A Time Series Forecasting Model Using Group Method Of Data Handling (GMDH)

Ruhaidah Samsudin
Department of Software Engineering
Faculty of Computer Science and
Information System
University Technology of Malaysia
607-5532228
ruhaidah@utm.my

Puteh Saad
Department of Software Engineering
Faculty of Computer Science and
Information System
University Technology of Malaysia
607-5532344
puteh@utm.my

ABSTRACT

The accuracy of time series forecasting is fundamental to many decisions processes. Forecasting can assist the organization to make a better planning and decision making. The selection of forecasting model is the important criteria that will influence to the forecasting accuracy. Hence, in order to cater these challenges, this study proposes a new time series forecasting model by using the Group Method of Data Handling(GMDH) technique. This model is an enhancement of conventional GMDH model that will improve the prediction accuracy of the traditional GMDH model. The benchmarked and non-benchmarked data are used in this study. The benchmarked data are five well-known data sets that always handled in real life time series application. Meanwhile, the non-benchmarked are rice yield data obtained from Muda Agricultural Development Authority (MADA) Kedah, Malaysia. The preliminary study was done by comparing the enhanced GMDH (Kondo model) with the conventional GMDH, neural network and ARIMA model. From the experimental results comparing the performance of four models, we can conclude that enhancement of GMDH model (Kondo model) in many cases is more competent in modeling and forecasting time series than the other models.

Categories and Subject Descriptors : I.6.5 [Model Development]:Modeling methodologies; I.2.8 [Problem Solving, Control Methods, and Search]: Heuristic methods

General Terms
Experimentation.

Keywords
Group Method of Data Handling (GMDH), Neural Network, ARIMA, rice yield, forecasting, time series.

1. INTRODUCTION

The agricultural sector has contributed significantly to the growth and development of the Malaysian economy even though the Malaysian economy has undergone significant structural changes over the last four decades. For the first three decades since independence, agriculture was the main contributor to the national economy. Rice yield production is among the top priority under the agricultural sector because of the global food crisis that affected to our country in 2007. According to Abdullah Ahmad Badawi at his opening remarks at the opening of the D8 Ministers' Meeting on Food Security, national rice production will be raised to meet up to 86% of the country's needs by 2010[1] .To achieve this, the Government has allocated an additional RM5.6bil to boost agricultural production. The Third Agriculture Policy (1998-2010) was established to meet at least 70% of Malaysia's demand a 5% increase over the targeted 65%. The remaining 30% comes from imported rice mainly from Thailand, Vietnam and China[11]. Hence, raising level of national rice self-sufficiency has become a strategic issue in the Agricultural Ministry of Malaysia.

The ability to forecast the future enables the farm managers to take the most appropriate decision in anticipation of that future. Predicting the future is important for the organization to plan the necessary policies or agenda. Forecasting can assist the organization to make a better planning and decision making. Therefore the selection of forecasting model is the important criteria that will influence to the forecasting accuracy. 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 [14][15]. The ARIMA model is one of the most popular models in traditional time-series forecasting and 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 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 [6][8][13][16][17][19], Group Method of Data Handling (GMDH)[5][7][9][10][11][15]. Hence, in order to cater these challenges, this study proposes a new time series forecasting model using Group Method of Data Handling (GMDH) technique. This model is an enhancement of conventional GMDH model that will improve the prediction accuracy of the traditional GMDH model.

2. OBJECTIVES

The major goal of time series forecasting is to get the best accuracy model in order to make a good decision in the organization. This research is focusing on the efforts of improving the accuracy of forecasting methods. As mentioned above, soft computing techniques or artificial intelligence have been extensively studied and used in time series forecasting. One of the technique is Group Method of Data Handling (GMDH).

GMDH is a heuristic self-organizing modeling method which are generated adaptively from data in the form of networks of active neurons in a repetitive generation of populations of competing models of growing complexity, corresponding validation, and selection model until an optimal complex model that is not too simple and not too complex have been realized.

The main idea of GMDH is to build an analytical function in a feedforward network based on a quadratic node transfer function whose coefficients are obtained using a regression technique [3]. In fact, a real GMDH algorithm in which the model coefficients are estimated by means of the least square method.

Hence, the goal of this study is to accurately develop a time series forecasting model using the Group Method of Data Handling (GMDH). The following objectives is considered in order to achieve the goal of this study:

i. To enhance a new design architecture of GMDH model in order to enrich and improve the prediction accuracy.

ii. To compare the performance of the new GMDH model with the benchmarked individual models such as ARIMA, ANN, and conventional GMDH

iii. To implement the new hybrid GMDH model with the benchmarked data such as Chemical process concentration reading:every 2 hour, Chemical process viscosity reading: every hour, IBM common stock closing prices, Wolf's Sunspots Numbers, International airline passengers and non-benchmarked data such as rice yield data

3. METHODOLOGY

A method of designing an enhanced of group method of data handling (GMDH) forecasting model consists four phases namely the data collection, pre-processing, design the architecture of GMDH model and compare the new model with others previous model. The detail of these phase are describe as below:

Phase 1: Data Collection

There are two type of data have been used in this study which are unbenchmarked and benchmarked data. The unbenchmarked data are collected from Muda Agricultural Development Authority (MUDA) Kedah, Malaysia ranging from 1995 to 2001. These time series come from different location and have different statistical characteristics.

The others data consists the benchmarked data series are employed by Hansen et al. (1994) and Hamzacebi et al. (2009). These data are utilized to forecast through an application aimed to handled real life time series. There are five well-known data sets - the chemical process concentration reading data, the IBM common stock closing prices, the chemical process temperature reading, the Wolf's sunspot data and the international airline passengers are used in this study to demonstrate the effectiveness of the hybrid model. These time series come from different areas and have different statistical characteristics. They have been widely studied in the statistical as well as the ANN literature. All series were classified into four main categories that encompass the majority of the time series types, namely seasonal and trended, seasonality, trended and nonlinear.

Phase 2: Data Pre-Procesing

All datasets are scaled during the data preprocessing phase. The two main advantages of scaling are to avoid attributes in greater numeric ranges from dominating those in smaller numeric ranges, and to prevent numerical difficulties during the calculation [4]. Generally, each data can be linearly scaled to the [0,1] range using the following formula $y_t = \frac{x_t}{x_{\max}}$ …………..(1)

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

Then, the scaled or normalized data is divided into training and testing. The training data set containing the first 85% values and the last 15% will be the testing data set. The training data will become the input to the model architecture. Meanwhile, the testing dataset will be used to validate the predictive model.

Phase 3: Design the architecture of GMDH model

The design step of the architecture of GMDH model is as follows:

[Step 1]: Consider $N_1 = n$ neurons in the first layer from of input vector $\mathbf{X} = (x_i, x_i, ..., x_n)$, where *n* is the number of inputs. Let say if n is 4. So, it has $\mathbf{X} = (x_1, x_2, x_3, x_4)$

[Step 2]: Set $M = \frac{N_k(N_k - 1)}{2}$ new variables $z_1, z_2, ..., z_M$ in the data set for all independent variables $\mathbf{X} = (x_i, x_i, ..., x_n)$. Construct the partial quadratic polynomials

$$z_k = f(x_i, x_j) = a_0 + a_1x_i + a_2x_j + a_3x_i^2 + a_4x_j^2 + a_5x_ix_j$$

…………………………………….(2)

(k = 1,2,.., M)

If n = 4 from step 1, then M is 6.

$$M = \frac{4(4-1)}{2} = 6$$

Then the partial quadratic polynomials are:

$$z_1 = f(x_1, x_2) = a_0 + a_1x_1 + a_2x_2 + a_3x_1^2 + a_4x_2^2 + a_5x_1x_2$$
$$z_2 = f(x_1, x_3) = a_0 + a_1x_1 + a_2x_3 + a_3x_1^2 + a_4x_3^2 + a_5x_1x_3$$
$$z_3 = f(x_1, x_4) = a_0 + a_1x_1 + a_2x_4 + a_3x_1^2 + a_4x_4^2 + a_5x_1x_4$$
$$z_4 = f(x_2, x_3) = a_0 + a_1x_2 + a_2x_3 + a_3x_2^2 + a_4x_3^2 + a_5x_2x_3$$
$$z_5 = f(x_2, x_4) = a_0 + a_1x_2 + a_2x_4 + a_3x_2^2 + a_4x_4^2 + a_5x_2x_4$$
$$z_6 = f(x_3, x_4) = a_0 + a_1x_3 + a_2x_4 + a_3x_3^2 + a_4x_4^2 + a_5x_3x_4$$

Based on Kondo(2006), the total equation of the model is increased according to eq. 3. Thus, the model become robust and suit to apply in different type of data.

$z_k = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + ... + a_rx_r$ (1<r<n) … (3)

Then, the additional equation become:

$$z_7 = a_1x_1$$
$$z_8 = a_1x_1 + a_2x_2$$
$$z_9 = a_1x_1 + a_2x_2 + a_3x_3$$
$$z_{10} = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4$$

In conventional GMDH models, many researcher considered the partial quadratic polynomials as the transfer function. However, there are many types of transfer functions that are available in the enhanced GMDH. Kondo [6] and Park et al. [10] showed that employing heterogeneous transfer function within a model gives better results rather than use units just of single transfer function. Hybrid transfer function produce more accurate result than single transfer function.

In this study 3 types of transfer function are used and shown in the Table 1. The transfer function polynomials with r nodes inputs have been implemented in this model. These were implemented not just to improve the accuracy of the algorithm but also to show that various alternatives can be incorporated into one network.

Table 1: The Transfer function for *r* node inputs

Type	Name	Transfer Function
1	Polynomial (PLY)	$y_k = z_k$
2	Sigmoid (NN)	$y_k = 1/(1+\exp(-z_k))$
3	Exponential (RBF)	$y_k = \exp(-z_k^2)$

For each transfer function, it will produce ten different equation (refer to z_k, k=10). So, for 3 transfer function, the total equation that will be produced is 30 different equation.

[Step 3]: Estimate the coefficients of the TF.

The coefficients of the TF are determined by least squared errors (MSE) in the form of

$$\mathbf{A}_i = (\mathbf{X}_i^T\mathbf{X}_i)^{-1}\mathbf{X}_i\mathbf{Y}$$

where $A = \{a_0, a_1, ..., a_5\}$ is the vector of unknown coefficients, $\mathbf{Y} = \{y_1, y_2, ..., y_M\}^T$ is the vector of output's value from observation and

$$\mathbf{X} = \begin{bmatrix} 1 & x_{1i} & x_{1j} & x_{1i}x_{1j} & x_{1i}^2 & x_{1j}^2 \\ 1 & x_{2i} & x_{2j} & x_{2i}x_{2j} & x_{2i}^2 & x_{2j}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{mi} & x_{mj} & x_{mi}x_{mj} & x_{mi}^2 & x_{mj}^2 \end{bmatrix}$$

[Step 4]: Choose the best variables and eliminate the weakest variable. The selection criteria of the best variables is based on some performance index (mean square error, absolute or relative error) that expresses how the values z_m follow the experimental output y. In some method, columns of $x_1, x_2, ..., x_n$ are replaced by the retained columns of $z_1, z_2, ..., z_m$, where m is the total number of the retained columns. In other versions, the best neuron of these m neurons is added to columns $x_1, x_2, ..., x_n$ to form a new set of the input variables.

[Step 5]: This step is to test whether the set of equations of the model can be further improved. The lowest value of the mean square error obtained in the current layer is compared with the smallest value obtained at the previous ones. If an improvement is achieved, one goes back and repeats steps 1 and 2, otherwise the iteration terminated and a realization of the network has been completed. The configuration of the conventional CGMDH structure is shown in Figure 1.

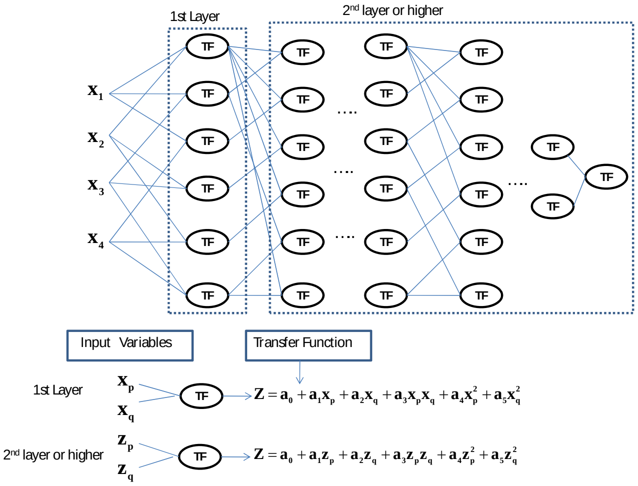

Figure 1: structure of conventional GMDH

Phase 4: Compare the result of GMDH model with other model

After the result is obtained from phase 3, then the result will be compared with the other previous model such as conventional GMDH, Neural Network and ARIMA model.

4. Empirical Result

There are two type of data have been used in this study which are unbenchmarked and benchmarked data. Both of these data will be divided into training set, containing the first 85% values and a test set, with the last 15%. Only the training set is used for model selection and parameter optimization, being the test set used to compare the proposed approach with other models. Information regarding the series distributed among the training and forecasting sets are given in Table 2.

Table 2: The series data that are used to compare forecast methods

Series	Data	Training Set	Forecasting Set
A	Rice Yields	301	50
B	Chemical process concentration reading:every 2 hour	167	30
C	IBM common stock closing prices	320	50
D	Chemical process viscosity reading: every hour	265	45
E	Wolf's Sunspots Numbers	85	15
F	International airline passengers	120	24

A. Rice Yields Data. The rice yields data contains the yields data from 1995 to 2001, giving a total of 351 observations. Given a set of 351 observations made at uniformly spaced time intervals, the locations of rice yield are rescaled to the time axis becomes the set of integers $\{1, 2, ..., 351\}$. For example the first location in 1995 is written as time 1, the second location in 1995 as time 2 and so on. The time series plot is given in Figure 2.

B. Chemical process concentration reading for every 2 hour: A very interesting problem is the prediction of data that are not generated from a mathematical expression and thus governed by a particular determinism. For this purpose, we used a natural phenomenon, the chemical process concentration reading for every 2 hour has a total of 197 data. The first 167 data pairs of the series were used as training data, while the remaining 30 (15% of the data) where used to validate the model identified. Fig. 3 shows a characterization of this time series. This series has been analysed by Box et al [2] and identified having an ARIMA(0,1,1) model.

C. IBM common stock closing prices: This time series is a real series of the daily data from May 17 1961 to November 2, 1962. The data shows a break in the last third of the series and no obvious trend or seasonality. This series has been analysed by Hamzacebi et al. (2009), and is identified as having an ARIMA (0, 1, 1)x(0,0,1)$_6$. Figure 4 shows this time series.

D. Chemical process viscosity reading for every hour: This series has 310 data and also has been analysed by Box et. al. (1994) and Hamzacebi et al. (2009), and is identified as having an ARIMA (0, 1, 1). Figure 5 shows this time series.

E. Wolf's Sunspots Numbers: The sunspot data contains the annual number of sunspots from 1770 to 1869, giving a total of 100 observations. The study of sunspot activity has practical importance to geophysicists, environment scientists, and climatologists. The plot of this time series (see Fig. 6) suggests that there is a cyclical pattern with the mean cycle of about 11 years. This series also has been analysed by Box and Jenkins, and is identified as having an ARMA(2,0) model Box et. al. ([2] and Hamzacebi et al. (2009).

F: International airline passengers: The series is an example of time series data with a clear trend and multiplicative seasonality. The data contains the monthly passengers from 1949 to 1960, has 144 data. The data has long served as a benchmark and has been well studied in statistical literature. Box et al. [2] show that the series is identified as having an (0,1,1) x (0,1,1)$_{12}$.

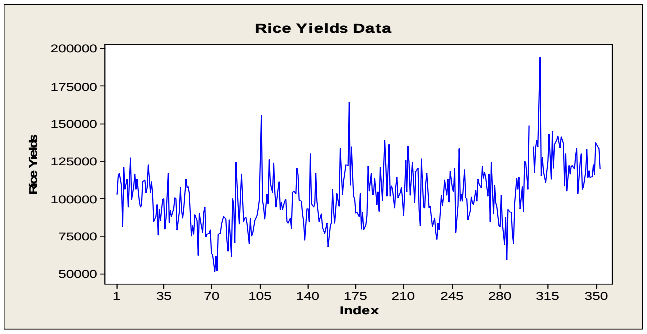

Figure 2: Series A Rice Yields Data

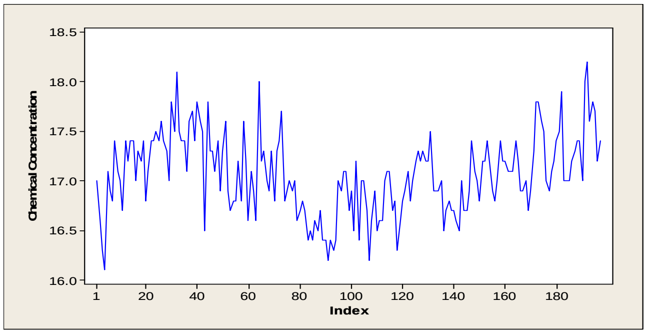

Figure 3: Series B Chemical process concentration reading(every 2 hour)

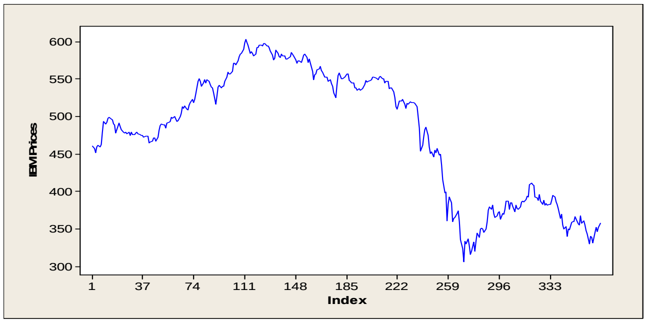

Figure 4 : Series C IBM common stock closing prices

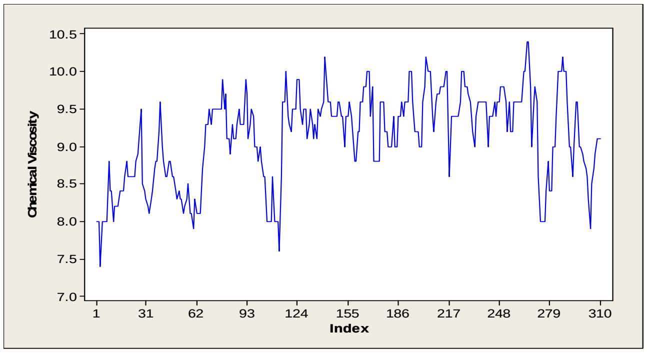

Figure 5: Series D Chemical process viscosity reading: every hour

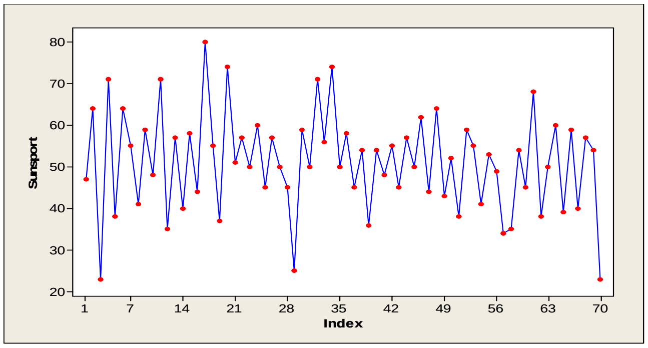

Figure 6: Series E Wolfer Sunspots Numbers

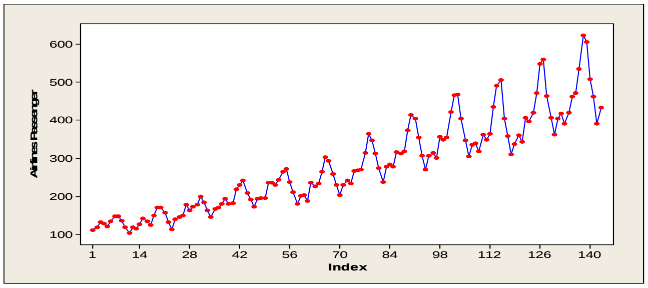

Figure 7: Series F: International airline passengers

In this section, we provide several numerical example to evaluate the advantages and the effectiveness of the KONDO approach. The performances of the each model for both the training data and forecasting data are evaluated and is selected according to the mean absolute error (MAE) and root-mean-square error (RMSE), which are widely used for evaluating results of time series forecasting. The MAE and RMSE are defined as

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

where y_t and $\hat{y}_t$ are the observed and the forecasted rice yields at the time *t*. The criterions to judge for the best

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

model are relatively small of MAE and RMSE in the modeling and forecasting.

4.1 Fitting ARIMA models to the data

Through the iterative model building process of identification, estimation, and diagnostic checking. Several models were identified and the statistical results during training and forecasting are compared and the criterions to judge for the best model are determined based on MAE and MSE. Two statistical functions are used for the graphical identification of ARMA model using the Box–Jenkins methodology, the autocorrelation function (ACF) and the partial autocorrelation function (PACF). Table 3 lists the possible ARMA structures of some real data series after being properly differenced with or without seasonal components. These models have also been used by many researchers such as Box et al. [2] and Hamzacebi et al. (2009).

Table 3 : Comparisons of ARIMA models' statistical results

Series	ARIMA	Training		Forecasting	
	Model	RMSE	MAE	RMSE	MAE
A	$(2,1,0)x(1,0,0)_{27}$	0.0631	0.0478	0.0834	0.0600
B	(0,1,1)	0.0176	0.0154	0.0248	0.0206
C	$(0,1,1)x(0,0,1)_{6}$	0.0118	0.0113	0.0787	0.1219
D	(0,1,1)	0.0282	0.0224	0.1578	0.1713
E	(2,0,0)	0.0999	9.0799	0.1393	1.0260
F	$(0,1,1)x(0,1,1)_{12}$	0.0187	0.0318	0.7946	0.7271

4.2 Fitting neural network models to the data

One of the key tasks in time series forecasting is the selection of the input variables and the number of neurons in the hidden layer. For the ANN models, there is no systematic approach which can be followed.The universal approximation theorem shows that a neural network with a single hidden layer with a sufficiently large number of neurons can in principle relate any give set of inputs to a set of outputs to an arbitrary degree of accuracy. As a result, the ANN designed in this study are equipped with one single hidden layer. The determination of the number of neurons in the hidden layer is more art than science.

Determining the size of the network (the number of neurons) has important consequences for its performance. Too small a network may not reach an acceptable level of accuracy. Too many neurons may result in an inability for the network to generalize (it may rote learn the training patterns).

Since the number of inputs varies depending on the input determination method used, it is not possible to use the same network architecture for each model. In this study the number inputs (I), 2, 4, 6, 8 and 12 were used for all data set. To help avoid the overfitting problem, some researchers have provided empirical rules to restrict the number of hidden nodes. To select an appropriate architecture, the guidelines using "2I" proposed by Maier Wong, "I" proposed by Tang & Fishwick [13] and "I/2" proposed by Kang [6].

The network was trained for 5000 epochs using the back-propagation algorithm with a learning rate of 0.001 and a momentum coefficient of 0.9. The networks that yielded the best results for the forecasting set were selected as the best ANN for the corresponding series. The best ANN structures and the best results for the training and forecast are shown in Table 4.

Table 4.3: Comparison of Neural Network Model

Series	Number of neurons			Training		Forecast	
	Input	Hidden	Output	RMSE	MAE	RMSE	MAE
A	12	6	1	0.0718	0.1140	0.0896	0.1019
B	12	24	1	0.0152	0.0127	0.0204	0.0157
C	12	24	1	0.0144	0.0133	0.1320	0.0181
D	4	4	1	0.0277	0.0233	0.0368	0.0348
E	4	4	1	0.0842	10.7032	0.0729	0.2426
F	12	6	1	0.0321	0.0688	0.1001	0.1155

4.3 Fitting GMDH and KONDO models to the data

The GMDH and KONDO models are typically composed of layers of nodes. In designing these models, one must determine the following variables: the number of input nodes, the number of hidden layers and the number of output. The selection of the number of input corresponds to the number of variables play important roles for many successful applications of these model. The issue of determining the optimal number of input nodes is a crucial yet complicated one. There is no theory that can used to guide the selection the number of input.

In this study, the 2, 4, 6, 8 and 12 are also used as the number of input nodes. Table 4 and Table 5 shows the comparison of modeling/forecasting precision among the two different approaches based on two statistical measures.

Table 4: Performance of GMDH

Series	Number of			Training			Forecasting	
	Input	Layer		RMSE	MAE		RMSE	MAE
A	12	3		0.0692	0.1064		0.0806	0.0913
B	8	2		0.0156	0.0135		0.0200	0.0149
C	8	2		0.0115	0.0111		0.0117	0.0161
D	4	2		0.0266	0. 0223		0.0351	0.0332
E	4	2		0.0785	3.7045		0.0727	0.2255
F	12	1		0.0237	0.0498		0.0638	0.0741

Table 5: Performance of KONDO approach

Series	Number of			Training			Forecasting	
	Input	Layer		RMSE	MAE		RMSE	MAE
A	6	3		0.0699	0.1079		0.0754	0.0786
B	12	3		0.0162	0.0132		0.0175	0.0135
C	8	3		0.0110	0.0106		0.0109	0.0150
D	12	3		0.0247	0.0206		0.0318	0.0276
E	8	3		0.0581	6.1845		0.0765	0.2434
F	12	3		0.0250	0.0540		0.0257	0.0319

For comparison purpose, the training and the forecast performance of KONDO were compared with the ARIMA, ANN and GMDH model. Table 6 shows the comparison of modeling/forecasting precision among the two approaches based on two statistical measures.

The results show the KONDO models in many cases is more competent in modeling and forecasting time series than the other models. The ARIMA model has a good performance for the two series (A and F) for testing.

Table 6: Comparative performance of the KONDO, ARIMA, ANN and GMDH for Six Series Data

		RMSE				MAE			
	Series	ARIMA	ANN	GMDH	KONDO	ARIMA	ANN	GMDH	KONDO
Testing	A	**0.0631**	0.0718	0.0692	0.0699	**0.0478**	0.1140	0.1064	0.1079
	B	0.0176	**0.0152**	0.0156	0.0162	0.0154	**0.0127**	0.0135	0.0132
	C	0.0118	0.0144	0.0115	**0.0110**	0.0113	0.0133	0.0111	**0.0106**
	D	0.0282	0.0277	0.0266	**0.0247**	0.0224	0.0233	0. 0223	**0.0206**
	E	0.0999	0.0842	0.0785	**0.0581**	9.0799	10.703	**3.7045**	6.1845
	F	**0.0187**	0.0321	0.0237	0.0250	**0.0318**	0.0688	0.0498	0.0540
Forecas-ting	A	0.0834	0.0896	0.0806	**0.0754**	**0.0600**	0.1019	0.0913	0.0786
	B	0.0248	0.0204	0.0200	**0.0175**	0.0206	0.0157	0.0149	**0.0135**
	C	0.0787	0.1320	0.0117	**0.0109**	0.1219	0.0181	0.0161	**0.0150**
	D	0.1578	0.0368	0.0351	**0.0318**	0.1713	0.0348	0.0332	**0.0276**
	E	0.1393	0.0729	**0.0727**	0.0765	1.0260	0.2426	**0.2255**	0.2434
	F	0.7946	0.1001	0.0638	**0.0257**	0.7271	0.1155	0.0741	**0.0319**

5 Conclusion

Forecasting is an important problem that spans many fields including business and industry, government, economics, environmental sciences, medicine, social science, politics, and finance. The reason that forecasting is so important is that prediction of future events is a critical input into many types of planning and decision making.

In this study, we conducted the experiment which involves two type of data, benchmarked data and non-benchmarked data. The non-benchmarked data obtained from Muda Agricultural Development Authority (MADA) Kedah, Malaysia ranging from 1995 to 2001. These time series come from different location and have different statistical characteristics. The others data consists the benchmark data series employed by Hansen et al. (1994) and Hamzacebi et al. (2009) are utilized to forecast through an application aimed to handled real life time series. There are five well-known data sets - the chemical process concentration reading data, the IBM common stock closing prices, the chemical process temperature reading, the Wolf's sunspot data and the international airline passengers are used in this study to demonstrate the effectiveness of the model. The **results** are compared with those the ARIMA, ANN, and GMDH.

From the experimental results comparing the performance of six models, we can conclude that KONDO models in many cases is more competent in modeling and forecasting time series than the other models.

6 ACKNOWLEDGMENTS

Our thanks to ACM SIGCHI for allowing us to modify templates they had developed and Soft Computing Research Group (SCRG) for supporting the comments and idea to this study.

7 REFERENCE

[1] Abdullah Ahmad Badawi. 2009. *Malaysia Aims To Increase Rice Production By 2010*. Bernama.

[2] Box GEP , Jenkins, GM. 1994. Time Series Analysis: Forecasting and Control, Prentice Hall PTR, Upper Saddle River, NJ.

[3] Farlow, S.J. (1984). Self-organizing Method in Modeling:GMDH Type Algorithm, Marcel Dekker Inc.

[4] Hsu, C.W., Chang, C.C and Lin, C.J. (2003). A practical guide to support vector classification. Available at : http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.

[5] Kalantary, F., Ardalan, H. and Nariman-Zadeh,N. (2009). An Investigation on the $S_u - N_{SPT}$ correlation using GMDH type neural networks and genetic algorithms. *Engineering Geology* 104 :144-155.

[6] Kang, S. An investigation of the Use of Feedforward Neural Network for Forecasting. Ph.D. Thesis, Kent State University, 1991.

[7] Kondo, T. and Ueno, J. (2009). Medical Image Recognition of Abdominal Multi-Organs by RBF GMDH-type Neural Network. *International Journal of Innovative Computing, Information and Control*. 5:1349-4198.

[8] Lai, K.K., Yu, L., Wang, S. & Huang, W. (2006). Hybridizing Exponential Smoothing and Neural Network for Financial Time Series Prediction. *ICCS 2006, Part IV, LNCS 3994*: 493-500.

[9] Oh, S.K., Pedrycz, W., and Roh, S.B. (2006).Genetically optimized fuzzy polynomial neural networks with fuzzy set-based polynomial neurons, *Inform. Sci.* 176(23) : 3490–3519.

[10] Onwubolu G.C. (2008). Design of hybrid differential evolution and group method of data handling networks for modeling and prediction. *Information Science*. 178:3616-3634

[11] Park, H.S., Pedrycz, W. and Oh, S.K. (2007). Evolutionary design of hybrid self-organizing fuzzy polynomial neural networks with the aid of information granulation, *Expert Syst. Appl.* 33 (4) : 830–846.

[12] Saad, P., Bakri, A., Kamarudin, S.S. & Jaafar, M.N. 2006. Intelligent Decision Support System for Rice Yield Prediction in Precision Farming. IRPA Report.

[13] Tang, Z. & Fishwick, P.A. Feedforward Neural Nets as Models for Time Series Forecasting. ORSA Journal on Computing, 5(4):374-385, 1993.

[14] Wong, F.S. Time Series Forecasting Using Backpropagation Neural Network. Neurocomputing, 2:147-159, 1991.

[15] Wongseree, N., Chaiyaratana,N., Vichittumaros, K., Winichagoon, P. and Fucharoen, S. (2007). Thalassaemia classification by neural networks and genetic programming, *Inform. Sci*. 177 (3) :771–786.

[16] Wu Jr. S., Han J, Annambhotla S. and Bryant, S., (2005), Artificial Neural Networks for Forecasting Watershed Runoff and Stream Flows, *Journal of Hydrologic Engineering*, 5(2), 216-222.

[17] Zhang, G., Patuwo, B.E. & Hu, M.Y. 2003. Forecasting with artifical neural networks: The state of the art. 14:35-62

[18] Zhang, G.P. 2003. Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. Neurocomputing 50: 159-175.

[19] Zou, H.F., Xia, G.P., Yang, F.T. & Wang, H.Y. (2007). An Investigation and Comparison of Artificial Neural Network

and Time Series Models for Chinese Food Grain Price Forecasting. Neurocomputing, 70: 2913-2923.