speed tends to fluctuate during the year, thus showing uncertainty issue related to the power of wind speed to produce electricity. Data for the average wind speed have been gathered from fifteen miles of surrounding area covering sea area, except the mountain.
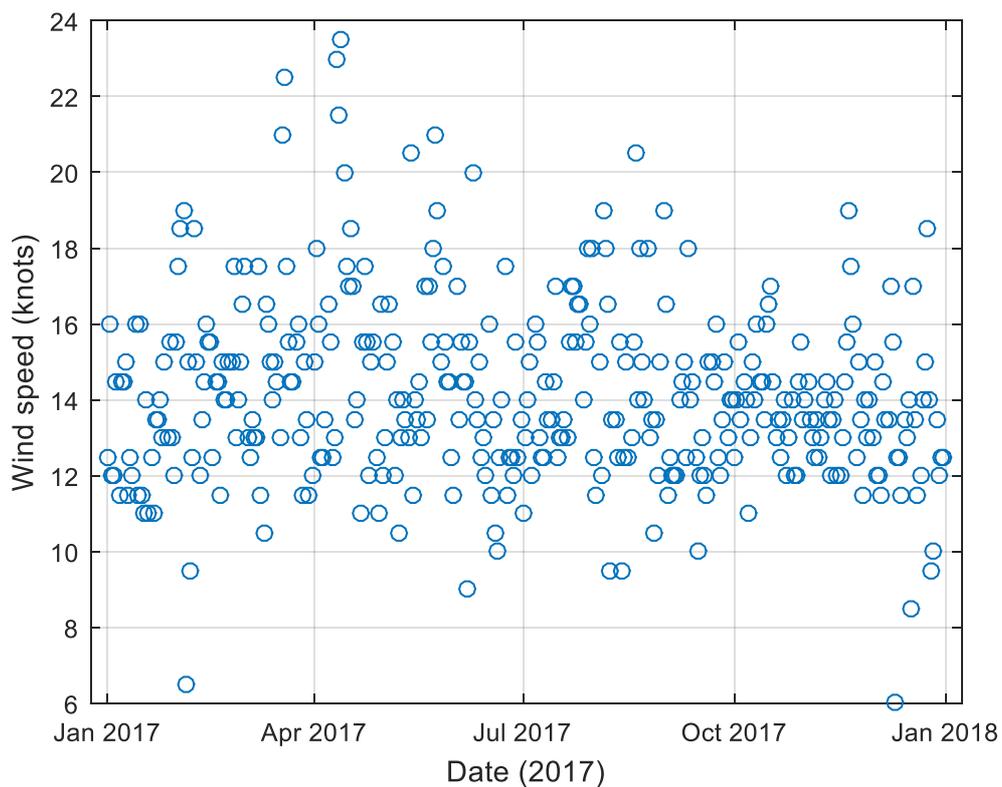
Figure 6-14 – Average minimum wind speed during 2017

For comparison purposes, Figure 6-15 shows the maximum wind speed in the same year.
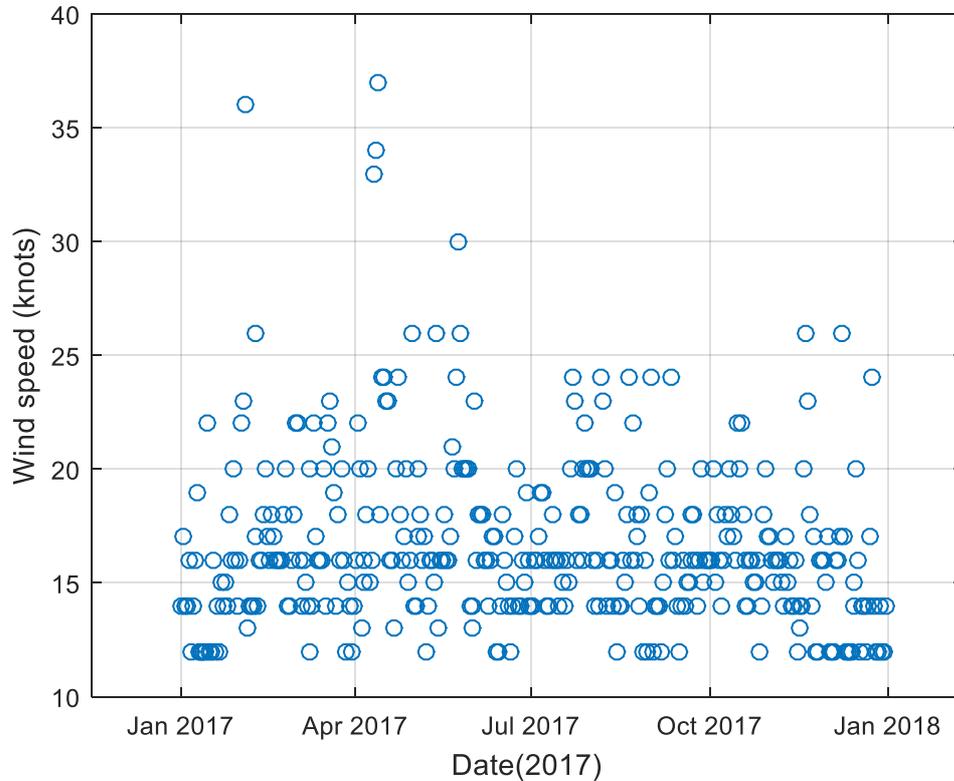
Figure 6-15 – Average maximum wind speed during 2017.

Relating data from previous figures, it can be seen that wind speed is maintained at 4-17 knots (2 – 8.7 m/s). On the other hand, wind turbines used in the region require 6 knots (3m/s) wind speed to start to produce electricity.

By relating the above figures showing wind speed at the region, one may note that maximum wind speed (at peak load time) is approximately 37 knots, indicating a safety margin of 31 knots to reach upper safety limits, identified to be 47 knots.

6.9. Data Modelling

In order to evaluate the possibility of integrating RES into the electric grid station data provided by SEC for summer, 2018 peak and base case final dates 07-11-2017 were interconnected. Data were then put under process via integrating wind platforms and solar panels farms into the grid transmission system of 380 kV. Integration then allowed running

of different power flow scenarios via integrating RES and testing the ability of grid station to meet the electricity load demand via ETAP® software.

Table 6.4 presents the amount of electricity generating capacity of three power plants in Jeddah, Rabigh and Shoiba using different technologies to produce electricity: steam, combined cycle and gas. A combined cycle (CC) power plant is one that can use both gas and steam turbine together to produce electricity using the same fuel as used by traditional simple-cycle plant (Petrakopoulou et al., 2012). Diesel and crude oil are the primary fuel types used in the three power stations and gas is used for steam units. The three PPs possess capacity to produce 18,820 MW when other grid stations, such as Yanbu, Mecca, Medina and Taif, are off services. However, these four grid stations possess capability to produce 1098 MW of electricity.

Table 6.4 - Electricity units generation capacity of grid stations

City	Steam unit	CC unit	Gas unit	Generation-based fuel (MW)	Generation on cycle combined (MW)
Jeddah	4	0	35	4026	0
Rabigh	10	15	44	7158	898
Shoiba	9	12	20	5538	1200

6.10. Simulation Results

This section presents investigation performed on the impact of RE in grid integration. Specifically, it evaluates the effect of switching off 17 and 33 gas units in Rabigh station, respectively, and integrates it with electricity produced from solar and wind sources. To this end, discussion evaluates the impact of two cases against five scenarios: 100%, 95%, 75%, 50% and 25% generation capacities of renewable energies during peak and off-peak times (Jeddah, Rabigh and Shoiba). For the purpose of simulation, data have been split into peak and off peak scenarios based on load forecasting data.

6.10.1. Case of Switching off 17 Gas Units

In this case, ETAP® software was fed with interconnected data of grid stations in Jeddah, Rabigh and Shoiba. Electricity generation capacity of two wind turbines in Rabigh and Shoiba and one solar panel in Medina was added. Table 6.5 below presents data when all units are in service during peak time with different load flows. Best-case scenario is at 100% generation from renewable energy sources. At this condition, 1,290MW of electricity can be produced.

Table 6.5 - Case of peak time from 13:00 to 21:00 next day: All units in service.

Peak incl. wind and PV	Jeddah (MW)	Rabigh (MW)	Shoiba (MW)	Wind turbine (MW)	PV (MW)	Total (MW)
100%	690	3,299	7,304	1,200	90	12,583
95%	690	3,299	7,304	1,140	90	12,523
75%	690	3,299	7,304	900	64	12,257
50%	690	3,434	7,304	600	42	12,070
25%	690	4,082	7,304	300	0.0098	12,376

With this scenario, the total energy produced is maintained above 12,000 MW by performing a trade-off between the energy produced by gas units in Rabigh and the energy from RES (wind and turbine). Each percentage value indicates a fraction of the potential RES load that can be inserted in the grid, being compensated by switching on/off gas units.

Table 6.6 presents data on similar conditions as presented above, but at off-peak time. It can be noticed that, at off-peak time, the capacity of Rabigh grid station reduces in

comparison to peak time, potentially due to low demand. Therefore, this would indicate that management has attempted to cool down the station.

Table 6.6 - Case of off-peak from 21:00-13:00 next day, all units in services.

Peak incl. wind and PV	Jeddah (MW)	Rabigh (MW)	Shoiba (MW)	Wind turbine (MW)	PV (MW)	Total (MW)
100%	690	2,301	7,304	1,200	90	11583
95%	690	2,301	7,304	1,140	90	11525
75%	690	2,301	7,304	900	64	12259
50%	690	2,301	7,304	600	42	10937
25%	690	2,301	7,304	300	0.0098	10595

Table 6.7 presents the impact of switching off 17 gas units at Rabigh station and substituting with RES.

Table 6.7 - Impact on generation capacity when 17 gas units of Rabigh are off, peak time 21:00-13:00 next day, MW

Peak incl. wind and PV	Jeddah (MW)	Rabigh (MW)	Shoiba (MW)	Wind turbine (MW)	PV (MW)	Total (MW)
100%	690	1,366	7,304	1,200	90	10650
95%	690	1,366	7,304	1,140	90	10590

75%	690	1,852	7,304	900	64	10810
50%	690	2,473	7,304	600	42	11109
25%	690	3,112	7,304	300	0.0098	11406

Gas units represent 12% of total power generation capacity. By changing the RES integration from 25% to 100%, the demand can be satisfied at low peak for values up to 11,406 MW, with no trade-off with traditional power plants It has been found that, in the event that RES are not able to meet the demand of electricity when 17 gas units are switched off, then gas units can get back into service within five minutes, thus this option is safer than using other electricity generation technologies such as steam.

Similarly, Table 6.8 shows that, during off-peak time when 17 gas units are switched off in Rabigh station, then at 100% capacity scenario a similar amount of electricity will be produced as in peak time. To achieve best safety measurement, it is recommended to put gas units in back-up mode so that they can get into the system within five minutes.

Table 6.8 - Case of off-peak during 21:00-13:00 next day, 17 gas units are out of service (MW)

Peak incl. wind and PV	Jeddah (MW)	Rabigh (MW)	Shoiba (MW)	Wind (MW)	PV (MW)	Total (MW)
100%	690	1,366	7,304	1,200	90	10,650
95%	690	1,366	7,304	1,140	90	10,590
75%	690	1,366	7,304	900	64	10,324
50%	690	1,366	7,304	600	42	10,002
25%	690	1,366	7,304	300	0.0098	9,660

Maintaining the RES in 100% can still satisfy 10,650 MW of load demand, even with all the 17 gas units out of service during 21:00 to 13:00 next day. Predictions provided by the artificial intelligent system will provide the required information for the decision management team to use this configuration. In the worst case, gas units can be easily be brought back to service in short time to satisfy increased demand.

6.10.2. Case of 33 Gas Units

In order to test the impact of switching off 33 gas units, ETAP® was used to integrate four turbines in Rabigh and Shoiba with two solar panels in Medina and Taif. Location of the solar panels was chosen because of favourable geography. Both Medina and Taif are in mountain with fewer sand deserts, thus, efficiency of producing clean solar energy is reported as high in these cities. Zell et al. (2015) argued that Taif and Medina are located 1518m and 610m elevation above the ground, thus meeting the requirements of global horizontal irradiance (GHI), diffuse horizontal irradiance (DHI), direct normal irradiance (DNI) and related meteorological parameters needed to capture solar energy.

Table 6.9 shows the scenarios for maximum electricity that can be produced when all the stations are in service at 100%, 95%, 75%, 50% and 25%. When compared with the results obtained for switching off 17 units, it can be clearly seen that the capability of wind and solar power doubles due to additional RES.

Table 6.9 - Case for peak time 13:00-21:00 next day, all units in service (MW)

Peak incl. wind and PV	Jeddah (MW)	Rabigh (MW)	Shoiba (MW)	Wind (MW)	PV (MW)	Total (MW)
100%	690	3,299	7,304	2,400	180	13,873
95%	690	3,299	7,304	2,280	180	13,573
75%	690	3,299	7,304	1,800	128	13,221
50%	690	3,434	7,304	1,200	84	12,712
25%	690	4,082	7,304	600	0.02	12,676

Table 6.10 presents the scenario of total electricity that can be produced when all the stations are in service at 100%, 95%, 75%, 50% and 25% generation capacity of RES at off-peak time.

Table 6.10 - Case for off-peak from 21:00-13:00 next day, all units in services (MW)

Peak incl. wind and PV	Jeddah (MW)	Rabigh (MW)	Shoiba (MW)	Wind (MW)	PV (MW)	Total (MW)
100%	690	2,301	7,304	2,400	180	12,875
95%	690	2,301	7,304	2,280	180	12,775
75%	690	2,301	7,304	1,800	128	12,223
50%	690	2,301	7,304	1,200	84	11,579
25%	690	2,301	7,304	600	0.02	10,895

Table 6.11 shows the impact of switching off 33 gas units during peak time from Rabigh station. This represents almost 25% of total production with generation capacity of 1,968MW. To overcome the deficit, energy produced from solar and wind turbines is injected into the services. Although RES were injected, it is important to consider that RES is an unreliable energy source. This is overcome by ensuring that gas units can get back into service within five minutes to meet the demand, if required. Such approach is only possible if gas units are in backup mode.

Table 6.11 shows results of five scenarios when 33 gas units are off during peak time. In this scenario, electricity produced through RES are injected at 100%, 95%, 75%, 50% and 25% capacity. This enables the RES supply of 2,580MW, 2,460 MW, 1,928MW, 1,284MW and 600MW of electricity into the system, respectively. Switching off 33 gas

114

units and keeping them in stand-still situation is a safer option than generating electricity from other sources, such as steam.

Table 6.11 - Case off-peak from 13:00 to 21:00 next day, switching off 33 gas units from service.

Peak incl. wind and PV	Jeddah (MW)	Rabigh (MW)	Shoiba (MW)	Wind (MW)	PV (MW)	Total (MW)
100%	690	1,538	7,304	2,400	180	12,112
95%	690	2,312	7,304	2,280	180	12,766
75%	690	2,573	7,304	1,800	128	12,495
50%	690	3,433	7,304	1,200	84	12,711
25%	690	4,067	7,304	600	0.02	12,661

Table 6.12 presents the scenario of switching off 33 gas units and injecting part of maximum potential (5,000 MW) of electricity from RES into the grid during off-peak time.

Table 6.12 - Case of off-peak from 21:00-13:00 next day, 33 units of gas units out of service (MW).

Peak incl. wind and PV	Jeddah (MW)	Rabigh (MW)	Shoiba (MW)	Wind (MW)	PV (MW)	Total (MW)
100%	690	1,538	7,304	2,400	180	12,112
95%	690	1,538	7,304	2,280	180	11,992
75%	690	1,538	7,304	1,800	128	11,460
50%	690	1,538	7,304	1,200	84	10,816
25%	690	1,538	7,304	600	0.02	10,123

6.11. Network Generation Quality

When integrating with RES, quality of network generation will depend on solar and wind intensity (Khadem and Conlon, 2010). Other factors that could influence network quality include PV modules, filters controlling mechanism and turbine geometry. Based on a 380kV generation and transmission network, it can be observed that the system has high technology, using cycle, combined and gas units (Ganjenhkaviri et al., 2015). However, having a backup plan is essential for emergency situations.

The results obtained from the investigated scenarios (switching off 17 and 33 gas units) suggest that, in the event of variation in electricity produced through RES, it may take merely five minutes for gas units to get back into service. This configuration provides safe and reliable network generation quality. Findings in this chapter prove that the system can switch off 33 gas units and use the grid station in combined cycle. This configuration may provide a high quality electricity network generation system.

6.12. Case Study – Forecasting Operation Plan

The complete energy management system is supposed to rely on a decision support framework called Forecasting Operation Plan and conceptualised as shown in Figure 6-16
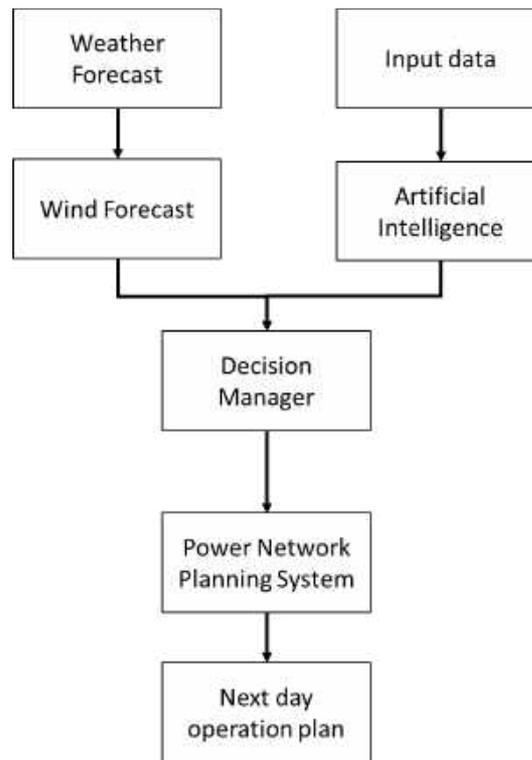
Figure 6-16 Complete operation plan for decision-making support.

The complete system consists of a series of stages, as can be seen above. A description of it follows.

The first step involves performing load forecasting using available data (time series - previous loads) and artificial intelligence algorithms, as shown in Chapter 4 as well as Chapter 5. As an example, the AI selected would FIS with ANN (obtained in Chapter 5), since this model was elected as the best model. Weather forecast data are used to predict wind and solar irradiation for the next day. Thus, the artificial intelligence block consists of both load and weather forecasting.

This data are then provided to the decision manager (or management system), which, by knowing the available energy sources and the forecasted load, uses a software (such as ETAP) to simulate the possibility of incorporating the RES units into the grid, as a replacement of the fossil fuel units.

The operation plan is completed by switching certain gas units into backup mode (according to forecasted availability and demand), thus using RES in the place of fossil fuel units.

118

This configuration of the system shows that the artificial intelligence algorithm works as a recommendation system, estimating the future load demand and providing the decision maker the possibility of configuring the system one-day-ahead, optimising the usage of RES and minimising the potential usage of fossil fuels.

Consider the following example of operation plan recommendation based on the available forecasts and power network planning.

For a future load of 12,000 MW (to be observed), the artificial intelligence algorithm provides load forecasting that may fall approximately in the range 1,100 and 1,300 MW, as shown in Figure 6-17.
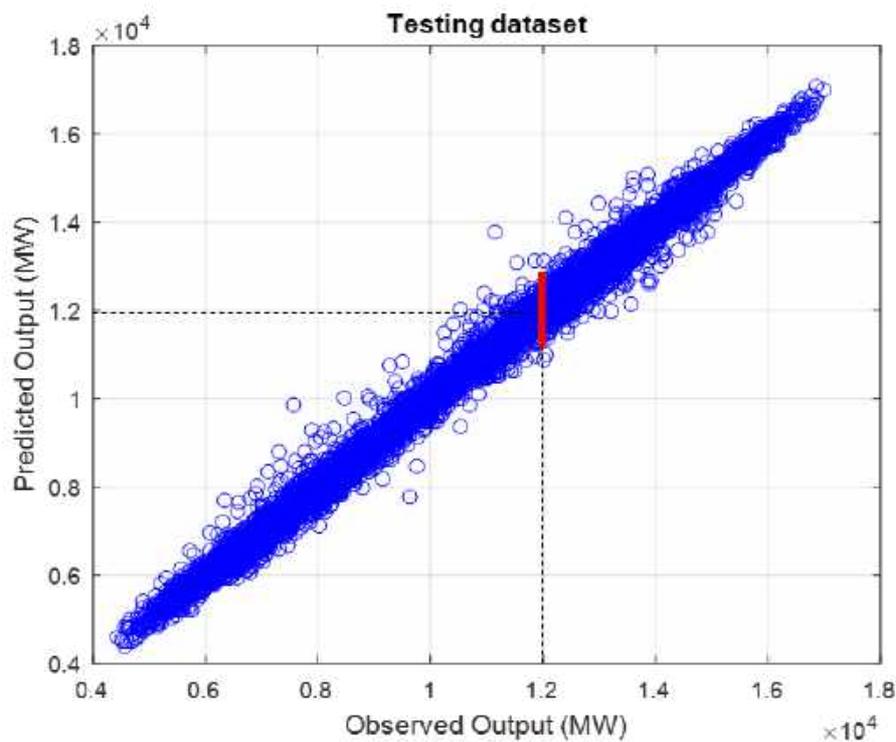
Figure 6-17 – Possible predicted values for an observed output of 12,000 MW using ANN with time series input.

For the same next day, the weather forecasting system predicted a potential production of 1,800 MW for station in service at 75% (peak hours).

Knowing this information, the decision manager can use power network modelling software, such as ETAP, to maximise the use of RES while minimising fossil fuels. With 1,800 MW of wind power, the fossil fuel usage can be reduced to 10,200 MW to meet the

load. Still, gas units are kept at backup mode to meet the 1,800 MW of required load in the case that wind power fails, i.e. the production is lower than the predicted.

The example above illustrates the operation steps of the decision-making system. It is important to highlight that such system may work with peak/off-peak information and wind/solar power depending on the data availability, prediction confidence and load levels.

6.13. Summary

This chapter presented the investigation of possibilities on injection of wind and solar energy in the transmission lines. Although it used real data collected from the western region of Saudi Arabia, the same analysis may be extended to other parts of the country to generate complete power system scheduling and optimisation with the inclusion of the renewable energy resources. An ideal configuration is to use incremental learning. In this way, knowledge is incremented rather than discarded. The model can be integrated with the weather conditions in order to predict the power requirements and the generated renewable energy for a particular day.

The software used for modelling was ETAP®. For the purpose of simulation, focus was given to seven grid stations operating on fossil fuels to produce electricity. These were located in Jeddah, Rabigh, Yanbu and Shoiba, with total generation capacity of 19,933 MW. Other stations, such as Yanbu, Makkah, Medina and Taif, were not used as authorities switched off these stations, since they run on diesel fuel, which has become unfeasible for the authorities. However, data from Medina and Taif solar and wind stations, respectively, were used for the purpose of simulation.

Investigation revealed that the region contains potential sites of solar and wind energy production. However, the arid nature of land complemented with dusty deserts and intermittence in solar sour causes hurdles in usage of RES for electricity production. To this end, electricity load forecast of one-day-ahead was used to identify possible fit of smart grid (grids working on combined cycle, such as using both RES and fossil fuels) to meet short-term load forecast predicted by the previous models presented in Chapter 4.

Solar energy and wind data were collected from the Saudi Meteorological Department and showed a promising possibility of using RES to meet electricity load

demand. For instance, it has been found that, on average, the western region of Saudi Arabia gets approximately nine hours of sunlight per day. Moreover, METAR and TAF wind data revealed consistent wind forecast and gust wind for one week.

Testing of two scenarios, such as switching off 17 and 33 gas units with RES, provided positive results for grids to meet electricity demand. In both cases it was suggested that gas units would be kept in backup mode to meet any eventual disturbance caused by variation in energy supply by RES. The backup mode enables the rapid reactivation of the gas units, providing safer electricity generation capability when compared to use of other sources for producing electricity, such as steam. Having backup gas units would ensure consistent provision of electricity to consumers, thus improving network generation quality and assisting management in better planning.

Limitations were identified for this approach. One of them is linked to the issues on wind data. Currently, wind power from mountain data, which is the responsibility of Jeddah airport, is not available. Moreover, the findings presented above do not include information from offshore stations. This occurred because the Meteorology and Environmental Protection Department did not provide such data, despite the fact that it could have been made available from coastal areas in Jeddah and Yanbu stations.

Chapter 7.

Conclusions and Future Work

7.1. Conclusions

The present research investigated different methodologies and techniques to predict energy load. The research was tested with data provided by SEC in the western region of Saudi Arabia. Two approaches were used:

- Time series as model input;

- Multivariate data as model input.

The first methodology evaluated was to investigate the applicability of Linear Regression model in energy load forecasting. The second methodology used Artificial Neural Network for the same task.

From this analysis, it was concluded that, using time series input, the Artificial Neural Network showed higher performance than Linear Regression. Specifically, the neural network using 15 neurons in one hidden layer showed MSE of 5.5×10^5, while LR had MSE of 7.5×10^5, both evaluated in the testing set. Additionally, among the neural network architectures evaluated, the one using 15 neurons in one hidden layer showed superior performance when compared with 10 and 20 neurons in one hidden layer, as well as 20-10 and 20-15 in two hidden layers. The difference in performance for all neural networks lies in a broad range, from values of MSE equal to 3.7×10^5 MW^2 to 6×10^5 MW^2. The model showing the lowest performance was 20-15 in two hidden layers, with MSE of approximately 6×10^5 MW^2 in testing set.

The second methodology investigated to predict energy load was the compilation of a combiner model, involving either a Weighted Average, PSO or Fuzzy as the superior state model combined with two different model structures:

- Linear Regression (LR);

- Artificial Neural Network (ANN);

122

It was noted that every combiner model was able to compensate the low performance of one model (in general, multivariate model) through the decrease on the performance of the best model. In this way it generated an averagely reasonable model.

Among the combiner models, the fuzzy model with ANN was the one which showed the highest performance, with relatively lower errors when compared with the other approaches, However, it is important to highlight that several modifications on the architecture may lead to a better performance, since such type of model possesses several degrees of freedom in its hyperparameters. Still, it was able to reach a performance on the testing set of MSE $5.8*10^5$ MW^2 with 93% of agreement (R^2) between training and testing set.

To investigate power system scheduling and optimisation, a model of network grid was developed on ETAP®. Data were provided by SEC for the western region of Saudi Arabia. The uniqueness of this data lies in the traits of this region compared with other parts of Saudi Arabia. Firstly, the climate is not desert, but mountainous and coastal, which brings different wind patterns and high humidity. In total, air condition affects 70% of the demand of the region. Also, this same factor influences people's behaviour, which is mostly indoor at day and outdoor at night, avoiding the hot weather. Also, in the summer, there is much use of air conditioners in local buildings, which highly increased energy demand.

Another unique trait of the region is the peak and off-peak times of temperature and wind. From the data, it was observed that both start to rise at 13:00 and decrease at 15:30, with a second peak on the day from 18:30 to 20:30. The same pattern repeats for mostly the whole year.

Finally, the event of Hajj attracts millions of visitors to Makkah-Medina, which deeply affects the energy demand in the area. Reports mention up to 15 million of visitors at the same epoch in the Muslim calendar due to religious tourism.

Therefore, with all these factors in consideration, the aim was to adapt the power grid according to the prediction of one-day-ahead load forecast. This means that, by knowing the one-day-ahead prediction, we can integrate RES and rely on the switch of the gas units to reduce carbon emission. Ultimately, the model may be reused for different regions by modifying the input data. This gives generalisation capabilities to this system. Instead of

implementing different systems, the main idea would be to have a unified structure, so that the knowledge derived from one application may be used in another case, thus characterising incremental learning.

Weather conditions can be integrated with the model for prediction of power generation, according to different weather conditions and the generated renewable energy for one-day-ahead.

With this purpose, the research compared the injection of wind with solar energy into transmission lines. It found various challenges faced by this region of Saudi Arabia, which may be attributed to a variety of local factors, namely high temperature, high humidity level and sandstorms. All of these may contribute to low performance of solar panels in the region. However, the wind potential is currently underestimated and there is reportedly potential for improvement in the use of green energy as a replacement of fossil energy sources. It should also be highlighted that the use of RES is easy to integrate due to the robustness of the system. By robustness it is meant that, even under load or adverse weather conditions, the generation by decision management can easily switch a number of gas units from backup mode to active mode within five minutes. Even if gas units are completely switched off, they can easily get back to the grid within 20 minutes in the worst scenario.

For a realistic scenario, it was tested different levels of RES integrated, from full load production (100%) to a more conservative approach, using only 25% and compensating in peak times with gas units. To implement this, the management system relies on the artificial intelligent systems presented in the previous chapters (ANN, LR and Hybrid models) to obtain a reliable prediction of future load demand.

7.2. Future Work

The current research focused on short-term (one-day-ahead) load forecast, applied in the Western region of Saudi Arabia. The aim is to integrate renewable energy sources to the grid. For future work, the expansion of this system is suggested, for application in different regions of the world, subjected to a variety of climate conditions and with different load demand patterns. Additionally, the investigation of long-term and not just short-term forecast may improve the decision support system, helping managers to rely on a decision process with a bigger time window.

Another suggestion for future work is to couple load forecasting with energy market dynamics. It is expected in the future that other companies may join the Saudi Arabian energy market, thus enforcing competition and pricing dynamics. Coupling the energy market with load forecasting would enable the balance between production at peak and off-peak times with energy pricing, thus increasing profit and reducing unnecessary costs.

Finally, investigation of the costs associated with wind farm building and weather data can improve the feasibility analysis of RES unit installation in both private and public sectors. Such feasibility can be associated with the negative impact of the CO_2 emission in order to upgrade decision process management.

References

[1].   Agatonovic-Kustrin, S. and Beresford, R. 2000. Basic concepts of artificial neural network (ANN) modelling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical analysis*. 22, pp. 717-727.

[2].   Ahmad, M. W., Mourshed, M. and Rezgui, Y. 2017. Trees vs neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy and Buildings*, 147, pp. 77–89.

[3].   Al-Ammar, E. and Al-Yousef, B. 2010, March. Feasibility study of establishing a wind farm to generate electricity in Saudi Arabia. In *Nuclear & Renewable Energy Conference (INREC), 2010 1st International* (pp. 1-5). IEEE.

[4].   Al-Arabia News. 2018. H*all sites of Mina register new record of high electric load during Eid.* [online] Available at

[5].   http://english.alarabiya.net/en/business/energy/2018/08/22/Hajj-sites-of-Mina-register-new-record-of-high-electric-load-during-Eid.html (Accessed:9th Jan 2019).

[6].   Al Busaidi, A.S., Kazem, H.A., Al-Badi, A.H. and Khan, M.F. 2016. A review of optimum siting of hybrid PV–Wind renewable energy systems in Oman. *Renewable and Sustainable Energy Reviews*, *53*, pp. 185-193.

[7].   Al Garni, H., Kassem, A., Awasthi, A., Komljenovic, D. and Al-Haddad, K. 2016. A multicriteria decision making approach for evaluating renewable power generation sources in Saudi Arabia. *Sustainable Energy Technologies and Assessments*, *16*, pp.137-150.

[8].   Al-Sharafi, A., Sahin, A.Z., Ayar, T. and Yilbas, B.S. 2017. Techno-economic analysis and optimisation of solar and wind energy systems for power generation and hydrogen production in Saudi Arabia. *Renewable and Sustainable Energy Reviews*, *69*, pp. 33-49.

[9].   Al-Shareef, A.J. and Abbod, M.F. 2005. Neural Network Complexity against Simplicity Electrical Load Forecast Case Study. In *Power Engineering Conf*, UPEC.

[10]. Al-Shareef, A.J., Mohamed, E.A. and Al-Judaibi, E. 2008. One hour ahead load forecasting using artificial neural network for the western area of Saudi Arabia. *International Journal of Electrical Systems Science and Engineering, 1*(1), pp. 35-40.

[11]. Al-Sughai, S. 2014. Energy waste cost local economy SR135bn annually. Arab news, Available at http://www.arabnews.com/economy/news/621821 [Accessed on 22nd Nov 2018].

[12]. AlDowsari, A., Bkayrat, R., AlZain, H. and Shahin, T. 2014. Best practices for mitigating soiling risk on PV power plants. In *2014 Saudi Arabia Smart Grid Conference (SASG)* (pp. 1-6). IEEE.

[13]. Alfares, H.K. and Nazeeruddin, M. 2002. Electric load forecasting: literature survey and classification of methods. *International Journal of Systems Science, 33*(1), pp. 23-34.

[14]. Alharthi, Y.Z., Siddiki, M.K. and Chaudhry, G.M. 2017. The New Vision and the Contribution of Solar Power in the Kingdom of Saudi Arabia Electricity Production. In *2017 Ninth Annual IEEE Green Technologies Conference (GreenTech)*(pp. 83-88). IEEE.

[15]. Almasoud, A.H. and Gandayh, H.M. 2015. Future of solar energy in Saudi Arabia. *Journal of King Saud University-Engineering Sciences*, *27*(2), pp. 153-157.

[16]. Aloughani, M. 2015. *Renewable energies management strategy challenges in the Arabian Gulf countries* (Doctoral dissertation, Brunel University London).

[17]. Amasyali, K. and El-Gohary, N.M. 2018. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews, 81*, pp. 1192-1205.

[18]. Arab News. 2018. *Hajj 2018: More than 2 million pilgrims begin journey of a lifetime*. [online] Available at http://www.arabnews.com/node/1358626/saudi-arabia (Accessed: 9th Jan 2019).

[19]. Aryaputera, A.W., Yang, D., Zhao, L. and Walsh, W.M. 2015. Very short-term irradiance forecasting at unobserved locations using spatio-temporal kriging. *Solar Energy*, *122*, pp. 1266-1278.

[20]. Azli, N.A., Nayan, N.M. and Ayob, S.M. 2013. Particle swarm optimisation and its applications in power converter systems. *International Journal of Artificial Intelligence and Soft Computing, 3*(4), pp. 372-386.

[21]. Azad, M.K., Uddin, S. and Takruri, M. 2018. Support vector regression based electricity peak load forecasting. In *Mechatronics and its Applications (ISMA), 2018 11th International Symposium on* (pp. 1-5). IEEE.

[22]. Badri, A., Ameli, Z. and Birjandi, A.M. 2012. Application of artificial neural networks and fuzzy logic methods for short term load forecasting. *Energy Procedia*, *14*, pp. 1883-1888.

[23]. Barak, S., Sadegh, S. S. 2016. Forecasting energy consumption using ensemble ARIMA–ANFIS hybrid algorithm. *International Journal of Electrical Power & Energy Systems, 82*, pp. 92-104.

[24]. Baras, A., Bamhair, W., AlKhoshi, Y., Alodan, M. and Engel-Cox, J. 2012. Opportunities and challenges of solar energy in Saudi Arabia. In *World Renewable Energy Forum", Denver* (1, p. 4721).

[25]. Barghinia, S., Kamankesh, S., Mahdavi, N., Vahabie, A. H., and Gorji, A. A. 2008. A combination method for short term load forecasting used in Iran electricity market by NeuroFuzzy, Bayesian and finding similar days methods. In *2008 5th International Conference on the European Electricity Market* (pp. 1-6). IEEE.

[26]. Barta, G., Nagy, G., Papp, G. and Simon, G. 2016. Forecasting framework for open access time series in energy. In *Energy Conference (ENERGYCON), 2016 IEEE International* (pp. 1-6). IEEE.

[27]. Baucells, M. and Borgonovo, E. 2013. Invariant probabilistic sensitivity analysis. *Management Science*, 59(11), pp. 2536–2549.

[28]. Bessa, R., Moreira, C., Silva, B. and Matos, M. 2014. Handling renewable energy variability and uncertainty in power systems operation. *Wiley Interdisciplinary Reviews: Energy and Environment, 3*(2), pp. 156-178.

[29]. Bhattay, A. 2017. *Hajj: Howe to make this thee million strong Muslim pilgrimage environmentally friendly*. The Independent, [online] Available at https://www.independent.co.uk/travel/asia/hajj-pilgrimage-green-guide-mecca-medina-muslims-environmental-friendly-travel-islam-sustainable-a7930676.html (Accessed: 9th Jan 2018).

[30]. Borgonovo, E. and Plischke, E. 2016. Sensitivity analysis: a review of recent advances. *European Journal of Operational Research*, *248*(3), pp. 869-887.

[31]. Box, G.E., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. 2015. *Time series analysis: forecasting and control*. John Wiley & Sons.

[32]. Braun, M.R., Altan, H. and Beck, S.B.M. 2014. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Applied Energy*, *130*, pp. 305-313.

[33]. Brown, J.J. and Foucher, S., 2008. A quantitative assessment of future net oil exports by the top five net oil exporters. *Energy Bulletin. January 8*.

[34]. Byun, J., Hong, I., Kang, B. and Park, S. 2011. A smart energy distribution and management system for renewable energy distribution and context-aware services based on user patterns and load forecasting. *IEEE Transactions on Consumer Electronics*, *57*(2).

[35]. Çevik, H.H. and Çunkaş, M. 2015. Short-term load forecasting using fuzzy logic and ANFIS. *Neural Computing and Applications*, *26*(6), pp. 1355-1367.

[36]. Che, J., Wang, J. and Wang, G. 2012. An adaptive fuzzy combination model based on self-organising map and support vector regression for electric load forecasting. *Energy*, 37(1), pp.657-664.

[37]. Chen, Y., Luh, P.B., Guan, C., Zhao, Y., Michel, L.D., Coolbeth, M.A., Friedland, P.B. and Rourke, S.J. 2010. Short-term load forecasting: similar day-based wavelet neural networks. *IEEE Transactions on Power Systems*, *25*(1), pp. 322-330.

[38]. Chen, J., Zhang, B., Wang, B., Qinglai, G. 2013. A Spinning Reserve Allocation Method for Power Generation, Dispatch Accommodating Large-Scale Wind Power Integration. *Energies, 6*, pp. 5357-538.

[39]. Cho, H., Goude, Y., Brossat, X. and Yao, Q., 2013. Modelling and forecasting daily electricity load curves: a hybrid approach. Journal of the American Statistical Association,108(501), pp. 7-21.

[40]. Climate Transparency. 2015. *Brown to green: G20 transition to a low carbon economy: Saudi Arabia.* [online] Available at: https://www.climate-transparency.org/wp-content/uploads/2016/09/Saudi-Arabia_Country-Profile.pdf (Accessed: 22nd Dec 2018).

[41]. Creswell, J. 2013. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches,* 4th edition. SAGE.

[42]. Cunningham, J., Asencio, M., Lent, S. and Brinton, C. 2014. The categorization of wind shift events using metar and personal weather station data. In *2014 Integrated Communications, Navigation and Surveillance Conference (ICNS) Conference Proceedings* (pp. S3-1). IEEE.

[43]. Deihimi, A. and Showkati, H. 2012. Application of echo state networks in short-term electric load forecasting. *Energy*, 39(1), pp. 327-340.

[44]. Deka, A., Hamta, N., Esmaeillian, B and Behdad, S. 2015. Predictive modelling techniques to forecast energy demand in the United States: A focus on economic and demographic factors. *Proceeding of the AME 2015 International design engineering technical conferences and computers and information in engineering conference*, IDETC/CIE.

[45]. Deng, J. 2010, August. Modelling and prediction of China's electricity consumption using artificial neural network. In *Natural Computation (Icnc), 2010 Sixth International Conference on* (4, pp. 1731-1733). IEEE.

[46]. Devenish, B.J., Francis, P.N., Johnson, B.T., Sparks, R.S.J. and Thomson, D.J. 2012. Sensitivity analysis of dispersion modelling of volcanic ash from Eyjafjallajökull in May 2010.*Journal of Geophysical Research: Atmospheres*, *117*(D20).

[47]. Diamantoulakis, P.D., Kapinas, V.M. and Karagiannidis, G.K. 2015. Big data analytics for dynamic energy management in smart grids. *Big Data Research*, *2*(3), pp. 94-101.

[48]. Ding, S. and WU, Q.H. 2012. Performance Comparison of Function Approximation Based on Improved BP Neural Network [J]. *Computer and Modernization*, *11*, pp.10-13.

[49]. Draper, N.R. and Smith, H. 2014. *Applied Regression analysis,* vol. 326. John Wiley & Sons.

[50]. Du, P., Wang, J., Yang, W. and Niu, T. 2018. Multi-step ahead forecasting in electrical power system using a hybrid forecasting system. *Renewable Energy, 122*, pp.533-550.

[51]. EIA, 2014. *Country Analysis Brief: Saudi Arabia. US Energy Information Administration*. Available at: http://www.eia.gov/countries/cab.cfm?fips=sa Accessed on 20th June 2019.

[52]. Ertugrul, Ö.F. 2016. Forecasting electricity load by a novel recurrent extreme learning machines approach. *International Journal of Electrical Power & Energy Systems*, *78*, pp. 429-435.

[53]. Estimo, R. C. Jr. 2016. *German delegation to promote solar power, water desalination*. Arab News [online] Available at: http://www.arabnews.com/news/465863 (Accessed: 22nd Dec 2018).

[54]. Fan, C., Xiao, F. and Wang, S. 2014. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Applied Energy*, *127*, pp. 1-10.

[55]. Faul, F., Erdfelder, E., Buchner, A. and Lang, A.G. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behaviour Research Methods, 41*(4), pp. 1149-1160.

[56]. Feilat, E.A. and Bouzguenda, M. 2011, December. Medium-term load forecasting using neural network approach. In *Innovative Smart Grid Technologies-Middle East (ISGT Middle East), 2011 IEEE PES Conference on* (pp. 1-5). IEEE.

[57]. Foucquier, A., Robert, S., Suard, F., Stéphan, L. and Jay, A. 2013. State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, *23*, pp. 272-288.

[58]. Fumo, N. 2014. A review on the basics of building energy estimation. *Renewable and Sustainable Energy Reviews*, *31*, pp. 53-60.

[59]. Fumo, N. and Biswas, M.R. 2015. Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews*, *47*, pp.332-343.

[60]. Gajowniczek, K. and Ząbkowski, T. 2016. Short term electricity forecasting based on user behaviour from individual smart meter data. *Journal of Intelligent & Fuzzy Systems, 30*(1), pp. 223-234

[61]. Ganguly, P., Kalam, A. and Zayegh, A. 2017. Short term load forecasting using fuzzy logic. In *International Conference on Research in Education and Science*.

[62]. Ganjehkaviri, A., Jaafar, M.M. and Hosseini, S.E. 2015. Optimisation and the effect of steam turbine outlet quality on the output power of a combined cycle power plant. *Energy Conversion and Management*, *89*, pp. 231-243.

[63]. Garulli, A., Paoletti, S. and Vicino, A. 2015. Models and techniques for electric load forecasting in the presence of demand response. *IEEE Transactions on Control Systems Technology*, *23*(3), pp. 1087-1097.

[64]. Government.ae. 2019. *Electricity*. The official portal of the UAE Government/ [online] Available at: https://government.ae/en/information-and-services/environment-and-energy/electricity (Accessed: 22nd Jan 2019_.

[65]. Groscurth, H.M. and Kress, K.P. 1998. Fuzzy data compression for energy optimisation models. *Energy*, *23*(1), pp. 1-9.

[66]. Hamdan, S., Binkhatim, S., Jarndal, A. and Alsyouf, I. 2017. On the performance of artificial neural network with sine-cosine algorithm in forecasting electricity load demand. In *Electrical and Computing Technologies and Applications (ICECTA), 2017 International Conference on* (pp. 1-5). IEEE.

[67]. Haykin, S. and Network, N. 2004. A comprehensive foundation. *Neural Networks*, *2*(2004), p. 41.

132

[68]. Henderson, J.C. 2011. Religious tourism and its management: The hajj in Saudi Arabia. *International Journal of Tourism Research*, *13*(6), pp. 541-552.

[69]. Herrmann, J., Klimm, E., Koehl, M., Slamova, K., Weiss, K.A., Farah, A.A. and Jaber, A.A. 2015, June. Dust mitigation on PV modules in western Saudi Arabia. In *2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC)* (pp. 1-5). IEEE.

[70]. Hobbs, B.F., Jitprapaikulsarn, S., Konda, S., Chankong, V., Loparo, K.A. and Maratukulam, D.J. 1999. Analysis of the value for unit commitment of improved load forecasts. *IEEE Transactions on Power Systems*, *14*(4), pp. 1342-1348.

[71]. Hong, T. and Fan, S. 2016. Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32*(3), pp. 914-938.

[72]. Huang, C.M., Huang, C.J. and Wang, M.L. 2005. A particle swarm optimisation to identifying the ARMAX model for short-term load forecasting. *IEEE Transactions on Power Systems*, *20*(2), pp. 1126-1133.

[73]. IESO. 2018. *Demand response: A smart approach to energy management*. [online] Available at: http://webcache.googleusercontent.com/search?q=cache:-Um1vuVF0m8J:www.ieso.ca/-/media/Files/IESO/Document-Library/publications/DemandResponseBrochure.pdf%3Fla%3Den+&cd=3&hl=en&ct=clnk&gl=ca (Accessed: 22nd Jan 2019).

[74]. IRENA, 2018. Renewable Power Generation Costs In 2018 [online]. Available at: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/May/IRENA_Renewable-Power-Generations-Costs-in-2018.pdf (Accessed 10th Feb 2020)

[75]. Jain, A.K., Mao, J. and Mohiuddin, K.M. 1996. Artificial neural networks: A tutorial. *Computer*, *29*(3), pp. 31-44.

[76]. Kanchev, H., Lu, D., Colas, F., Lazarov, V. and Francois, B. 2011. Energy management and operational planning of a microgrid with a PV-based active generator for smart grid applications. *IEEE transactions on industrial electronics*, *58*(10), pp. 4583-4592.

[77]. Kaytez, F., Taplamacioglu, M.C., Cam, E. and Hardalac, F. 2015. Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power & Energy Systems*, *67*, pp. 431-438.

[78]. Kennedy, J. and Eberhart, R. 1942. Particle swarm optimisation. *Proceedings of IEEE International Conference on Neural Networks.* Perth, Australia, 1948, p.1995.

[79]. Khadem, S.K., Basu, M. and Conlon, M., 2010. Power quality in grid connected renewable energy systems: Role of custom power devices.

[80]. Khashei, M. and Bijari, M. 2010. An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with Applications*, *37*(1), pp. 479-489.

[81]. Khamis, M.F.I., Baharudin, Z., Hamid, N.H., Abdullah, M.F. and Nordin, F.T. 2011. Short term load forecasting for small scale power system using fuzzy logic. In *2011 Fourth International Conference on Modelling, Simulation and Applied Optimisation* (pp. 1-5). IEEE.

[82]. Koprinska, I., Rana, M., Troncoso, A. and Martínez-Álvarez, F. 2013. Combining pattern sequence similarity with neural networks for forecasting electricity demand time series. In *Neural Networks (IJCNN), The 2013 International Joint Conference on* (pp. 1-8). IEEE.

[83]. Kranzberg, D.G. 2012. *Submitted in Partial Fulfillment of the Requirements for the Degree of Masters in Engineering and Public Policy* (Doctoral dissertation, McMaster University Hamilton).

[84]. Kulkarni, S. and Sontakke, S. 2015. Power system analysis of a microgrid using ETAP. *International Journal of Innovative Science and Modern Engineering (IJISME)*, *3*(5).

[85]. Kumar, U. and Jain, V.K. 2010. Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. *Energy*, *35*(4), pp. 1709-1716.

[86]. Kuo, R.J., Hong, S.Y. and Huang, Y.C. 2010. Integration of particle swarm optimisation-based fuzzy neural network and artificial neural network for supplier selection. *Applied Mathematical Modelling*, *34*(12), pp. 3976-3990.

[87]. Lang, K., Zhang, M., Yuan, Y. and Yue, X. 2018. Short-term load forecasting based on multivariate time series prediction and weighted neural network with random weights and kernels. *Cluster Computing*, pp. 1-9.

[88]. Lee, W.J. and Hong, J. 2015. A hybrid dynamic and fuzzy time series model for mid-term power load forecasting. *International Journal of Electrical Power & Energy Systems*,*64*, pp. 1057-1062.

[89]. Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), pp. 764-766.

[90]. Li, H.Z., Guo, S., Li, C.J. and Sun, J.Q. 2013. A hybrid annual power load forecasting model based on generalised regression neural network with fruit fly optimisation algorithm. *Knowledge-Based Systems*, *37*, pp. 378-387.

[91]. Liserre, M., Sauter, T. and Hung, J.Y. 2010. Future energy systems: Integrating renewable energy sources into the smart power grid through industrial electronics. *IEEE Industrial Electronics Magazine*, *4*(1), pp. 18-37.

[92]. Mahmud, K., Amin, U., Hossain, M.J. and Ravishankar, J. 2018. Computational tools for design, analysis, and management of residential energy systems. *Applied Energy*, *221*, pp. 535-556.

[93]. Mares, J.J. and Mercado, K.D. 2017. A methodology for short-term load forecasting. *IEEE Latin America Transactions*, *15*(3), pp. 400-407.

[94]. Martín, L., Zarzalejo, L.F., Polo, J., Navarro, A., Marchante, R. and Cony, M. 2010. Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning. *Solar Energy*, *84*(10), pp. 1772-1781.

[95]. Matar, W., Murphy, F., Pierru, A., Rioux, B. and Wogan, D. 2017. Efficient industrial energy use: The first step in transitioning Saudi Arabia's energy mix. *Energy Policy*, *105*, pp. 80-92.

[96]. Menge, E.E., Wilreker, V.F. and Northcote-Green, J.E.D. 1977, January. Electrical loads can be forecasted for distribution planning. [Time-series and multivariate spatial analyses]. In *Proc. Am. Power Conf.;(United States)* (vol 39, No. CONF-770403-). Westinghouse Electric Corp., Pittsburgh.

[97]. Minhas, D.M., Khalid, R.R. and Frey, G. 2017. Short term load forecasting using hybrid adaptive fuzzy neural system: The performance evaluation. *In 2017 IEEE PES Power Africa* (pp. 468-473). IEEE.

[98]. Mitchell, J., Marcel, V. and Mitchell, B. 2012. *What next for the oil and gas industry?* London: Chatham House.

[99]. Murata, N., Yoshizawa, S. and Amari, S.I. 1994. Network information criterion-determining the number of hidden units for an artificial neural network model. *IEEE Transactions on Neural Networks*, *5*(6), pp. 865-872.

[100]. Murtagh, F. and Heck, A. 2012. *Multivariate data analysis, v*ol. 131. Springer Science & Business Media.

[101]. Mwasilu, F., Justo, J.J., Kim, E.K., Do, T.D. and Jung, J.W. 2014. Electric vehicles and smart grid interaction: A review on vehicle to grid and renewable energy sources integration. *Renewable and Sustainable Energy Reviews*, *34*, pp.501-516.

[102]. Nematollahi, O., Hoghooghi, H., Rasti, M. and Sedaghat, A. 2016. Energy demands and renewable energy resources in the Middle East. *Renewable and Sustainable Energy Reviews*, *54*, pp.1172-1181.

[103]. Neusser, L. and Canha, L.N. 2013. Real-time load forecasting for demand side management with only a few days of history available. In *Power Engineering, Energy and Electrical Drives (POWERENG), 2013 Fourth International Conference on* (pp. 911-914). IEEE.

[104]. OPEC. 2017. *Saudi Arabia facts and figures*. [online] Available at: https://www.opec.org/opec_web/en/about_us/169.htm [Accessed on 22nd Dec 2018].

[105]. Ozerdem, O.C., Olaniyi, E.O. and Oyedotun, O.K., 2017. Short term load forecasting using particle swarm optimisation neural network. *Procedia computer science*, *120*, pp. 382-393.

[106]. Pai, P.F. and Hong, W.C. 2005. Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms. *Electric Power Systems Research, 74*(3), pp. 417-425.

[107]. Palensky, P. and Dietrich, D. 2011. Demand side management: Demand response, intelligent energy systems, and smart loads. *IEEE Transactions on Industrial Informatics*, *7*(3), pp. 381-388.

[108]. Palma, W. 2016. Time series analysis. Wiley series in probability and statistics.

[109]. Pan, X., Lee, B. and Zhang, C.= 2013, November. A comparison of neural network backpropagation algorithms for electricity load forecasting. In *Intelligent Energy Systems (IWIES), 2013 IEEE International Workshop on* (pp. 22-27). IEEE.

[110]. Panapakidis, I.P. and Dagoumas, A.S. 2016. Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy*, *172*, pp. 132-151.

[111]. Panagiotelis, A. and Smith, M. 2008. Bayesian density forecasting of intraday electricity prices using multivariate skew t distributions. *International Journal of Forecasting*, *24*(4), pp. 710-727.

[112]. Pandian, S.C., Duraiswamy, K., Rajan, C.C.A. and Kanagaraj, N.,2006. Fuzzy approach for short term load forecasting. *Electric Power Systems Research, 76*(6-7), pp. 541-548.

[113]. Panjwani, M. K., & Narejo, G. B. (2014). Effect of humidity on the efficiency of solar cell (photovoltaic). *International Journal of Engineering Research and General Science*, 2(4), 499-503.

[114]. Pannell, D.J. 1997. Sensitivity analysis: strategies, methods, concepts, examples. *Agricultural Economics*, *16*, pp. 139-152.

[115]. Papalexopoulos, A.D. and Hesterberg, T.C. 1990. A regression-based approach to short-term system load forecasting. *IEEE Transactions on Power Systems*, *5*(4), pp. 1535-1547.

[116]. Paton, F.L., Maier, H.R., Dandy, G.C. 2013. Relative magnitudes of sources of uncertainty in assessing climate change impacts on water supply security for the southern Adelaide water supply system. *Water Resource Research, 49*(3), pp. 1643-1667.

[117]. Pazheri, F.R., Malik, N.H., Al-Arainy, A.A., Ottukulotk, S., Othman, M.F., Al-Ammar, E.A. and TP, I.A. 2012. Use of renewable energy sources in Saudi Arabia through smart grid. *Journal of Energy and Power Engineering*, *6*(7), p. 1065.

[118]. Petrakopoulou, F., Tsatsaronis, G., Morosuk, T. and Carassai, A. 2012. Conventional and advanced exergetic analyses applied to a combined cycle power plant. *Energy*, *41*(1), pp. 146-152.

[119]. Phuangpornpitak, N. and Tia, S. 2013. Opportunities and challenges of integrating renewable energy in smart grid system. *Energy Procedia*, *34*, pp. 282-290.

[120]. Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B. and Wagener, T. 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, *79*, pp. 214-232.

[121]. Ponds, K., Arefi, A., Sayigh, A. and Ledwich, G. 2018. Aggregator of demand response for renewable integration and customer engagement: Strengths, weaknesses, opportunities, and threats. *Energies*, *11*(9), p. 2391.

[122]. Quan, H., Srinivasan, D. and Khosravi, A., 2014. Short-term load and wind power forecasting using neural network-based prediction intervals. *IEEE transactions on neural networks and learning systems*, *25*(2), pp. 303-315.

[123]. Qiu, J., Men, K., Zhao, J. and Zheng, Y. 2017. Power network planning considering trade-off between cost, risk and reliability. *International Journal of Trans Electronic Energy Systems, 27*, pp. 1-16.

[124]. Rafique, M., Rehman, S., Alam, M. and Alhems, L. 2018. Feasibility of a 100 MW Installed Capacity Wind Farm for Different Climatic Conditions. *Energies*, *11*(8), p. 2147.

[125]. Rahman, F., Rehman, S. and Abdul-Majeed, M.A. 2012. Overview of energy storage systems for storing electricity from renewable energy sources in Saudi Arabia. *Renewable and Sustainable Energy Reviews*, *16*(1), pp. 274-283.

[126]. Ramli, M.A., Twaha, S. and Al-Hamouz, Z. 2017. Analysing the potential and progress of distributed generation applications in Saudi Arabia: The case of solar and wind resources. *Renewable and Sustainable Energy Reviews*, *70*, pp. 287-297.

[127]. Rani, U. and Dalal, S. 2016. Neural Network Applications in Design Process of Decision Support System.

[128]. Ratikainen, K.W. 2017. *Transitioning to renewable energy in Saudi Arabia-A multi-level perspective analysis of the Saudi renewable energy policies* (Master's thesis).

[129]. Raza, M.Q. and Khosravi, A. 2015. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews, 50*, pp. 1352-1372.

[130]. Reed, S. 2018. *From oil to solar: Saudi Arabia plots a shift to renewables*. The New York Times. [online] Available at: https://www.nytimes.com/2018/02/05/business/energy-environment/saudi-arabia-solar-renewables.html (Accessed: 2nd Jan 2019).

[131]. Rehman S, Bader, M. A., and Al-Moallem, S. A. 2007. The cost of solar energy generated using PV panels. *Renewable and Sustainable Energy Reviews, 11*(8), pp. 1843–1857.

[132]. Ryu, S., Noh, J. and Kim, H. 2016. Deep neural network-based demand side short term load forecasting. *Energies*, *10*(1), p. 3.

[133]. Said, S. A. M., El-Amin, I. M., Al-Shehri, A. M. 2008. Renewable energy potentials in Saudi Arabia (lecture). American University of Beirut, Faculty of Engineering and Architecture, Beirut.

[134]. Saidur, R., Rahim, N.A., Islam, M.R. and Solangi, K.H. 2011. Environmental impact of wind energy. *Renewable and sustainable energy reviews*, *15*(5), pp. 2423-2430.

[135]. Sangrody, H., Sarailoo, M., Zhou, N., Tran, N., Motalleb, M. and Foruzan, E. 2017. Weather forecasting error in solar energy forecasting. *IET Renewable Power Generation*, *11*(10), pp. 1274-1280.

[136]. Sarabia, Y.R., Lorenzo, M.G.C., Pérez, R.B. Falcon, R. 2005. Extending CBR-ANN Hybrid Models Using Fuzzy Sets. *2005 International Conference on Neural Networks and Brain* 3, pp. 1755-1760.

[137]. Saravanan, S., Kannan, S. and Thangaraj, C. 2012. Forecasting India's electricity demand using artificial neural network. In *Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on* (pp. 79-83). IEEE.

[138]. Scott, M. 2018. *Saudi Arabia plans to source 10% of its power from renewable energy within 5 years*. Forbes. [online] Available at: https://www.forbes.com/sites/mikescott/2018/01/18/saudi-arabia-plans-to-source-10-of-its-power-from-renewable-energy-within-5-years/#196e0a54485f .

[139]. Shrestha, B., Paudyal, B., Karki, N.R. and Gyawali, N.P. 2017. Steady state analysis of PV integration in hydro dominated INPS grid using ETAP. In *2017 7th International Conference on Power Systems (ICPS),* pp. 743-748. IEEE.

[140]. Simao, H.P., Powell, W.B., Archer, C.L. and Kempton, W. 2017. The challenge of integrating offshore wind power in the US electric grid. Part II: Simulation of electricity market operations. *Renewable Energy*, *103*, pp. 418-431.

[141]. Schalkoff, R.J. 1997. *Artificial Neural Networks*. Vol. 1. New York: McGraw-Hill.

[142]. Setiawan, A., Koprinska, I. and Agelidis, V.G. 2009, June. Very short-term electricity load demand forecasting using support vector regression. In *Neural Networks, 2009. IJCNN 2009. International Joint Conference on* (pp. 2888-2894). IEEE.

[143]. Shalwala, R.A. and Bleijs, J.A.M. 2009. Impact of Grid-Connected PV systems in residential areas in Saudi Arabia. In *Universities Power Engineering Conference (UPEC), 2009 Proceedings of the 44th International*, (pp. 1-5). IEEE.

[144]. Shalwala, R.A. and Bleijs, J.A.M. 2010. Impact of Grid-Connected PV systems on voltage regulation of a residential area network in Saudi Arabia. In *Nuclear & Renewable Energy Conference (INREC), 2010 1st International*, (pp. 1-5). IEEE.

[145]. Sietsma, J. and Dow, R.J. 1991. Creating artificial neural networks that generalize. *Neural Networks*, *4*(1), pp. 67-79.

[146]. Singh, N.K., Singh, A.K. and Tripathy, M. 2014. A comparative study of BPNN, RBFNN and ELMAN neural network for short-term electric load forecasting: A case study of Delhi region. In *Industrial and Information Systems (ICIIS), 2014 9th International Conference on* (pp. 1-6). IEEE.

[147]. The Straits Times. 2018. *From oil to solar energy: Saudi Arabia plots a shift to renewables*. [online] Available at https://www.straitstimes.com/world/middle-east/from-oil-to-solar-energy-saudi-arabia-plots-a-shift-to-renewables.

[148]. Suganthi, L. and Samuel, A.A. 2012. Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews*, *16*(2), pp. 1223-1240.

[149]. Tagle, F., Castruccio, S., Crippa, P. and Genton, M.G. 2017. Assessing potential wind energy resources in Saudi Arabia with a skew-t distribution. *arXiv preprint arXiv:1703.04312*.

[150]. Taha, A. M. 2012. *Report: Half of Saudi Arabia's power generation comes from natural gas.* [online] Available at: http://www.arabnews.com/report-half-saudi-arabia%E2%80%99s-power-generation-comes-natural-gas)Accessed: 22nd Dec 2018).

[151]. Taher, N. and Al-Hajjar, B. 2013. *Energy and Environment in Saudi Arabia: Concerns & Opportunities*. Springer Science & Business Media.

[152]. Tanaka, H. and Watada, J. 1988. Possibilistic linear systems and their application to the linear regression model. *Fuzzy Sets and Systems*, *27*(3), pp. 275-289.

[153]. Taylor, J.W. and McSharry, P. E. 2007. Short-term load forecasting methods: An evaluation based on European data. *IEEE Transactions on Power Systems*, *22*(4), 2213-2219.

[154]. Teixeira, J., Macedo, S., Gonçalves, S., Soares, A., Inoue, M. and Cañete, P. 2017. Hybrid model approach for forecasting electricity demand. *CIRED-Open Access Proceedings Journal*, *2017*(1), pp. 2316-2319.

[155]. Vortex, 2014. Wind map of Saudi Arabia. https://vortexfdc.com/knowledge/saudi-arabia-wind-map/ Retrieved 2010-02-04

[156]. Yan, Q., Ye, Z., Ji, F., & Gao, X. (2019). Influence of different time series on load recovery following faults in the vessel regional distribution system. The Journal of Engineering, 2019(16), 2717-2720.

[157]. Wan, C., Zhao, J., Song, Y., Xu, Z., Lin, J. and Hu, Z. 2015. Photovoltaic and solar power forecasting for smart grid energy management. *CSEE Journal of Power and Energy Systems*,*1*(4), pp. 38-46.

[158]. Wang, D., Luo, H., Grunder, O., Lin, Y. and Guo, H. 2017. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimised by firefly algorithm. *Applied Energy,* 190, pp. 390-407.

[159]. Wang, L., Zeng, Y. and Chen, T. 2015. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Systems with Applications*, *42*(2), pp.855-863.

[160]. Wang, X., Chen, Z., & Yang, S. (2012). Forecasting modeling and simulation analysis of a power system in China, based on a class of semi-parametric regression approach. South African Journal of Industrial Engineering, 23(3), 154-168.

[161]. Warrior, K.P., Shrenik, M. and Soni, N. 2016. Short-term electrical load forecasting using predictive machine learning models. In *India Conference (INDICON), 2016 IEEE Annual* (pp. 1-6). IEEE.

[162]. Weather Online. 2019. *Saudi Arabia*. [online] Available at: https://www.weatheronline.co.uk/reports/climate/Saudi-Arabia.htm (Accessed: 9th Jan 2019).

[163]. Wei, Y.M., Li, D., Tan, Z. and Zhou, J. 2018. Short term electricity load forecasting using a hybrid model. *Energy,* 158, pp. 774-781.

[164]. World Data. 2015. *Energy consumption in Saudi Arabia.* [online] Available at: https://www.worlddata.info/asia/saudi-arabia/energy-consumption.php.

[165]. Xing, H., Cheng, H. and Zhang, L. 2015. Demand response based and wind farm integrated economic dispatch. *CSEE Journal of Power and Energy Systems, 1*(4), pp. 37-41.

[166]. Yaïci, W. and Entchev, E. 2016. Adaptive Neuro-Fuzzy Inference System modelling for performance prediction of solar thermal energy system. *Renewable Energy*, *86*, pp. 302-315.

[167]. Yamin, H.Y. 2004. Review on methods of generation scheduling in electric power systems. *Electric Power Systems Research*, *69*(2-3), pp. 227-248.

[168]. Yang, D., Gu, C., Dong, Z., Jirutitijaroen, P., Chen, N. and Walsh, W.M. 2013. Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging. *Renewable Energy*, *60*, pp. 235-245.

[169]. Yang, Y., Che, J., Li, Y., Zhao, Y. and Zhu, S, 2016. An incremental electric load forecasting model based on support vector regression. *Energy*, 113, pp. 796-808.

[170]. Yang, H., Liu, L., Zhou, H. and Gao, T. 2016. Multivariate statistic methods for predicting electricity consumption of Beijing. In *Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE* (pp. 7197-7202). IEEE.

[171]. Yang, S.L. and Shen, C. 2013. A review of electric load classification in smart grid environment. *Renewable and Sustainable Energy Reviews*, *24*, pp. 103-110.

[172]. Ying, L.C. and Pan, M.C. 2008. Using adaptive network based fuzzy inference system to forecast regional electricity loads. *Energy Conversion and Management*, *49*(2), pp. 205-211.

[173]. Yu, S.W. and Zhu, K.J. 2012. A hybrid procedure for energy demand forecasting in China. *Energy, 37*(1), pp. 396-404.

[174]. Yuan, J., Farnham, G., Azuma, C. and Emura, K. 2018. Predictive artificial neural network models to forecast the seasonal hourly electricity consumption for a university campus. *Sustainable Cities and Society*, *42*, pp. 82-92.

[175]. Zell, E., Gasim, S., Wilcox, S., Katamoura, S., Stoffel, T., Shibli, H., Engel-Cox, J. and Al Subie, M. 2015. Assessment of solar radiation resources in Saudi Arabia. *Solar Energy*, *119*, pp. 422-438.

[176]. Zhang, C., Shao, H. and Li, Y., 2000. Particle swarm optimisation for evolving artificial neural network. In *Smc 2000 conference proceedings. 2000 IEEE international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions',* (vol 4, pp. 2487-2490). IEEE.

[177]. Zhang, J., Wei, Y.M., Li, D., Tan, Z. and Zhou, J., 2018. Short term electricity load forecasting using a hybrid model. Energy, 158, pp. 774-781.

[178]. Zhou, K., Fu, C. and Yang, S. 2016. Big data driven smart energy management: From big data to big insights. *Renewable and Sustainable Energy Reviews*, *56*, pp. 215-225.