FORECASTING TIME SERIES DATA USING HYBRID GREY RELATIONAL ARTIFICIAL NEURAL NETWORK AND AUTO REGRESSIVE INTEGRATED MOVING AVERAGE MODEL

*Roselina Sallehuddin, Siti Mariyam Hj. Shamsuddin, Siti Zaiton Mohd. Hashim,***
Ajith Abraham†

Abstract: In business, industry and government agencies, anticipating future behavior that involves many critical variables for nation wealth creation is vitally important, thus the necessity to make precise decision by the policy makers is really essential. Consequently, an accurate and reliable forecast system is needed to compose such predictions. Accordingly, the aim of this research is to develop a new hybrid model by combining a linear and nonlinear model for forecasting time series data. The proposed model (GRANN_ARIMA) integrates nonlinear Grey Relational Artificial Neural Network (GRANN) and linear ARIMA model, combining new features such as multivariate time series data as well as grey relational analysis to select the appropriate inputs and hybridization succession. To validate the performance of the proposed model, small and large scale data sets are used. The forecasting performance was compared with several models, and these include: individual models (ARIMA, Multiple Regression, Grey Relational Artificial Neural Network), several hybrid models (MARMA, MR_ANN, ARIMA_ANN), and Artificial Neural Network (ANN) trained using Levenberg Marquardt algorithm. The experiments have shown that the proposed model has outperformed other models with 99.5% forecasting accuracy for small-scale data and 99.84% for large-scale data. The empirical results obtained have proved that the GRANN_ARIMA model can provide a better alternative for time series forecasting due to its promising performance and capability in handling time series data for both small and large scale data.

Key words: *Hybrid model, GRANN_ARIMA, multivariate time series, linear model, nonlinear model*

*Roselina Sallehuddin, Siti Mariyam Hj. Shamsuddin, Siti Zaiton Mohd. Hashim,
Faculty of Computer Science and Information System, University Technology Malaysia, 81310, Skudai, Johor, E-mail: `roselina, mariyam, sitizaiton@utm.my`

†Ajith Abraham
Centre for Quantifiable Quality of Service, Norwegian University of Science and Technology, Norway, E-mail: `ajith.abraham@ieee.org`

Received: November 26, 2007
Revised and accepted: December 15, 2007

1. Introduction

Predicting the future is important for the organization to plan or adopt the necessary policies. Forecasting can assist them to make a better development and decision-making for the country. There are various forecasting techniques available in the academic literature. However, the selection of these techniques normally depends on the availability of data, the quality of available models and some predefined assumptions. According to Makridakis et al. [33], each method is different in terms of accuracy, scope, time horizon and cost. To facilitate an adequate level of forecasting accuracy, the developer has to be responsive to the characteristics of different methods, and determine if a particular method is appropriate for the undertaken situation before embarking its usage in real application. As a result, the choice of a forecasting model is one of the important factors that will influence the forecasting accuracy.

Forecasting methods can be broadly divided into two categories: Statistical and Artificial Intelligence (AI) based techniques. Box-Jenkins or Auto Regressive Integrated Moving Average (ARIMA), Multiple Regressions and Exponential Smoothing are examples of statistical methods, whilst AI paradigms include fuzzy inference systems, genetic algorithm, neural networks, machine learning etc. [55]. Statistical methods are usually associated with linear data, while neural networks are usually associated with nonlinear data. Statistical methods have been used successfully in time series forecasting for several decades. As well being simple and easy to interpret, statistical methods also have several limitations. One of the major limitations of statistical methods is it is merely depicted as a linear model, also known as model driven approach. Thus, they have to fit the data with the available data and the prior knowledge about the relationships between the inputs and outputs before modeling is highly desired.

Due to the limitations of statistical methods, nonlinear statistical time series models have been proposed, with the aim to improve the forecasting performance of nonlinear systems. These include bilinear model, threshold autoregressive model (TAR), smoothing transition autoregressive model (STAR), autoregressive conditional heteroscedastic model (ARCH), and generalized autoregressive conditional heteroscedastic model (GARCH). These models are known as the second generation of time series models. However, limited success or gain has been found during the last two decades using nonlinear models since most of them are developed specifically for particular problems without broad-spectrum applicability for other situations. In addition, the formulations of these models are more complex and difficult to develop compared to linear models [21, 22].

Hence, a different approach has been proposed and engaged successfully in time series forecasting. Artificial neural network (ANN) has been applied in solving numerous time series forecasting problems such as stock, electricity prices, breast cancer, rainfall-runoff [1, 15, 18, 23, 41, 45] and others. One of the main reasons that ANN performs better than the statistical method is due to its influential feature in handling nonlinear time series data. In addition, ANN has also been shown to be effective in modeling and forecasting nonlinear time series with or without noise [55]. ANN also does not require any knowledge nor prior information about systems of interest [31]. Hippert et al. [19] and Zhang [57] have claimed that forecasting is a major application area of ANN.

Zhang et al. [56] has compiled the results achieved by previous researchers. Even though most of published researches indicate the superiority of the ANN model in comparison to a simpler linear model, but quite a few studies give disparity comments on the ANN performance. Denton [16] and Gorr et al. [20] showed that the ANN perform about the same as the linear model. Several other researchers [7, 8, 24, 46] also reported the pessimistic findings about ANN in forecasting daily electric load for one step ahead forecast. In their study, they showed that ANN was not as effective as the linear time series model in forecasting performance even if the data is nonlinear. However, Kang [27] had shown that ANN always performs well compared to ARIMA, and even better when the forecasting horizon is increased.

Some researchers [6, 11, 33, 55] have reported that there is no such a single forecasting method that gives an appropriate result in all situations. The reason for this is due to the characteristic of the model itself, in which the statistical model is usually a linear model, and ANN is a nonlinear model. Each of them will perform well in linear and nonlinear data respectively. Therefore, it is hard for us to determine whether the time series problems under study are linear or nonlinear, particularly when we are dealing with real world time series data.

With the intention to improve the forecasting accuracy, the combination of forecasting approaches has been proposed by many researchers [4, 5, 9, 36]. From their studies, they indicate that the integrated forecasting techniques outperform the individual forecasts.

The remainder of this paper is organized as follows. In Section 2, further discussion on hybrid models is presented. Section 3 describes the methodology that will be implemented throughout this study. Then the proposed hybrid model for forecasting time series is discussed in detailed in Section 4. Meanwhile, Section 5 and 6 will describe the results of the experiment to test the usefulness of the proposed hybrid model (GRANN_ARIMA). Finally, the conclusion that describes the contribution of this paper is summarized and several future researches are listed in Section 7.

2. Literature Review on Hybrid Models

Hybrid models have been introduced to overcome the deficiency of using a individual model such as statistical methods (ARIMA, Multiple Regression and etc.) and AI methods. Hybrid models merge different methods to improve the prediction accuracy. Hybrid models can be also referred as combined models or ensemble models and often used synonymously. Hybrid methods can be implemented in three different ways; linear models, nonlinear models and both linear and nonlinear models.

In linear hybridization, two or more linear models are combined together using the same data set or a different data set to gain an ultimate forecasting value. Shamsuddin and Arshad [42] had used multivariate autoregressive moving average (MARMA) model to predict natural rubber prices for Malaysian Domestic market.

Shamsuddin and Arshad's work differs from Shamsuddin [43] in terms of techniques, which is implemented using different model based on the different set of data. Authors combined auto regressive moving average (ARMA) and econometric model (multiple regression), where the ARMA model is used to explain the residual yield from the multiple regression model. The findings show that the forecasting errors produced by the MARMA model is reduced by 4.5 per cent compared to the individual econometric models. This result indicates that the hybrid model has the potential to improve forecasting accuracy.

Hybrid forecasting has also been implemented using a nonlinear model, for instance hybridizing ANN with genetic algorithm (GA), fuzzy logic (FL) and rough sets (RS) [17, 25, 52, 58]. Authors found that by hybridizing ANN with these methods could improve the forecasting accuracy. Hou et al. [25] combined ANN and Rough set to predict air conditioning load. They used both univariate and multivariate time series data. Authors findings illustrate that the empirical results are better compared to the result given by the ANN model alone. The result also indicates that if more relevant data are used in the study, the forecasting accuracy could be better. In this hybridization, GA, RS or FL are embedded in ANN as preprocessing tools to improve the ANN forecasting performance by extracting important and significant features in time series data.

However, most of the hybridization methods, which have been proposed in the previous literature [9, 17, 43] having major drawbacks. Most of them are designed to combine similar methods; linear model with linear model, and nonlinear model with nonlinear model. In reality, time series data typically contain both linear and nonlinear patterns. Therefore, neither linear nor nonlinear model can be sufficient in modeling time series data since the linear model cannot deal with nonlinear relationship. Additionally, nonlinear model also cannot handle both linear and nonlinear pattern equally well.

To overcome this drawback, several studies have suggested the combining linear model and the nonlinear model. Previous studies have showed that combining different relevant methods could improve the forecasting accuracy. The merging of this structure can help the researchers in modeling complex autocorrelation structures in time series data more efficiently. Furthermore, by using different models or models that contradict with each other significantly; lower generalization variance or error could be generated [54].

Five studies have suggested in using the hybrid models, i.e., combining the ARIMA model and ANN. Zhang [54] used this combination and implemented on three data sets; Wolfs sunspot data, Canadian lynx data and British pound/US dollar change. They used a backpropagation learning algorithm to model the residual yield from the ARIMA model. Pai and Lim [37] used a hybrid model to forecast daily stock by using the support vector machine and ARIMA model. Lu et al. [28] used the hybrid model to forecast daily load data and Tseng and Tzeng [49] the combined seasonal auto regressive integrated moving average (SARIMA) and the backpropagation model to forecast seasonal time series data. Their results showed that the hybrid model produced a better forecasting result compared to the SARIMA model or the ANN model alone. Jain and Kumar [26] found that more accurate results could be obtained by hybridizing the ARIMA model and ANN in forecasting hydrologic time series.

However, several researchers have argued that the predictive performance improves when using hybrid models [2, 46, 47]. For example, Taskaya and Casey [47] showed that an individual model outperformed five of nine data sets used. These inconsistent results indicate the need for further research on how to obtain a good forecasting result from a hybrid linear and a nonlinear model. It is observed that there are three weaknesses in the previous studies such as the type of data used, redundancy factors and implementation of hybridization sequence.

First, most of the studies used univariate time series data. Most of them are solely based on one historical data such as previous sales and previous income. However, Hou et al. [25] showed that considering more significant input can improve the forecasting accuracy. Furthermore, Makridakis et al. [32, 33] illustrated that the accuracy of the time series methods can be improved by incorporating multivariate information that will affect the future behavior of the series so that the prediction can be improved.
Second, in most of the works utilizing ANN for prediction didn't look at the possibility of input redundancy. For an ordinary user, ANN appears like a black box processor that does not have any capability to recognize insignificant inputs. Improper selection and redundancy of inputs can lead to instability that will affect the accuracy of prediction [30]. Several methods have been introduced to eliminate the redundancy inputs such as grey relational analysis [53]. Markov blanket model, decision trees [59], genetic programming [60] and adaptive genetic algorithm [10].
Third, the hybrid sequence in the conventional hybrid is normally started with a linear model and followed by a nonlinear model to model the residual. This is due to the ANN's capability to deal with linear data that tend to be usually over fitting. But, Heravi et al. [24] had showed that the linear ARIMA model outperformed the ANN in forecasting nonlinear stock data. This result indicated that over the fitting problem essentially occurred in both the linear and nonlinear model. Nevertheless, the issue is to choose one that will suffer from over a fitting problem acutely.

Hence, in this study, a new hybrid approach for combining a nonlinear model and a linear model is proposed to overcome the drawbacks of previous studies by including more additional features; these include multivariate time series, feature selection in removing and selecting significant input data and altering the sequence of combination execution. In this study, the grey relational analysis (GRA) is integrated with ANN (GRANN) to remove the redundancy inputs. The grey relational analysis is employed due to its adaptability in dealing with small or large data sets [53].

3. Research Methodology

We address four research questions as listed below based on the various issues mentioned above. In practice, multiple regression (MR) is usually used in modeling multivariate time series data due to its simplicity [12, 35, 50]. In this study, we attempt to use the grey relational neural network (GRANN) instead of MR. Thus, the issues that need to be addressed are:

- Could the GRANN model outperform the multi regression model in handling multivariate time series analysis?
- Would forecasting accuracy increase when a hybrid model is used instead of an individual model in forecasting multivariate time series data?
- Would the sequence changes of hybridization implementation affect the forecasting accuracy?
- Would the proposed hybrid model GRANN_ARIMA give the best forecasting accuracy compared to others?

To address the above research questions, two experiments are conducted. The first experiment compares the performance between the multiple regression model (MR) and the grey relational neural network model (GRANN) in handling multivariate time series analysis, and the second experiment examines the accuracy of the combination between linear and nonlinear time series forecasting models in predicting multivariate time series analysis. Although several studies have shown that the combinations of the linear and nonlinear models could improve the accuracy, but most of the studies employed univariate time series data [28, 37, 51]. Therefore, in this study, the experiments are conducted to see whether the same result will be formed whenever the multivariate time series data are employed. Furthermore, to investigate the effect of changing hybrid sequence, two types of hybrid model (Hybrid I and Hybrid II) are developed. Hybrid I consists of MR and ANN using the conventional hybrid sequence, and Hybrid II integrates GRANN and ARIMA with an altered hybrid sequence. Hybrid I is used as a comparative model in order to evaluate the performance of the proposed hybrid model. As benchmark, a conventional hybrid method as proposed in the previous study (ARIMA_ANN) for handling univariate time series data is also developed. To find out whether the GRANN_ARIMA is the best model for forecasting multivariate time series data, several comparisons are conducted as stated below:

- Comparison between GRANN_ARIMA with individual models and hybrid model.
- Comparison between GRANN_ARIMA and the traditional ARIMA and MARMA models (each of them represents the standard benchmark for the univariate individual model and the hybrid multivariate model respectively).
- Comparison between the GRANN_ANN with the ANN trained by second order error using the Levenberg-Marquardt (LVM) approach.

3.1 A Framework of the proposed Hybrid methods

Figs. 1 and 2 illustrate the framework of the conventional hybrid model [28, 51, 54]; Hybrid I and the proposed Hybrid II. The conventional hybrid model and Hybrid I use the same sequence of hybridization in which a linear model is applied primarily to find the linear relationship in the data. Subsequently, ANN is utilized to model the residual derived from the linear model. In this case, we assume that the linear components have been fully identified by the linear model. Consequently, the residual left in the data presents the nonlinear component. To confirm that the assumption is true, the McLeod and Li tests are used to verify the nonlinearity of the residual data, before the modeling process using ANN is carried out by Hybrid I.

Meanwhile, the proposed Hybrid II model and the conventional hybrid model are reversed from each other in terms of the model used and the sequence of hybridization. In Hybrid II, GRANN is initially applied, and followed by linear model, ARIMA. In this model, GRA is used to select the significant inputs before the forecasting is implemented using ANN. McLeod and Li test is conducted to verify the linearity of the residual data; both of these steps are not included in conventional methods.

Tab. I summarizes the similarities and differences of each model. Hybrid I is a hybrid model using conventional approach with multivariate time series data. Meanwhile Hybrid II (GRANN_ARIMA) is the proposed approach for forecasting multivariate time series data. Grey Relational analysis is used as feature selection tools to extract the significant factors that have effect on China crop yield and daily KLSE close price.

Type of Hybrid	Data	Model	Sequence of hybridization	Feature Selection
Conventional	Univariate	ARIMA, ANN	Linear, Non-linear	none
Hybrid I	Multivariate	MR, ANN	Linear, Non-linear	Goodness Fit test
Hybrid II (Proposed model)	Multivariate	GRANN, ARIMA	Nonlinear, Linear	Grey Relational Analysis

Tab. I *Similarities and differences between Conventional Model and Proposed Hybrid Model.*

3.2 Experimental data setup

To assist the analysis for benchmarking of the proposed model, two different datasets are used. The first sample contains 13 observations and represents small scale data. These data are obtained from [53] and they had specified annual China gross grain crop yields with their affecting factors. Zhang and He [53] has combined two methods; neural network and rough set to predict the national gross grain crop yield from 1990 to 2003.

Tab. II shows yearly data for gross grain crop yields and its affecting factors in China during 1990 till 2003. There are ten factors that affect the production of gross grain crop in China, and these include total power of agricultural, (a), electricity consumed in rural areas, (b), irrigation area, (c), consumption of chemical fertilizer, (e), areas affected by natural disaster, (f), budgetary expenditure for agriculture,

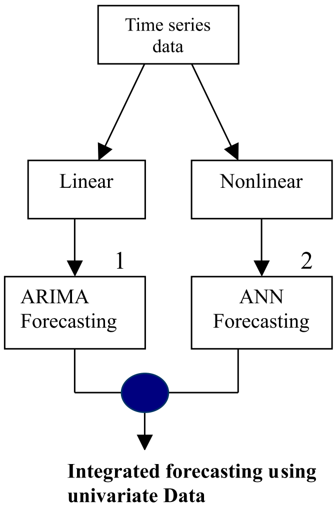
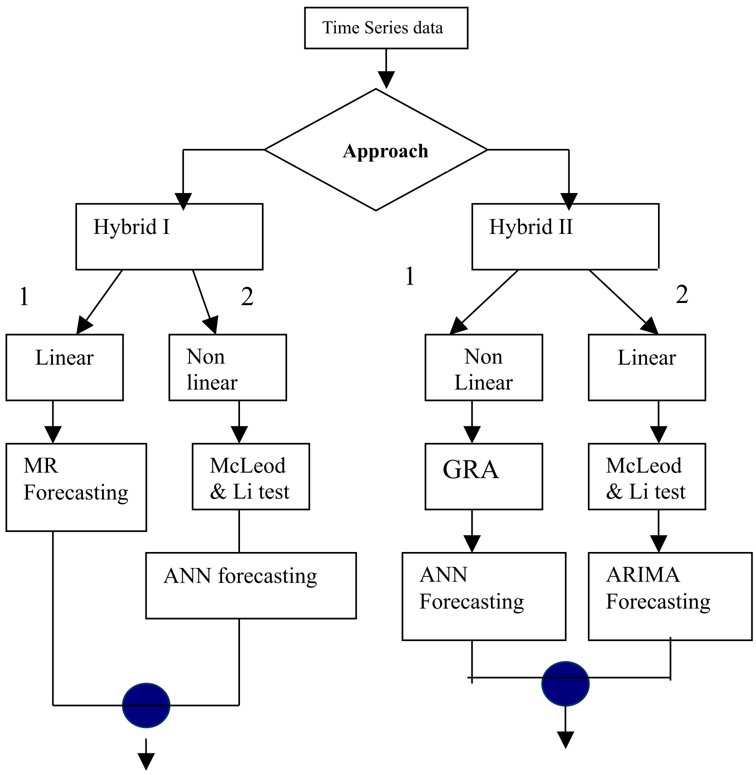


Fig. 1 *A Conventional Hybrid Model.*



Integrated forecasting using multivariate data

Fig. 2 *A Proposed Hybrid Model.*

(g), sown area of grain crops, (h), consumption of pesticide, (l), consumption of agricultural film, (m), and agriculture laborers, (n). The total production of grain crop yield is denoted by (d).

KLSE data contains 200 observations of daily Kuala Lumpur Stock Exchange (KLSE) close price from 4th January 2005 till 21st October 2005, and it represents large scale data. Tab. III illustrates fraction of the stock market data set that was used in this research. They are daily close price for KLSE ($Close\_KLSE$), consumer index (CI), construction index (CoI), gold index (GI), finance index (FI), product index (PI), Mesdaq index (MI), mining index ($MinI$), plantation index (PlI), property index ($ProI$) ($ProI$), syarian index (SI), technology index (TI), trading/service index (TSI), composite index ($CptI$) and industrial index (II).

Essentially, each data set is divided into two parts: in-of-sample and out-of-sample data. In-of-sample data refers to the training data set and is used exclusively for model development. While out-of-sample refers to the test data and is used for evaluation of the unseen data. In other words, the test data is used for an independent measure of how the model might be expected to perform on untrained data. However, in ANN, training data are usually divided further into training and validation set, where the validation set is used to monitor network performance during training with the intention that early stopping criteria will be met if the network attempts to over fit the training data.

3.2.1 Grey relational analysis

Grey relational analysis (GRA) is an analysis method that has been introduced in Grey System by Deng Julong [13, 14]. GRA is used to evaluate the degree of correlation for different data sequences. The degree of correlation between a data sequence (x) and the reference sequence (y) is expressed by a scalar within bound of 0 and 1. If the degree of correlation is near to 1, it indicates the high correlation between x and y.

There are 3 main steps in GRA. The first step is data pre-processing. Data pre-processing is normally required due to the range, and unit of one data sequence might differ from others. Therefore, data must be normalized, scaled and polarized initially into a comparable sequence before proceeding to other steps. There are few equations for data preprocessing using grey relational analysis. In this study, Equation (1) is employed (Tosun [61]):

$$x^*(k) = \frac{x_i^0(k) - \min x_i^0(k)}{\max x_i^0(k) - \min x_i^0(k)}, \tag{1}$$

where

$$i = 1, \dots m; \qquad k = 1, \dots n.$$

m	is the number of experimental data items,
n	is the number of parameters,
$x_i^0(k)$	is the original sequences,
$x_i^*(k)$	is the sequences after data preprocessing,
$\min x_i^0(k)$	and $\max x_i^0(k)$ is the smallest and the largest value of $x_i^0(k)$.

Year	a	b	c	e	f	g	h	l	m	n	d
1990	28707.7	844.5	47403.1	2590.3	17819	9.98	113466	73.3	481982	33336.4	44624
1991	29388.6	963.2	47822.1	2805.1	27814	10.26	112314	76.1	642145	34186.3	43529
1992	30308.4	1106.9	48590.1	2930.2	25859	10.05	110560	79.5	780611	34037.2	44265.8
1993	31816.6	1244.8	48727.9	3151.9	23133	9.49	110509	84.5	707321	33258.2	45648.8
1994	33802.5	1473.9	48759.1	3317.9	31383	9.2	108544	97.9	887064	32690.3	44510.1
1995	36118.1	1655.7	49281.2	3593.7	22267	8.43	110060	108.7	915487	32334.5	46661.8
1996	38546.9	1812.7	50381.4	3827.9	21233	8.82	112548	114.1	1056151	32260.4	50453.5
1997	42015.6	1980.1	51238.5	3980.7	30309	8.3	112912	119.5	1161532	32434.9	49417.1
1998	45207.7	2042.1	52295.6	4083.7	25181	10.69	113787	123.2	1206867	32626.4	51229.5
2000	48996.1	2173.4	53158.4	4124.3	26731	8.23	113161	132.2	1258674	32911.3	50838.6
2001	52573.6	2421.3	53820.3	4146.4	34374	7.75	108463	128	1335446	32797.5	46217.5
2002	55172.1	2610.8	54249.4	4253.8	31743	7.71	106080	127.5	1449286	32451	45263.7
2003	57929.9	2993.4	54354.8	4339.4	27319	7.17	103891	131.2	1539484	31990.6	45705.8

Tab. II *Grain crop yield and its affecting factors.*

Date	CI	CoI	GI	FI	PI	MI	$MinI$	PlI	$ProI$	SI	TI	TSI	$CptI$	II	$Close\_klse$
4-Jan-05	231.61	169.94	213.01	7460.1699	84.1	121.94	361.24	2415.1001	710.18	132.72	42.07	131.04	902.49	1946	902.49
5-Jan-05	232.22	169.78	214.07	7511.6602	84.16	122.06	365.64	2400.5	711.62	133.5	41.79	132.19	907.96	1949.29	907.96
6-Jan-05	232.78	169.66	214.76	7544.4902	84.32	122.35	352.42	2402.3101	714.54	133.88	41.84	132.7	910.42	1960.54	910.42
7-Jan-05	234.42	170.47	216.46	7630.5098	85.25	123.65	352.42	2416.8899	721.89	134.66	42.08	133.52	916.28	1974.95	916.28
10-Jan-05	235.9	171.61	217.17	7650.6201	85.48	124.13	361.24	2420.1399	725.84	135.13	42.32	133.87	919.02	1980.6801	919.02
11-Jan-05	237.2	173.49	219.79	7817.7002	86.19	125.78	359.03	2427.01	734.76	136.36	42.34	135.38	930.63	1997.16	930.63
12-Jan-05	237.75	172.93	220.19	7846.1899	86.13	125.52	361.24	2418.9099	736.34	136.56	41.84	135.77	933.33	2017.97	933.33
13-Jan-05	237.74	172.96	220.49	7823.4199	85.91	124.43	361.24	2424.23	736.25	136.71	41.76	136.51	934.1	2021.42	934.1
14-Jan-05	238.04	177.51	220.43	7941.6001	85.66	126.68	378.86	2404	735.01	136.03	41.64	135.19	935.53	2030.74	929.74
17-Jan-05	237.45	176.54	220.31	7844.52	85.8	125.65	352.42	2420.9199	733.58	136.29	41.76	135.83	932.26	2023.1899	932.26
18-Jan-05	238.3	176.14	220.99	7884.2202	86.24	125.35	370.05	2407.1799	733.87	136.7	41.76	136.4	937.56	2040.34	937.56
19-Jan-05	238.04	177.51	220.43	7941.6001	85.66	126.68	378.86	2404	735.01	136.03	41.64	135.19	935.53	2030.74	935.53
20-Jan-05	237.43	177.13	219.62	7919.7402	85	125.82	387.67	2400.26	734.97	135.32	41.64	134.53	929.72	2012.4399	929.72
24-Jan-05	236.86	176.26	218.32	7820.79	85.51	125.43	387.67	2410.21	732.82	134.79	41.46	133.4	923.33	2016.41	923.33
25-Jan-05	236.45	175.65	217.4	7839.4199	85.36	126.84	385.46	2406.1299	730.81	134.09	41.32	132.09	919.75	2000.63	919.75
26-Jan-05	236.11	176.44	217.86	7823.8999	85.12	125.87	381.06	2397.1299	731.46	134.68	41.25	133.12	923.37	1999.89	923.37

Tab. III *Fraction of daily KLSE close price with 14 affecting factors.*

The range of data is adjusted so as to fall within [0,1] range. The second step is to locate the grey relational coefficient by using Equation (2), (Tosun [61]):

$$\xi_i(k) = \frac{\Delta \min + \zeta \Delta \max}{\Delta_{0,i}(k) + \zeta \Delta \max,} \tag{2}$$

where
$\xi_i(k) =$ grey relational coefficient at any data point (k),
$\Delta_{0i} =$ deviation sequences of the reference sequence and comparability sequence,
$\Delta_{0,j} = \left\| x_0^*(k) - x_i^*(k) \right\|,$
$\Delta \min = \min\limits_{\forall j \in i} \min\limits_{\forall k} \left\| x_0^*(k) - x_j^*(k) \right\|,$
$\Delta \max = \max\limits_{\forall j \in i} \max\limits_{\forall k} \left\| x_0^*(k) - x_j^*(k) \right\|,$
$x_0^*(k) =$ the reference sequence, and
$x_i^*(k) =$ the comparative sequence.

ζ is known as identification coefficient with $\zeta \in [0, 1]$, and normally $\zeta = 0.5$ is used.

Finally, to obtain the grey relational grade, the average value of the grey relational coefficient is computed and is defined as (Tosun [61]):

$$\gamma_i = \frac{1}{n} \sum_{k=1}^{n} \xi_i(k), \tag{3}$$

where n is the number of the objective function or the reference sequence, $x_0^*(k)$.

The grey relational grade γ_i represents the level of correlation between the reference sequence and the comparability sequence.

3.2.2 Application of grey relational analysis

Tab. IV and V illustrate the affecting factors that yield by applying GRA, which has the greatest influence for annual grain crop yield and daily close price for KLSE. Based on the calculated values of the grey relational grade, only six factors; a, b, c, e, h and l are selected as the inputs to ANN to predict the grain crop yield. This result is alike to the previous study by Zhang and He [53]. While, for KLSE daily close price, out of 14 affecting factors being observed, only four factors were identified as the most influential factors; $SI, TSI, CmpI$ and II. Therefore, these four factors are used as the inputs to ANN to predict the next day close price for KLSE.

On the other hand, in the multiple regression analysis, goodness fit test is used to recognize the proper inputs. The significant inputs are identified based on t-values and p-values. If the t-values are less than 1 and the p-values are above some accepted level, such as 0.05, then this variable is excluded from the list.

3.2.3 Nonlinearity test: McLeod and Li test

Nonlinearity test is implemented to examine the degree of linearity of time series data in our study. The McLeod and Li Test [34] is based on the autocorrelations of the squared residuals produced by the ARIMA model using the n observations as shown below:

$$Q = n(n+2)\sum_{i=1}^{q}\frac{r^2(i)}{n-i}, \tag{4}$$

where
$r(i)$ an autocorrelation of the squared residual, and
n is the sample size.

Year	a	b	c	e	h	l	d
1990	1.0000	1.0000	1.0000	1.0000	0.0324	1.0000	0.8578
1991	0.9767	0.9448	0.9397	0.8772	0.1488	0.9525	1.0000
1992	0.9452	0.8779	0.8293	0.8057	0.3261	0.8947	0.9043
1993	0.8936	0.8137	0.8094	.6789	0.3312	0.8098	0.7247
1994	0.8257	0.7071	0.8049	0.5840	0.5298	0.5823	0.8726
1995	0.7464	0.6225	0.7298	0.4263	0.3766	0.3990	0.5932
1996	0.6633	0.5494	0.5716	0.2924	0.1252	0.3073	0.1008
1997	0.5446	0.4715	0.4483	0.2051	0.0884	0.2156	0.2354
1998	0.4354	0.4427	0.2962	0.1462	0.0000	0.1528	0.0000
2000	0.3057	0.3816	0.1721	0.1230	0.0633	0.0000	0.0508
2001	0.1833	0.2662	0.0769	0.1103	0.5380	0.0713	0.6509
2002	0.0944	0.1780	0.0152	0.0489	0.7788	0.0798	0.7747
2003	0.0000	0.0000	0.0000	0.0000	1.0000	0.0170	0.7173

Tab. IV *Affecting factors for grain crop yield selected by GRA.*

Date	SI	TSI	$CmpI$	II	$Close\_KLSE$
4-Jan-05	0.3298	0.5428	0.5454	0.7306	0.5454
5-Jan-05	0.2653	0.4615	0.4858	0.7114	0.4858
6-Jan-05	0.2339	0.4254	0.4591	0.6456	0.4591
7-Jan-05	0.1694	0.3675	0.3953	0.5614	0.3953
10-Jan-05	0.1306	0.3428	0.3654	0.5279	0.3654
11-Jan-05	0.0289	0.2360	0.2391	0.4316	0.2391
12-Jan-05	0.0124	0.2085	0.2097	0.3100	0.2097
13-Jan-05	0.0000	0.1562	0.2013	0.2898	0.2013
14-Jan-05	0.0562	0.2495	0.1857	0.2354	0.2487
17-Jan-05	0.0347	0.2042	0.2213	0.2795	0.2213
18-Jan-05	0.0008	0.1640	0.1636	0.1793	0.1636
19-Jan-05	0.0562	0.2495	0.1857	0.2354	0.1857
20-Jan-05	0.1149	0.2961	0.2490	0.3423	0.2490
24-Jan-05	0.1587	0.3760	0.3185	0.3191	0.3185
25-Jan-05	0.2165	0.4686	0.3575	0.4113	0.3575
26-Jan-05	0.1678	0.3958	0.3181	0.4157	0.3181

Tab. V *Affecting factors for KLSE close price selected by GRA.*

The null hypothesis of linearity, the above statistic is asymptotically distributed by χ^2_q, where q is the number of autocorrelations (readers may consult [34] for details).

3.2.4 Performance measurement

To evaluate the performance of the proposed hybrid model; GRANN-ARIMA, four statistical tests are carried out. These tests are Root Mean Square Error (RMSE), Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Deviation (MAD). Following are mathematical formulas for each test:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(observed_t - predicted_t)^2}. \quad (5)$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(observed_t - predicted_t)^2 \quad (6)$$

$$MAPE = \sum_{t=1}^{n}\left|\frac{observed_t - predicted_t}{observed_t}\right| \quad X \ \tfrac{100}{n} \quad (7)$$

$$MAD = \sum_{t=1}^{n}\frac{|observed_t - predicted_t|}{n} \quad (8)$$

where n is the number of forecasting periods, $observed_t$ is the actual time series values and $predicted_t$ is the forecasting time series values.

GRANN-ARIMA is the best alternative model for forecasting multivariate time series data if it gives the lowest values for RMSE, MSE, MAD and MAPE compared to other models. The RMSE, MSE, MAD and MAPE are calculated based on the out-of-sample data.

4. Proposed Hybrid Method

Most of the real world problems consist of linear and nonlinear patterns. Even though, there are a lot of methods that can be applied to solve time series forecasting problems, but none of them can handle both patterns simultaneously. To tackle these two patterns uniformly well, hybridizing the linear and nonlinear model is proposed to improve the forecasting accuracy.

4.1 Conventional hybrid method for univariate time series data

In the literature, several works related to hybrid nonlinear and linear models (ARIMA model as linear model and ANN as an nonlinear model) to forecast univariate time series data could be found [28, 44, 49, 51, 54]. Authors applied the ARIMA model initially to the data and followed by ANN for data residual. This hybrid model, (Y_t), may be illustrated as:

$$\begin{aligned} Y_t &= L_t + N_t \\ &= ARIMA + ANN \end{aligned} \tag{9}$$

where L_t and N_t are linear and nonlinear components of the hybrid model Y_t using univariate time series data. Here, we used subscript t and tm to represent univariate time series and multivariate time series respectively.

4.2 Hybrid model for multivariate time series data using conventional approach

In this hybridization, Multiple linear regression (MR) is used to represent the linear model, L_{tm} and ANN represents nonlinear model, N_{tm} in handling multivariate time series data. The proposed Hybrid I model $\left(Y_{tm}^1\right)$ is given as,

$$\begin{aligned} Y_{tm}^1 &= L_{tm} + N_{tm} \\ &= MR + ANN. \end{aligned} \tag{10}$$

Let $\hat{L}_{tm}$ is the forecast value of the MR model at time t, and e_{tm}^1 represents the residual at time t as obtained from the MR model, then:

$$e_{tm}^1 = Y_{tm}^1 - \hat{L}_{tm}. \tag{11}$$

These residuals, e_{tm}^1 represent the nonlinear component of multivariate time series data. Therefore, ANN are used to model e_{tm}^1 and can be represented as follows:

$$e_{tm}^1 = f(\varepsilon_{tm-1}, \varepsilon_{tm-2}, \ldots, \varepsilon_{tm-n}, w) + \Delta_{tm}, \tag{12}$$

where f is a nonlinear function determined by the ANN structures along with connection weights, and Δ_{tm} is the random error.

Hence, the hybridized forecast given by Hybrid I model,

$$\overset{\frown}{Y}{}_{tm}^{1} = \hat{L}_{tm} + \hat{N}_{tm}, \tag{13}$$

where $\hat{N}_{tm}$ is the forecast value of Equation (12).

Consequently, Hybrid I model can be determined by two steps. First MR model is used to analyze the linear part of the multivariate time series problem. After that, ANN model is build to model the residuals produce by linear model, MR.

4.3 Proposed Hybrid II model

In the proposed method, ARIMA is used as a linear model, L_{tm} and GRANN is used as a nonlinear model, N_{tm}. Both models are hybridized as shown by Equation (14);

$$\begin{aligned} Y_{tm}^{II} &= N_{tm} + L_{tm} \\ &= GRANN + ARIMA, \end{aligned} \tag{14}$$

where Y_{tm}^{II} denoted the Hybrid II model that composed of nonlinear GRANN model and linear ARIMA model. Assume that $\widehat{N}_{tm}$ is the forecast value of the GRANN model at time t, and let e_{tm}^{II} represent the residuals at time t as obtained from the GRANN model, then

$$e_{tm}^{II} = Y_{tm}^{II} - \widehat{N}_{tm}. \tag{15}$$

The residuals now represent the linear part of the data, which enables us to employ ARIMA to model the residual and can be represented as below:

$$e_{tm}^{II} = f(e_{tm-1}, e_{tm-2,\dots,} e_{tm-n}) + \delta_{tm}, \tag{16}$$

where f is a linear function modeled by the ARIMA model, and δ_{tm} is the random error.

Therefore, the hybridized forecast obtained from Hybrid II, ($\hat{Y}_{tm}^{II}$) model can be written as

$$\hat{Y}_{tm}^{II} = \hat{N}_{tm} + \hat{L}_{tm}, \tag{17}$$

where $\hat{L}_{tm}$ is the forecast value obtained from Equation (16) above.

Similar to Hybrid I model, Hybrid II model can be determined through two steps. But both hybrid models are differs in terms of the hybridization execution where in Hybrid II model, nonlinear model is implemented first than followed by the linear model.

5. Experimental Results

Discussions for the results of this study are divided into two parts; Part 1 and Part 2. Part 1 discusses the result that is produced by Experiment I, and Part 2 discusses the result from Experiment II.

5.1 Results from Experiment I

As we mentioned earlier, the aim of this experiment is to investigate the capability of GRANN model in analyzing multivariate time series in searching the relationships between the independent and dependent variables. Before the modeling process is done using ANN, grey relational analysis is employed to obtain the significant affecting factors that affect the production of the crop yield in China and the KLSE close price.

Tab. VI depicts the structure and learning parameter used in developing GRANN and the error produced in training and testing phase for both data sample. For example, based on the calculated value of grey relational grade, only six factors; a, b, c, e, h, and l are selected as the inputs to grey relational artificial neural network (GRANN) to predict the grain crop yield. Thus, a three-layer feedforward neural network with a single output unit, 12 hidden units and 6 input units are used in this study with the learning rate (α) and momentum (β) are (0.5, 0.9) respectively. The network structures and learning parameters are determined by trial and error. In this study, we only consider the situation of one-step-ahead forecasting. Therefore, only one output node is employed. The RMSE for the best GRANN model are 232 for the training phase and 417 for the testing phase.

	Crop Yield	KLSE close price
ANN structure		
Number of input nodes (affecting factors determined by GRA)	Six nodes (a, b, c, e, h, l)	Four nodes ($CptI, TSI, SI, II$)
Number of hidden nodes	Twelve nodes	Nine nodes
Number of output nodes	One node	One node
Learning parameter		
Learning rate (α)	0.5	0.5
Momentum (β)	0.9	0.9

Tab. VI *Grey relational artificial neural network (GRANN) structure.*

Ten independent variables are used to build a multi regressions model for grain crop yield. Several models are built and evaluated based on statistical goodness fit. However, the final model only used 4 independent variables; a, c, h and n.

The equation for the grain crop yield is given as:

$$\begin{aligned} Crop_{yield_t} &= f(a_t, c_t, h_t, n_t) \\ &= -2.45a_t^{+}2.83c_t + 0.44h_t - 0.61n_t - 92298 \end{aligned} \tag{18}$$

where

a_t is a total power of agricultural at period t,
c_t is an irrigation area at period t,
h_t is a sown area of grain crops at period t, and
n_t is an agriculture labors at period t.

For KLSE close price, 14 variables are used initially; but only three variables are used finally; finance index (FI), trading/service index (TSI) and composite index ($CptI$). The equation for KLSE close price is shown below:

$$Close_{KLSE_t} = 2.20402 - 0.05FI_t - 0.03TSI_t + 1.05CptI_t \tag{19}$$

where

FI_t is the value of finance index at period t,
TSI_t is the value of trading/service index at period t, and
$CptI_t$ is the value of composite index at period t.

Various statistical tests can be used to validate the models. In this study, R^2, adjusted R^2, standard error, S_e, F-test and p-value are used to validate the model. R^2 and adjusted R^2 are the square of the correlation between the observed values of the response variable and the fitted values from the regression equation. Therefore, they are used to indicate the robustness of the model in explaining the actual consumption of the data. Both models are tolerable since R^2 and adjusted R^2 have high values with small standard error, S_e (refer to Tab. VII). In the interim, the F-test and p-values are used to determine the importance of the model. From Tab. VII, the F-test is considered significant since the p-value is approaching to zero. Therefore, Equation (18) and (19) is considered reasonable enough for predicting grain crop yield.

	R^2	adjusted R^2	S_e	F-values	P-values
Crop Yield	0.9760	0.9590	584.8000	60.1200	0.00006
Close_KLSE	0.9998	0.9996	42257.0000	254400.0000	0.00001

Tab. VII *Statistical test for the Grain crop and KLSE close price.*

Comparing the input parameter used by GRANN and MR in the forecasting crop yield, it is found that three variables selected by both models are equal, thus indicating the importance of these variables $\{a, c,$ and $h\}$. The result shows that the total power of agricultural (a), irrigation area (c) and sown area of grain crops (h) are the most affecting factors to grain crop yield in China since they are preferred by both MLR and GRANN. While, composite index ($CptI$) and trading/service index (TSI) are the most influential factors that affect the movement of KLSE close price.

5.2 Comparison between GRANN model and MR model

In order to examine the performance of GRANN in forecasting multivariate time series, the result from the multiple regression (MR) model and the GRANN model are compared. RMSE, MSE, MAD and MAPE are used to observe the forecasting performance between the GRANN and MR models. Tab. VIII gives the performance measure of the GRANN models and the MR models, and the prediction outputs of each model are shown in Fig. 3(a) and (b). Tab. VIII shows the predicted error values of the crop grain yield for the next two years (2002 and 2003) given by the GRANN model and MR. The application of the GRANN model gives the smallest error (RMSE, MAPE, MSE, MAD) compared to the MR model.

Error	Crop yield		KLSE close price	
	MR	GRANN	MR	GRANN
MSE	3207537.00	174165.00	2245.37	540.89
RMSE	1790.96	417.33	12.66	6.22
MAD	1321.50	369.45	12.67	6.17
MAPE	2.89	0.81	1.37	0.67

Tab. VIII *Forecasting values of GRANN versus MR.*

Fig. 3(a) shows an increasing production of the grain crop in 2003. The output from the GRANN model shows that there is a slight increment (0.1%) for the grain crop yield in 2003. However, the MR model predicts the production of the grain crop is going to be decreased about 4.8% in 2003. The actual values shows about 0.9% increment in 2003 for the crop grain yield. Hence, we can conclude that the GRANN forecasting result is more reliable than MR.

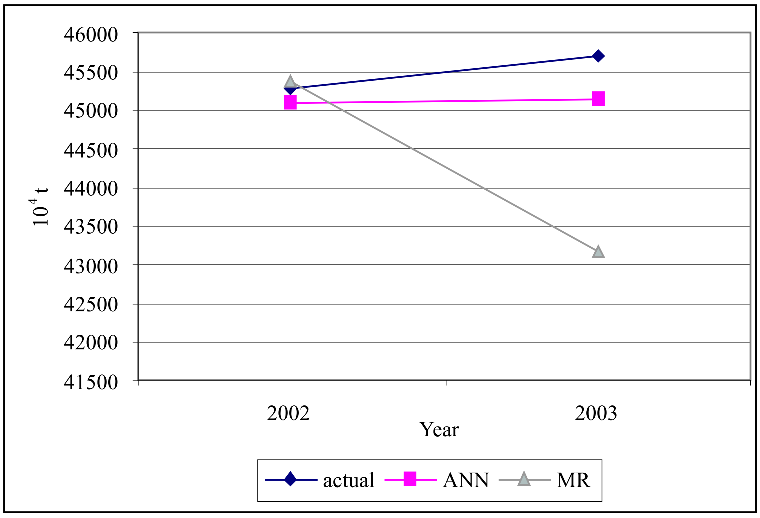


Fig. 3 *(a) Forecasting values for each model (Crop yield).*

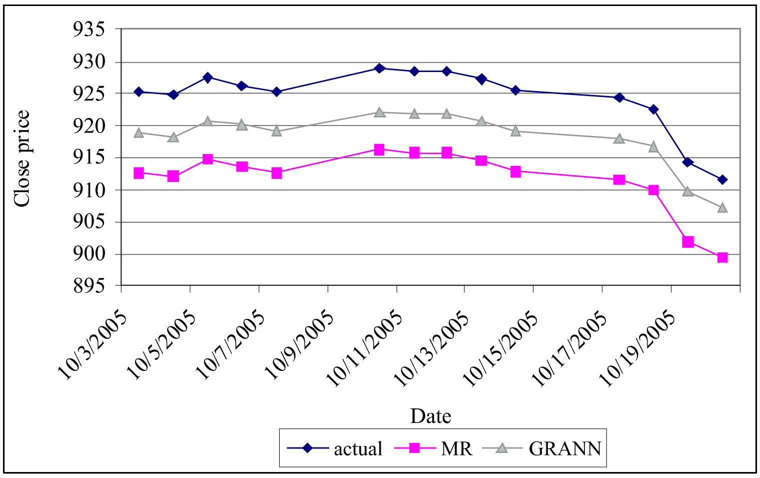


Fig. 3 *(b) Forecasting values for each model (KLSE close price).*

Tab. VIII depicts the forecasting error generated by GRANN and MR in forecasting KLSE daily close price for the next 14 days. The values of RMSE, MAPE, MSE, MAD given by GRANN is smaller than MR. The forecasting values given by GRANN is closer to the actual value as illustrated in Fig. 3(b). This indicates that the forecasting results yield from GRANN are more accurate compared to MR.

From the study, we found that GRANN performs better than MR, because ANN provides superior methodology than multivariate analytical [3, 38, 40]. This is conformed with the results of the previous study that ANN is suitable for the multivariate data analysis [39]. This result also strengthened our justification of implementing ANN as multivariate model in our study.

5.3 Results of Experiment II

The second experiment is conducted with two main objectives. The first objective is to compare the performance of individual and proposed hybrid model in forecasting time series. In order to achieve this objective, the proposed hybrid model that consists of nonlinear and linear models is developed. [42] proposed hybrid linear model known as MARMA to predict natural rubber price. In their study, they applied the MR and ARMA method to model the residual. However, in our study, we developed two hybrid models; Hybrid I and Hybrid II. Both methods pooled linear and nonlinear models using multivariate time series data. In Hybrid I, MR is used as a multivariate tool in conjunction with ANN to model the residual; in Hybrid II, GRANN is employed as a multivariate tool in cooperation with ARIMA to model the residual. MR and GRANN models are used as a yardstick to measure the performance of the proposed hybrid model respectively.

The second objective is to investigate the stream of implementation that can affect the performance of the proposed model in forecasting multivariate time series data. Previous studies by [28, 44, 49, 51, 54] have employed the linear model originally to the data, and followed by the nonlinear model (ANN) to predict the residual. In this study, the sequence is changing, i.e., the nonlinear model primarily, and followed by the linear model for residual. Time series forecasting can be obtained by integrating the values from linear and nonlinear models.

5.3.1 Experimental results for Hybrid I model (MR_ANN)

Results obtained from this test verify that the data are nonlinear and ANN is suitable to model the residual. Tab. IX shows the predicted residual value from the ANN model. The network structure for the crop yield is 2-4-2; 2 inputs unit, 4 hidden units and 2 output units respectively. The 5-11-4 is the network structure used for the KLSE close price. The learning and momentum rates are set between [0.5, 0.9].

The predicted value from the MR model (from Experiment I) is integrated with a predicted value from the ANN model to get the ultimate forecasting values for the China total grain crop yield in year 2002 and 2003, and 14 days ahead for the KLSE close price.

Tab. IX shows that the RMSE, MSE, MAPE and MAD for the Hybrid I model, which combines MR and ANN models have the lowest value compare to the MR model alone.

Hybrid I Error	Crop Yield		KLSE close price	
	MR	MR_ANN	MR	MR_ANN
RMSE	1790.00	1290.00	12.66	12.28
MSE	3207537.00	1664384.00	2245.37	2112.34
MAD	1322.00	1279.00	12.67	12.27
MAPE	2.89	2.81	1.37	1.33

Tab. IX *Results for Hybrid I Model.*

Based on RMSE, the prediction accuracy in the Hybrid I model increases about 28% and 30% for the crop yield and KLSE close price. As a result the combination of 2 different models that have dissimilar characteristics (linear and nonlinear) can improve the forecasting accuracy in multivariate time series analysis.

5.3.2 Result from proposed Hybrid II

In this experiment, we alter the sequence order of implementation proposed by previous researchers, in which a nonlinear model is employed initially and followed by a linear model for the residual. However, in this study GRANN is used as a nonlinear model and ARIMA as a linear model. McLeod and Li test is employed to check the nonlinearity degree of the data. The results from this test verify that the residuals data are linear; hence, ARIMA is suitable to model the residual. ARIMA (1,0,1) is used to model the residual and the predicted residual for 2 years ahead and ARIMA (0,1,3) is used to model the residual daily KLSE close price (Tab. X).

The predicted value from the GRANN model (from Experiment I) is integrated with predicted value from the ARIMA model to get the final forecasting value for both samples. Tab. X shows the result for the proposed Hybrid II models which gives a smaller prediction error (GRANN_ARIMA). This result indicates that the forecasting ability of the multivariate time series data is improved even further if the GRANN_ARIMA model is adopted. The forecasting accuracy of GRANN_ARIMA in both samples is increased for about 33% and 73% compared to the prediction given by the GRANN model.

Hybrid II Error	Crop Yield		KLSE close price	
	GRANN	GRANN_ARIMA	GRANN	GRANN_ARIMA
RMSE	417.33	278.19	6.22	1.67
MSE	174165.20	154777.00	540.89	39.33
MAD	369.45	212.49	6.17	1.52
MAPE	0.81	0.46	0.67	0.16

Tab. X *Results for Proposed Hybrid II.*

The results of these experiments have proven that the performance of the proposed model is better to compare the individual model in analyzing and forecasting multivariate time series data, thus conform to the previous results obtained from Hybrid I.

6. A Comparison of Hybrid I, Hybrid II and Conventional Hybrid models

Experimental result yield from this study reveals that both the blend methods of Hybrid I and Hybrid II can improve the forecasting accuracy. Nonetheless, the arising issue is to choose which one is better, Hybrid I or Hybrid II. Consequently, a comparison of the forecasting performance for both hybrid models (Hybrid I and Hybrid II) is given in this section to examine the effect of forecasting performance by varying the order of implementation on multivariate time series data.

As depicted in Tab. XI, the RMSE, MSE, MAPE and MAD given by the Hybrid II model is lower than Hybrid I with the forecasting accuracy increases 78% and 86% in each sample data respectively. These results show that the best method for combining models is Hybrid II. At the same time, this result indicates that altering the sequence of hybridization will affect the accuracy. For illustration, the error given by the proposed approach, GRANN_ARIMA is better compared to MR_ANN with a conventional hybridization approach.

To further evaluate the performance of the GRANN_ARIMA model, a comparison with conventional hybrid model (ARIMA_ANN) that proposed by the previous studies are carried out. The ARIMA_ANN model cannot be developed for crop yield data since the sample size is too small and insufficient for developing the ARIMA model. From the experiment, we found that the forecasting performance improves better if Hybrid II is used instead of using conventional hybrid. This shows that the type of data being used will also influence the forecasting performance. The more relevant data being considered in the experiment, the better the performance of the forecasting model; GRANN_ARIMA used multivariate time series instead of univariate time series used in ARIMA_ANN.

Fig. 4(a) and 4(b) summarize the results of examining the affect of altering the sequence of hybridization. It is clear that the GRANN_ARIMA performs better than the other two hybrid models. As shown in Fig. 4(a) and 4(b) the forecasting value from hybrid GRANN_ARIMA is more accurate compared to MR_ANN and ARIMA_ANN since the values are approaching to the actual value. The value of R^2, adjusted R^2 in MR and the value of Ljung-box test in ARIMA are high since these values are based on in-sample data. Therefore, there is no guarantee that the model can also give a high performance for out sample data.

Tab. XI showed that the RMSE, MAD, MAPE and MSE that are calculated based on out of sample data for MR_ANN or ARIMA_ANN is worse compare GRANN_ARIMA in both samples data used. The difference performance between in sample and out sample indicates that over fitting exists. For example, over fitting occurs when determining the appropriate parameters to be included in the MR model. Goodness fit test, which used to check the adequacy of the MR model, has probably excluded the significant factors that should be considered.

According to [54], we need to apply a linear model first to avoid over-fitting in artificial neural network. However, the result from this study shows that the over-fitting also occurred in the linear model. This result supports the findings given by [24]. Therefore, we conclude that the over-fitting can also occur in the linear model. Based on the MR and ARIMA performance, the probability of the over-fitting arises in the linear model but higher compared is the nonlinear model.

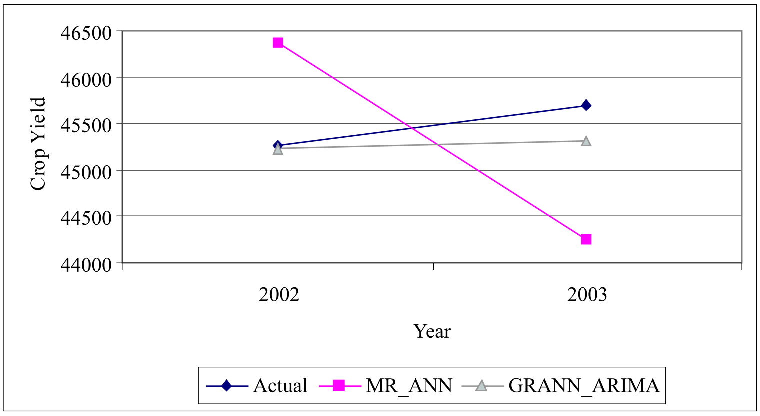


Fig. 4 *(a) Comparison of Hybrid I and Hybrid II (Crop yield).*

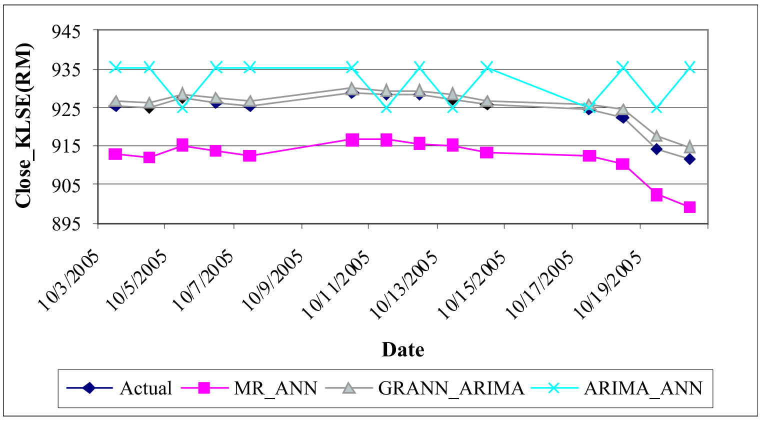


Fig. 4 *(b) Comparison of Hybrid I and Hybrid II (KLSE close price).*

This result supports our claims to change the sequence of hybridization for better forecasting.

To obtain a more accurate forecasting result, Hybrid II methods using the GRANN_ARIMA model is suggested because it can work well in both sample data that represent small scale and large scale data. This result may be explained by the fact that: i) ANN is capable to deal equally well with linear and nonlinear data; (ii) ANN is accepted as universal approximator; (iii) furthermore, past studies have shown that over fitting problem in ANN can be avoided by using the cross validation and optimal learning parameters [29].

Error	Crop Yield		KLSE close price		
	Hybrid I (MR_ANN)	Hybrid II (GRANN_ARIMA)	Hybrid I (MR_ANN)	Hybrid II (GRANN_ARIMA)	Conventional Hybrid (ARIMA_ANN)
RMSE	1290.00	278.19	12.28	1.67	10.12
MSE	1664384.00	154777.00	2112.34	39.33	1434.52
MAPE	2.81	0.47	1.33	0.16	0.92
MAD	1279.00	212.49	12.27	1.52	8.48

Tab. XI *Comparison performance of Hybrid I, Hybrid II and conventional hybrid model.*

Error	Crop Yield			KLSE close price		
	GRANN_ARIMA	MARMA	ARIMA	GRANN_ARIMA	MARMA	ARIMA
RMSE	278.19	1742.51	na	1.67	12.68	6.33
MSE	154777.00	3036334.00	na	39.33	2252.32	560.45
MAPE	0.47	2.53	na	0.16	1.37	0.44
MAD	212.49	1158.12	na	1.52	12.67	4.07

** na: not applicable: model cannot be developed

Tab. XII *Comparison of GRANN_ARIMA with MARMA and ARIMA.*

6.1 Comparison of GRANN_ARIMA (Hybrid II) model with benchmark model

From the previous experiment, we found that GRANN_ARIMA is the best model for forecasting multivariate time series data. To further validate our findings, assessment performance with the benchmark model for univariate and multivariate model are conducted. In this study, MARMA and ARIMA models are selected because they are linear statistical models. Furthermore, ARIMA or known as Box-Jenkins model has dominated time series forecasting for more than half a century. Additionally, ARIMA modeling has been used in univariate framework as a sophisticated benchmark for evaluating alternative proposal [1, 48, 54]. MARMA is used in this study because it has been used as a statistical modeling technique for the hybrid model in previous study [42, 43]. Tab. XII shows the performance of each model. We found that accuracy of GRANN_ARIMA is always better than the MARMA and ARIMA models. This result is not surprising because ARIMA is a linear model and MARMA model is a combination of two linear models. Due to this reason, it cannot handle nonlinear data as well as the nonlinear model.

Tab. XII depicts that the comparison with the individual linear ARIMA cannot be implemented for crop yield data. Due to small sample data set being used and the non stationary time series data, the ARIMA model cannot be developed. Theoretically, the minimum of data that need to apply the ARIMA model is about 40 to 50 periods of data. This model needs more data for tracking the pattern or component in time series data prior to the modeling process. Before the model estimation can be done, time series data must be in a stationary form. Otherwise, the differencing process need to be implemented and it will reduce the size of the data. In this study, the data is an annual data with approximately 11 periods, and non stationary. These data need to be transformed into a stationary form. The result shows that the propose GRANN_ARIMA model can also perform well in non-stationary and small size of time series data.

6.2 Comparison of the GRANN_ARIMA model with ANN using the second order error (Lavemberg Marquet)

Tab. XIII demonstrates the result given by the GRANN_ARIMA model and ANN model using the second order error; Lavemberg Marquet. In this study, the ANN model using Lavemberg Marque (LVM) is developed using the default parameters given by MATLAB. Tab. XIII shows the results of the experiments. For the first data set (crop yield - small scale data), the performance of LVM is slightly better than GRANN_ARIMA with 5% growth. On the other hand, GRANN_ARIMA performs much better than LVM in the second data set (KLSE close price which represents large scale data) with more than 80% growth.

6.3 Summary of the experimental result

In this section, a summary of the experimental result that has been conducted in this study is presented. Tab. XIV(a) and XIV(b) depict the errors generated by the individual model and the proposed Hybrid I, Hybrid II model, Conventional hybrid model, Benchmark model and ANN with the second order error. ARIMA

Error	Crop Yield		KLSE close price	
	GRANN_ARIMA	LVM	GRANN_ARIMA	LVM
RMSE	278.19	260.79	1.67	8.51
MSE	154777.009	125791.91	39.33	1014.45
MAD	212.499	210.65	1.52	6.27
MAPE	0.47	0.46	0.16	0.68

Tab. XIII *Comparison of GRANN_ARIMA with Lavenberg Marquet (LVM).*

and MARMA are used as yardstick models which represent traditional statistical individual and hybrid model accordingly. RMSE, MSE, MAD, and MAPE are used to quantify of the difference between the actual output and the predicted output given by each model.

Tab. XIV(a) shows that the proposed hybrid models always give better results compared to an individual model regardless of the sequence of the proposed hybrid model. For instance, an error produced by an individual model is greater than an error produced by the proposed hybrid models (MR and MR _ANN; GRANN and GRANN_ARIMA) except for the ANN trained with the second order error (LVM).

As shown in Tab. XIV(a), the difference of RMSE, MSE, MAPE and MAD produced from the proposed Hybrid II model that comprises of GRANN and ARIMA and LVM are small, where the difference of accuracy percentage is 0.02, almost approaching to zero. As a consequence, we can conclude that both of them are comparable.

Fig. 5(a) shows that the proposed Hybrid II (GRANN_ARIMA), LVM and GRANN models are able to give prediction values that are closed to actual value compared to Hybrid I, MARMA and MR models. Moreover, they can also predict the behavior of the data better than MR, MARMA and Hybrid I model. They manage to predict the increment pattern in the grain crop in 2003. However, Hybrid I, MARMA and MR model failed to predict this pattern. Fig. 5(a) shows an increasing production of the grain crop in 2003. The output of GRANN, GRANN_ARIMA and LVM models show that there are increments about 0.1%, 0.2% and 0.04% respectively for the grain crop yield in 2003. The actual values show about 0.9% increment in 2003 for the crop grain yield. Hence, we can say that the prediction given by our proposed hybrid model, GRANN_ARIMA, is more significant and reliable since it manages to forecast the augmentation more proper compared to LVM and GRANN than GRANN and LVM.

In view, of GRANN_ARIMA is ranked the first, followed by the LVM and GRANN. The hybrid model that used Hybrid I approach is ranked as the fourth, and followed by MARMA and finally followed by the MR model. Based on the results obtained from the experiment, we conclude that the best predictor for the grain crop yield for China is the proposed Hybrid II model that consists of GRANN and ARIMA models.

Tab. XIV(b) shows the statistical test results and accuracy percentage obtained from each individual and hybrid model used in modeling forecasting model to predict daily close price for KLSE. From the Table, it shows that RMSE, MSE, MAPE, MAD for the Hybrid II model, which represents the GRANN_ARIMA,

	Individual	Model	Conventional Hybrid	Hybrid I	Hybrid II	Benchmark	Model	Second order error
Model	MR	GRANN	ARIMA_ANN	MR_ANN	GRANN_ ARIMA	MARMA	ARIMA	LVM
RMSE	1790.96	417.33	-na	1290.11	278.19	1742.51	-na	260.79
MSE	3207537.00	174165.00	-na	1664384.00	77389.00	3036334.00	-na	125791.91
MAPE	2.89	0.81	-na	2.81	0.47	2.53	-na	0.46
MAD	1321.50	369.45	-na	1278.98	212.49	1158.12	-na	210.65
Accuracy	97%	99 %	-na	97.3%	99.5%	97.2%	-na	99.52%

** na: not applicable; model cannot be developed

Tab. XIV *(a) Statistical test and accuracy percentage for an individual model and a hybrid model (Crop Yield).*

	Individual	Model	Conventional Hybrid	Hybrid I	Hybrid II	Benchmark	Model	Second order error
Error	MR	GRANN	ARIMA_ANN	MR_ANN	GRANN_ ARIMA	MARMA	ARIMA	LVM
RMSE	12.66	6.22	10.12	12.28	1.67	12.68	6.33	8.51
MSE	2245.37	540.89	1434.52	2112.34	39.33	2252.32	560.45	1014.45
MAPE	1.37	0.67	0.92	1.329	0.16	1.37	0.44	0.68
0.68 MAD	12.67	6.17	8.48	12.27	1.52	12.67	4.07	6.27
Accuracy	98.60	99.33	99.08	98.67	99.84	98.62	99.56	99.32

Tab. XIV *(b) Statistical test and accuracy percentage for individual models and hybrid models (KLSE close price).*

is the lowest compared to the individual model or benchmark model or Hybrid I model, conventional model and LVM.

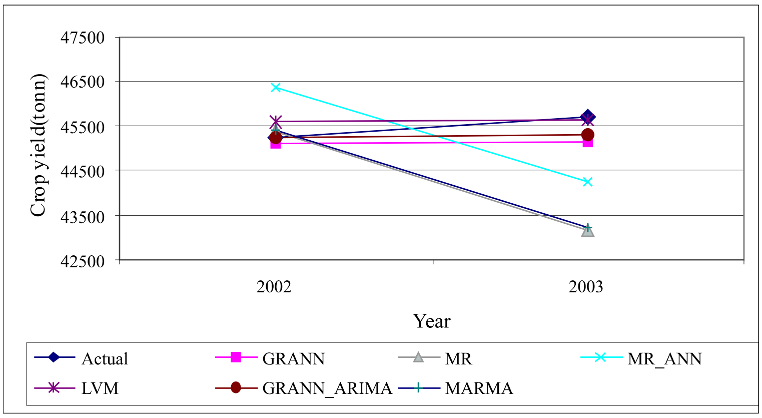


Fig. 5 *(a) Comparison of Forecasting Values Yield from Each Model (Crop yield).*

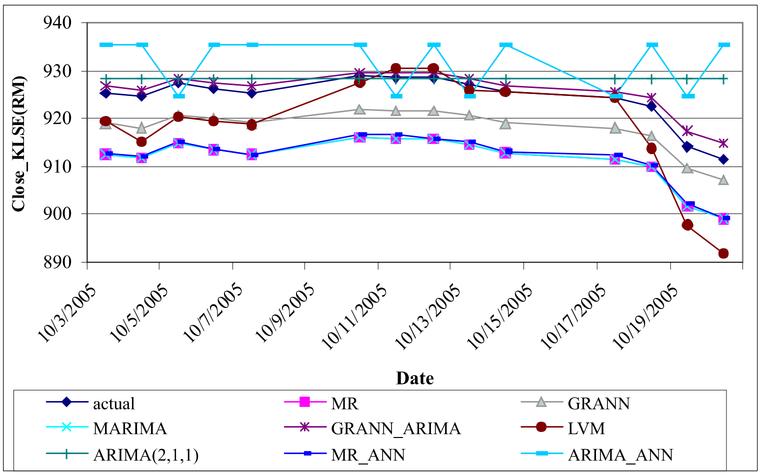


Fig. 5 *(b) Comparison of the Forecasting Values Yield from Each Model (KLSE close price).*

Unlike LVM, the forecasting accuracy percentage of the GRANN_ARIMA is increased for about 0.5% . The results of this study also are depicted that forecasting performance yield from the individual model (ARIMA) outperformed the conventional hybrid (ARIMA_ANN), and it conforms to the previous study. This discrepancy may be due to the insufficient data information, over the fitting problems in a linear model and redundancy information while modeling conventional hybrid model, and ARIMA_ANN. The outcome from this study also revealed that these discrepancies can be solved by using our proposed hybrid model, GRANN_ARIMA.

Fig. 5(b) summarizes the results given in Tab. XIV(b). It is clear that the

GRANN_ARIMA performs better than the other models. As shown in Fig. 5(b) the forecasting value from the hybrid GRANN_ARIMA is more accurate due to its behavior in approaching nearer to actual value.

From the result, we wind up that there are three factors that will influence the accuracy of the hybrid model. The first factor is the type of a model used in combination. As shown in Fig. 5(a) and 5(b), the performance of the hybrid model that consists of two dissimilar models which have different characteristics (linear and nonlinear models) gives better results compared to the hybrid model that combine both linear models. The second factor is the sequence of the implementation for hybridization. From Tab. XIV(b) and Fig. 5(b), it shows that by altering the conventional sequence of hybridization, the forecasting error is decreasing and at the same time the forecasting accuracy is increasing. The third factor is the type of the data being used; the forecasting performance is increasing when the multivariate data is used in modeling the time series data.

7. Conclusion

In this study, GRANN_ARIMA is proposed as a new approach for hybridizing linear and nonlinear models. Unlike conventional hybrid model, the proposed model has few integrated features such as engaged with multivariate time series data, GRA as feature selection to remove irrelevant input data and altering the sequence of hybridization. To verify the effectiveness of a proposed hybrid model, several comparisons have been conducted. The outcomes from the experiment revealed that:

- GRANN can perform better than the multiple regression in handling multivariate time series data due to ANN effectiveness in modeling and forecasting nonlinear time series with or without noise [9, 39, 50]. Therefore, ANN is used as multivariate modeler in our proposed model.
- Forecasting accuracy of the proposed hybrid model, GRANN_ARIMA is better compared to the individual model such as GRANN, ARIMA, MR and the second order error, LVM. This result supports the outcome from the previous studies because hybridizing two dissimilar models will reduce the forecasting error [53, 54].
- Altering the sequence of a hybrid model will improve the forecasting accuracy. The Hybrid model with changing sequence of hybridization (Hybrid II) is outperformed Hybrid I and conventional hybrid.
- Forecasting value of GRANN_ARIMA is more accurate compared to hybrid linear MARIMA model since it can handle equally well both linear and nonlinear patterns in time series data.
- The forecasting error produced by GRANN_ARIMA is the smallest compared to other models that are tested in this study.
- GRANN_ARIMA also performs well in both small scale data and large scale data.

In conclusion, the proposed approach for hybridizing linear and nonlinear models, GRANN_ARIMA can be used as an alternative tool for forecasting time series data for better forecasting accuracy. Prior studies concealed that ANN learning algorithm was time consuming and tended to trap into local minima solution. Therefore, more studies will be conducted on a new concept of ANN learning algorithm, i.e., a biologically inspired algorithm to speed up the learning time and the accuracy of our proposed hybrid model. In this study, the type of time series data is limited to complex time series; however, it can be extended to the simple and seasonal time series data. Meanwhile, compared with individual models, the modeling process of a hybrid model is slightly difficult and time consuming. However, due to rapid improvement of computer technologies, this limitation does not seem to be significant.

References

[1] Abraham A., Nath B.: A Neourofuzzy Approach For Modeling Electricity Demand In Victoria. Applied Soft Computing Journal, **1**, 2001, pp. 127-138.

[2] Amstrong J. S.: Long Range Forecasting: From Crystal Ball to Computer (2nd Ed.), Wiley, New York, 2001.

[3] Ali J., Al-Mahmeed M.: Neural Networks: Predicting Kuwaiti KD Currency Exchange Rates Versus US Dollar, Arab Journal of Administrative Sciences. **6**, pp. 17-35. In: C. P. Rao, Jafar Ali, Neural Network Model For Database Marketing in the New Global Economy, Marketing Intelligence & Planning, **20/1**, 2002, pp. 35-43.

[4] Bates J. M., Granger C. W. J.: The Combination of Forecasts. Operational Research Quarterly, **20**, 1969, pp. 451-468.

[5] Bessaler D. A., Brandt J. A.: Forecasting Livestocks Prices with Individual and Composite Methods, Applied Economics, **13**, 1981, pp. 513-522.

[6] Box, GEP, Jenkins G.: Some Practical Aspects of Forecasting in Organization, Journal. Forecasting, **I**, 1982, pp. 3-21.

[7] Brace M. C., Schmidt J., Hadlin M.: Comparison of The Forecasting Accuracy of Neural Networks With Other Established Techniques. In: Proceedings of the First Forum On Application of Neural Network to Power Systems, Seattle, WA., 1991, pp. 31-35.

[8] Caire P., Hatabian G., Muller C.: Progress in Forecasting by Neural Networks, In: Proceedings of The International Joint Conference on Neural Networks, **2**, 1992, pp. 540-545.

[9] Chi F-L.: Forecasting Tourism; A Combined Approach, Tourism Management, **19**, 6, 1998, pp. 515-520.

[10] Chaivivatrakul S., Somhom S.: Vegetable Product Forecasting System by Adaptive Genetic Algorithm, TENCON 2004. 2004 IEEE Region 10 Conference, **B**, 21-24 Nov. 2004, pp. 211-214.

[11] Chatfield C.: Time Series Forecasting. Bocaraton, FL: Chapman & Hall/CRC, 2001.

[12] Chu C. W., Zhang G. P.: A Comparative Study of Linear and Nonlinear Models for Aggregates Retail Sales Forecasting. International Journal of Production Economics, **86**, 3, 2003, pp. 217-231.

[13] Deng J. L.: Control Problems of Grey Systems, Systems and Control Letters, **5**, 1982, pp. 288-294.

[14] Deng J. L.: Introduction to Grey System Theory, the Journal of Grey System, **1**, 1989, pp. 1-24.

[15] Delen D., Waller G., Kadam A.: Predicting Breast Cancer Survivability: A Comparison of Three Data Mining Methods, Artificial Intelligence in Medicine, **34**, 2005, pp 113-127.

[16] Denton J. W.: How Good are Neural Networks for Causal Forecasting? The Journal of Business Forecasting, **14**, 2, Summer, 1995, pp. 17-20.

[17] Dorganis P., Alexandidi A., Patrinos P., Sarinevers H.: Time Series Sales Forecasting for Short Shelf-Life Food Products Based on Artificial Neural Networks and Evolutionary Computing, Journal of Food Engineering, **75**, 2006, pp. 196-204.

[18] Ganeta R., Romeo L. M., Gill A.: Forecasting of Electricity Prices With Neural Networks. Energy Conversion and Management, **47**, 2006, pp. 1770-1778.

[19] Hippert H. S., Pedreira C. E., Souza R. C.: Neural Networks for short-term load forecasting: A review and evaluation. IEEE Transactions on Power Systems, **16**, 1, 2001, pp. 44-45.

[20] Gorr W. L., Nagin D., Szcypula J.: Comparative Study of Artificial Neural Network and Statistical Models for Predicting Student Grade Point Averages, International Journal of Forecasting, **10**, 1994, pp. 17-34.

[21] Granger C. W. J.: Strategies For Modeling Nonlinear Time Series Relationships. The Economic Record, **69**, 206, 1993, pp. 233-238. In: G. Zhang et al. (Eds.), Forecasting with Artificial Neural Networks: The State of Art, International Journal of Forecasting, **14**, 1998, pp. 35-62.

[22] Granger C. W. J., Terasvirta T.: Modeling Nonlinear Economic Relationships, Oxford University Press, Oxford. In: G. Zhang et al. (Eds.), Forecasting with Artificial Neural Networks: The State of Art, International Journal of Forecasting, **14**, 1998, pp. 35-62.

[23] Hamid S. A., Iqbal Z.: Using Neural Networks for Forecasting Volatility of S&P 500 Index Futures Prices, Journal of Business Research, **57**, 10, October 2004, pp. 1116-1123.

[24] Heravi S., Osborn D. R., Birchenhallc R.: Linear Versus Neural Network Forecasting for European Industrial Production Series. International Journal of Forecasting, **20**, 2004, pp. 435-446.

[25] Hou Z., Lian Z., Yao Y., Yuan X.: Cooling-Load Prediction by the Combination of Rough Set Theory and Artificial Neural Network Based on Data Fusion Technique, Applied Energy, **83**, 2006, pp. 1033-1046.

[26] Jain A., Kumar A. M.: Hybrid Neural Network Models for Hydrologic Time Series Forecasting, Applied Soft Computing, **7**, 2, March 2007, pp. 585-592.

[27] Kang S.: An Investigation of the Used of Feedforward Neural Networks for Forecasting. Phd Thesis. Kent State University, 1991.

[28] Lu J.-C., Niu D.-X., Jia Z.-L.: A Study of Short Term Load Forecasting Based on ARIMA-ANN, Proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai, August, 2004, pp. 26-29.

[29] Lachtermacher G., Fuller J. D.: Backpropagation in Time Series Forecasting, Journal of Forecasting, **14**, 1995, pp. 381-393.

[30] Li Y., Li A., Li R., Wei L., Qin S.: Study on Characteristics of Biomass Pyrolysis Gas in View of Grey Relation Analysis and BP Neural Network, Acta Energies Solaris Sinica, **24**, 6, Dec. 2003, pp. 776-780.

[31] Ma L., Khorasani K.: New Training Strategies for Constructive Neural Networks with Application to Regression Problem, Neural Network, **17**, 2004, pp. 589-609.

[32] Makridakis S., Wheelwright S. C.: Forecasting Methods For Management (5th Edition) New York: John Wiley and Sons, 1989.

[33] Makridakis S., Anderson A., Carbone R., Fildes R., Hibdon M., Lewandowski R., Newton J., Parzen E., Winkler R.: The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition, J. Forecasting, **1**, 1982, pp. 111-153.

[34] McLeod A. I., Li W. K.: Diagnostic Checking ARMA Time Series Model Using Squared Residual Autocorrelations Journal of Time Series Analysis, **4**, pp. 269-273. In: S. Heravi et al. (Eds.), Linear Versus Neural Network Forecasts for European Industrial Production Series, International Journal of Forecasting, **20**, 2004, pp. 435-446.

[35] Mentzer J. T., Bienstock C. C.: Sales Forecasting Management. Sage Publications, Thousand Oaks, CA, 1998.

[36] Newbold P., Granger C. W. J.: Experience with Forecasting Univariate Time Series and Combination of Forecast, Journal of the Royal Statistical Society Series, **A 137**, 1974, pp. 131-149.

[37] Pai P. F., Lim C.-S.: A Hybrid ARIMA and Support Vector Machines Model in Stock Price Forecasting, Omega, **33**, 2005, pp. 497-505.

[38] Parzen E., Winkler R.: The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition, J. Forecasting, **1**, 1982, pp. 111-153.

[39] Prerez S. M., Ortega M. H., Gonzalez M. L., Boger Z.: Comparative Study of Artificial Neural Network and Multivariate Methods to Classify Spanish DO Rose Wines, Talanta, **62**, 2004, pp. 983-990.

[40] Rao C. P., Ali J.: Neural Network Model for Database Marketing in the New Global Economy, Marketing Intelligence & Planning **20/1**, 2002, pp. 35-43.

[41] Ganeta R., Romeo L. M., Gill A.: Forecasting of Electricity Prices With Neural Networks. Energy Conversion and Management, **47**, 2006, pp. 1770-1778.

[42] Shamsuddin M. N.: A Short Note on Forecasting Natural Rubber Prices Using A MARMA. The Malaysian Journal of Agricultural Economic, **9**, 1992, pp. 59-68.

[43] Shamsuddin M. N., Arshad F. M.: Composite Model For Short Term Forecasting for Natural Rubber Prices. Pertanika, **12**, 2, 1990, pp. 283-288.

[44] Shouyang W., Lean W. Y., Lai. K. K.: Crude Oil Price Forecasting With Tei @Imethodology, Journal of Systems Science and Complexity, **18**, 2, April 2005.

[45] Srinivasulu S., Jain A.: A Comparative Analysis of Training Methods for ANN Rainfall-Runoff Models, Applied Soft Computing, **6**, 2006, pp. 295-306.

[46] Taskaya-Termizel T., Ahmad K.: Are ARIMA Neural Network Hybrids Better than Single Model? Proceeding of International Joint Conference on NNs (IJCN 2005), July 31-August 4, 2005, Montreal, Canada.

[47] Taskaya-Termizel T., Casey M. C.: A Comparative Study of Autoregressive Neural Network Hybrids, Neural Networks, **18**, 5-6, July-August 2005, pp. 781-789.

[48] Taylor J. W.: Short Term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing. Journal of the Operational Research Society, **54**, 2003, pp. 799-805.

[49] Tseng F.-M. Y. H.-C., Tzeng. G.-H.: Combining Neural Network Model With Seasonal Time Series ARIMA Model, Technological Forecasting and Social Change, **69**, 2002, pp. 71-87.

[50] Uysal M., Roubi M.: Artificial Neural Network Versus Multiple Regression in Tourism Demand Analysis, Journal of Travel Research, **38**, November, 1999, pp. 111-118.

[51] Voort V. D., Dougherty M., Watson M.: Combining Kohonen Maps with ARIMA Time Series Models to Forecast Traffic Flow, Transp. Resc. Circ. (Emerge Techno) **4C**, 5, 1996, pp. 307-308, **8**, 18-21 Aug. 2005, pp. 4764-4769.

[52] Yang H. Y, Ye H., Wang G., Khan J., Hu T.: Fuzzy Neural Very Short Term Load Forecasting Based on Chaotic Dynamics Reconstruction, Chaos, Solitons and Fractals, **29**, 2006, pp. 462-469.

[53] Zhang Y., He Y.: Study of Prediction Model on Grey Relational BP Neural Network Based on Rough Set Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21 August 2005.

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

[55] Zhang G. P., Patuwo B. E., Hu M. Y.: A Simulation Study of Artificial Neural Networks For Nonlinear Time-Series Forecasting, Computers & Operations Research, **28**, 2001, pp. 381-396.

[56] Zhang G. P., Patuwo B. E., Hu M. Y.: Forecasting with Artificial Neural Networks: The State of the Art, International Journal of Forecasting, **14**, 1, 1 March 1998, pp. 35-62.

[57] Zhang G. P.: Neural Networks in Bussiness Forecasting, IRM, Press, 2004.

[58] Abraham A.: Intelligent Systems: Architectures and Perspectives, Recent Advances in Intelligent Paradigms and Applications, Abraham A., Jain L. and Kacprzyk J. (Eds.), Studies in Fuzziness and Soft Computing, Springer Verlag, Germany, ISBN 3790815381, Chapter 1, 2002, pp. 1-35.

[59] Chebrolu S., Abraham A., Thomas J.: Feature Deduction and Ensemble Design of Intrusion Detection Systems, Computers and Security, Elsevier Science, **24/4**, 2005, pp. 295-307.

[60] Abraham A., Grosan C., Martin-Vide C.: Evolutionary Design of Intrusion Detection Programs, International Journal of Network Security, **4**, 3, 2007, pp. 328-339.

[61] Tosun N.: Determination of optimum parameters for multi-performance characteristics in drilling by using grey relational analysis, Int. J. Adv. Manuf. Technol, 2005.