COMBINING GROUP METHOD OF DATA HANDLING MODELS USING ARTIFICIAL BEE COLONY ALGORITHM FOR TIME SERIES FORECASTING

NURHAZIYATUL ADAWIYAH BINTI YAHYA

UNIVERSITI TEKNOLOGI MALAYSIA

COMBINING GROUP METHOD OF DATA HANDLING MODELS USING ARTIFICIAL BEE COLONY ALGORITHM FOR TIME SERIES FORECASTING

NURHAZIYATUL ADAWIYAH BINTI YAHYA

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Philosophy (Computer Science)

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

JUNE 2019

To those lovely people

who waited years for me to finish this thesis

ACKNOWLEDGEMENT

All praise is due to Allah S.W.T, where without His help, the completion of this thesis would be impossible. So above all else, I thank Allah S.W.T for guiding me every step of the way.

There is no amount of “Thank you’s” that can express how appreciative I am to the overwhelming support and guidance I received from my supervisor, Dr. Ruhaidah Samsudin and my co-supervisor, Associate Prof. Dr. Ani Shabri. Thank you for being patient with me and for never giving up on me despite my constant misbehaviour. Furthermore, thank you for always being considerate towards me and for your invaluable financial support. May Allah S.W.T bless both of you for all your kind deeds. To my precious family, thank you for being there for me especially during those tough times. Without them pushing me to the core, I would never get the momentum needed to finish writing my thesis. Their continuous support and encouragement are truly amazing!

A very big thanks to my dear friend, Najah Mohd Nasir for always being there for me. She played a major role in helping me complete my thesis and has helped me more than she needs to. In this lifetime, I can never repay all the things she has done for me. May Allah S.W.T make everything easy for her and her family. Najah, I am grateful for you and for all the time we shared together. I am also truly grateful for all the help I received from my lab mate, Kak Nad. Thank you for all the foods, support and patience, especially for always lending us a hand whenever we need it the most. And thank you Kak Amalina for always assisting us with our grants. Your kindness will never be forgotten. Finally, thank you to all my friends and acquaintances who have helped me in various ways throughout my journey in completing my master’s degree. May Allah S.W.T bless all of you.

ABSTRACT

Time series forecasting which uses models to predict future values based on some historical data is an important area of forecasting, and has gained the attention of researchers from various related fields of study. In line with its popularity, various models have been introduced for producing accurate time series forecasts. However, to produce an accurate forecast is not an easy feat especially when dealing with nonlinear data due to the abstract nature of the data. In this study, a model for accurate time series forecasting based on Artificial Bee Colony (ABC) algorithm and Group Method of Data Handling (GMDH) models with variant transfer functions, namely polynomial, sigmoid, radial basis function and tangent was developed. Initially, in this research, the GMDH models were used to forecast the time series data followed by each forecast that was combined using ABC. Then, the ABC produced the weight for each forecast before aggregating the forecasts. To evaluate the performance of the developed GMDH-ABC model, input data on tourism arrivals (Singapore and Indonesia) and airline passengers' data were processed using the model to produce reliable forecast on the time series data. To validate the evaluation, the performance of the model was compared against benchmark models such as the individual GMDH models, Artificial Neural Network (ANN) model and combined GMDH using simple averaging (GMDH-SA) model. Experimental results showed that the GMDH-ABC model had the highest accuracy compared to the other models, where it managed to reduce the Root Mean Square Error (RMSE) of the conventional GMDH model by 15.78% for Singapore data, 28.2% for Indonesia data and 30.89% for airline data. As a conclusion, these results demonstrated the reliability of the GMDH-ABC model in time series forecasting, and its superiority when compared to the other existing models.

ABSTRAK

Peramalan siri masa yang menggunakan model untuk meramalkan sesuatu nilai masa depan berdasarkan beberapa data masa lampau merupakan bidang ramalan yang penting, dan telah menarik perhatian penyelidik daripada pelbagai bidang pengajian yang berkaitan. Selaras dengan popularitinya, pelbagai model telah diperkenalkan bagi tujuan menghasilkan ramalan siri masa yang tepat. Namun begitu, bagi menghasilkan ramalan yang tepat bukanlah satu perkara mudah terutamanya apabila berurusan dengan data yang tidak linear disebabkan oleh sifat data yang abstrak. Dalam kajian ini, model untuk ramalan siri masa yang tepat berdasarkan model algoritma Koloni Lebah Buatan (ABC) dan Model Kaedah Kumpulan Pengendalian Data (GMDH) dengan fungsi pemindahan varian, iaitu fungsi polinomial, sigmoid, radial dan tangen telah dibangunkan. Pada awal kajian, beberapa model GMDH telah digunakan untuk meramalkan data siri masa dan setiap ramalan tersebut kemudiannya digabungkan menggunakan ABC. ABC akan menghasilkan pemberat bagi setiap ramalan tersebut sebelum mengagregatkan ramalan. Untuk menilai prestasi model GMDH-ABC yang dibangunkan, data siri masa, iaitu ketibaan pelancong (Singapura dan Indonesia) dan data penumpang penerbangan akan diinput dan diproses dengan menggunakan model GMDH-ABC untuk menghasilkan ramalan yang tepat. Bagi mengesahkan penilaian, prestasi model tersebut dibandingkan dengan model penanda aras seperti model GMDH individu, model Rangkaian Neural Buatan (ANN) dan gabungan GMDH menggunakan model purata sederhana (GMDH-SA). Hasil eksperimen menunjukkan bahawa model GMDH-ABC mempunyai ketepatan yang tinggi berbanding dengan model lain, yakni dapat mengurangkan Ralat Punca Min Kuasa Dua (RMSE) model GMDH konvensional sebanyak 15.78% bagi data Singapura, 28.2% bagi data Indonesia dan 30.89% bagi data penumpang penerbangan. Kesimpulannya, hasil kajian ini menunjukkan kebolehpercayaan model GMDH-ABC dalam ramalan siri masa, dan keunggulannya jika dibandingkan dengan model sedia ada yang lain.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF TERMINOLOGY

ABC	-	Artificial bee colony
AI	-	Artificial intelligence
AIC	-	Akaike information criterion
ANN	-	Artificial neural network
BDS	-	Brock-Dechert and Scheinkman
GA	-	Genetic algorithm
GMDH	-	Group method of data handling
MAE	-	Mean absolute error
MAPE	-	Mean absolute percentage error
PACF	-	Partial autocorrelation function
PD	-	Partial description
PSO	-	Particle swarm optimisation
RBF	-	Radial basis function
RMSE	-	Root mean square error
SA	-	Simple averaging
SSE	-	Sum of squared error

CHAPTER 1

INTRODUCTION

1.1 Overview

Forecasting of data is an important aspect that can assist in any modern organisational decision-making process and planning. A forecast can be defined as a prediction of future events based on some past or present data. Simply put, forecasting is an act of deriving or estimating what will happen in the future. Among the important area of forecasting is time series forecasting.

In time series forecasting, past data (historical data) of the same variables are collected over a duration of fixed time to be used as inputs to make a forecast. These data are called time series data. In general, the activity of time series forecasting involves developing and applying a forecasting model on data where an ordered relationship between observations exists.

Forecasting a time series is a challenging problem that has gained popularity over the years, making it an active area of research. Its popularity has led to the emergence of various forecasting models which are applied to arrays of time series problems such as in tourism forecasting (Palmer et al., 2006; Claveria and Torra, 2014), wind speed forecasting (Hu et al., 2013), hydrology (Jain and Kumar, 2007) and others. Due to its wide range of applications, research on this area is constantly carried out in order to discover more accurate methods for time series forecasting.

1.2 Background of Study

When forecasting a time series, the problem a forecaster would face is in dealing with the pattern in the data. In real-world problems, the time series data are rarely pure linear or nonlinear and are usually a mixture of both linear and nonlinear patterns (Zhang, 2003). While linear models such as Auto-Regressive Integrated Moving Average are able to capture the linearity in the data (Box and Jenkins, 1970), understanding the complex underlying nonlinear relationships in the data is not an easy feat. Various models have been used by past researchers in conducting a time series forecast, but recently a lot of attention has been given to the development of Artificial Intelligence (AI) models. One of the most popular AI models which has frequently been used in time series forecasting is the Artificial Neural Network (ANN) model due to it having a flexible nonlinear modelling capability (Zhang, 2003).

Inspired by the human brain, ANN has the ability to learn from the past data to make assumptions about the future. There are several features in ANN that makes it attractive for forecasting practitioners such as its powerful pattern recognition and classification capabilities (Zhang et al., 1998). However, the most prominent attribute of ANN lies in its ability at handling nonlinear data. Despite ANN's reliability, in practice, the implementation of ANN is quite tricky as ANN has a lot of parameters such as the number of layers and neurons which needs to be set prior to forecasting (Zhang et al., 1998).

Group Method of Data Handling (GMDH) is an AI model which has a relatively similar structures to ANN. In previous literatures, comparisons between ANN and GMDH have frequently been made where their performances vary. For example, in the research done by Ugrasen et al. (2014) in comparing the performance of GMDH and ANN, ANN was said to be superior than GMDH. On the other hand, according to Ghazanfari et al. (2017), GMDH was far more successful than ANN in terms of prediction. Nevertheless, Varahrami (2012) stated that GMDH is more reliable than ANN when the system at hand is very complex, and the underline input-output relationship are not completely comprehensible or if the system exhibits a chaotic pattern. As for Dorn et al (2012), their empirical results found that GMDH networks are simpler and can be trained faster than ANN. However, the performance of ANN and GMDH in prediction task were relatively the same. Based on the empirical findings of the past researchers, it can be seen that the predictive ability of GMDH model is comparable or at least as good as ANNs. Furthermore, GMDH has its own strength which makes it an attractive model in forecasting area.

GMDH is a nonlinear regression model introduced by Prof. Alexey G. Ivakhnenko in the late 1960s as a mean for identifying the nonlinear relationships between the input and output variables, modelling of complex systems, prediction, pattern recognition and data clustering (Ivakhnenko, 1971; Ivakhnenko and Ivakhnenko, 2000). The main idea behind GMDH is to implement a “survival-of-the-fittest” concept where models of gradually increasing complexity are sorted and estimated according to some external criterion. Similar to ANN, GMDH consists of an input layer, hidden layers and an output layer. In each hidden layer, simple neurons (or nodes) generated through a combination of two variables will perform its own quadratic polynomial transfer function and its outputs will be passed on to the neurons in the next layer. Nevertheless, prior to this process, pruning of useless neurons will take place based on some threshold value. The neurons which performs best will be kept, and the least performing neurons will be discarded. In the last layer, there is only one neuron and the output of this layer is the output of the whole net.

A basic GMDH process is based on the forward propagation of signals through neurons which is similar to the principle of classical neural network. However, the strength of GMDH lies in its ability to self-organize its own structures heuristically (AlBinHassan and Wang, 2011). Not only are the number of neurons in a layer and the number of layers generated automatically without human’s intervention, the self-organizing feature of GMDH also allows it to find the optimal solution for a given problem while avoiding bias and misjudgements (AlBinHassan and Wang, 2011). Consequently, this feature of GMDH also contributed to it having a small number of parameters to be tuned i.e. maximum number of neurons, maximum number of layers and selection pressure, making it a simple and reliable AI model (Ghazanfari et al., 2017).

Furthermore, while ANN is prone to overfitting, GMDH is reported to be resistant to the issue of overfitting (Tauser and Buryan, 2011). Perhaps this is due to

GMDH's salient feature of dividing the data into two subsets (Training and Validation sets) before initiating the learning process. Once the structure of the model has been established using both training and validation sets, the model is tested on an entirely new separate data called the testing set.

Despite GMDH's interesting potential, it does have its own drawbacks. One of its prominent limitation is its tendency to produce quite complex polynomials even for a relatively simple system, and more so if it is dealing with a highly nonlinear system due to its limited transfer function, i.e. quadratic two-variable polynomial (Hu et al., 2013). The complexity of GMDH effects the accuracy of the model in forecasting. According to Jirina (1994), as the complexity of the model increases, the degeneration of GMDH's accuracy could be due to the polynomial transfer function which causes multilayerness error to occur in GMDH's network. Meanwhile, Ivakhnenko and Ivakhnenko (1995) also mentioned that the low accuracy in GMDH might be owing to the insufficient functional variety of the model.

In order to alleviate the problems with the basic GMDH model, various modifications on GMDH has been proposed. Additionally, due to its similarities with ANN in terms of structures, numerous researches have incorporated some of ANN's features in GMDH so that the model has both characteristic of ANN and GMDH such as in the notable work by Kondo (1998). In his early works, he proposed applying many types of neurons in the GMDH model such as logistic sigmoid and polynomial transfer functions (or objective functions). Transfer function can generally be defined as the input-output explanation of the system and it expresses how the input variables are transferred through the system. In ANN model especially, it is common to apply different transfer functions from one layer to another. In GMDH however, nearly all known GMDH algorithms applied linear polynomial functions (Ivakhnenko and Ivakhnenko, 1995). Nevertheless, Ivakhnenko and Ivakhnenko (1995) mentioned that other functions can also be used such as harmonic or logistic function.

Over the years, Kondo has applied several transfer functions in GMDH such as polynomial, logistic sigmoid and Radial Basis Function (RBF) as seen in his works (Kondo et al., 1999; Kondo, 2002; Kondo and Ueno, 2009; Kondo et al., 2017). According to Kondo and Pandya (2003), employing heterogenous transfer functions within a model gives better results than using homogenous transfer function and it can fit the complexity of the nonlinear system. Other than Kondo, there are also other researchers which implemented heterogenous functions in GMDH so as to improve the problems in GMDH. For example, a research which was based on Kondo's work also implemented multiple transfer functions in GMDH's network such as polynomial, logistic sigmoid, RBF and tangent function (Dag and Yozgatligil, 2016). Additionally, Oh and Pedrycz (2002) proposed a new class of Polynomial Neural Network whereby the proposed model exploited different order of polynomials such as linear, quadratic, cubic, etc. This approach is useful in handling various nonlinear characteristics of the systems. A research carried out by Tauser and Buryan (2011) also applied seven types of transfer functions such as polynomial, harmonic (cosine), square root, inverse polynomial, logarithmic, exponential, arc tangent and rounded polynomial. The introduction of non-polynomial transfer functions is done to increase the flexibility of GMDH in modelling nonlinear system (Tauser and Buryan, 2011).

Even though many variations of GMDH has been proposed, it is a well-known fact that there is no model that can perform best in every situation. This is mainly due complex nature of real-world problem which makes it difficult for any single model to capture the different patterns in it equally well (Zhang, 2003). Nevertheless, the research done by Bates and Granger (1969) found that combining several models could significantly produce a better forecast. Hence, the combination of forecasts has since been an active area of research and has been applied by private sectors forecasters (Aiolfi et al., 2010). There are several ways to combine two or more models, but the simplest and the most flexible approach is called the weight based combining method. This approach essentially requires the practitioners to decide on two important things; a) which forecasts to include, b) how to weight them.

In previous researches, assignment of weights to the included models are often done using traditional mathematical calculations such as Simple Averaging (SA), Weight Average, or applying the inverse of Mean Square Error. The most widely used combining method is the SA method. However, this method does not exploit the past information regarding the precision of the forecasts or the dependence among the forecasts (De Gooijer and Hyndman, 2006).

A more sophisticated method of finding weights is the implementation of heuristic optimization algorithms. In this method, optimization algorithms are used to find the optimal weights for each forecast through successive iterations. Past researches that applied this method includes Xiao et. al (2015) who proposed using the Chaos Particle Swarm Optimization algorithm and the Genetic Algorithm (GA) to determine weights for combining several models. Based on that study, the implementation of optimization algorithm in weights assignment for combining forecast seems to yield a promising result and has the potential to be further explored.

1.3 Problem Statement

A nonlinear time series is different from a linear time series, whereby the changes in the output data is not proportional to the changes in the input data. Various real-world systems exhibit nonlinear characteristics, or a mixture of linear and nonlinearity. Unlike the simple linear data, the input-output mappings of a nonlinear data are difficult and hence cannot be treated satisfactorily using linear means. Therefore, GMDH model which belonged to the family of universal approximators provides an ideal means for the modelling of the complex nonlinear systems. Furthermore, the execution of GMDH is simpler than the notable ANN model, promoting it as a powerful tool for time series forecasting.

The main issue in every forecasting process is to obtain as much accurate forecast result as possible. Since the reliability of a model is measured on its accurate forecast, the choice of a correct model is of paramount importance. In terms of modelling nonlinear data, according to Granger (1993), a good nonlinear model is a model that would be able to approximate any systems and should be highly flexible such that it is inclined to pick up any subtle nuance in the data. However, such model is unrealistic as there is no model that can perform best for all types of data. As such, a combination of models with variations and modifications (if the need arises), should be explored and tested in order to achieve the best end result.

1.4 Research Questions

The research questions which can be derived from the previous statements are;

1. In the case where the polynomial transfer function in the GMDH model fails to fully describe the input-output relationship of a system, could a better function replace the conventional transfer function to improve the forecasting accuracy of GMDH model?
2. In the event where no functions could perform well in every situation provided, could combining several GMDH models assist GMDH in achieving a better accuracy than the individual models?
3. Owing to its heuristic nature, can optimization algorithm assign appropriate weights for each model to ensure an improvement in the accuracy of GMDH model?

1.5 Research Aim

The aim of this research is to propose a methodology which combines several GMDH with different transfer functions using an optimization algorithm to contribute to the improvement and enhancement of the existing GMDH model in time series forecasting.

1.6 Objectives

The objectives of this research are as follows;

1. To improve the performance of the conventional GMDH model in time series forecasting by substituting the quadratic polynomial transfer function with other transfer functions.
2. To propose a combination of several individual GMDH models with different transfer functions using a heuristic weight based combination method.

3. To compare the performance of the proposed model with benchmark models and to evaluate the models using real and benchmark data.

1.7 Scope of Research

The scope of this research are as follows;

1. For the purpose of evaluating the accuracy of the forecasting models, two types of data will be used in this research, namely real data and benchmarked data. The real data used is tourism data, that is, data of tourist arrivals to Malaysia, while the benchmark data is the well-known monthly airline passengers' arrival data.
2. The benchmark models are used in this study to evaluate the performance of the proposed model. The models are the individual GMDH models itself, the individual GMDH models combined using SA technique and an ANN model.
3. The performance of the models will be evaluated using two criteria; predictive ability and statistical significance of the models. The predictive abilities which focuses on the forecasting accuracies of the models will be estimated using three well-known performance measurement; Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE). Meanwhile, the statistical significance of the models will be evaluated using the notable paired sample t-test.

1.8 Significance of Research

The main focus of this research is in contributing to the advancement of the existing GMDH models, especially in the area of time series forecasting. Even though nonlinear time series is common in real life systems, forecasting it successfully is not such an easy feat due to the complexity of its input-output relationships. Albeit being able to approximate complex nonlinear systems, GMDH model still has many rooms for improvements particularly in the enhancement of its transfer functions.

REFERENCES

Abbod, M., & Deshpande, K. (2008). Using intelligent optimization methods to improve the group method of data handling in time series prediction. In *International Conference on Computational Science* (pp. 16-25). Springer, Berlin, Heidelberg.

Ahmad, A. S., Hassan, M. Y., Abdullah, M. P., Rahman, H. A., Hussin, F., Abdullah, H., & Saidur, R. (2014). A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renewable and Sustainable Energy Reviews*, *33*(1), 102-109.

Aiolfi, M., Capistran, C., & Timmermann, A. (2010). Forecast combinations. *CREATES research paper*, (2010-21).

AlBinHassan, N. M., & Wang, Y. (2011). Porosity prediction using the group method of data handling. *Geophysics*, *76*(5), O15-O22.

Anastasakis, L. & Mort, N. (2001). The development of self-organization techniques in modelling: a review of the group method of data handling (GMDH). *ACSE Research Report 813*. Department of Automatic Control and Systems Engineering, University of Sheffield.

Armstrong, J. S. (2001). Combining forecasts. In *Principles of forecasting* (pp. 417-439). Springer, Boston, MA.

Attoh-Okine, N. O. (1999). Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance. *Advances in Engineering Software*, *30*(4), 291-302.

Ayoub, M. A. & Elraies, K. A. (2014). Development of a Universal Pressure Drop Model in Pipelines Using Group Method of Data Handling-Type Neural Networks Model. *Journal of Applied Sciences*, *14*(23), 3220-3227.

Basheer, H., & Khamis, A. B. (2016). A hybrid group method of data handling (GMDH) with the Wavelet Decomposition for Time Series Forecasting: A review. *ARPN Journal of Engineering and Applied Sciences*, *11*(18), 10792-10800.

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, *20*(4), 451-468.

Benediktsson, J. A., Swain, P. H., & Ersoy, O. K. (1989). Neural network approaches versus statistical methods in classification of multisource remote sensing data. In *12th Canadian Symposium on Remote Sensing Geoscience and Remote Sensing Symposium*, (Vol. 2, pp. 489-492). IEEE.

Box, G. E., & Jenkins, G. M. (1970). *Time series analysis: forecasting and control*. San Francisco, CA: Holden-Day.

Brown, R. G. (2004). *Smoothing, forecasting and prediction of discrete time series*. Courier Corporation.

Bueno, E. I., Pereira, I. M., & Silva, A. T. E. (2011). Group method of data handling and neural networks applied in temperature sensors monitoring. *International Journal of Nuclear Knowledge Management*, *5*(3), 260-271.

Buryan, P. (2007). Enhanced MIA-GMDH Algorithm. In *Proceedings of the International Workshop on Inductive Modelling IWIM 2007* (pp. 144-155). IRTC ITS NASU.

Claveria, O., & Torra, S. (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series models. *Economic Modelling*, *36*, 220-228.

Clemen, R. T., & Winkler, R. L. (1986). Combining economic forecasts. *Journal of Business & Economic Statistics*, *4*(1), 39-46.

Dag, O., & Yozgatligil, C. (2016). Gmdh: An R package for short term forecasting via gmdh-type neural network algorithms. *The R Journal*, *8*(1), 379-386.

De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International journal of forecasting*, *22*(3), 443-473.

Dorn, M., Braga, A. L., Llanos, C. H., & Coelho, L. S. (2012). A GMDH polynomial neural network-based method to predict approximate three-dimensional structures of polypeptides. *Expert Systems with Applications*, *39*(15), 12268-12279.

Duffy, J. J., & Franklin, M. A. (1975). A learning identification algorithm and its application to an environmental system. *IEEE Transactions on Systems, Man, and Cybernetics*, (2), 226-240.

Durbin, J. (1960). The fitting of time-series models. *Revue de l'Institut International de Statistique*, 233-244.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Farlow, S. J. (1981). The GMDH algorithm of Ivakhnenko. *The American Statistician*, *35*(4), 210-215.

Farlow, S. J. (1984). *Self-organizing methods in modeling: GMDH type algorithms* (Vol. 54). CrC Press.

Ghazanfari, N., Gholami, S., Emad, A., & Shekarchi, M. (2017). Evaluation of GMDH and MLP Networks for Prediction of Compressive Strength and Workability of Concrete. *Bulletin de la Société Royale des Sciences de Liège, 86*(special edition), 855-868.

Ghosh-Dastidar, S., & Adeli, H. (2003). Wavelet-clustering-neural network model for freeway incident detection. *Computer-Aided Civil and Infrastructure Engineering*, *18*(5), 325-338.

Godfrey, L. B., & Gashler, M. S. (2018). Neural decomposition of time-series data for effective generalization. *IEEE transactions on neural networks and learning systems*, *29*(7), 2973-2985.

Graettinger, T. J., & DuBose, P. A. (1998). Computer-based neural network system and method for medical diagnosis and interpretation. *U.S. Patent No. 5,839,438*. Washington, DC: U.S. Patent and Trademark Office.

Granger, C. W. J. (1993). Strategies for modelling nonlinear time-series relationships. *Economic Record*, *69*(3), 233-238.

Granger, C. W. J., & Anderson, A. P. (1978). An introduction to bilinear time series models. *Angewandre Statistik und Okonometrie*, *8*. Vandenhoeck & Rpurecht, Gottingen.

Granger, C. W. J., & Hatanaka, M. (2015). *Spectral Analysis of Economic Time Series.(PSME-1)*. Princeton university press.

Granger, C. W. J., & Terasvirta, T. (1993). Modelling non-linear economic relationships. *OUP Catalogue*. Oxford University Press.

Green, D. G., Reichelt, R. E., & Bradbury, R. H. (1988). Statistical behaviour of the GMDH algorithm. *Biometrics*, 49-69.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, *31*(4), 337-350.

Gupta, M. (2010). *Intelligent System Using GMDH Algorithm*. Florida Atlantic University.

Haykin. S. (1999). *Neural Networks: A Comprehensive Foundation* (2nd ed.). Prentice Hall, New Jersey.

He, C., & Xu, X. (2005). Combination of forecasts using self-organizing algorithms. *Journal of forecasting*, *24*(4), 269-278.

Hepner, G., Logan, T., Ritter, N., & Bryant, N. (1990). Artificial neural network classification using a minimal training set- Comparison to conventional supervised classification. *Photogrammetric Engineering and Remote Sensing*, *56*(4), 469-473.

Hong, W. C. (2011). Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. *Energy*, *36*(9), 5568-5578.

Howland, J. C., & Voss, M. S. (2003). Natural gas prediction using the group method of data handling. In *Proceedings of the Seventh IASTED International Conference on Artificial Intelligence and Soft Computing*. ACTA Press.

Hu, J., Wang, J., & Zeng, G. (2013). A hybrid forecasting approach applied to wind speed time series. *Renewable Energy*, *60*, 185-194.

Huang, G., & Wang, L. (2011). Hybrid neural network models for hydrologic time series forecasting based on genetic algorithm. In *2011 Fourth International Joint Conference on Computational Sciences and Optimization* (pp. 1347-1350). IEEE.

Hung, N. Q., Babel, M. S., Weesakul, S., & Tripathi, N. K. (2009). An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrology and Earth System Sciences*, *13*(8), 1413-1425.

Ikeda, S., Ochiai, M., & Sawaragi, Y. (1976). Sequential GMDH algorithm and its application to river flow prediction. *IEEE Transactions on Systems, Man, and Cybernetics*, (7), 473-479.

Ismail, S. (2016). Empirical Mode Decomposition with Least Square Support Vector Machine Model for River Flow Forecasting (Doctoral dissertation, Universiti Teknologi Malaysia).

Ismail, S., Shabri, A., & Samsudin, R. (2011). A hybrid model of self-organizing maps (SOM) and least square support vector machine (LSSVM) for time-series forecasting. *Expert Systems with Applications*, *38*(8), 10574-10578.

Ivakhnenko, A. G. (1971). Polynomial theory of complex systems. *IEEE transactions on Systems, Man, and Cybernetics*, (4), 364-378.

Ivakhnenko, A. G., & Ivakhnenko, G. A. (1995). The review of problems solvable by algorithms of the group method of data handling (GMDH). *Pattern Recognition And Image Analysis C/C Of Raspoznavaniye Obrazov I Analiz Izobrazhenii*, *5*, 527-535.

Ivakhnenko, A. G., & Ivakhnenko, G. A. (2000). Problems of further development of the group method of data handling algorithms. Part I. *Pattern Recognition and Image Analysis c/c of raspoznavaniye obrazov i analiz izobrazhenii*, *10*(2), 187-194.

Ivakhnenko, A. G., Ivakhnenko, G. A., Savchenko, E. A., & Wunsch, D. (2002). Problems of further development of GMDH algorithms: Part 2. *Pattern Recognition and Image Analysis*, *12*(1), 6-18.

Jain, A., & Kumar, A. M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, *7*(2), 585-592.

Jirina, M. (1994). The Modified GMDH: Sigmoidal and Polynomial Neural Net. *IFAC Proceedings Volumes*, *27*(8), 611-613.

Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization. *Technical Report-TR06*. Computer Engineering Department, Erciyes University.

Karaboga, D., & Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of global optimization*, *39*(3), 459-471.

Karaboga, D., & Basturk, B. (2008). On the performance of artificial bee colony (ABC) algorithm. *Applied soft computing*, *8*(1), 687-697.

Khosravi, A., Koury, R. N. N., Machado, L., & Pabon, J. J. G. (2018). Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system. *Sustainable Energy Technologies and Assessments*, *25*, 146-160.

Kondo, T. (1998). GMDH neural network algorithm using the heuristic self-organization method and its application to the pattern identification problem. In *Proceedings of the 37th SICE Annual Conference. International Session Papers* (pp. 1143-1148). IEEE.

Kondo, T. (2002). Identification of radial basis function networks by using revised GMDH-type neural networks with a feedback loop. In *Proceedings of the 41st SICE Annual Conference. SICE 2002.* (Vol. 5, pp. 2672-2677). IEEE.

Kondo, T., & Pandya, A. S. (2003). Structural identification of the multi-layered neural networks by using revised GMDH-type neural network algorithm with a feedback loop. In *SICE 2003 Annual Conference (IEEE Cat. No. 03TH8734)* (Vol. 3, pp. 2768-2773). IEEE.

Kondo, T., & Ueno, J. (2006). Medical image recognition of the brain by revised GMDH-type neural network algorithm with a feedback loop. *International Journal of Innovative Computing, Information and Control*, *2*(5), 1039-1052.

Kondo, T., & Ueno, J. (2009). Medical image recognition of abdominal multi-organs by RBF GMDH-type neural network. *International Journal of Innovative Computing, Information and Control*, *5*(1), 225-240.

Kondo, T., Kondo, S., Ueno, J., & Takao, S. (2017). Medical image diagnosis of kidney regions by deep feedback GMDH-type neural network using principal component-regression analysis. *Artificial Life and Robotics*, *22*(1), 1-9.

Kondo, T., Pandya, A. S., & Zurada, J. M. (1999). GMDH-type neural networks and their application to the medical image recognition of the lungs. In *SICE'99. Proceedings of the 38th SICE Annual Conference. International Session Papers (IEEE Cat. No. 99TH8456)* (pp. 1181-1186). IEEE.

Kondo, T., Ueno, J., & Takao, S. (2012). Hybrid multi-layered GMDH-type neural network using principal component-regression analysis and its application to medical image diagnosis of lung cancer. In *2012 ASE/IEEE International Conference on BioMedical Computing (BioMedCom)* (pp. 20-27). IEEE.

Kostenko, A. V., & Hyndman, R. J. (2008). Forecasting without Significance Test. *manuscript, Monash University, Australia*.

Li, G., Shi, J., & Zhou, J. (2011). Bayesian adaptive combination of short-term wind speed forecasts from neural network models. *Renewable Energy*, *36*(1), 352-359.

Li, Q. M., Tian, Y. X., & Zhang, G. X. (2013). GMDH Modeling Based on Polynomial Spline Estimation and Its Applications. *International Journal of Mathematical and Computational Sciences*, *7*(3), 458 - 462.

Li, Q. M., Tian, Y. X., & Zhang, G. X. (2014). An Elman Model Based on GMDH Algorithm for Exchange Rate Forecasting. *Management Science and Engineering*, *8*(4), 92.

Lin, Z. S., Liu, J., & He, X. D. (1994). The self-organizing methods of long-term forecasting (I)—GMDH and GMPSC model. *Meteorology and Atmospheric Physics*, *53*(3-4), 155-160.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., ... & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of forecasting*, *1*(2), 111-153.

Mohanty, R., Ravi, V., & Patra, M. R. (2013). Hybrid intelligent systems for predicting software reliability. *Applied Soft Computing*, *13*(1), 189-200.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons.

Moosavi, V., Talebi, A., & Hadian, M. R. (2017). Development of a hybrid wavelet packet-group method of data handling (WPGMDH) model for runoff forecasting. *Water resources management*, *31*(1), 43-59.

Mustaffa, Z., Yusof, Y., & Kamaruddin, S. S. (2014). Application of LSSVM by ABC in energy commodity price forecasting. In *2014 IEEE 8th International Power Engineering and Optimization Conference (PEOCO2014)* (pp. 94-98). IEEE.

Najafzadeh, M., & Azamathulla, H. M. (2013). Group method of data handling to predict scour depth around bridge piers. *Neural Computing and Applications*, *23*(7-8), 2107-2112.

Najafzadeh, M., & Barani, G. A. (2011). Comparison of group method of data handling based genetic programming and back propagation systems to predict scour depth around bridge piers. *Scientia Iranica, 18*(6), 1207-1213.

Najafzadeh, M., Barani, G. A., & Hessami-Kermani, M. R. (2015). Evaluation of GMDH networks for prediction of local scour depth at bridge abutments in coarse sediments with thinly armored beds. *Ocean Engineering*, *104,* 387-396.

Oh, S. K., & Pedrycz, W. (2002). The design of self-organizing polynomial neural networks. *Information Sciences, 141*(3-4), 237-258.

Onwubolu, G. C., & Sharma, A. (2004). Particle Swarm Optimization for the assignment of facilities to locations. In *New Optimization Techniques in Engineering* (pp. 567-584). Springer, Berlin, Heidelberg.

Onwubolu, G. C., Buryan, P., & Lemke, F. (2008). Modeling tool wear in end-milling using enhanced GMDH learning networks. *The International Journal of Advanced Manufacturing Technology*, *39*(11-12), 1080-1092.

Palmer, A., Montano, J. J., & Sesé, A. (2006). Designing an artificial neural network for forecasting tourism time series. *Tourism management*, *27*(5), 781-790.

Park, H. S., Oh, S. K., Ahn, T. C., & Pedrycz, W. (1999). A study on multi-layer fuzzy polynomial inference system based on an extended GMDH algorithm. In *FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems. Conference Proceedings (Cat. No. 99CH36315)* (Vol. 1, pp. 354-359). IEEE.

Pereira, I. M., & Bueno, E. I. (2011). Variable identification in group method of data handling methodology. In *Proceedings of the 2011 International Nuclear Atlantic Conference - INAC 2011, ENAN 10, ENFIR 17, ENIN 2*. ABEN, Sao Paulo.

Rasouli, S., Tabesh, H., & Etminani, K. (2016). A study of input variable selection to artificial neural network for predicting hospital inpatient flows. *Br. J. Appl. Sci. Technol*, *18*(4), 1-8.

Rayegani, F., & Onwubolu, G. C. (2014). Fused deposition modelling (FDM) process parameter prediction and optimization using group method for data handling (GMDH) and differential evolution (DE). *The International Journal of Advanced Manufacturing Technology*, *73*(1-4), 509-519.

Sakaguchi, A., & Yamamoto, T. (2000). A GMDH network using backpropagation and its application to a controller design. In *SMC 2000 Conference Proceedings. 2000 IEEE International Conference on Systems, Man & Cybernetics (IEEE Cat. No. 00CH37166)*. (Vol. 4, pp. 2691-2696). IEEE.

Samsudin, R., Saad, P., & Shabri, A. (2009). Combination of forecasting using modified gmdh and genetic algorithm. *International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM), 1*, 170-176.

Shabri, A. (2015). A novel hybrid ensemble learning paradigm for tourism forecasting. In *AIP Conference Proceedings* (Vol. 1643, No. 1, pp. 192-200). AIP.

Shabri, A., & Samsudin, R. (2014). A hybrid GMDH and box-jenkins models in time series forecasting. *Applied Mathematical Sciences*, *8*(62), 3051-3062.

Sharda, R. (1994). Neural networks for the MS/OR analyst: An application bibliography. *Interfaces*, *24*(2), 116-130.

Shi, X., Li, Y., Li, H., Guan, R., Wang, L., & Liang, Y. (2010). An integrated algorithm based on artificial bee colony and particle swarm optimization. In *2010 Sixth international conference on natural computation* (Vol. 5, pp. 2586-2590). IEEE.

Shokouhifar, M., & Abkenar, G. S. (2011). An artificial bee colony optimization for MRI fuzzy segmentation of brain tissue. In *2011 International Conference on Management and Artificial Intelligence IPEDR* (Vol. 6). IACSIT Press.

Srinivasan, D. (2008). Energy demand prediction using GMDH networks. *Neurocomputing*, *72*(1-3), 625-629.

Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of forecasting*, *23*(6), 405-430.

Sun, W. Z., Wang, J. S., & Gao, S. Z. (2017). Soft-Sensor Modeling of PVC Polymerizing Process Based on F-GMDH-Type Neural Network Algorithm. *Journal of Sensors*, *2017*.

Tamura, H., & Kondo, T. (1978). Revised GMDH algorithm using prediction sum of squares (PSS) as a criterion for model selection. *Transactions of the Society of Instrument and Control Engineers*, *14*(5), 519-524.

Tamura, H., & Kondo, T. (1980). Nonlinear Steady State Modeling for River Quality by a Revised GMDH. *Transactions of the Society of Instrument and Control Engineers*, *16*(2), 189-194.

Tascikaraoglu, A., & Uzunoglu, M. (2014). A review of combined approaches for prediction of short-term wind speed and power. *Renewable and Sustainable Energy Reviews*, *34*, 243-254.

Tauser, J., & Buryan, P. (2011). Exchange Rate Predictions in International Financial Management by Enhanced GMDH Algorithm. *Prague Economic Papers*, *20*(3), 232-249.

Teng, G., Xiao, J., He, Y., Zheng, T., & He, C. (2017). Use of group method of data handling for transport energy demand modeling. *Energy Science & Engineering*, *5*(5), 302-317.

Terasvirta, T., & Kock, A. B. (2010). Forecasting with nonlinear time series models. *CREATES Research Paper*, (2010-1).

Tong, H., & Lim, K. S. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society. Series B (Methodological), 42*(3). pp. 245-292.

Ugrasen, G., Ravindra, H. V., Prakash, G. N., & Keshavamurthy, R. (2014). Estimation of machining performances using MRA, GMDH and artificial neural network in wire EDM of EN-31. *Procedia materials science*, *6*, 1788-1797.

Varahrami, V. (2012). Good Prediction of Gas Price between MLFF and GMDH Neural Network. *International Journal of Finance and Accounting*, *1*(3), 23-27.

Wang, W. (2006). *Stochasticity, nonlinearity and forecasting of streamflow processes*. IOS Press.

Wang, L., Wang, Z., Qu, H., & Liu, S. (2018). Optimal forecast combination based on neural networks for time series forecasting. *Applied Soft Computing*, *66*, 1-17.

Willmott, C. J., & Matsuura, K. (2006). On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. *International Journal of Geographical Information Science*, *20*(1), 89-102.

Winkler, R. L., & Makridakis, S. (1983). The combination of forecasts. *Journal of the Royal Statistical Society: Series A (General)*, *146*(2), 150-157.

Xiao, J., Li, Y., Xie, L., Liu, D., & Huang, J. (2018). A hybrid model based on selective ensemble for energy consumption forecasting in China. *Energy*, *159*, 534-546.

Xiao, L., Wang, J., Dong, Y., & Wu, J. (2015). Combined forecasting models for wind energy forecasting: A case study in China. *Renewable and Sustainable Energy Reviews*, *44*, 271-288.

Xu, C., & Duan, H. (2010). Artificial bee colony (ABC) optimized edge potential function (EPF) approach to target recognition for low-altitude aircraft. *Pattern Recognition Letters*, *31*(13), 1759-1772.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159-175.

Zhang, G. P., & Berardi, V. L. (2001). Time series forecasting with neural network ensembles: an application for exchange rate prediction. *Journal of the Operational Research Society*, *52*(6), 652-664.

Zhang, G. P., Patuwo, B. E., & Hu, M. Y. (2001). A simulation study of artificial neural networks for nonlinear time-series forecasting. *Computers & Operations Research*, *28*(4), 381-396.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, *14*(1), 35-62.

Zhang, M., He, C., & Liatsis, P. (2012). A D-GMDH model for time series forecasting. *Expert Systems with Applications*, *39*(5), 5711-5716.

Zhang, Q. (1997). Using wavelet network in nonparametric estimation. *IEEE Transactions on Neural Networks*, *8*(2), 227-236.

Zhang, X., Liang, Y., Zhou, J., & Zang, Y. (2015). A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement*, *69*, 164-179.