FUZZY RULE BASED AND ANFIS CLASSIFICATION FOR RAINFALL DISTRIBUTION

NOOR HAYATI BINTI MOHD. ZAIN

A dissertation submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia

JANUARY 2012

"Dedicated to my beloved family and friends, without their understanding, supports, and most of all love, the completion of this work would not have been possible."

ACKNOWLEDGEMENT

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, my most gratitude to my supervisor Associate Professor Dr. Siti Zaiton Mohd. Hashim whose sincerity and encouragement I will never forget. She has been my inspiration as I hurdle all the obstacles in the completion this research work.

Secondly, thank you also to all my beloved friends that always gave me courage to finish my master. I will never forget your supports in any situation that I faced in finishing this job.

Last but not the least, my family; especially to my beloved mum and dad for their continuous support and the one above all of us, the omnipresent God, for answering my prayers for giving me the strength to plod on despite my constitution wanting to give up and throw in the towel, thank you so much Dear Allah.

ABSTRACT

Throughout history, weather has always been a form of frightful uncertainty to mankind. Other than that, it also has some negative impact on certain industries and jobs including the clothing industry and food production, and it affect peoples' lives in certain situations especially in terms of making travel plans. Hence, it can be said that weather is always related to the daily human activities. Therefore, weather prediction should be classified as an important factor in achieving better living. The data regarding the distribution of weather is usually in a dynamic pattern and hard to predict, which is why the weather classification be done to improve the prediction. The rainfall classification model based on soft computing is feasible to implement and could produce desirable result through the training and testing of the available dataset. Soft computing techniques such as Fuzzy Rule Based (FRB) and Adaptive-Neural Fuzzy Inference System (ANFIS) are investigated in this study to determine which technique is most effective and can achieve higher percentage of accuracy for the purpose of rainfall classification. A number of 720 Senai weather hourly datasets are used in order to test the result of weather classification. The results of the experiments done using both methods show that ANFIS is capable of producing better result for classification with the accuracy percentage of 97.22% in first experiment with FRB which only produce an accuracy of 2.08%. In second experiment, FRB gave 53.47% classification rate, lower than ANFIS which produced higher classifier rate 97.22% thus proving that is it better than FRB.

ABSTRAK

Menurut sejarah, cuaca dikatakan menjadi salah satu kebimbangan dalam kehidupan manusia. Selain itu, cuaca juga dikatakan memberi impak negatif dalam pekerjaan harian, seperti industri pakaian dan pengeluaran makanan, yang memberi kesan kepada kehidupan manusia dan juga aktiviti perjalanan tempat dan situasi. Oleh itu, boleh dikatakan cuaca adalah berkait dengan kehidupan manusia. Keadaan cuaca adalah perlu dikelaskan bagi kemudahan masa depan. Data taburan cuaca biasanya adalah dalam corak dinamik dan sukar untuk diramal, yang dimana proses pengelasan cuaca dilakukan bagi meningkatkan kadar peramalan. Perkiraan model cuaca berdasarkan perkomputeran lembut mudah untuk dilaksanakan dan kebiasaannya akan menghasilkan pengelasan yang dikehendaki berdasarkan latihan yang dijalankan pada set data. Kaedah perkomputeran lembut seperti *Fuzzy Rule Based (FRB)* dan *Adaptive-Neural Fuzzy Inference System (ANFIS)* digunakan di dalam kajian untuk menentukan kaedah yang paling berkesan dan dapat mencapai peratusan ketepatan yang lebih tinggi untuk tujuan kajian pengelasan taburan hujan. 20 set data dari pangkalan data Senai digunakan di dalam kajian ini bagi mendapatkan hasil kajian terhadap pengelasan cuaca. Selepas ujikaji dilakukan terhadap kedua-dua kaedah, *ANFIS* telah memberi keputusan yang lebih baik untuk pengelasan dengan mempamirkan 97.22 peratus ketepatan dalam ujian pertama dengan *FRB* yang hanya menghasilkan ketepatan 2.08 peratus. Dalam ujian kedua. *FRB* memberi kadar pengelasan 53.47 peratus iaitu lebih rendah berbanding *ANFIS* yang menghasilkan kadar pengelasan 97.22 peratus dan membuktikan bahawa *ANFIS* adalah lebih baik daripada FRB

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligent
FRB	Fuzzy Rule Based
ANFIS	Adaptive-Neural Fuzzy Inference System
FIS	Fuzzy Inference System
NN	Neural Network
ANN	Artificial Neural Network
CP	Circulation Patterns
MIQ	machine intelligence quotient
FAM	fuzzy associative memories
EM	Expectation Maximization
SSC	Spatial Synoptic Classification
MLP	Multilayer perceptron network
SMD	Senai Meteorology Department
MMD	Malaysia Meteorology Department
NR	No Rain
LR	Low Rain
MR	Moderate Rain
HR	Heavy Rain
Kmph	kilometre per hour
Mm	milimetre

CHAPTER 1

INTRODUCTION

1.1 Overview

Theoretically, weather can be described as the condition of the air at a particular time and place. There are various conditions of weather where it can be in warm, cold, wet, dry, cloudy or windy. Throughout history, the condition of the weather has always been a form of fear to mankind due to its uncertainty. Other than that, it also has some negative impact on human lives as unpredictable weather can be hazardous to certain occupations, especially those in the clothing and food production industries, and it can affect peoples' lives in terms of travelling. Consequentially, weather can be considered always related to the daily human activities. With this knowledge, meteorologists who are experts in understanding weather, gather information which can hopefully help humanity cope with weather unpredictability. The weather on earth is situated at the lowest level of the atmosphere which is called the troposphere level, the lowest level of the sky.

Classification refers to identifying the sub-population to which new observations belong, where the identity of the sub-population is unknown, on the basis of a training set of data containing observations whose sub-population is known. Thus the requirement is that new individual items are placed into groups based on quantitative information on one or more measurements, traits or characteristics, etc.

In relation with the classification of weather, the synoptic weather-typing or the classification of weather conditions or patterns into categories is known to be popular where numerous methods have been developed over the past century. The recent increase of interest in the procedure is attributed to its utility in solving a wide array of applied climatologically problems. Concern over the impacts of weather, especially for the purpose of understanding possible implications of climate change, has driven the search for more, and better, weather-typing schemes.

One researcher (Yarnal, 1993) notes several different subdivisions of synoptic classifications; among these is the distinction of *manual* versus *automated* classification schemes. Manual procedures involve the subjective classification of circulation patterns or weather types from visual analyses of individual synoptic maps. Schemes such as these have several benefits. The investigator is in full control of the process and classification. The classification system can thus be tailored precisely to the researcher's needs but manual classification schemes quite time consuming and can be difficult to export to other locations.

It is found in related literatures, in particular soft computing techniques; Artificial Intelligent (AI) has provided an alternative solution in doing weather classification. This study investigates and compares the performance of Fuzzy Rule Based (FRB) and the hybrid technique of Adaptive-Neural Fuzzy Inference System (ANFIS) in weather classification.

Acquiring the expert knowledge of a skilled domain specialist in the form of fuzzy values for each fuzzy rule is an arduous step in the design procedure of FRB for modelling and control. It is desirable that the rules are formed in such a way that each rule shows a dominant effect of the input-output mapping. Consequently, the maximum efficiency of rules must be taken into account ensuring that the fuzzy rule based is optimal. There are essentially two different approaches in the definition of fuzzy rules for modelling or control employing FRB. The rules are either defined using empirical experience or applying an adaptive scheme to adjust the parameters of FRB. There are mainly two different types of FRB, such as (Mamdani,1974) and

(Takagi and Sugeno, 1985). In this work, Mamdani is applied due to its simple of 'min-max' operations.

It is discarding in literatures that hybrid techniques have been widely used among researchers in many areas such as patterns recognition, making predictions, medical diagnosis and such other applications. A hybrid technique is a potentially powerful tool that may enable us to address and solve problems that are just too complex for conventional approaches (Jackson, 1999). Some of the hybrid techniques that have already been used by researchers include Neuro-fuzzy with knowledge-based, fuzzy logic with case-based reasoning, fuzzy with artificial immune system (Salfarina, 2008).

Therefore, it can be concluded that FRB and ANFIS are suitable techniques for the purpose of weather classification and are capable of producing accurate results.

1.2 Problem Background

The weather classification has always been known as one of the most popular fields of study which require critical analysis. It is not an easy task to achieve complete accuracy. Usually, there always will be an error either big or small cause by chaotic weather data. This is one of the problems that humans hope to solve in order to ensure better future and living. However, it is impossible to accurately classify the weather all the time. Tiny changes in the atmospheric conditions can have huge effects on weather that happens days later. These changes can be too small to detect (Towe, 2004).

In meteorology, the small changes in the initial condition of the atmosphere will lead to big changes in future weather classification. It is considerably really sensitive as small changes in the state of the weather can cause big differences in the future weather. These small differences can sometimes be too small to detect. That is

why scientists say that day-to-day classification of the weather for more than about two weeks ahead will never be possible. Nowadays, based on previous experimental results, five days is the limit of the time range to which the day-to-day classification can be effectively done (Salfarina, 2008). The technique or method to be used solving the problem of classification has to be a technique with strong computations and complexity restrictions because the changes in the weather are always chaotic and uncertain.

Due to the problems in weather classification and the possible suggested solutions, there is a need to produce better classification of weather in order to obtain more accurate classification results.

1.3 Problem Statement

In doing weather classification, the expert must have broad knowledge about weather and the variability related with it and expert should be able to deal with chaotic atmosphere. Based on that, weather classification requires massive computational power in solving problems as it deals with weather classification calculations.

First method in understanding the underlying pattern is recognizing that the fuzzy logic provides the means of dealing with uncertainty; for example, the extended concept of industrial process control is that it is expected to be applied to systems in which the precise measurements of the state variables could not be obtained. Fuzzy logic and rule based systems are a means of dealing with imprecision, a method of modelling human behaviour, and a means of achieving control of industrial systems that cannot be modelled rigorously (Mamdani and Assilian, 1975).

Currently, such facility is possible by using a combination of a neural network and Fuzzy logic system (Aldrian and Djamil, 2008). On the other hand, ANFIS method, which is the integration of NN and fuzzy logic methods, has the potential to capture the benefits of both these methods in a single framework. ANFIS eliminates the basic problem in fuzzy system design through defining the membership function parameters and design of fuzzy if–then rules by effectively using the learning capability of NN for automatic fuzzy rule generation and parameter optimization (Bacanli *et al*., 2009).

Hence, FRB and ANFIS can be considered techniques which are suitable to be used in this investigation for classification of rainfall distribution and the techniques will be compared to find out which technique is capable of providing more accurate results.

1.4 Dissertation Aim

The aim of this study is to investigate whether rainfall classification problems can be solved using the Fuzzy Rule Based technique and the ANFIS technique and to evaluate the techniques in order to discover which technique can achieve higher accuracy.

1.5 Dissertation Objectives

The objectives of this dissertation are:

i. To investigate rainfall classification problems and how they can be solved using computational intelligence techniques.
ii. To develop a Fuzzy Rule Based and ANFIS techniques for classification of rainfall distribution.

iii. To compare and evaluate which method will produce better results for classification of rainfall in terms of percentage of accuracy.

1.6 Dissertation Scopes

The scopes of the study include:

i. The techniques to be used for rainfall distribution calculation are Fuzzy Rule Based and ANFIS.
ii. The type of data is hourly data from 1 November 2010 to 30 November 2010 obtained from the Senai Meteorology Department, Johor Bahru.
iii. The input data for Fuzzy Rule Based and ANFIS are Temperature, Humidity, and Mean Wind Speed.
iv. The target output will be the rainfall data of the next hour (target rainfall).

1.7 Significance of the Dissertation

The importance of this study is to prove how the classification of rainfall distribution can be improved with the use of Fuzzy Rule Based and ANFIS techniques with the percentage of accuracy produced by each method compared to each other. This study uses Malaysia dataset for the purpose of weather classification and will hopefully contribute to the future works in the study of rainfall distribution based on soft computing approaches.

1.8 Organization of the Dissertation

This dissertation consists of five chapters. The first chapter presents the introduction on the dissertation which includes the background of problem, the objectives of the study, the scope as well as the significance of the dissertation. Chapter two reviews the issues ~~of~~ regarding weather and weather classification and the techniques that have been widely used ~~to~~ in providing good classification of rainfall distribution. The third chapter discusses about methodology used in this dissertation and elaborates on the detailed process of the dissertation. Chapter four presents the experimental results and the analysis of the experiments conducted. Last but not least, chapter five concludes the dissertation and provide suggestions for future works and further enhancements.

REFERENCES

Christopher Towe (2004), *Scalia Laboratory for Atmospheric Analysis*, Ohio University. 25 April 2004.

Clouds R Us.com (2009). Available at: http://www.rcn27.dial.pipex.com/cloudsrus/whatis.html. Retrieved on 8 September 2009.

Blurtit.com (2009). Available at: http://www.blurtit.com/q103191.html. Retrieved on 8 September 2009.

Wikipedia (2009), *Weather_forecasting*. Available at: http://en.wikipedia.org/wiki/Weather_forecasting. Retrieved on 9 September 2009.

usatoday.com (2009). Available at: http://www.usatoday.com/weather/wforcst0.htm. Retrieved on 9 September 2009.

emc.ncep.noaa.gov (2009). Available at: http://www.emc.ncep.noaa.gov/gmb/ens/info/ens_detbak.html. Retrieved 10 September 2009.

Myles Allen, Jamie Kettleborough and David Stainforth, *Department of Physics, University of Oxford, Space Science and Technology Department, Rutherford Appleton Laboratory.*

bus.ucf.edu (2005). Available at: www.bus.ucf.edu/rszymanski/ISM3530/CP2forecasting.doc. Retrived 5 October 2009.

Jackson, P. (1999), *Introduction to Expert System*, 3^{rd} *Edition.* Addison-Wesley.

Wan Salfarina Wan Husain (2008), *"Fuzzy Case-based Reasoning for Weather Forecasting".* 31 October 2008.

James N. K. Liu and Raymond S. T. Lee (1999), *"Rainfall Forecasting from Multiple Point Sources Using Neural Networks"*. Department of Computing, Hong Kong Polytechnic University, Hong Kong.

J. L. McGregor, K. J. Walsh, K. J. and J. J. Katzfey (1993), *"Climate simulations for Tasmania".* In Fourth Int'l Conference on Southern Hemisphere

Meteorological and Oceanography, American Meteorological Society, pp. 514-515.

L. A. Zadeh (1994), *"Soft Computing and Fuzzy Logic"*. University of California at Berkeley, 1994.

Artificial Neural Network. Available at: http://www.learnartificialneural znetworks .com/

Jyh-Shing Roger Jang (1993), *"ANFIS: Adaptive-Network-Based Fuzzy Inference System"*. University of California at Berkeley.

L. A. Zadeh (1965), *"Fuzzy sets"*. Information and Control, pp. 8:338-353.

T. Takagi and M. Sugeno (1983), *"Derivation of fuzzy control rules from human operator's control actions*". Proc. IFAC Symp. Fuzzy Inform, Knowledge Representation and Decision Analysis, pp. 55-60, July 1983.

P. Werbos (1974), *"Beyond regression: New tools for prediction and analysis in the behavioral sciences".* Ph.D. dissertation, Harvard Univ., Cambridge, MA, 1974.

Mathworks Inc. (2001). MATLAB 6 user's guide. Available at: http://www. mathworks.com

R. Clint Whaley, A. Petitet, and J. J. Dongarra (2001), *"Automated empirical optimization of software and the ATLAS project"*. *Parallel Comput.*, vol. 27, no. 1–2, pp. 3–35.

M. Frigo and S. Johnson (1998), *"FFTW: an adaptive software architecture for the FFT".* Proc. Int. Conf. Acoustics, Speech and Signal Processing, vol. 3, p. 1381.

Ron Choy and Alan Edelman (2005), *"Parallel MATLAB: Doing It Right".* Invited Paper.

Wikipedia (2009). Available at: http://wiki.answers.com/Q/What_is_orographic_ rainfall.

Mamdani EH (1974). *Application of fuzzy algorithm for control of simple dynamic plant 1974*. IEEE Proc 1974; 121: 715–719.

Takagi T, Sugeno M (1985). *Fuzzy identification of systems and its application to modelling and control*. IEEE Trans Systems, Man and Cybernetics 1985; 15: 116–132.

Hasan, M.Tsegaye, D.T., Shi, F., Schaefer, L.G.,Taylor, H.G. (2009), *Improvement of the Model for Predicting Rainfall by Fuzzy Set Theory Using USDA Scan*

Data. http//:www. usawaterquality.org/conferences/2009/PDF/Other-poster/Hasan09.pdf.

Sivanandam S.N., Sumathi S. and Deepa S. N. (2007), *Introduction to Fuzzy Logic using MATLAB*. Springer, pp.113-145.

M. Hellmann, Fuzzy Logic Introduction. Universite de Rennes 1, France.

E. H. Mamdani and S. Assilian, *"An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller"*. Int. J. Man-Machine Studies, vol. 7, pp. 1-13, 1975.

Mamdani EH (1974). Application of fuzzy algorithm for control of simple dynamic plant 1974. IEEE Proc 1974; 121: 715–719.

Takagi T, Sugeno M (1985). *Fuzzy identification of systems and its application to modelling and control*. IEEE Trans Systems, Man and Cybernetics 1985; 15: 116–132.

Tools for computational intelligence. Available at: http://www.data machine.nl/ fuzzy1.htm. Retrieved 30 May 2011.

Malik Shahzad K. A. and Mian Muhammad Awais (2009). *Prediction weather events using fuzzy rule based system*. Applied Soft Computing, pp. 56-63.

Rita Pongracz, Eotvos Lorand University, Budapest, Hungary; and J. Bartholy, A. Philipp, C. Beck, and A. Kern (2010), *"Application of different weather pattern classifications to simulated future climate conditions for Central Europe"*. 18th Comference on Applied Climatology, 2010.

Philip Sallis, Mary Jarur and Marcelo Trujillo (2009), *"Frost Prediction Characteristics and Classification Using Computational Neural Networks"*. Springer Berlin/Heidelberg, pp 1211-1220.

Gamal A. M. Al-Shawadfi and Hindi A. Al-Hindi (2003*), " Automatic Classification Using Neural Networks"*. International Journal Of The Computer, The Internet and Managment, Vol. 11, No.3, pp. 72 -82.

D Panagoulia, A Grammatikogiannis, A Bardossy (2006*), "An Automated Classification Method of Daily Circulation Patterns for Surface Climate Data Downscaling Based On Optimized Fuzzy Rules"*. Global NEST Journal , Vol 8, No 3, pp 218-223.

eMathTeacher (2008). Available at: http://www.dma.fi.upm.es/java/fuzzy/fuzzyinf/ mamdani_en.htm

Jason Gluckman (2006), Rain Gauges. Available at: http://EzineArticles.com /269029

Membership Functions, Fuzzy Logic Toolbox. Available at: http://radio.feld.cvut.cz/ matlab/toolbox/fuzzy/fuzzytu3.html