PARAMETERS ESTIMATION OF HOLT-WINTER SMOOTHING METHOD USING GENETIC ALGORITHM

NUR INTAN LIYANA BINTI MOHD AZMI

A dissertation submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Science (Mathematics)

Faculty of Science
Universiti Teknologi Malaysia

JANUARY 2013

I dedicated with greatest love and affection to my parents. Their love, joy and wisdom have not only graced my life but shine in every line of this Thesis.

ACKNOWLEDGEMENTS

Alhamdulillah, with the resources and guidance of Allah S.W.T, I managed to complete this technical report as scheduled.

I would like to take this opportunity to convey my sincere appreciation to my supervisor, Prof. Dr. Hj. Zuhaimy Hj. Ismail for his excellent supervision, motivation and valuable advice given through the course of this project even though he has a tight schedule. I am particularly indebted to Assoc. Prof. Dr. Ismail Mohamad, the mathematics lecturer for the course MSM1280 for his thoughtful and constructive input towards this project.

Without a doubt, my part to make this report a reality would not be possible without the support of my parents and siblings especially to my brother Muhammad bin Mohd Azmi. They all have a tremendous amount of patience and to Mohd Azmi bin Mohd Noor that is my father. You are a constant source of inspiration to me.

To my entire friend especially Ahmad Zharif Salami and whomever I interact regularly, Thank you so much. Finally, I hope that my work will be beneficial to all and can be used as a reference in the future.

ABSTRACT

A powerful technique based on adaptive heuristic namely Genetic Algorithm is widely used in many fields. This technique is very popular for solving global optimization problems. In this thesis, the Genetic Algorithm approach is used to estimate the parameters of Holt-Winter Exponential Smoothing method. The value of a combination of three parameters to be optimized, namely α, β and γ must lie between 0 and 1 by minimizing the one-step ahead forecasting accuracy of Mean Absolute Percentage Error (MAPE). Moreover, the difference of the initialization method, population size and crossover probability were also used, so that the comparative study of minimum value of MAPE can be done. The overall results of the Genetic Algorithm are compared with the conventional methods. From this study, it was found that the genetic algorithm outperformed the conventional method by giving the lowest value of MAPE. Hence, this proved that the genetic algorithm is effective for estimating Holt-Winter parameters. The data used in this study are monthly data set for the total number of tourist arrivals to Langkawi from 2002 until 2011. This investigation is done using computer simulations programmed by Microsoft Visual Studio 2010.

ABSTRAK

Satu teknik yang kuat berdasarkan adaptif heuristik iaitu Algoritma Genetik digunakan secara meluas dalam banyak bidang. Teknik ini sangat popular untuk menyelesaikan masalah pengoptimuman sejagat. Dalam tesis ini, pendekatan Algoritma Genetik digunakan untuk menganggar parameter dalam kaedah Pelicinan Eksponen Holt-Winter. Nilai gabungan tiga parameter yang akan dioptimumkan, iaitu α, β dan γ mesti berada antara 0 dan 1 dengan meminimumkan ketepatan ramalan satu langkah ke hadapan iaitu Peratusan Ralat Min Mutlak (MAPE). Selain itu, perbezaan kaedah pengawalan, saiz penduduk dan kebarangkalian pindah silang juga digunakan, supaya kajian perbandingan nilai minimum MAPE boleh dilakukan.Keputusan keseluruhan Algoritma Genetik dibandingkan dengan kaedah konvensional. Daripada kajian ini, ia telah mendapati bahawa Algoritma Genetik mengatasi kaedah konvensional dengan memberi nilai terendah MAPE. Oleh itu, ini membuktikan bahawa Algoritma Genetik adalah berkesan untuk menganggar parameter Holt-Winter. Data yang digunakan dalam kajian ini adalah set data bulanan bagi jumlah bilangan pelancong yang tiba ke Langkawi dari tahun 2002 sehingga 2011. Penyiasatan ini dilakukan dengan menggunakan simulasi komputer yang diaturcara oleh Microsoft Visual Studio 2010.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

GA	-	Genetic Algorithm
HW	-	Holt-Winter
MAPE	-	Mean Average Percentage Error
MSE	-	Mean Square Error
MAD	-	Mean Absolute Deviation
LSE	-	Least Square Estimation
TSP	-	Traveling Ssaleman Problem
NLSE	-	Nonlinear Least Square
STES	-	Smooth Transition Exponential Smoothing
N-N	-	Nearest Neighbor
SA	-	Simulated Annealing
TS	-	Tabu Search
LADA	-	Lembaga Pembangunan Langkawi
ACPES	-	Automated Computerized Parameters Estimation System

LIST OF SYMBOLS

α	-	Alpha
β	-	Beta
γ	-	Gamma
Θ	-	Theta
P_i	-	Probability of selection
P_c	-	Probability of crossover
P_m	-	Probability of mutation
ε_t	-	Error
r	-	Random numbers

CHAPTER 1

INTRODUCTION

1.1 Introduction

In most real world problem, it is necessary to predict what will the future look like. This prediction of future events is called forecasting. Forecasting is the estimation or prediction of the future events based on past and present data. It is an aid to decision making and planning that can provide potential information about future and consequences for the organization. There are two broad classes of forecasting techniques namely qualitative and quantitative.

The qualitative techniques are subjective forecast techniques whereby there are no available past data. It relies mostly on judgment and opinion of an expert via their experience and knowledge. This class of forecasting is specially applied to cases in intermediate-long range decisions. Some examples of this subjective forecast are the Delphi method, the executive opinion method, market survey, historical analogy etc. The opposite of qualitative techniques is quantitative techniques. It is known as objective forecast that rely mostly on the collection and analysis of numerical data and statistics. This class of technique usually applied to short-intermediate range decisions. There are two types of objective forecast techniques that are univariate and explanatory model.

Forecasting has been used since last century. This knowledge started by fortune tellers, astrologers, priest, and prophets have sought to fulfil man's need to predict the future and reduce its certainties (Makridakis and Wheelwright, 1979). Nowadays, the conventional method of forecasting has been used widely in many fields such as supply chain management, weather forecasting, transport planning, economic, technology, earthquake prediction, product forecasting, politic, energy consumption and many more. But, still until today, people are looking for best prediction. Therefore, the optimization technique can give best future value in term of minimizing the forecasting error. The combination between the two knowledge of forecasting and optimization can help to gain the objective for better result in the future.

Optimization is the process of making something better or finding the best solution. The terminology best solution implies that there are more than one solution and the solution are not equal values (Haupt and Haupt, 2004). The optimization problems are complex in nature and hard to solve by conventional optimization techniques (Gen *et al.*, 2008). One of the optimization techniques that can solve complex and hard problem is Genetic Algorithm (GA). GA can outperform conventional optimization method when applied to real world problem.

Genetic Algorithm was first introduced by John Holland in the early 1970s at the University of Michigan. It is an adaptive heuristic search method based on population genetics. This concept was developed from the Darwinian theory of evolution which is the survival of the fittest. The basic concept of GAs is designed to simulate processes in natural system necessary for evolution which consists of a population of chromosomes that evolve over a number of generations and are subject to genetic operators at each generation. The operators of GAs are reproduction, mutation, crossover and inversion. Recently, many applications such as transportation, scheduling, vehicle routing and many more have used GAs. In this research, the researcher introduces GA as an approach for estimating the parameters in the exponential smoothing method of forecasting.

1.2 Background Of Study

According to Gardner Jr and Diaz-Saiz (2008), Exponential smoothing was originally suggested by Robert G. Browns in 1956 with the idea of simple exponential smoothing with continuous data. Brown then extended the research to discrete data and his early application was in forecasting the demand for spare parts in Navy inventory systems. Later in 1957, Charles C. Holt continued developed the exponential smoothing of adaptive trend. Holt's method is widely used in many applications since 1960s. Exponential smoothing implies the decreasing weights whenever the observations get older. It consists of several methods that have common property of recent values which are relatively more weight in forecasting the older observation (Makridakis *et al.*, 1998). These methods are categorized as single, double and triple exponential smoothing.

Single exponential smoothing also known as simple exponential smoothing is described by no trend equation which involved single parameter. This method was the idea of Browns which give the changes of parameters in slowly rate over time. Basically, simple exponential smoothing is based on averaging or smoothing past values in an exponentially decreasing manner. This can be described as the most recent observation receives the largest weight, α, where α is between 0 and 1. The next recent observation receives less weight, $\alpha(1-\alpha)$; the observation two time periods in the past receives even less weight, $\alpha(1-\alpha)^2$; and so forth. In one representation, the single exponential smoothing equation can be written as

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t \tag{1.1}$$

where

$\hat{Y}_{t+1}$ = The new smoothed value or the forecast value for the next period

α = The smoothing constant $(0 < \alpha < 1)$

Y_t = The new observation or the actual value of the series in period, t

$\hat{Y}_t$ = The old smoothed value or the forecast for period, t

Double exponential smoothing was the extended idea by Holts in 1950s which allowed trend in data series. It consists of two parameters to be estimated for next future values. Double exponential smoothing also known as Holts method allowed trend data to be estimated. When trend component anticipated, current slope and level need to be estimated as well. Kalekar (2004) explained that Holts method works much like single exponential smoothing except that two components must be updated each period. The level is smoothed estimate of the value of the data at the end of each period. The trend is a smoothed estimate of average growth at the end of each period. The equations involved in this method are described as follows:

(i) The current level estimate

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1}) \tag{1.2}$$

(ii) The trend estimate

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \tag{1.3}$$

(iii) The forecast for p period into the future

$$\hat{Y}_{t+p} = L_t + pT_t \tag{1.4}$$

where

L_t = The estimate of current level

α = The smoothing constant for the level $(0 < \alpha < 1)$

Y_t = The new observation or the actual value of the series in period, t

β = The smoothing constant for the trend $(0 < \beta < 1)$

T_t = The trend estimate

p = The periods to be forecast into the future

$\hat{Y}_t + p$ = The forecast for p periods into the future

While triple exponential smoothing was suggested by Holts student, Peter Winters, in 1960. Winter named the method as Holt-Winter (HW) method or often simply just Winters method. This method involved level, trend, and seasonal components that each of them contains three parameters to be estimated.They are alpha,α beta,β and gamma,γ. HW models exist in two types that are additive and multiplicative seasonality model. Usually, time series data display seasonal behavior. Seasonality is defined as the behavior in the time series data that repeats itself for every L periods whereby L is known as the seasonal length in periods. Additive seasonality model is the model that consist seasonal component is added to level and trend components. The additive model is appropriate for time series in which the growth is constant. While multiplicative seasonality model is the model that the seasonal component is multiplied by the total level and direction components. This model is appropriate for time series in which there is a growth of data. The models are described below.

- Additive Seasoanlity

  (i) Level

$$L_t = \alpha(Y_t - St - s) + (1 - \alpha)(L_{t-1} + T_{t-1}) \tag{1.5}$$

  (ii) Trend

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \tag{1.6}$$

  (iii) Seasonal

$$S_t = \gamma(Y_t - L_{t-1}) + (1 - \gamma)S_{t-s} \tag{1.7}$$

  (iv) Forecast

$$\hat{Y}_t + p = L_t + pT_t + S_{t-s+p} \tag{1.8}$$

- Multiplicative Seasoanlity

  (i) Level

$$L_t = \alpha\frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \tag{1.9}$$

(ii) Trend

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \tag{1.10}$$

(iii) Seasonal

$$S_t = \gamma\frac{Y_t}{S_t} + (1-\gamma)S_{t-s} \tag{1.11}$$

(iv) Forecast

$$\hat{Y}_t + p = (L_t + pT_t)S_{t-s+p} \tag{1.12}$$

where

L_t = The estimate of current level

α = The smoothing constant for the level $(0 < \alpha < 1)$

Y_t = The new observation or the actual value of the series in period, t

β = The smoothing constant for the trend $(0 < \beta < 1)$

T_t = The trend estimate

γ = The smoothing constant for the seasonality estimate$(0 < \beta < 1)$

S_t = The seasonal estimate

p = The periods to be forecast into the future

s = The length of seasonality estimate

$\hat{Y}_t + p$ = The forecast for p periods into the future

The value of smoothing parameters in single, double, and triple exponential smoothing must lie between 0 and 1. These parameters need to estimate in order to update the value of forecasting. Conventionally, the value of parameters α, β and γ for level, trend and seasonal components respectively are estimated using maximum likelihood estimation. In state space approach, these unknown parameters α, β and γ are estimate using maximum likelihood by minimizing

$$\mathfrak{L}^*(\theta, x_0) = n \log \left(\sum_{t=1}^{n} \varepsilon_t^2\right) + 2\sum_{t=1}^{n} \log |r(x_{t-1})| \tag{1.13}$$

where parameters $\theta = (\alpha, \beta, \gamma, \phi)$, initial states $x_0 = (L_0, T_0, S_0, S_{-1}, S_{-m+1})$ and n is the number of observation. The parameters θ and the initial states x_0 are estimate by minimizing $\mathfrak{L}^*$. Alternatively, estimate can be obtained by minimizing the one-step-ahead mean square error (MSE), minimizing the residual variance, σ^2 , or via some of other criterion for measuring forecast error. This minimizing will begin with the optimization of initial state x_0 using heuristic scheme that will discuss more detail in Chapter 3.

Instead of using maximum likelihood estimation for determining the value of three parameters in exponential smoothing, there are others approaches such as non-linear optimization or grid search. Grid search works with giving a limitation value for α, β and γ with increment. For example, giving $\alpha = 0.1$ to $\alpha = 0.9$ with an increment of 0.2. Then α is chosen if it produce the smallest forecast error such as one-step-ahead mean square error (MSE) or one-step-ahead mean absolute percentage error (MAPE). However these approaches are consuming a lot of time.

1.3 Problem Statement

There are 941,192 possible combination of selecting α, β and γ when we set these parameters to two decimal places, so the problem is to find the best possible combination of α, β and γ to minimize the objective function. To overcome this problem, we proposed this research for estimating Holt-Winter's parameters using one of the heuristic technique namely Genetic Algorithm (GA). GA is an adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. It is popular technique in heuristic and also a powerful technique based on the robustness and ease of use.

1.4 Objectives Of Study

The objectives of this research are:

(i) To estimate the parameters of Holt-Winter smoothing method by using Genetic Algorithm.

(ii) To analysis the influences of operators in Genetic Algorithm.

(iii) To compare the parameters obtained between the conventional method of least square estimation and Genetic Algorithm.

1.5 Scope Of Study

This study focused on designing Genetic Algorithm for finding possible combination of three parameters α, β and γ of level, trend and seasonal components respectively. The models that belong to these components are known as Holt-Winter exponential smoothing. In order to optimize these parameters, we estimated them by minimizing the one-step-ahead Mean Absolute Percentage Error (MAPE). In Genetic Algorithm, we used a rank selection for selecting parents. While the operator, we used is crossover probabilities of 0.5, 0.75 and 0.9. When creating a new population, there will be a possibility that we will have chances of losing best chromosome or best combination of parameters. To avoid this problem, we implement elitism as a selection to keep the best minimum MAPE's value. For the simulation, we used Microsoft Visual Studio 2010 for solving Iterative Holt-Winter together with Genetic Algorithm. The data used for this study is real data of visitor arrival into Langkawi from year 2002 until 2010. The data were obtained from Lembaga Pembanguna Langkawi (LADA).

1.6 Significance Of Study

The finding of this study is important in highlighting the application of heuristic method namely Genetic Algorithm in estimating the parameters of Holt-Winter smoothing method. This study can also be the references for solving real life problem in forecasting.

1.7 Thesis Organization

This project consists of six chapters. The first chapter of Introduction includes topics on the background of study, problem statement, objective of study, scope of study and the significance of the study. Second Chapter is discussed in the literature reviewed whereby the analysis required in two main topics of this study. The first topic is regarding Holt-Winter Smoothing and the second one is regarding Genetic Algorithm. The third chapter briefly explained on how this study will be conducted whereby all procedure needed for Holt-Winter and Genetic Algorithm are discussed. The fourth chapter is about the implementation of Holt-Winter and Genetic Algorithm in solving the problem of this study. The fifth chapter is the chapter discussed about the software development for estimating parameters. Lastly, the final chapter that is chapter six are discussed about the analysis of the result obtained. The conclusion and recommendation for further study were also being discussed in here.

REFERENCES

Agapie, A. (1997). Forecasting the economic cycles based on an extension of the Holt-Winters model. A genetic algorithms approach. 96–99.

Billah, B., King, M., Snyder, R. and Koehler, A. (2006). Exponential smoothing model selection for forecasting. *International Journal of Forecasting*. 22(2): 239–247.

Bowerman, B., O'Connell, R. and Koehler, A. (2005). *Forecasting, time series, and regression: an applied approach*. Duxbury Applied Series. Thomson Brooks/Cole.

Corberán-Vallet, A., Bermúdez, J. and Vercher, E. (2011). Forecasting correlated time series with exponential smoothing models. *International Journal of Forecasting*. 27(2): 252–265.

Davis, L. et al. (1991). *Handbook of genetic algorithms*. Vol. 115. Van Nostrand Reinhold New York.

Falkenauer, E. (1999). Applying genetic algorithms to real-world problems. *IMA Volumes In Mathematics And Its Applications*. 111: 65–88.

Freitas, A. (2002). *Data Mining and Knowledge Discovery with Evolutionary Algorithms*. Natural Computing Series. Springer.

Gardner, E. (2006). Exponential smoothing: The state of the artPart II. *International Journal of Forecasting*. 22(4): 637–666.

Gardner Jr, E. and Diaz-Saiz, J. (2008). Exponential smoothing in the telecommunications data. *International Journal of Forecasting*. 24(1): 170–174.

Gen, M. and Cheng, R. (1997). *Genetic Algorithms and Engineering Design*. Engineering Design and Automation. Wiley.

Gen, M., Cheng, R. and Lin, L. (2008). *Network Models and Optimization: Multiobjective Genetic Algorithm Approach*. Decision Engineering. Springer.

Goldberg, D. (1989). *Genetic algorithms in search, optimization, and machine learning*. Artificial Intelligence. Addison-Wesley.

Goldberg, D., Korb, B. and Deb, K. (1989). Messy genetic algorithms: Motivation, analysis, and first results. *Complex systems*. 3(5): 493–530.

Goodwin, P. (2010). The Holt-Winters Approach to Exponential Smoothing: 50 Years Old and Going Strong. *The International Journal of Applied Forecasting, Forthcoming*.

Hanke, J. and Wichern, D. (2009). *Business Forecasting*. 9 Edition. Pearson Prentice Hall.

Haupt, R. and Haupt, S. (2004). *Practical Genetic Algorithms*. Wiley-Interscience publication. Wiley.

Hedar, A. and Ahmed, A. (2004). Studies on Metaheuristics for continuous global optimization problems.

Hsu, L. (2010). A genetic algorithm based nonlinear grey Bernoulli model for output forecasting in integrated circuit industry. *Expert Systems with Applications*. 37(6): 4318–4323.

Kalekar, P. (2004). Time series forecasting using Holt-Winters exponential smoothing. *Kanwal Rekhi School of Information Technology*.

Kumar, R. and Jyotishree (2012). Blending Roulette Wheel Selection & Rank Selection in Genetic Algorithms. In *International Journal of Machine Learning and Computing, Vol. 2, No. 4*.

Makridakis, S. G., Wheelwright, S. and Hyndman, R. (1998). *Forecasting: methods and applications*. Wiley series in management. Wiley.

Makridakis, S. and Wheelwright, S. (1979). Forecasting: framework and overview. *TIMS Studies in the Management Sciences*. 12: 1–15.

Michell, M. (1998). *An Introduction to Genetic Algorithms*. Complex Adaptive Systems Series. Mit Press.

Mitra, S. and Mitra, A. (2011). A genetic algorithms based technique for computing the nonlinear least squares estimates of the parameters of sum of exponentials model. *Expert Systems with Applications*.

Montgomery, D., Jennings, C. and Kulahci, M. (2008). *Introduction to Time Series Analysis and Forecasting*. Wiley Series in Probability and Statistics. John Wiley & Sons.

Sivanandam, S. and Deepa, S. (2008). *Introduction to Genetic Algorithms*. Springer London, Limited.

Weiers, R. (2010). *Introduction to Business Statistics (Book Only)*. Cengage South-Western.