# COMBINE HOLTS WINTER AND SUPPORT VECTOR MACHINES IN FORECASTING TIME SERIS

ALFA MOHAMMED SALISU

A thesis is submitted infulfilment of the
requirements for the award of the degree of
Master of Science (Mathematics)

Faculty of Sciences
Universiti Teknologi Malaysia

DECEMBER 2013

# DEDICATION

IN MEMORY OF MY PARENTS

MAY ALLAH HAVE MERCY ON THEM

AMEEN

## ACKNOWLEDGEMENT

In the name of Allah, the most beneficent and the most merciful, my special appreciation and praises are due to Allah (SWT) for the opportunity and protection to undertake this programme. May His name be glorified always and may the peace and blessings of Allah continue to be on the seal of prophets (SAW).

I hereby express my heartfelt gratitude to my Supervisor, Dr.Ani binShabri for his support, guidance and patience which made it possible for me to begin and complete this work in record time. I also thank him for introducing me to a new area in time series forecasting, which is the support vector machines (SVM). It is my prayers that Allah will bless and guide him in this world and in the hereafter.

Special thanks go to examiners who gave useful direction during my first assessment/ proposal defense,Assoc Prof Dr. Maizah Hura Ahamad (Examiner) and Assoc. Prof. Dr. Robiah Adnan (Chairperson). Also, my special thanks go to the panel of examiners during my Viva voce, Prof. Dr. Mohd Nor bin Mohamad (Chairperson), Prof. Dr. Hj Zhuaimy bin Hj Ismail (Internal Examiner) and Assoc. Prof. Dr. Zazli Chik, Universiti Malaya (External Examiner). I would also like to thank Madam Shalizah and Ms. Siti Amirah in the faculty of Science office for all their assistance and cooperation during my programme.

To my parents, may their souls rest in perfect peace and I thank the members of my family for their encouragement, patience and prayers.

To my employers, I thank the Rector, the Deputy Rector and the other management staff of the Federal polytechnic, Bida for releasing and supporting me for the programme. Other members of staff and students are not left out.

Furthermore, I will like to express my special appreciation to Messrs Victor Okolobah, Mustapha Adamu and Abdullahi A. Ndawancin who facilitated and encourage me for this programme, may God see them through in their endeavors. Besides that, I thank my colleagues both in Federal polytechnic, Bida, Universiti Teknologi Malaysia (UTM) and others who always give my family support while undertaking this programme.

## ABSTRACT

This study proposes on a combine methodology that exploits the Holts-Winter (HW) model and the Support Vector Machines (SVM) model in forecasting time series. Problems of forecasting using time series data have been and still being addressed at every sphere of research using different approaches. The performance of the forecast was compared among the three models, the HW model, the SVM model and the combine model (HW and SVM). Four different data sets namely, airline passengers' data, machinery industry production data, clothing industry data and sugar production data were considered in the study. The statistical measures such as mean squared error (MSE), mean average error (MAE) and correlation coefficient, R, were used to evaluate the performance of the propose model. The result of this study indicated that the combine model shows an improvement of 149.3% over HW model and 35.9% improvement over the SVM model for the airline passengers' data. The result of the machinery industry presented that the combine model shows an improvement of 93.3% over HW model and 42.8% improvement over the SVM model. In the case of the clothing industry the result shows the combine model gives an improvement of 61.6% over HW model and 12.0% improvement over SVM model. Lastly, with respect to the sugar production, the result shows that the combine model indicated an improvement of 34.4% over HW model and 25.1% improvement over SVM model. Therefore the results of the experiments suggest that the proposed combine model is more reliable in time series when compared with the individual models.

# ABSTRAK

Kajian ini mencadangkan satu gabungan kaedah yang mengeksploitasi model Holts – Winter (HW) dan model *Support Vector Machines* (SVM) dalam peramalan siri masa. Masalah ramalan menggunakan data siri masa telah dan masih ditangani di setiap bidang penyelidikan yang menggunakan pendekatan yang berbeza. Prestasi ramalan telah dibandingkan antara tiga model, model HW, model SVM dan model gabungan (HW dan SVM). Empat set data yang berbeza iaitu, data penumpang syarikat penerbangan, data pengeluaran industri jentera, data industri pakaian dan data pengeluaran gula telah dipertimbangkan dalam kajian. Pengukuran statistik seperti min ralat kuasa dua (MSE), min ralat purata (MAE) dan pekali korelasi, R, digunakan untuk menilai keupayaan model yang dicadangkan. Hasil kajian ini menunjukkan bahawa gabungan model menunjukkan peningkatan sebanyak 149.3% berbanding model HW dan peningkatan 35.9% berbanding model SVM untuk penumpang syarikat penerbangan. Hasil daripada industri jenteramendapati model gabungan menunjukkan peningkatan sebanyak 93.3% berbanding model HW dan peningkatan 42.8% berbanding model SVM. Dalam kes industri pakaian, hasil kaj ian menunjukkan bahawa model gabungan memberikan peningkatan sebanyak 61.6% berbanding model HW dan peningkatan 12.0% berbanding model SVM. Akhir sekali, berkaitan dengan pengeluaran gula, hasilnya menunjukkan bahawa model gabungan menunjukkan peningkatan sebanyak 34.4% berbanding model HW dan peningkatan 25.1% berbanding model SVM. Oleh itu, keputusan kajian menunjukkan bahawa model gabungan yang dicadangkan adalah lebih dipercayai dalam siri masa berbanding dengan model individu.

# TABLE OF CONTENT

## REFERENCES


Alexandra K. (2011). Very short- term load forecasting using exponential smoothing and ARIMA models. Journal of information, control and management systems. Vol.9 no.2, 85-92.

Amit, K., Hari, S. (2011). Optimization of proceeding parameters in induction hardening using response surface methodology. Sadhana, Indian Academy of Sciences. Vol. 36. Pp141-152.

Asefa, T., Kemblowski, M., Mckee, M., Khalil, A. (2006). Multi-time scale stream flow predictions: the support vector machines approach. Journal of hydrology. 318, 7-16.

Auria, L.Moro, A.R. (2008). Support Vector Machines (SVM) as a technique for solvency analysis. DIW Berlin discussion papers 811. http://www.diw.de.

Billah, B., King, M. L., Snydar, R. D. and Koelhler, A. B. (2006). exponential smoothing model selection for forecasting. *International journal of forecasting*. 22, 239-247.

Bozic, M. and Stojanovic, M. (2011). application of SVM method for methods for mid-term loard forecasting. *Serbian journal of electrical engineering*. 8(1), 73-83.

Brown, R. G. (1959). Statistical forecasting for inventory control. New York, McGraw-Hill.

Gardner, E. S. (1985). Exponential Smoothing: The state of the Art. Journal of forecasting. 4, 1-28.

Cao, D. Z., Pan, S. L. and Y.H, B. (2005). Forecasting exchanging rate using support vector machines. *International conference on machine learning and cybernetics,*.

Cao, L. and Tay, F. E. H. (2001). financial forecasting using support vector machines. *Neurocomputing and application*. 10, 184-192.

Cao, J.L. and Tay, H.E.F. (2003). Support Vector Machines with adaptive parameters in financial time series forecast. IEEE transaction ons on networks 14(6), 1506-1518

Chang, C., Lin, C. (2011). LIB SVM. A library for support vector machines. ACM Transactions on intelligent Systems and Technology 2(3): 1-27

Chang, K.W. Hsieh, C.J. Lin, C.J. (2008). Co-ordinate Decent method for large-scale L2-loss linear support Vector machines. Journal of machine learning Research 9, 1369-1398.

Changxiu, Z. C. and Xu xiaoling, W. W. (2002). A novel approach based on support vector machine to forecasting the quality of friction welding. *4th World congress on intelligent control and automation*.

Chen, K.-Y. and Wang, C.-H. (2007). A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. *Expert Systems with Applications*. 32(1), 254-264.

Chen, K.Y.(2011). Combining linear and nonlinear model in forecasting. Expert Systems with Applications. 38, 10368-10376.

Chen, W-H., Shih, J-Y(2006). A study of Taiwan's issuer credit rating systems using support vector machines. Expert systems with application, ELSEVIER. 30: 427-435.

Clara, C., Neves M. M. (2009). Forecasting time series with boot. EXPOS procedure. RENTSTAT-Statistical journal vol. 7(2), 135-149.

Colojoar, A. (2012). modeling seasonal time series. *Surveys in mathematics and its application*. 1, 1842-6298.

Cook, C.N., Carter, B.W.R., Fuller, A.R., Hockings, M. (2012). Managers consider multiple lines of evidence important for biodiversity management decisions. Journal of enviromental management xxx 1-6

Cortes, C., Vapnik, V. (1995). Support vector networks. Maching learning 20 (3): 273-279.

Cristianini, N., Guyon, I., Weston, J., Barnhill, S. (2002). Gene selection for Cancer clssification using support Vector machines. Kluwer Academic publishers 46, 389-422.

Crone, S. F. and Dhawan, R. (2007). Forecasting Seasonal Time Series with Neural Networks:A Sensitivity Analysis of Architecture Parameters. *International Joint Conference on Neural Networks,*.

Dawson, C. W., Abrahart, R. J., See, L. M.(2007). Hydro Test: A web-base toolbox of evaluation metrics for the standardized assessment of hydrological forecasts. Enviromental model and software, 22, 1030-1052.

Denton J. W. (1995). How Good are Newtworks for Causal Forecasting? Journal of Business Forecasting. 14, 17-20

Dong, D., Fang, P., Bock, Y., Cheng, M.K., Muyazaki, S. (2002). Anotomy of apparent seasonal variations from GPS- derived site position time series. Journal of Geophysical Research 107(B4 10.1029)

Duan, K., Keerthi, S. S., Poo, A. N. (2003). Evaluation of simple performance measures for turning SVM hyper parameters. Neurocomputing. Vol.51, pp 41- 59

Everette, S. Gardner, Jr. Elizabeth, A. Anderson, F. Angela, M.W.(2001). Further results on focus forecasting vs. exponential smoothing. International journal of forecasting. 17, 287-293.

Fatima, S. (2007). Hybrid system of simple expontential smoothing and neural network for KSE 100 Index. *Market force: Journal of management thought*. 4(2).

Feng, L., Yao., Jin,B. (2010). Research on credit scoring model with SVM for network management. Journal of computational information system 6:11, 3567-3574.

Francis, E. H. and Cao, T. L. (2001). Application of support vector machines in F!nancial time series forecasting *International journal of management science*. 29, 309-317.

Froglich, H., Zell, A. (2005). Efficient parameter selection for support vector machines in classification and regression via model based global optimization. proceedings of the 2005 IEEE international conference on neural networks (IJCNN '05). Vol.3 pp 1431-1436

Gao, S., Zhang, Z. and Cao, C. (2010). Road traffic frieght volume forecasting using support vector machine. *International conference on computer science and computational technology (ISCSCT'10)*. 329-332

Gelper S., Fried R. and Croux C. (2008). Robust Forecasting with exponential and Holt-Wnters smoothing. International journal of forecasting. No.22 pg 637-666

Han, D., Chan, L. and Zhu, N. (2007). flood forecasting using support vector machines. *Journal of hydroinformatics*. 9(4), 267-276.

Hsu C., Chang C., Lin C. (2010). A practical guide to support vector clssification. National Taiwan university, Taipei 106, Taiwan: http://www.csie.ntu.edu.tw/cjlin

Huang, W., Nakamori, Y., Wang S. (2005). Forecasting stock market movement direction with support vector machines. Computer and operation research. 32 (10): 2513-2522

Ismail S., Shabri A., Samsudin R, (2012). A hybrid model of self organising maps and least square support vector machine for river flow forecasting.Hydrology and Earth system sciences.16,4417-4433

Joseph, J. and Viola (2003). Double exponential smoothing: An alternative to kalman filter-basedpredictive tracling. International Immersive projection Technologis workshop.

Kandananond, K. (2012). A comparison of various forecasting methods for Autocorrelation time series. International journal of Engineering Business Management. 4(4) 1-6.

Kalekar S. P. (2004). Time series forecasting using Holt Winters exponential smoothing. Kanwal Rekhi school of information Technology

Khashei, M. and Bijari, M. (2011). Which method is better for combining linear and nonlinear models for time series forecasting. Journal of industrial and systems engineers 4(4) 265-285.

Khashei, M. and Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing*. 11(2), 2664-2675.

Kim, K. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*. 55, 307-319.

Kim, D., Cho, S,. Lim, Y. (2005). Performance improvement in traffic vision system using SVM. Journalof Eastern Asia society for transportation studies. 5, 2589-2599.

Lapedes, A., Farber, R. (1987). Nonlinear Signal Processing Using Neural Network Prediction and System Modeling. Theoretical Division, Los Alamos National Laboratory, NM Report. No. LA-UR-87-2662.

Leung, M. T., Daouk, H., Chen, A. S. (2000). Forecasting Stock Indices: A Comparison of Classification and level Estimation Models. International Journal of forecasting. 16, 173-190

Lilien, G. L., Kotler, P. (1983). Marketing Division Making: A Model Building Approach. New York, Harper and Row Publishers

Liu, Z., Ren, G, Shi, X. (2006). Research Chaotic SVM with Incorporated Intelegence Algorithm forecasting model in SCM. IJCSNS international journal of Computer Science and Network Security. 6(10) 136-141.

Lu, W. z. and Wang, W. J. (2005). potential assessment of the support vector machine method in forecasting ambient air pollutant trends. *Chemosphere* 59, 693-701.

Makridakis, S., Wheelwright, S. C., Mcgee, V. E. (1983). Forecasting: Methods and Applications. New York: Wieley and Sons. 2nd Ed.

Makridakis, S., Wheelwright, S. C. (1989). Forecasting Methods for Management. New York: Wieley and Sons. 5nd Ed.

Mavroforakis, E.M. and Theodoridis, S. (2006). A Geometric Approach to support vector machine (SVM) Classification. IEEE Transaction on Neura Networks. 17(3) 671-682.

Merh,N., Saxena,P.V., Pardasani, R.K. (2010). A Comparison between Hybrid approaches of ANN and ARIMA for Indian stock trend forecasting. Business Intellegence journal. 23-43

Mezghani, A.B.D., Boujelbene, Z.S., Ellouze, N. (2010). Evaluation of SVM kernels and conventional machine learning Algorithm for speaker identification. International journal of Hybrid information technology. 3(3) 23-34.

Mojaveri, S.R.M.,Mousavi, S.S., Heydar, M., Aminian, A. (2009). Validation and selection between machine learning techniques and traditional methods to reduce building effect: A data mining approach. International journal of Human and Social Sciences. 4(16) 1164-1170.

Pai, P.F., Lin, C.S. Hong, W.C., Chen, C.T. (2006). Hybrid Support Vector Machine Regression for exchange rate prediction. Information and management sciences. 17(2) 19-32.

Pai, P. F. and Lin, C. S. (2005). Using support vector machines to forecast the production values of the machinery industry in Taiwan. *International Journal of Advance Manufacturing Technology*. 27, 205-210.

Paul, S. K., Azeem, A. (2011). determination of exponential smoothing constant to minimize mean square error and mean absolute deviation. *Global journal of research in engineering*. 11(3).

Phadke, A. C., Rege, P. P. (2013). Classification of Architectural Distortin from other Abnormalities in Mammograms. Internaional journal of Application or Innovation in Engineering and Management (IJAIEM) Vol.2, Issue 2. ISSN 2319 - 4847.

Prasert, C., Chukiat, C., Ratchancee. M. (2009). Time series models for forecasting international visitor arrivals to Thailand. Interantional conference on Applied Economics. 159-163.

Qi, W.(2010). A hybrid forecasting model based on Gaussian support vector machine and chaotic particle swarm optimization. Expert systems with applications, ELSEVIER. 37, 2388-2394.

Refenes, A. N., Azema-Barac, M., Chen, L., Karoussos, S. A. (1993). Currency Exchange Rate Prediction and Neural Network Design Strategies. Neural Computing Appling. 1, 46-58.

Reid M.J. (1968), Combining three estimates of gross domestic product; *Economica* 35; 431–444.

Saeid, A., Dietrich, V. R., Silvelyn, Z. (2009). The SVM approach for the Box-Jenkins models. REVSTAT- Statistical journal v.7 (1), pg 23-36.

Samson, D.C. Downs, T. Saha, T.T.(2004). Evaluation of support vector machine based forecasting tool in electricity price forecasting for Austrian national electricity market participants. Journal of electrical and electronics engineering, Austria. vol.22 no.3.

Samsudin, R., Shabri, A., Saad, P.(2010). A Comparison of time series forecasting using Support vector machine and Artificial neural network model.Journal of applied sciences 10(11): 950-958.

Shabri, A. and suhartono (2012). streamflow forecasting using least- squares support vector machines. *Hydrological sciences journal* 57(7), 1275-1293.

Shin, J.H. and Cho, S. (2006). Response modeling with support vector machines. Expert System with applications. 30, 746-760.

Shik, K., Taik, S. L., Kim, H-J.(2005). An application of support vector machines in bankrupcy prediction model. Expert systems with applications. ELSEVIER. 28, 127-135.

Shumway, H.R., Stoffer, S.D. (1982). An approach to time series smoothing and forecasting using the E.M. Algorithm. Journal of time series Analysis. 3(4) 253-264.

Smits, G., F., Jordaan, E., M. (2002). Improved SVM Regression using mixtures of Kernels. IEEE. o - 7803 - 7278, pg 2785 - 2790.

Smits, G., F., Jordaan, E., M. (2002). Estimation of the regularization parameter for suppor vector regression. Poceedings of international joint conference on Neural Networks (IJCNN '02) Vol.3 pp 2192-2197

Štepnicka, M., Peralta, J., Cortez, P., Vavrícková, L. and Gutierrez, G. ( 2011). Forecasting seasonal time series with computational intelligence: contribution of a combination of distinct methods. *EUSFLAT-LFA*.

Suhartona, Muhammad, H. L.(2011). A hybrid approach based on Winters model and weighted fuzzy time series for forecasting trend and seasonal data. Journal of Mathematics and Statistics. 7(3). 177-183. ISSN 1549-3644.

Tang, X., Yang, C., Zhou, J. (2009). Stock price forecasting by combining news mining and Time Series Analysis. IFEEE/WIC/ACM international conference on web intellegence and intellegent agent technology-workshops.

Tay, F. E. H., Cao, L. (2001). Application of support vector machines in financial time series forecasting. Omega, 29 (4) : 309-317.

Taylor, W.J.(2004). Smooth Transition Experimental Smoothing. Journal of forecasting. 23, 385-394.

Valeriy, V. G., Supriya, B.(2006). Support vector machine as an efficient frame work for stock market volatility forecasting. Science applications international corporation Springer CMS, 3: 147-160.

Vapnik, V. (1998). Statistical learning theory New york. Wiley.

Vicram, P. and Veer, R.P. (2011). Rainfall forecasting using Nonlinear SVM based PSO. International journal of computer science and information technology. 2(5) 2309-2313.

Vladimir, C., Yunqian, M. (2004). Selection of SVM parameters and noise estimation for SVM regression. ELSEVIER, Neural networks, 17, 113-126.

Vladimir, C., Yunqian, M. (2003). Comparison of model selection for regression. Neural computation, vol.15, issue 7, pp 1691-1714

Wei, H., Yoshiteru, N., Shou, Y. W.(2004). Forecasting stock market movement direction with support vector machines research, 32:2513-2522.

Xuemei, L., Lixing, L., Yuyuan, L. and lanlan, L. (2010). Hybrid Support Vector Machine and ARIMA Model in Building Cooling Prediction. *International Symposium on Computer, Communication, Control and Automation*.

Zhang, G., Patuwo, E.B., Hu, Y.M. (1998). Forecasting with artificial neural network the state of the art. International journal of forecasting. 14, 35-62.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neuro computing.ELSEVIER vol.50:159-175.

Zakaria, Z. A. (2012). Streamflow forecasting at ungaged sites using support vector machines. *Applied Mathematical sciences*. 6(60), 3003-3014.

Zeng, J. and Qiao, W. (2011). Support Vector Machine based short-term wind power forecasting. IFEEE/PES Power system conference and Exposition 10, 1-9.

Zuhaimy, I. and Foo, F. Y. (2011). Genetic Algorithm for parameter Estimation in double exponential smoothing. Australian Journal of Basic and Applied Sciences. 5(7): 1174-1180. ISSN 1991-8178.

# LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

## 1.1 Introduction

Time series forecasting is an important practical problem with a diverse range of applications in many observational disciplines, such as finance, economics, meteorology, biology, medicine, hydrology, oceanography, physics, engineering, and geomorphology. Forecasting can assist them to make a better development and decision-making for most of the organization. The identification of highly accurate and reliable time series forecasting models for future time series is a precondition for successful planning and management for applications in variety of areas.

Trend and seasonal variations are the two most commonly encountered phenomena in many sectors of business and economics. Seasonality is periodic and recurrent pattern caused by factors such as weather, holidays, repeating promotions as well as the behavior of economic agents.

A seasonal time series is assumed to be a series which is influenced by seasonal factors. A seasonal factor is a characteristic of time series in which the data experiences regular and predictable changes which reoccur every calendar year. Example of this are the annual data of monthly industrial production, monthly airline passengers' data. Seasonal variation on the other hand is a component of a time series which is a repetitive movement around the trend line in one year or less. Some organizations facing seasonal variation should be interested in knowing their performance relative to the normal seasonal variation. Organizations that are affected by seasonal variation need to identify and measure this seasonality to help in planning for temporal increase or decrease in organizational productions (Dong et al 2002).

A very difficult task for any organization is to forecast using the seasonal time series data. Accurate forecasting of seasonal and trend time series is very important for effective decisions in planning, marketing, production, personnel, inventory control and other related decision making activities. Problems of forecasting using seasonal time series data have been addressed and still being addressed by different researchers using different approaches.

In this chapter, we will give a brief discussion on the research background of seasonal time series, in addition to the problem statement along with the objectives and scope of the current study. Further to that we will discuss the significance of the study of seasonal time series to organizations.

## 1.2 Background of the Study

Modeling seasonal and trend time series has been one of the main researches for several decades. Thus, various kinds of forecasting methods have been developed by many researchers and business practitioners. Of the various forecasting models, the work by Box and Jenkins (1976) on the seasonal ARIMA model has the most significant and popular on the practical applications to seasonal time series modeling. This model has performed well in many real world applications and is still one of the most widely used seasonal forecasting methods. The popularity of ARIMA model is due to its statistical properties as well as the well-known Box-Jenkins methodology in the building process. Although ARIMA models are quite flexible in that they can represent several different types of time series, seasonal or non-seasonal time series, their major limitation is the pre-assumed linear form of the model. The ARIMA model is, therefore, unable to find subtle nonlinear patterns in the time series data. Obviously, the approximation of linear models to complex real-world problem is not always satisfactory (Zhang, 2003).

Recently, artificial neural network (ANN) models have been extensively studied and shown their nonlinear modeling capability in time series forecasting. Being a flexible modeling tool, ANN can in principle, model any type of relationship in the data with the high accuracy. One of the most important advantages of ANN is their flexible nonlinear modeling in solving many complex real world forecasting problems. In addition, although ANN is inherently nonlinear models, the model is capable of modeling linear process as well (Zhang, 1988).

Although ANN has been successfully used for numerous forecasting applications, several issues in ANN model building still have not been solved. One of the most critical issues is how to select appropriate network architecture for finding the good forecasting. Due to a typically large number of parameters to be estimated, ANN often suffers overfitting problems. That is, ANN sometimes can fit in-sample data vary well but forecast poorly out of sample (Zhang, 2001). ANN is also used to study seasonal time series, although many of the studies reach different conclusions (Hamzacebi 2008).

More advanced, Artificial Intelligence (AI) is support vector machine (SVM) is proposed by Vapnik and his co-workers in 1995 through statistical learning theory, has gained the attention of many researchers. SVM is nonlinear model which is one of the soft computational techniques is a powerful methodology and has been successfully applied to solve various problems (Thiessen Wang et al. 2009; Asefa et al. 2006; Lin et al. 2006; Liong and Sivapragasam 2002; Yu et al. 2006). The standard SVM is solved using quadratic programming methods. This method needs very few assumptions to learn patterns of input variables for prediction variables since it deals with nonlinear models. According to Auria and Moro (2008) SVM on the other hand has advantage of not necessarily creating dummy variables while dealing with categorical variables but also there is no limit in the number of independent variables. Han et al. (2007) stated that the foundation of SVMs was developed by Vapnik, who was a Russian Mathematician at the beginning of 1960s popularly referred to as “Vapnik 1965” whose principle was based on the structural Risk Minimization principle from the statistical learning theory and this gained popularity due to its many attraction features and promising empirical performance. Han et al (2007) also viewed that SVM has been proved to be effective in classification by many researchers in different fields of Science, Engineering, Medical and Financial studies (Vapnik 1998) and extended to the domain of regression problems, (Kecman 2001).

SVM offers remarkable generalization performance in many areas such as pattern recognition, text classification and regression estimation (Asefa et al 2006) and also that some researchers deal with the application of SVM in time series forecasting. According to Feng et al (2010) in recent years, SVM has become a popular tool for pattern recognition and machine learning. SVM is used for classification problems and its goal is to optimize ‘‘generalization’’ (Cristianini et al 2002). However, applications of the SVM models in seasonal time series data have not been widely studied (Saeid et al, 2009). Therefore, this study attempts to use the SVM model in the seasonal time series forecasting problems.

Although the SVM models achieve success in time series forecasting, they have some disadvantages. Since the seasonal time series is a complex and nonlinear problem, there exist some linear and nonlinear patterns in the time series simultaneously. It is not sufficient to use only a nonlinear model for time series because the nonlinear model might miss some linear features of time series data. In such situations, it is necessary to combine the linear model and nonlinear model for seasonal time series forecasting.

In this study, the exponential smoothing model rather than other linear models such as ARIMA is chosen as SVM model’s complement for several reasons. First of all, the major advantage of exponential smoothing methods is that they are simple, intuitive, and easily understood. These methods have been found to be quite useful for short term forecasting of large numbers of time series. At the same time, exponential smoothing techniques have also been found to be appropriate in such applications because of their simplicity. Second, the exponential smoothing model has less technical modeling complexity than the ARIMA model and thus makes it more popular in practice.

## 1.3 Problem Statement

Exponential smoothing models have been found to be amongst the most effective forecasting models. It has been applied in many fields of human endeavors. However, it suffers from the limitation of being able to capture only linear features in time series data. Support vector machines (SVM) on the other hand though new has made remarkable in roads in the field of time series forecasting.

The current trend in forecasting practice is the combination of the two or more technique into one. This is done to harness the advantages of the techniques involved. It is this trend that has motivated this study in which we propose to develop a combine model which is a mix of Holt winters (HW) and SVM to help overcome the shortcomings of using either of the two techniques.

## 1.4 Research Questions

Therefore, our main research questions can be summarized as follows:

(a) How reliable and accurate are Holts' winter and SVM predictions for time series forecasting?

(b) When should each of the methods (exponential smoothing and SVM) be preferred, and what are the strengths and weaknesses of each procedure?

(c) Can we increase the accuracy and reliability of predictions in seasonal time series forecasting by combine exponential smoothing and SVM?

We therefore, propose in this study to apply a forecasting model that combines the traditional technique exponential smoothing with intelligent technique-SVM using time series to provide a better forecast.

## 1.5 Objectives

This research is aimed at studying seasonal time series and to develop a model which will be used to make forecast for an organization. Other objectives are:

i. To study the time series by (a) obtaining Holts' Winter model and (b) the SVM model of the time series data.
ii. To design and build a combine model for forecasting the time series data.
iii. To compare the forecasting performance of the proposed model with the individual HW and SVM models for the time series involving airlines, machinery, clothing and sugar data.

## 1.6 Scope of the Study

In this study, four different data involving airlines, machinery, cloth and sugar are analyzed to obtain the most efficient model to enhance the understanding of forecasting with time series using H-W and SVM to improve the existing methods. The forecasting performances of the models are evaluated using the performance measures like MSE, MAE and R (which is the correlation coefficient).

## 1.7 Significance of the Study

The study of seasonal time series is very important in achieving a good and reliable forecast so as to make a significant impact for future plan of organizational decisions. To study the behaviour of different time series data will also enable the decision makers to improve the organization in the areas of management and policy formulation.

## 1.8 Summary of the Research

The purpose of this work is to conduct a study to build a model for forecasting the time series data and compare the performance of this model with the H-W and SVM models for the most efficient. For this purpose, the data used are airlines, machinery, cloth and sugar.

Chapter 1 gives the introduction of this study which involves the framework; it begins with the introduction of seasonal time series data and the models to be included, the background of the study, problem statement, aim and objectives, scope of the study and the significance of the study. Chapter 2 presents the literature review in which the works of other researchers involving time series; H-W, SVM and combine methods are reviewed. Chapters 3 and 4 discussed the methodology which involves the various methods used in the study and the data analysis which describes the analysis carried out in the study respectively. Chapter 5 spelt out the results and the various discussions arising from the analysis and finally chapter 6 gave the conclusion of the study and the possible direction in the future research.

# LIST OF TABLES

# LIST OF FIGURES