VOL. 11, NO. 18, SEPTEMBER 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

A HYBRID GROUP METHOD OF DATA HANDLING (GMDH) WITH THE WAVELET DECOMPOSITION FOR TIME SERIES FORECASTING: A REVIEW

Huma Basheer and Azme Bin Khamis
Department of Mathematics and Statistics, Faculty of Science, Technology and Human Development, Universiti Tun Hussein Onn Malaysia Parit Raja, Batu Pahat, Johor, Malaysia
E-Mail: hw140007@siswa.uthm.edu.my

ABSTRACT

Hybridization of existing competitive modeling methodologies is now an active area of research. The GMDH algorithm is a heuristic and computer oriented method which provides the foundation for the construction of high order regression models of complex system. The research for improving the effectiveness of forecasting models has never been stopped. Currently it was reported that a hybrid system in prediction and classification achieved a higher performance level against the traditional system. The selection of the forecasting model is the important criteria that will influence to the forecasting accuracy. So the enhancement of conventional GMDH model through hybridization will improve the prediction accuracy of the traditional GMDH for time series forecasting. This paper presents a short overview of Group Method of Data Handling (GMDH), its modification and hybridization for time series forecasting. The overview will aim to provide further investigation on the hybrid Group Method of Data Handling (GMDH) with the Wavelet Decomposition for Time Series Forecasting. This modification and hybridization will be proposed a new hybrid GMDH that will be helped for the computational constrain to become more flexible as well as robust/efficient than the conventional GMDH.

Keywords: GMDH, hybridization, modification, wavelet.

INTRODUCTION

The accuracy of time series forecasting is fundamental to many decisions processes and hence the research for improving the effectiveness of forecasting models has never been stopped(Zhang *et al*., 1998). The ARIMA model is one of the most popular models in traditional time-series forecasting and is often used as a benchmark model to compare with other models. However, the ARIMA model is only a class of linear model and thus it can only capture linear feature of a data time series. Consequently, lots of researches had tried to apply the artificial intelligent techniques to improve the accuracy of the time series forecasting issues such as Artificial Neural Network(Wu *et al*., 2005), (Zou *et al*., 2007) and recently Group Method of Data Handling (Kalantary *et al*., 2009), (Wongseree *et al*., 2007), (Onwubolu, 2008). Currently, it was reported that a hybrid system in prediction and classification achieved a higher performance level against the traditional system (Min *et al*., 2006), (Lekakos and Giaglis, 2007), (H. Kim and Shin, 2007).

In this paper we discuss the Group Method of Data Handling (GMDH), its modification and hybridization for times series forecasting. This literature review will help to identify and provide further investigation on the Hybrid Group Method of Data Handling (GMDH) with the Wavelet Decomposition for Time Series Forecasting to improve the performance of the standard GMDH approach.

Group method of data handling (GMDH)

GMDH is a method of developing nonlinear systems with many input variables. The GMDH algorithm was first presented by a Ukrainian scientist Ivakhnenko and his colleagues in 1968 to produce mathematical models of complex systems by handling data samples of observations (Ivakhnenko, 1971). His intent was to develop a rival method to stochastic approximation. The GMDH theory or polynomial networks are called Statistical Learning Networks (Baron and Barron, 1988) in the United States of America. The GMDH method was originally formulated to solve for higher order regression polynomials specifically for solving modeling and classification problems (Park and Pedrycz, 2007). GMDH algorithm has also been extensively used for prediction and modeling complex nonlinear process (Kim and Park, 2005). The proposed algorithm is based on a multilayer structure using, for each pair of variables, a second order polynomial of the form:

$$y = a_0 + \sum_{i=1}^{M} a_i x_i + \sum_{i=1}^{M}\sum_{j=1}^{M} a_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M} a_{ijk} x_i x_j x_k + \ldots\ldots\ldots\ldots (1)$$

In this case, x represents the input to the system, M is the number of inputs, and a are the coefficients or weights of the terms (Ivakhnenko, 1971). He was looking for a computational instrument allowing him to model real world systems characterized by data with many inputs (dimensions) and few records. Such ill-posed problems could not be solved traditionally (ill-conditioned matrixes) and therefore different approach was needed. Prof. Ivakhnenko proposed the GMDH method which avoided solution of ill-conditioned matrixes by decomposing them into submatrices of lower dimensionality that could be solved easily. The more important idea behind the GMDH is the adaptive process of combination of these submatrices back to the final solution. The original GMDH method is called Multilayered Iterative Algorithm

(MIA GMDH). There are four major advantages to this algorithm: A small training set is required, the multiple layer structure of the designed system results in a feasible way of implementing high degree multinomial, the computation burden is reduced, and inputs/functions of inputs that have little impact on the output are automatically filtered out (Chang and Hwang, 1999).

The GMDH method are applied in many fields in order to model and predict the behaviors of unknown and/or very complex systems based on given input-output data. Theoretically, in develop GMDH method; it is required to understand the explicit mathematical input-output relationship precisely. Such explicit mathematical modeling is however very difficult and is not readily tractable in poorly understood time series data.

The GMDH algorithm has been successfully used to deal with uncertainty and nonlinearity of systems in a wide range of disciplines such as economy, ecology, medical diagnostics, signal processing and control systems (Voss and Feng, 2002). Some simplified approximations, such as the two-direction regressive GMDH (Li *et al*., 2006) and the revised GMDH algorithms (Chang and Hwang, 1999), have been introduced to model dynamic systems in flood forecast and petroleum resource prediction with some success. However, owing to its limited generic structure, GMDH tends to generate an overly complex polynomial when it is applied to estimate highly nonlinear systems. Networks developed using methods based on GMDH concepts tend to have fewer, but far more flexible, nodes than a typical artificial neural network (Pachepsky *et al*., 1998). Robinson (1998) introduced a Multi-Objective GMDH (MOGMDH) algorithm in which the regularity criterion was used as well as three selectors in the selection process. This resulted in a significant improvement in the performance of the GMDH algorithm

The revised GMDH-type neural networks with sigmoid functions were applied for the identification of the liver in CAT (Computer Aided Tomography) scan images and it is shown that this algorithm is very useful for medical image recognition (Kondo and Pandya, 2003). The GMDH-type neural networks with sigmoid functions have the abilities of self-selecting useful input variables and of self-organizing optimum neural network architecture. Such regularization takes into account just the complexity of GMDH network. Outputs from neurons in a layer can be highly correlated resulting to a redundant GMDH network.

In recent years, however the use of such self-organizing networks leads to successful application of the GMDH type algorithm in a broad range of areas in engineering, science and economics (Ivakhnenko, 1971), (Ahmadi *et al*., 2007). The inherent complexity in the design of feed forward neural networks in terms of understanding the most appropriate topology and coefficients has a great impact on their performance. In the case of weight or coefficient training procedures, the most commonly used learning algorithm is the gradient descent algorithm, e.g. back propagation.

There have been many efforts in recent years to deploy population based stochastic search algorithms such as evolutionary methods to design artificial neural networks since such evolutionary algorithms are particularly useful for dealing with complex problems having large search spaces with many local optima. Consequently, genetic algorithms have been used in a feed forward GMDH type neural network for each neuron searching its optimal set of connections with the preceding layer (Mottaghitalb, 1996). Such an advantage of evolutionary algorithms is very fruitful to solve many real world optimal design or decision making problems that are, indeed, multi-objective. Hiassat *et al.* (1975) (Hiassat and Mort, 1975) introduced the GP-GMDH algorithm, which uses genetic programming to find the best function that maps the input to the output in each layer of the GMDH algorithm, and showed that it performs better than the conventional GMDH algorithm. Kim *et al.* (2009) (Kim *et al*., 2009)proposed a new heuristic approximation method for identification nonlinear relationships which is a hybrid of combines neural networks and a Polynomial Neural Network. Huang *et al.* (2002) (Huang *et al*., 2002) applied the enhancement of group method of data handling for short - term load forecast of a power system. In addition, Hwang *et al.* (2006) (Hwang *et al*., 2006) proposed neural - fuzzy GMDH which have several advantages compared with conventional multi layered GMDH models.

Hybridization of existing competitive modeling methodologies is now an active area of research. For example, Sakaguchi *et al.* (2003)(Sakaguchi *et al*., 2003) proposed the use of GMDH and two kinds of genetic algorithm where the binary coded and adjusted the system parameters. Also Kondo *et al.* (2005) (Kondo *et al*., 2005) proposed the GMDH type neural network algorithm with Prediction Sum of Squares (PSS) to minimize the prediction error criterion by using the heuristic self - organization method. Ahmadi *et al.* (2007) (Ahmadi *et al*., 2007) indicating addition presented the GMDH type neural network algorithm with radial basis functions (RBF) to three dimensional medical image recognition of the liver. Different restrictions are applied to the system, which considerably reduce the complexity of the inference mechanism. Hence, efficient implementations can be developed.

Consequently, it could be inferred that lots of researcher have been hybridized artificial intelligent techniques with GMDH to improve the performance of the standard GMDH approach. Thus, the main focus of this research is to extend the hybridization spectrum to include wavelet decomposition into the GMDH for modeling and prediction of complex data. Wavelet transform provides time - frequency localization during the feature selection to select the optimum features or input node from GMDH model. Table-1 shows summary of Previous Studies on GMDH Model.

VOL. 11, NO. 18, SEPTEMBER 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Table-1. Summary of previous studies on GMDH model.

Study	Type polynomial	Method
Kim & Park (2005); Oh &Pedrycz (2006)	$y_1 = c_0 + c_1x_1 + c_2x_2$ $y_2 = c_0 + c_1x_1 + c_2x_2 + c_3x_1^2 + c_4x_2^2 + c_5x_1x_2$ $y_3 = c_0 + c_1x_1 + c_2x_2 + c_3x_1x_2$	GA (Genetic Algorithm)
Zadeh *et al.*, (2003,2005, 2007); Atashkari *et al.* (2007)	$y = c_0 + c_1x_1 + c_2x_2 + c_3x_1^2 + c_4x_2^2 + c_5x_1x_2$	GA, Singular value Decomposition, Least Square
Shinohara *et al.*, (1999)	$y = c_0 + c_1x_1 + c_2x_2 + c_3x_1^2 + c_4x_2^2 + c_5x_1x_2$	Least Square
Oh *et al.* (2006)	$y = c_0 + c_1x_1 + c_2x_2 + c_3x_1^2 + c_4x_2^2 + c_5x_1x_2 + c_7x_1x_3 + c_8x_2x_3$	GA
Kondo & Ueno (2006); Kondo (2004,2006); Kondo *et al.*, (1999)	$z_1 = a_1x_i + a_2x_j + a_3x_ix_j + a_4x_i^2 + a_5x_j^2 + a_6x_i^3$ $+ a_7x_j^3 + a_8x_ix_j^2 + a_9x_jx_i^2 - a_0\theta_1$ $z_2 = c_1x_{i1} + c_2x_{i2} + c_3x_{i3} + ... + c_rx_{ir} - c_0\theta_1$ $y = \frac{1}{1+e^{-z}}$, $y = z$, $y = e^{-z^2}$	Least Square, AIC(Akaike's, Information Criterion) PSS(Prediction sum of squares)
Wang *et al.* 2005	$y = c_0 + c_1x_1 + c_2x_2 + c_3x_1^2 + c_4x_2^2 + c_5x_1x_2$ &ANN	Least Square, BP
Vaseckina&Yarin (1999);Pappas &Ekonomou, (2006)	$y = c_1x_1+c_2x_2+c_3x_3+\dots\dots\dots\dots\dots\dots\dots\dots+c_rx_r$ r = the number of layer	Least Square,GA
Abbod& Deshpande (2008)	$y = c_0 + c_1x_1 + c_2x_2 + c_3x_1^2 + c_4x_2^2 + c_5x_1x_2$	GA & PSO (Particle swarm optimization)
R.E.Abdel-Aal *et al.*2009, Vida Varahrami *et al* 2011,Prerna Mishra *et al* 2014	$y = c_0 + c_1x_1 + c_2x_2 + c_3x_1^2 + c_4x_2^2 + c_5x_1x_2$	AIM(Abductory Inductive Mechanism) network, GMDH Neural Networks and Intermittent Demand Forecast

The modified group method of data handling

The architecture of GMDH is similar to feed forward neural networks whose neurons used polynomial nodes. The polynomial is also called partial descriptions (PD). GMDH have fewer nodes than ANN but the nodes are flexible. The basic building block of GMDH is a quadratic polynomial of two variables consists of many layers and each layer consists of a bank of quadratic polynomial functions that requires input from the previous layer after having passed a certain selection. Although the GMDH is a structured systematic design procedure but the GMDH has some disadvantages. GMDH procedure has a tendency to produce overlay complex network if there are sufficiently large number of input variables and data point. It tends to generate quite complex polynomials if there are less than three inputs variables owing to its limited generic structure (quadratic polynomial).

In this study, addressing the problems with the conventional GMDH based on quadratic polynomial, we propose that the GMDH are based on three other functions. The basic quadratic polynomial could be replaced by other simple function forms(Lin *et al.*, 1994). The three kinds of functions, namely radial basis, sigmoid and tangent hyperbolic functions are introduced. The main purpose of this study is to investigate the applicability and capability of the GMDH using the other functions compared with the conventional GMDH and ANN methods for modeling.

Conventional GMDH is developed using the polynomial in equation (1) as PD. In this study, three kinds of PD structures, namely radial basis, sigmoid and tangent hyperbolic are proposed to replace the polynomial in the construct of the modified GMDH. Table-2 summarizes all of the suggested PD function where M is the total number of input.

Table-2. The architecture of the suggested functions.

Function	$y = f(x)$	Transform
Polynomial	$y = z$	$z = y$
Sigmoid	$y = \frac{1}{1+e^{-z}}$	$z = \ln\left(\frac{y}{1-y}\right)$, $y \neq 1$
Radial Basis Function	$y = e^{-z^2}$	$z = \sqrt{-\ln(y)}$
Tangent Hyperbolic	$y = \tanh(z)$	$z = \tanh^{-1}(y)$

ARPN Journal of Engineering and Applied Sciences

where $z = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \ldots + a_Mx_M$
(linear function)

$$z = a_0 + a_1x_i + a_2x_j + a_3x_ix_j + a_4x_i^2 + a_5x_j^2$$

(polynomial function)

The hybrid wavelet decomposition and GMDH model for time - series forecasting

Combining the GMDH with wavelet decomposition to forecast the future data by using the function of wavelet as provides time frequency localization. As a result, through decomposition, the proper input is constructed for the GMDH algorithm. Through analysis, this method has some advantages compared with other method. This hybrid is carrying out to overcome the difficulty for GMDH model in choosing the proper input.

At present, many research organizations and universities have been studying on run - off time series prediction. Run - off series is complex. It contains a wide range frequency component. Therefore, it is required to decompose it into different frequency component. The wavelet method provides a convenient time - frequency analysis way. Wavelet decomposition is a powerful tool to break down the problem and feature extraction is widely used in various fields. However, there are different kinds of time series data in specific areas. We decompose these data by different kinds of wavelet and the sequence generated by decomposition will affect the performance in prediction model establishment in some degrees. As the wavelet filter has different supporting length and self - filtering properties, it requires an appropriate wavelet to decompose specific time series and thus improves the prediction performance.

There are certain limitations in improving prediction accuracy with their own information provided by prediction time series(Samsudin *et al*., 2011). Time series data is different areas bears different characteristics. It is not appropriate to decompose with only one wavelet (Yuelong *et al*., 2009). Yuelong *et al* (2009) employed Monte Carlo method to explore the influential factors in wavelet function selecting by threshold low and high frequency coefficients for the purpose of sequence reconstruction. Results showed that the variation characteristic of series is the key factor affecting the selection of wavelet. Their purpose was for time series reconstruction but not prediction. Liu *et al.* (2003) (Shum, 2003) compared the impact of multi - scale framework run - off prediction with wavelet functions by training multi - scale prediction model. It is time - consuming to determine optimal wavelet function with different data, especially for large amounts of data.

The wavelet transform has been proposed for time series analysis in many papers in recent years. Wavelet would appear to be very appropriate for analyzing non stationary signals (Swee *et al* 1999) and a link between wavelets and the different operator was made.

Several approaches have been proposed for time - series filtering and prediction by the wavelet transform, based on a neural network(Bashir & El-Hawary, 2000), (Chen *et al*., 2006), (Dash *et al*., 2007) or an autoregression (AR) model (Soltani *et al*., 2000). Antoniadis *et al*., (2003) (Antoniadis *et al*., 2003) proposed the forecast method for continuous time series by combining the wavelet method with Hilbert auto regression stochastic process then compared to the ARIMA model. Numari *et al*., (2005) used the genetic algorithm and the wavelet function approximation to forecast the air pollution time series. Chang-il *et al*., (2006) applied the wavelet transformation and the regression model to forecast the system marginal price (SMP) of the electrical power system. Ruy *et al*., (2006) combined the wavelet transforms with mixes expert model (MEM) to forecast the time series. Table-3 shows summary of Previous hybridized using wavelet decomposition.

The function of wavelet method in the researches to minimized the error criterion and decomposed the sequence data to correlate with the objectives.

Wavelet analysis

Wavelet analysis which is known as mathematics microscope with the advantage of the same time frequency localization is a landmark in Fourier analysis development history. Wavelet transform provides time - frequency localization at the same time and overcomes the limited resolution by short time Fourier transforms.

The continuous wavelet transform of a continuous function produces a continuum of scales as output. Input data, however, is usually discretely sampled and a furthermore two fold relationship between resolution scales is both practical and adequate. The latter two issues lead to the discrete wavelet transform.

Discrete wavelet decomposition can be carried out on time -series multi scale decomposition and extracting components of different frequency interval in order to realize the research on time series frequency division. The output of a discrete wavelet transform can take various forms. Traditionally, a triangle is often used to represent all that we have to consider in the sequence of resolution scales. Such a triangle comes about as a result or the retaining of one sample out of every two. The major advantage of decimation is that just enough information is kept allow exact reconstruction of the input data. Therefore decimation is ideal for an application such as compression. A major disadvantage of the decimated form of output is that we cannot simply relate information at a given time point at the different scales.

A redundant transform based on an *N* length input time series, and then has an N length resolution scale for each of the resolution levels that we consider. It is easy, under these circumstances, to relate information each resolution scale for the same time point. We do have shift invariance. Finally, the extra storage requirement is by no means excessive. The torus wavelet transform decomposes a signal $X = (X_1, \ldots X_N)$ as a superposition of the form

VOL. 11, NO. 18, SEPTEMBER 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

$$X_t = c_{J,t} + \sum_{j=1}^{J} w_{j,t} \quad (2)$$

Where c_J is a coarse or smooth version of the original signal X and w_j represents the details of X at scale 2^{-j}. Thus, the algorithm outputs $J + 1$ subbands of size N. The indexing is such that here, $j = 1$ corresponds to the finest scale.

We derive it from the input data by convolving the latter with h. Then:

$$C_{j+1,t} = 0.5\ (c_{j,t-2^j} + c_{j,t}) \quad (3)$$

and

$$w_{j+1,t} = c_{j,t} - c_{j+1,t} \quad (4)$$

At any time point, t, we never use information after t in calculating the wavelet coefficient. This algorithm is having the following advantages:

- It is simple to implement. The computational requirement is O (N) per scale and in practice the number of scales is set as a constant.
- Because we do not shift the signal, the wavelet coefficients at any scale j of the signal ($X_1,\ldots, X_t$) are strictly equal to the first t wavelet coefficients at scale j of the signal ($X_1,\ldots,X_N$) ($N > t$).

Results of this filtering stage are the approximation and detail coefficient series for each resolution levels. The number of resolution levels is experimentally chosen from the low frequency coefficient series at the level such that the original data distribution is preserved.

The application of the wavelet transform provides an effective method of analyzing and synthesizing this variable structure of a signal in time and provides a means of localizing events of interest at their exact temporal location (Newland *et al.*, 1993., Strang and Nguyen, 1996).

Table-3. Summary of Previous Studies on the hybridized using wavelet decomposition.

Study	Type hybridized
Sang and Wang (2008)	Monte Carlo and the wavelet function
Zheng *et al*., (1999);Bashir and El-Hawary (2000); Chen *et al*., (2006); Dash and Nayak (2007)	Neural network and the wavelet function
Soltani *et al*., (2000)	Autoregression (AR) model and the wavelet function
Anestis *et al*. (2003)	The wavelet method with Hilbert auto regression stochastic process
Numari *et al*., (2005)	Genetic algorithm and the wavelet function
Chang-il *et al*., (2006)	The wavelet transformation and the regression model
Ruy *et al*., (2006)	The wavelet transforms with mixes expert model (MEM)
Bao *et al* 2007	Hybridizing Wavelet and Least Squares Support Vector Machines for Crude Oil Price Forecasting
C.Stolojescu et. Al. 2010	A Wavelet Based Prediction Method for TimeSeries
Khashei *et al*., 2011	A novel hybridization of artificial neural networks and ARIMA models for time series forecasting
Pandhiani *et al*., 2013	Time Series Forecasting Using Wavelet-Least Squares Support Vector Machines and Wavelet Regression Models for Monthly Stream Flow Data
Okkan *et al*., 2013	The combined use of wavelet transform and black box models in reservoir inflow modeling

The method allows specific features of a signal into a family basis functions called wavelets. Although its application in signal processing is relatively new, wavelet analysis has already been successfully applied in many fields including studies in geophysics (Foufoulla *et al*., 1997). These applications maximize the qualitative use of

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

the wavelet analysis technique in transforming one dimensional time series into two dimensional time frequency image.

EXPECTED FINDINGS

The expected finding from the literature review is shown by the research flow chart of Figure-1.The purpose of the flow chart to summarize the idea or observation through a well-focused research on the literature search and provide guidance to focus developing research ideas. The data mentioned in the flow chart of rice yield will be taken from Muda Agricultural Development Authority (MUDA), Malaysia ranging from 1978 to 2014. As the agricultural sector has contributed significantly to the growth and development of the Malaysia economy(Kingdom *et al.*, 2015). So the rice yield production is among the top priority under the agricultural sector due to global food crisis. Therefore we choose rice yield data from Muda Agricultural Development Authority (MUDA), Malaysia

Research flow chart

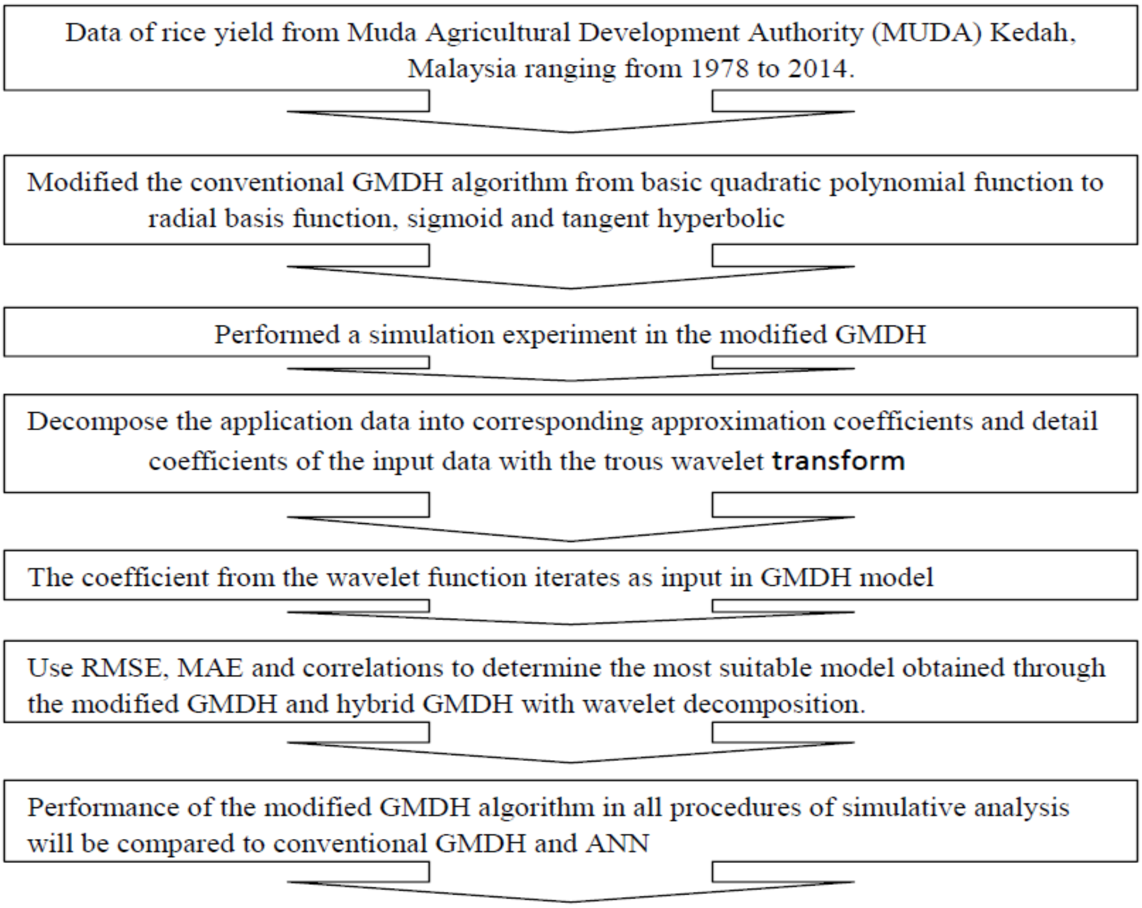

Figure-1. The research flow chart.

Performance measurement

The performance of the each model for both the training data and forecasting data will be evaluated and selected according to the mean absolute error (MAE), root mean square error (RMSE) and correlation coefficient (R), which are widely used for evaluating results of time series forecasting. The MAE, RMSE and R are defined as

$$\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N}\left|y_t - \hat{y}_t\right| \tag{5}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(y_t - \hat{y}_t\right)^2} \tag{6}$$

$$R = \frac{\frac{1}{N}\sum_{t=1}^{N}(y_t-\hat{y}_t)(o_t-\hat{o}_t)}{\frac{1}{N}\sum_{t=1}^{N}(y_t-\hat{y}_t)^2\sqrt{\frac{1}{N}\sum_{t=1}^{N}(o_t-\hat{o}_t)}} \tag{7}$$

Where o_t and y_t are the observed and forecasted values at data point t, respectively, $\hat{y}_t$ is the mean of the forecasted values at the time *t*, ô represents the mean of the observed values and *N* is the number of data points. The criterions to judge for the best model are relatively small of MAE and RMSE in the modeling and forecasting.

Correlation coefficient evaluates the flows predicted correlate with the flows observed. Clearly, the R value close to unity indicates a satisfactory result, while a low value or close to zero implies an inadequate result.

Before the training process begins, data normalization is performed. The linear transformation formula to [0, 1] is used

$$x_t = \frac{h_t}{h_{\max}} \tag{8}$$

where x_t and h_t represent the normalized and original data; and $h_{\max}$ represent the maximum values among the original data.

Assumptions and limitations

This study limits to the modified GMDH algorithm as modified the basic polynomial function transform to radial basis function, sigmoid and tangent hyperbolic. We also introduced a new hybrid GMDH and wavelet decomposition using self - organizing model, shown in Figure-2. The propose model is based on the idea of combining the GMDH algorithm with the wavelet analysis into one methodology.

CONCLUSIONS

This paper has reviewed Group Method of Data Handling (GMDH), its modification and hybridization for times series forecasting. In conventional GMDH models, many researchers considered the partial quadratic polynomials as the transfer function. However, owing to its limited generic structure, GMDH tends to generate an overly complex polynomial when it is applied to estimate highly nonlinear systems.

The modified GMDH algorithm is proposed by using the basic theory of the GMDH algorithm. Better partial descriptions (PDs) to replace the quadratic polynomial to fit the regression technique and generate a function that described the input – output more accurately, leading to an improvement in the prediction accuracy of GMDH algorithm. In these algorithms, it is similar to the conventional GMDH algorithm except that the transfer function polynomials are fitted using simplified function such as radial basis, sigmoid and tangent hyperbolic. Moreover, the modified GMDH can generate different nonlinear combinations of input and select useful combinations so that the high - order effects of input suiting the complexity of the nonlinear system.

In the future work, the researcher should tackle to produce a new architecture of the GMDH model using hybrid Group Method of Data Handling with wavelet decomposition, because no further research investigating on GMDH algorithm combining with the wavelet decomposition thus far. As a result, through wavelet decomposition, the proper input will be constructed for the GMDH algorithm. This hybrid wavelet decomposition will be carried out to overcome the difficulty for GMDH model in choosing the proper input and to improve the performance of the standard GMDH approach.

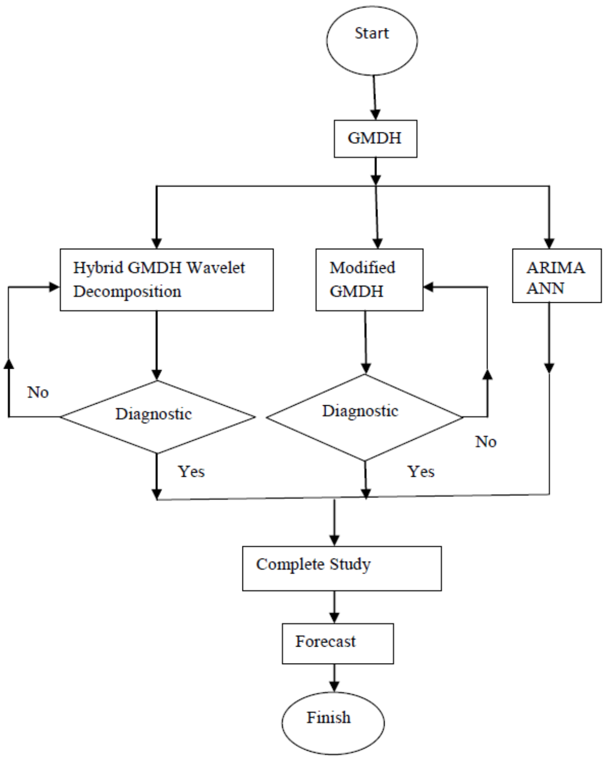

Figure-2. Hybrid GMDH wavelet decomposition model for times series forecasting.

ACKNOWLEDGEMENT

This work was supported by Research and Innovation Centre (R&D), Research, Innovation, Commercialization and Consultancy Office (ORICC) Universiti Tun Hussein Onn Malaysia. This support is gratefully acknowledged. The author would also like to thanks ORICC for the financial support through Post Graduate Incentive Research Grant (GIPS) in making the project a success.

REFERENCES

Ahmadi, H., Mottaghitalaband, M. and Nariman-Zadeh, N. 2007. Group method of data handling-type neural network prediction of broiler performance based on dietary metabolizable energy, methionine, and lysine. Journal of Applied Poultry Research, 16(4), 494-501. http://doi.org/10.3382/japr.2006-00074.

Antoniadis, A. and Sapatinas, T. 2003. Wavelet methods for continuous-time prediction using Hilbert-valued autoregressive processes. Journal of Multivariate Analysis, 87(1), 133-158. http://doi.org/10.1016/S0047-259X(03)00028-9.

Bashir, Z. and El-Hawary, M. E. 2000. Short term load forecasting by using wavelet neural networks. 2000 Canadian Conference on Electrical and Computer Engineering. Conference Proceedings. Navigating to a New Era (Cat. No.00TH8492), 1, 163–166. http://doi.org/10.1109/CCECE.2000.849691.

Chang, F. and Hwang, Y. 1999. A self-organization algorithm for real-time ¯ ood forecast Abstract:, 138(April 1998), 123-138.

Chen, Y., Yang, B. and Dong, J. 2006. Time-series prediction using a local linear wavelet neural network. Neurocomputing, 69(4-6), 449-465. http://doi.org/10.1016/j.neucom.2005.02.006.

Dash, P. K., Nayak, M., Senapati, M. R., & Lee, I. W. C. 2007. Mining for similarities in time series data using wavelet-based feature vectors and neural networks. Engineering Applications of Artificial Intelligence, 20(2), 185–201. http://doi.org/10.1016/j.engappai.2006.06.018.

Hiassat, M. and Mort, N. 1975. An evolutionary method for term selection in the Group Method of Data Handling 2 the Group Method of Data Handling 3 Genetic Programming (GP), 1-4.

Ivakhnenko, A. G. 1971. Polynomial Theory of Complex Systems Polynomial Theory of Complex Systems, (4), 364-378.

Kalantary, F., Ardalan, H. and Nariman-Zadeh, N. 2009. An investigation on the Su-NSPT correlation using GMDH type neural networks and genetic algorithms. Engineering Geology, 104(1-2), 144-155. http://doi.org/10.1016/j.enggeo.2008.09.006.

Kim, D. and Park, G. 2005. GMDH-Type Neural Network Modeling in Evolutionary Optimization, 563-570.

Kim, D., Seo, S.-J., & Park, G.-T. 2009. Hybrid GMDH-type modeling for nonlinear systems: Synergism to intelligent identification. Advances in Engineering Software, 40(10), 1087–1094. http://doi.org/10.1016/j.advengsoft.2009.01.029.

Kim, H. and Shin, K. 2007. A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. Applied Soft Computing, 7(2), 569-576. http://doi.org/10.1016/j.asoc.2006.03.004.

Kingdom, U., Abas, N. M., Junaini, N., Yaacob, A., Nurul, S., Mohd, A., Azidin, R. A. 2015. An overview of primary sector in Malaysia, III(2), 1–13.

Kondo, T. and Pandya, a. S. 2003. Structural identification of the multi-layered neural networks by using revised GMDH-type neural network algorithm with a feedback loop. SICE 2003 Annual Conference (IEEE Cat. No.03TH8734), 3, 1051-1059.

Kondo, T., Ueno, J. and Kondo, K. 2005. Revised GMDH-Type Neural Networks Using AIC or PSS Criterion and Their Application to Medical Image Recognition.

Lekakos, G. and Giaglis, G. M. 2007. A hybrid approach for improving predictive accuracy of collaborative filtering algorithms. User Modeling and User-Adapted Interaction, 17(1-2), 5–40. http://doi.org/10.1007/s11257-006-9019-0.

Li, L., Mphil, B. E. and Lockington, D. 2006. Department of Civil Engineering Sydney NSW 2006 Environmental Fluids / Wind Group Self-Organizing Polynomial Neural Network for Modelling Complex Hydrological Processes Research Report No R861 D-S Jeng, BE ME PhD Self-Organizing Polynomial Neural Network. Civil Engineering, (December 2005).

Lin, Z., Liu, J. I. and He, X. 1994. Meteorology and Atmospheric Physics The Self-Organizing Methods of Long-Term Forecasting (I) - GMDH and GMPSC Model *, 160, 155-160.

Min, S.-H., Lee, J. and Han, I. 2006. Hybrid genetic algorithms and support vector machines for bankruptcy prediction. Expert Systems with Applications, 31(3), 652-660. http://doi.org/10.1016/j.eswa.2005.09.070.

Mottaghitalb, M. (1996). Genetic Algorithms and Group Method of Data Handling-Type Neural Networks Applications in Poultry Science.

Onwubolu, G. C. 2008. Design of hybrid differential evolution and group method of data handling networks for modeling and prediction. Information Sciences, 178(18), 3616-3634. http://doi.org/10.1016/j.ins.2008.05.013.

Pachepsky, Y., Rawls, W., Gimenéz, D. and Watt, J. P. 1998. Use of soil penetration resistance and group method of data handling to improve soil water retention estimates. Soil and Tillage Research, 49(1-2), 117–126. http://doi.org/10.1016/S0167-1987(98)00168-8.

Park, H. and Pedrycz, W. 2007. Evolutionary design of hybrid self-organizing fuzzy polynomial neural networks with the aid of information granulation, 33, 830-846. http://doi.org/10.1016/j.eswa.2006.07.006.

Sakaguchi, a. and Yamamoto, T. 2003. A study on system identification using GA and GMDH network. IECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society (IEEE Cat. No.03CH37468), 3. http://doi.org/10.1109/IECON.2003.1280618.

Samsudin, R., Saad, P. and Shabri, a. 2011. River flow time series using least squares support vector machines. Hydrology and Earth System Sciences, 15(6), 1835-1852. http://doi.org/10.5194/hess-15-1835-2011.

Shum, C. L. H. 2003. Kullback-Leibler Boosting.

Soltani, S., Boichu, D., Simard, P. and Canu, S. 2000. Long-term memory prediction by multiscale decomposition. Signal Processing, 80(10), 2195-2205. http://doi.org/10.1016/S0165-1684 (00)00077-3.

Swee, E. G. T. and Elangovan, S. 1999. Applications of symlets for denoising and load forecasting. Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics. SPW-HOS '99. http://doi.org/10.1109/HOST.1999.778717.

Voss, M. S. and Feng, X. 2002. {ARMA} Model Selection Using Particle Swarm Optimization and {AIC} Criteria. IFAC 15th Triennial World Congress, 2(2).

Wongseree, W., Chaiyaratana, N., Vichittumaros, K., Winichagoon, P. and Fucharoen, S. 2007. Thalassaemia classification by neural networks and genetic programming. Information Sciences, 177(3), 771-786. http://doi.org/10.1016/j.ins.2006.07.009.

Wu, J. S., Han, J., Annambhotla, S. and Bryant, S. 2005. Artificial Neural Networks for Forecasting Watershed Runoff and Stream Flows. Journal of Hydrologic Engineering, 10(3), 216-222. http://doi.org/10.1061/(ASCE)1084-0699(2005)10:3(216).

Yuelong, Z., Jihong, Q., Qingsong, F., Dingsheng, W. and Shijin, L. 2009. Study on hydrology time series prediction based on wavelet-neural networks. Proceedings of the 2009 8th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2009, 411-415. http://doi.org/10.1109/ICIS.2009.116.

Zhang, G., Patuwo, B. E. and Hu, M. Y. 1998. Forecasting with artificial neural networks: The state of the art, 14, 35-62.

Zou, H. F., Xia, G. P., Yang, F. T. and Wang, H. Y. 2007. An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. Neurocomputing, 70(16-18), 2913-2923. http://doi.org/10.1016/j.neucom.2007.01.009.