UNIVERSITI TEKNOLOGI MARA

SHORT TERM LOAD FORECASTING USING ARTIFICIAL NEURAL NETWORK (ANN) WITH MULTIPLE TIME LAGS, STATIONARY TIME SERIES AND PRINCIPAL COMPONENT ANALYSIS (PCA)

MOHD HAFEZ HILMI HARUN

Thesis submitted in fulfillment of the requirements

for the degree of

Master of Science

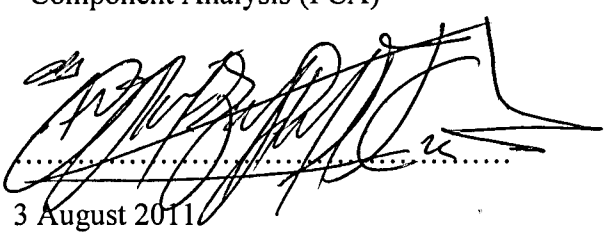
Faculty of Electrical Engineering

August 2011

Candidate's Declaration

I declare that the work in this thesis was carried out in accordance with the regulation of University Teknologi Mara. It is original and is the result of my own work, unless otherwise indicated or acknowledged as referenced work. This topic has been not submitted to any other academic institution or non-academic institution for any other degree or qualification.

In the event that my thesis is found to violate the conditions mention above, I voluntarily waive the right of conferment of my degree and agree to be subjected to the disciplinary rules and regulations of University Teknologi Mara.

Name of Candidate	Mohd Hafez Hilmi Harun
Candidate's ID no	2007131329
Programme	Master of Science in Electrical Engineering
Faculty	Electrical Engineering
Thesis Title	Short Term Load Forecasting Using Artificial Neural Network (ANN) with Multiple Time Lags, stationary Time Series and Principal Component Analysis (PCA)
Signature of Candidate	
Date	3 August 2011

ABSTRACT

Load forecasting has always been an essential task for the electric utilities in which it may assist to an effective operational planning and security assessment of a power system. This is important to ensure that the electric utilities are operating in an economic, reliable and uninterrupted service to the customers. With the advent of deregulation in electric utilities, load forecasting becomes even more important especially to the system operators and market participants in whom this may assist towards organizing appropriate strategies of risk management and competitive energy trading. Load power forecasting is classified into four categories which are the very short-term, short-term, medium-term and long-term load forecasts. This thesis only presents Short-term load forecasting (STLF).

Generally, in STLF, there are two approaches to forecast load power which are statistical technique and AI based technique. ANN and Box Jenkins models are chosen to represent AI based technique and statistical technique respectively. The ANN and the Box Jenkins models are used to perform STLF for the next 24 hours for 29 days. The load data of the year 2002 is used as case study in this research.

The Box-Jenkins models are comprised of the autoregressive (AR), moving average (MA) and autoregressive integrated moving average (ARIMA) techniques. In order to identify an appropriate Box-Jenkins model for STLF the sample autocorrelation (SAC) and sample partial autocorrelation (SPAC) functions are used to analyze the behavior of the past univariate load time series. Generally, the AR Box-Jenkins model forecasts the future load based on linear function of the past peak loads, the MA Box-Jenkins model forecasts the future load based on linear combination of the past peak load errors and the ARIMA Box-Jenkins model forecast the future peak load based on the combination of AR Box-Jenkins and MA Box-Jenkins models.

The Artificial Neural Network (ANN) is adaptable and robust in forecasting the highly non-linear data. So that, ANN is used to do the STLF for next 24 hours .The ANN input data is generated with multiple time lags of load data in time series whereby the ANN output data is convert to stationary time series. The Principle Component Analysis (PCA) is used to optimize the ANN input data that generated by multiple time lag of time series. PCA also help to reduce the time taken during the ANN training procedure. The cross validation of ANN training procedure is performed to ensure accuracy of the ANN. The input data of ANN is divided into three sets for training, validation and training of ANN session.

The forecasting results from both models of the ANN and the Box-Jenkins is compared to the actual value by calculating Mean Average Percentage Error (MAPE). The research shows that ANN model forecast better and provide less MAPE error than Box-Jenkins model. It shows that ANN model using multiple time lags of time series optimized with PCA ANN input and stationary ANN output is more robust and reliable for STLF even though historical data and variable factors of load data are limited.

ACKNOWLEDGEMENTS

I would like to express my grateful to Allah the almighty Who has given me the strength to complete this thesis. I also want to express my sincere appreciation and deepest gratitude to my main supervisor, Dr Muhammad Murtadha bin Othman. Without his motivations, supports, guidance and instructions, I will not be able to complete this thesis. May Allah bless you Dr Muhammad Murtadha bin Othman and his family. Last not least, my thanks to my co-supervisor, Ass Prof Dr Ismail bin Musirin.

I also appreciate the valuable support and precious prayers from my beloved parents Hajah Aminah and Haji Harun; my siblings Haslena, Faraheida, Nurulain Izzati, Muhammad Hadri Hanafi and Muhammad Hariz Hussaini. Finally, I wish my deepest appreciation to my beloved wife, Hasliza binti Hassan, who is always give me the inspiration and together with me in the moment of joy and sorrow to complete this thesis. Last but but least, Muhammad Haris Hakimi our beloved son who is always together with us to cheerful our daily lives.

Thank you.

Yours Truly,
Mohd Hafez Hilmi bin Harun
3 August 2011.

TABLE OF CONTENTS