FORECASTING SECTORAL INDICES IN THE KUALA LUMPUR STOCK EXCHANGE

[illegible] Leng

[illegible] Keng

ABSTRACT

This paper evaluates the use of (i) Box-Jenkins autoregressive-moving average model, (ii) vector autoregression model that incorporates short-run intersectoral relationship, and (iii) vector error correction model that incorporates long-run intersectoral relationship, for forecasting the daily Finance, Industrial, Plantation, Mining and Property Index of the Kuala Lumpur Stock Exchange. Given its dominance in explaining the behaviour of the stock prices, the random walk was used as a benchmark. Incorporation of the long-run equilibrium sectoral relationship was found to track rather closely the one-step-ahead forecasts of a random walk. The autoregressive-moving average model follows next, and the vector autoregression model has the poorest performance.

INTRODUCTION

The random walk is often used to explain stock market behaviour. Among others, evidence of random walk in the Malaysian stock market has been reported by Laurence (1986), Saw and Tan (1989), Mansor (1994), and Kok and Goh (1994a, 1994b, 1996). Of these, the studies by Saw and Tan (1989) and Kok and Goh (1994a, 1994b) further suggest that the movements of sectoral indices in the Kuala Lumpur Stock Exchange (KLSE) conform to a random walk.

Although evidence of random walk is over whelming, an issue remains to be investigated is the usefulness of the model for forecasting the Malaysian stock market performance. This paper seeks to evaluate its forecast performance and makes a comparison with that of other models. The outcome of such an evaluation bears not only practical importance for forecasting purposes, but is also pressing to answer a few challenges that arise from recent developments in the literature related to financial markets.

These challenges are of many facets. On modeling techniques, many univariate (e.g. Box-Jenkins ARIMA integrated autoregressive-moving average) representations) and multivariate time-series models (e.g. vector autoregressions and error correction mechanism) have been developed and they offer alternatives to the random walk. On empirical evidence, Meese and Rogoff (1983) showed that a random w model is difficult to beat in their assessment of the forecast performance of various exchange rate mod Chinn and Meese (1995), by imposing the long-run fundamentals using the error correction mod found that this model for exchange rate forecasts no better than a random walk for short-term predicti horizons, although it can explain exchange rate movements better in some cases for longer horizons. analytical work, Yoon (1998) proved that a random walk model can produce more accurate foreca than the true model (assumed known in this theoretical study) in the presence of a structural change

Given the dominance of a random walk, which is essentially a naïve model of no change, one has ponder upon its practical usefulness. The fact that a random walk premises on today's value of variable of interest as the best predictor for tomorrow's value, limits its practical significance considera If the random walk indeed has the best forecast performance, the immediate challenge is to find alternative that comes as close as possible to this performance. This forms the motivation to this pa

The focus of this paper is on five sectoral indices of the KLSE, namely, the Finance, Industrial, Planta Mining and Property Index. In this forecast performance evaluation exercise, the Box-Jenkins ARI models are considered besides the random walk for univariate modeling. Interestingly, Kok and G (1997) reported that the sectoral indices in KLSE are inter-related in the short run, and in addition, G (1999) found long-run relationship among the same indices. Information from such relationships ca exploited for forecasting the sectoral indices. With this motivation, the short-run relationship among five sectors is modeled using a vector autoregression (VAR). The vector error correction (VEC) appro is adopted to model the long-run dynamics of the sectoral index movements. The forecast performa of these univariate and multivariate models is compared to that of a random walk.

At the outset, the scope of this paper must be defined. The paper is confined to modeling the mean o process. The aim is to use the models for forecasting purposes, and not for analyzing the struc relationship of the sectoral behaviour. All models considered are linear in nature, and the non-li dynamics are not investigated. As the focus is to forecast the mean of the process, we leave the issu volatility modeling to future research.

This paper is organized as follows. After this introduction, Section 2 discusses the data and methodo employed in this study. The results and analysis are reported in Section 3. Section 4 presents s discussions to conclude this study.

DATA AND METHODOLOGY

This study uses the daily closing levels of the Finance Index, Industrial Index, Plantation Index, Mining Index and Property Index of the KLSE.[1] The period included for estimation is from 1 April 1993 to 30 June 1999. The out-of-sample forecast period is 1 July to 30 July 1999. The data are available from the local daily newspapers, and the Daily Diary and Investors' Digest published by the KLSE.

Unit Root Tests

Let X_{it} denote the logarithm of index for sector-i, where i = 1, 2, …, 5. We first establish the order of integration of X_{it} before proceeding to modeling the series. A widely applied procedure is the augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979). This involves testing for presence of a unit root in X_{it} in the following hypothesis

$$H_0: \alpha = 0 \quad \text{against} \quad H_1: \alpha < 0$$

in the equation

$$\Delta X_{it} = \mu + \beta t + \alpha X_{i,t-1} + \sum_{j=1}^{m} \theta_j \Delta X_{i,t-j} + \varepsilon_{it}, \quad t = 1, 2, \ldots, N, \qquad (1)$$

where Δ is the difference operator, t is the trend term, m is the number of lags included, N is the sample size and $\varepsilon_{it} \sim IN(0, \sigma^2)$. The Dickey-Fuller $t\alpha$ statistic is used and its empirical distribution is tabulated by MacKinnon (1991). If H_0 is not rejected, X_{it} is non-stationary and contains at least one unit root. The series will then have to be differenced before it is tested for stationarity again by repeating the same procedure, but with X_{it} and ΔX_{it} in equation (1) replaced with ΔX_{it} and $\Delta^2 X_{it}$, respectively.

The assumption of independently and identically distributed error term underlying the ADF test may not be true. The Phillips-Perron (Phillips and Perron, 1988) test of unit roots, which uses nonparametric corrections to improve on the ADF test statistic, is also employed.

Univariate Processes

A random walk process is defined as

$$X_{it} = X_{i,t-1} + \varepsilon_{it}. \qquad (2)$$

This model implies that the best predictor for the sectoral index tomorrow is its value today. Often, this is referred to as a naïve forecast of no change.

---

[1] The same sectoral indices were used in the studies by Kok and Goh (1994a, 1994b, 1997). Gui (Chapter 4, 1999) provides the explanation to the selection of these indices.

A random walk belongs to a wider class of the commonly used Box-Jenkins ARIMA models. To facilit the discussion that follows, we make use of the results that are presented in the next section that X_{it} integrated of order one, i.e., ΔX_{it} is stationary. The general autoregressive-moving average model order (p,q) for ΔX_{it} is represented by

$$\Delta X_{it} = \mu + \sum_{j=1}^{p} \phi_{ij}\, \Delta X_{i,t-j} + \sum_{k=1}^{q} \delta_{ik}\, \varepsilon_{i,t-k}\; \varepsilon_{it} \qquad (3)$$

and this model is denoted by ARMA(p,q).

2.3 Vector Autoregression (VAR) Model

A VAR of order p is used to model the short-run relationship among the five sectors of the KLSE. T model is chosen for examining the usefulness of short-run information for forecasting purposes. T system of five equations is given by

$$\Delta \mathbf{x}_t = \mathbf{a}_0 + \mathbf{a}_1 \Delta \mathbf{x}_{t-1} + \ldots + \mathbf{a}_p \Delta \mathbf{x}_{t-p} + \varepsilon_t \qquad (4)$$

where

$$\Delta \mathbf{x}_t = \begin{bmatrix} \Delta \mathbf{x}_{1t} \\ \Delta \mathbf{x}_{2t} \\ M \\ \Delta \mathbf{x}_{5t} \end{bmatrix}, \quad \mathbf{a}_0 = \begin{bmatrix} a_{01} \\ a_{02} \\ M \\ a_{05} \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ M \\ \varepsilon_{5t} \end{bmatrix} \quad a_j = \begin{bmatrix} a_{11,j} & a_{12,j} & \Lambda & a_{12,j} \\ a_{21,j} & a_{22,j} & \Lambda & a_{25,j} \\ M & M & & M \\ a_{51,j} & a_{52,j} & \Lambda & a_{55,j} \end{bmatrix}$$

$j = 1, 2, ..., p$ and $\varepsilon_t \sim$ i.i.d. $N(\mathbf{0},\Omega)$.

To determine the lag order p, we use the Schwarz (1978) criterion for a system of equations give

$$SC = \frac{-2l}{N} + \frac{W \log N}{N}$$

where W is the number of parameters in the model and l is the value of the log-likelihood func evaluated at the estimates for these parameters. As this model involves five equations, the full sys log likelihood is used to compute SC. Assuming a multivariate normal distribution,

$$\ell = \frac{-5N}{2}(1 = \log 2\Pi) - \frac{N}{2}\log|\hat{\Omega}|$$

where $|\hat{\Omega}| = \det\left(\sum e_t e'_t / N\right)$ and $\mathbf{e}_t$ is the vector of residuals for period t.

Cointegration and Vector Error Correction (VEC) Model

The VAR model does not take into account possible long-run relationships among the five sectors. The cointegration test is used to detect the existence of such relationships. If a linear combination of more than one non-stationary series result in a stationary relationship, the series are said to be cointegrated. The stationary relationship is the long-run equilibrium relationship that is incorporated in a VEC model. A maximum likelihood test procedure that estimates the cointegrating vectors in a multivariate framework was introduced by Johansen (1991). To discuss this procedure, consider a VEC model given by

$$\Delta\mathbf{x}_t = \mu + \Pi\mathbf{x}_{t-1} + \Gamma_1\Delta\mathbf{x}_{t-1} + \Gamma_2\Delta\mathbf{x}_{t-2} + \ldots\ldots + \Gamma_p\Delta\mathbf{x}_{t-p} + \varepsilon_t \qquad (5)$$

where

$$\mathbf{x}_t = \begin{bmatrix} x_{1t} \\ x_{2t} \\ M \\ x_{5t} \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ M \\ \mu_5 \end{bmatrix} \quad \Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} & \Lambda & \Pi_{15} \\ \Pi_{21} & \Pi_{22} & \Lambda & \Pi_{25} \\ M & M & & M \\ \Pi_{51} & \Pi_{52} & \Lambda & \Pi_{55} \end{bmatrix} \quad \Gamma_j = \begin{bmatrix} C_{11,j} & C_{12,j} & \Lambda & C_{15,j} \\ C_{21,j} & C_{22,j} & \Lambda & C_{25,j} \\ M & M & & M \\ C_{51,j} & C_{52,j} & \Lambda & C_{55j} \end{bmatrix}$$

and j = 1, 2, ..., p.

If the rank of Π is r, where $r < 5$, then there exists r linear independent cointegrating vectors. Initially, the hypothesis H_0: r = 0 (no cointegrating equation) is tested against a general alternative of H_1: r > 0. If the null hypothesis is rejected, we proceed to test for H_0: r = 1 against H_1: r > 1 for existence of one cointegrating equation. If this null hypothesis is not rejected, it means that the system has one cointegrating equation. If the null hypothesis is rejected, the process is repeated until a non-rejection is found.

For this test, we use the likelihood ratio trace test statistic given by

$$Q_r = -N\sum_{j=r+1}^{5}\log\left(1-\lambda_j\right) \qquad (6)$$

where r is the hypothesized number of cointegrating vector under H_0, and λ_j is the j-th largest eigenval

for C = 0 where

$$C = | \lambda S_{11} - S_{10}\, S_{00}^{-1}\, S_{01} |$$

$$S_{00} = N^{-1}\, \Sigma r_{0t}\, r'_{0t}$$

$$S_{01} = N^{-1}\, \Sigma r_{0t}\, r'_{1t}$$

$$S_{10} = N^{-1}\, \Sigma r_{1t}\, r'_{0t}$$

$$S_{11} = N^{-1}\, \Sigma r_{1t}\, r'_{1t}$$

and r_{0t} and r_{1t} are the residuals from the regression of Δx_t and x_{t-1} on μ and the lags of Δx_t, respectiv The critical values for the trace test have been computed by Osterwald-Lenum (1992). If the test indica existence of r cointegrating vectors, the 5 x r matrix of eigenvectors corresponding to the r larg eigenvalues gives the long-run relationship. The relationship enters the VEC model through t term Πx_{t-1}.

2.5. Forecast Performance

Three different measures, namely, the root mean squared error (RMSE), mean absolute deviation (MA and mean absolute percent error (MAPE), are used to evaluate the forecast performance of all the mo considered above. These measures for sector-i are given as follows:

$$RMSE_i = \sqrt{MSE_i}\,, \qquad MSE_i = \frac{\sum_{t=N+1}^{N+s} (X_{it} - \hat{X}_{it})^2}{s}\,,$$

$$MAD_i = \frac{\sum_{t=N+1}^{N+s} | X_{it} - X_{it} |}{s}\,,$$

$$MAPE_i = \frac{\sum_{t=N+1}^{N+s} | (X_{it} - \hat{X}_{it}) / X_{it} |}{s} \times 100$$

where n is the number of observations in the out-of-sample forecast period and $\hat{X}_{it}$ is the predicted value of X_{it}.

RESULTS

Table 1 presents the results for the ADF test for presence of unit roots. To arrive at these results, equation (1) was initially fitted for m = 1 to 12. The Schwarz criterion was used to select the 'optimal' lag length[2]. The lag length reported is for the equation that minimizes the Schwarz criterion, where m = 1 is used for all the indices except for the Finance Index where m = 3, when the null hypothesis of a unit root is tested on the level of the data (panel (a) of Table 1). It is clear that α is very close to zero, and the null hypothesis cannot be rejected. This indicates that the sectoral indices contain at least a unit root. The same hypothesis, however, is strongly rejected when first differences of the data were used (panel (b) of Table 1), suggesting that they are stationary. The finding is that the sectoral indices are integrated of order one.

[2] The Schwartz criterion for a single equation is given by

$$SC_i = \frac{W \log N}{N} + \log\left(\frac{1}{N}\sum_{t=1}^{N} e_t^{\,2}\right)$$

Table 1: The Augmented Dickey-Fuller Test for the Presence of a Unit Root, H_0: $\alpha = 0$

$$\Delta X_{it} = \mu + \beta t + \alpha X_{i,\,t-1} + \sum_{j=1}^{m} \theta_j \,\Delta X_{i,\,t-j} + \varepsilon_{it}$$

Index	Lag (m)	μ(t-stat)		β(t-stat)		α(t_α)		Serial Correlation LMTest Lag 1	Lag 2	Lag 3	ARCH LM test	White test
(a) Logarithm of Sectoral Indices												
Finance	3	0.0290**	(2.28)	-2.22×10^{-6}*	(-1.58)	-0.0031	(-2.19)	0.3778	0.6715	0.1071	0.0000***	0.0000***
Industrial	1	0.0284**	(2.12)	-2.06×10^{-6}**	(-1.73)	0.0036	(-2.05)	0.4740	0.7307	0.8804	0.0000***	0.0000***
Plantation	1	0.0463***	(3.02)	-2.94×10^{-6}***	(-2.41)	-0.0057	(-2.89)	0.3370	0.4914	0.3737	0.0000***	0.0000***
Mining	1	0.0465***	(3.23)	-7.00×10^{-6}***	(-2.92)	-0.0069	(-3.08)	0.1789	0.3227	0.3231	0.0000***	0.0000***
Property	1	0.0321***	(2.43)	-5.00×10^{-6}***	(-2.50)	-0.0038	(-2.35)	0.3742	0.5845	0.3315	0.0000***	0.0000***
(b) First-Difference of Logarithm of Sectoral Indices												
Finance	3	0.0013	(1.11)	-1.21×10^{-6}	(-0.91)	-0.8003***	(-18.17)	0.5911	0.1109	0.1337	0.0000***	0.0000***
Industrial	4	0.0009	(0.91)	-1.01×10^{-6}	(-0.92)	-0.9753***	(-17.18)	0.3450	0.6403	0.5713	0.0000***	0.0000***
Plantation	1	0.0020	(1.90)	-2.28×10^{-6}**	(-1.89)	-0.9645***	(-27.61)	0.1405	0.2664	0.2796	0.0000***	0.0000***
Mining	1	0.0026*	(1.44)	-2.97×10^{-6}*	(-1.49)	-1.0465***	(-28.92)	0.2498	0.4400	0.6404	0.0000***	0.0000***
Property	5	0.0009	(0.76)	-1.45×10^{-6}	(-1.03)	-0.8773***	(-15.97)	0.6258	0.2833	0.3713	0.0000***	0.0000***

Notes:

t_α is the ADF test statistic for Ho : $\alpha = 0$ and the critical values are -3.9692, -3.4152 and -3.1295 a 1%, 5% and 10%, respectively (MacKinnon, 1991).The p-values are reported for the Serial Correl LM Test, and the ARCH LM and White tests for heteroscedasticity.

***Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

These results must be interpreted with caution. Although the LM test shows that serial correlation is not a problem, the ARCH and White tests indicate presence of heteroscedasticity. The Phillips-Perron test that allows for fairly mild assumptions concerning the distribution of the disturbances provides an alternative. The results of the test are reported in Table 2. This test utilizes estimated autocorrelation of the residuals in a non-parametric correction to the ADF t_α statistic for autocorrelation and heteroscedasticity. The number of autocorrelations included was set to be same as the lag length used in the ADF test. In addition, lag lengths of 5, 10 and 20 were also considered. The results are basically robust to the lag length. The null hypothesis of a unit root cannot be rejected for the level of the data, but is rejected in the case of the first differences. This reaffirms the findings of the ADF test that sectoral indices are integrated of order one.

Table 2: The Phillip-Perron Test Statistics for Testing the Presence of A Unit Root

Index	Lag Length			
	Same as Table 1	5	10	20
(a) Logarithm of Sectoral Indices				
Finance	-2.2019 (3)	-2.2110	-2.2157	-2.2532
Industrial	-2.0576 (1)	-2.0517	-2.0478	-2.0669
Plantation	-2.8725 (1)	-2.8857	-2.8989	-2.9058
Mining	-3.0679 (1)	-3.0689	-3.0798	-3.1044
Property	-2.2941 (1)	-2.3270	-2.3339	-2.3632
(b) First-Difference of Logarithm of Sectoral Indices				
Finance	-34.3031*** (3)	-34.3939***	-34.3913***	-34.9250***
Industrial	-38.0858*** (4)	-38.0866***	-38.0841***	-38.0935***
Plantation	-37.0246*** (1)	-37.0844***	-37.1813***	-37.2339***
Mining	-39.6550*** (1)	-39.6538***	-39.6648***	-39.7532***
Property	-36.0523*** (5)	-36.0523***	-36.0794***	-36.3434***

Notes:

Lag length refers to the number of estimated autocorrelations used in the non-parametric corrections on the ADF t_α statistic to account for weakly dependent and heterogenerously distributed disturbances.

The critical values are -3.9692, -3.4152 and -3.1295 at the 1%, 5% and 10%, respectively (MacKinnon, 1991).

The figures in parentheses show the lag length used in Table 1.

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level.

3.1. The models

In view of the results from the unit root tests, the Box-Jenkins model was fitted to the first differences. Autocorrelation and partial autocorrelation coefficients[3] were examined to identify the model. Generall the coefficients for lags beyond two are not significant and hence, the highest order considered i ARMA(2,2). Subsequently, terms that are not significant were dropped from this model. This resulte in ARMA(2,2) for the Finance Index, ARMA(1,1) for the Industrial and Plantation Index, and ARMA(1,2 for the Property Index. ARMA models of different orders were considered for the Mining Index, bu none have terms which are significant. The model ARMA(0,1) that has a minimum Schwarz criteri was ultimately selected. All the models are given in Table 3.

Table 3: The Box-Jenkins Autoregressive-Moving Average Models

Independent Variable	Dependent Variable				
	ΔX_{1t} (Finance)	ΔX_{2t} (Industrial)	ΔX_{3t} (Plantation)	ΔX_{4t} (Mining)	ΔX_{5t} (Property)
Constant	0.0005 (0.0007)	0.0001 (0.0005)	0.0003 (0.0006)	0.0002 (0.0009)	-0.0003 (0.0009)
AR1	1.3519*** (0.1350)	-0.7744*** (0.1308)	-0.6510*** (0.1720)		0.9557*** (0.0443)
AR2	-0.6781*** (0.0983)				
MA1	-1.2403*** (0.1380)	0.8192*** (0.1185)	0.7108*** (0.1594)	-0.0089 (0.0254)	-0.8712*** (0.0513)
MA2	0.6172*** (0.0969)				-0.0639** (0.0273)

Notes:

AR1 and AR2 refer to the first and second autoregressive terms, and MA1 and MA2 refer to the firs second moving average terms in the model. A cell without entry in the table indicates that the term o corresponding row is not included in the model for the variable of the corresponding column.

Figures in parentheses are standard errors.

***Significant at the 1% level.**Significant at the 5% level.*Significant at the 10% level

[3] Results are not reported but available on request

The VAR model was fitted to incorporate the short-run relationship among the five sectors. As this [illegible] focused on the short-run dynamics, only data for the period 2 September 1998 to 30 June 1999 [illegible] used. This is a period of market recovery after the decline due to the financial crisis. The VAR models of order p = 1 to 10 were fitted, and the model that has the smallest Schwarz criterion (discussed in Section 2.3) was chosen. In all cases, the order selected is p = 1, which is similar to the study of Kok [illegible] (1997). The results are reported in Table 4.

Table 4: The Vector Autoregression Model

Independent	Dependent Variable				
Variable	ΔX_{1t} (Finance)	ΔX_{2t} (Industrial)	ΔX_{3t} (Plantation)	ΔX_{4t} (Mining)	ΔX_{5t} (Property)
Constant	0.0069*** (0.0025)	0.0054** (0.0022)	0.0027 (0.0018)	0.0083** (0.0041)	0.0050* (0.0027)
$\Delta X_{1,t-1}$	-0.1395 (0.2231)	-0.1473 (0.2045)	-0.1652 (0.1646)	-0.5967 (0.3749)	-0.2083 (0.2425)
$\Delta X_{2,t-1}$	0.4158* (0.2147)	0.1801 (0.1968)	0.3098* (0.1584)	0.6888* (0.3608)	0.5044** (0.2333)
$\Delta X_{3,t-1}$	-0.2182 (0.2208)	-0.1361 (0.2024)	-0.1056 (0.1628)	-0.2361 (0.3710)	-0.3398 (0.2399)
$\Delta X_{4,t-1}$	-0.1292* (0.0744)	-0.1009 (0.0682)	-0.0960* (0.0548)	-0.2933** (0.1250)	-0.1106 (0.0808)
$\Delta X_{5,t-1}$	0.0119 (0.1559)	0.0644 (0.1429)	0.0784 (0.1150)	0.2951 (0.2620)	0.1420 (0.1694)

Notes:

Figures in parentheses are standard errors.

*** Significant at the 1% level.

** Significant at the 5% level.

* Significant at the 10% level.

Similarly, VEC models of order p = 1 to 10 were fitted for every sectoral index, and the model with the lowest Schwarz criterion was chosen. It is assumed that intercepts are present in the cointegrating vector. In this case, the model with p = 1 has the smallest SC and the cointegration test is based on this model. This test takes into consideration that the sectoral indices contain a deterministic trend, as suggested by the results in Table 1. The results of the test including the likelihood ratio trace statistic defined in (6) are given in Table 5. The null hypothesis of no cointegration is rejected in favour of at least one cointegrating equation. At the subsequent stages of testing, no evidence was found to indicate that there is more than one cointegrating equation. Normalizing on the Financial Index, the cointegrating equation is given by

$$\hat{X}_{1t} = 6.97\, X_{2t} - 14.61\, X_{3t} + 10.54\, X_{4t} - 6.29\, X_{5t} + 54.46 \quad (7)$$

where X_{1t}, X_{2t}, X_{3t}, X_{4t} and X_{5t} represent the Finance, Industrial, Plantation, Mining and Property Indices respectively.

Table 5: Results of Cointegration Test

Eigenvalue	Likelihood Ratio Trace Statistic	5 percent critical value	1 percent critical value	Hypothesized Number of Cointegrating equation(s)
0.0201	78.4841	68.52	76.07	0 ***
0.0144	47.0663	47.21	54.46	1
0.0115	24.5879	29.68	35.65	2
0.0029	6.6391	15.41	20.04	3
0.0014	2.21329	3.76	6.65	4

Note:

Critical values are obtained from Osterwald-Lenum (1992).

The vector error correction model on which the test is based on is reported in Table 6.

*** Significant at the 1% level.

The estimated VEC model is reported in Table 6. This model incorporates the long-run dynamics by inclusion of the cointegrating equation (7). The error correction terms for all five sectoral indices are significant. We see that the changes in the sectoral rate of return adjust by between 0.07 to 0.16 per cent to a 1 per cent deviation from the long-run equilibrium relationship. The fastest speed of adjustment is found for the financial sector, and the slowest for the industrial sector.

Table 6: The Vector Error Correction Model

Independent Variable	Dependent Variable				
	ΔX_{1t} (Finance)	ΔX_{2t} (Industrial)	ΔX_{3t} (Plantation)	ΔX_{4t} (Mining)	ΔX_{5t} (Property)
Constant	0.0004 (0.0006)	0.0001 (0.0005)	0.0003 (0.0005)	0.0001 (0.0009)	-0.0002 (0.0006)
ΔX_{1t-1}	0.2449*** (0.0558)	0.1166** (0.0460)	0.0968 (0.0506)	0.1823** (0.0836)	0.1668*** (0.0586)
ΔX_{2t-1}	-0.0820 (0.0636)	-0.0934* (0.0524)	-0.0708 (0.0577)	-0.0125 (0.0953)	-0.0158 (0.0668)
ΔX_{3t-1}	0.0006 (0.0526)	0.0066 (0.0433)	0.0068 (0.0477)	0.0691 (0.0789)	-0.0921* (0.0553)
ΔX_{4t-1}	-0.0555** (0.0270)	-0.0064 (0.0222)	-0.0101 (0.0245)	-0.1019** (0.0404)	0.0569** -(0.0284)
ΔX_{5t-1}	-0.0072 (0.0507)	0.0004 (0.0417)	0.0350 (0.0460)	-0.0451 (0.0759)	0.0784 (0.0532)
Z_{t-1}	-0.0016*** (0.0004)	-0.0007*** (0.0003)	-0.0010*** (0.0003)	-0.0009** (0.0005)	-0.0015*** (0.0004)

Notes:

Figures in parentheses are standard errors.

Z_t, the error correction term, is given by

$$Z_t = X_{1t} - 6.97\,X_{2t} + 14.61\,X_{3t} - 10.54\,X_{4t} + 6.29\,X_{5t} - 54.46$$

*** Significant at the 1% level.

** Significant at the 5% level.

* Significant at the 10% level.

3.2. Forecast Performance

The models discussed above were used to forecast the daily sectoral indices for the month of July 1999. The random walk is also included to provide the benchmark for comparison. A one-day ahead forecast is performed in all cases. The forecast and actual values of the five sectoral indices are plotted in Figures 1 to 5. In all cases, the forecast series track the actual series rather closely.

Table 7 presents the results for measuring the forecast performance of each model by comparing the forecast values to the actual sectoral indices. The relative measure is also reported. This is the ratio of the measure for each forecast method to that for the worst method. In this case, the worst forecast method for a particular sectoral index is normalized to a value of 1.00, and used as a benchmark to assess the other methods. Hence, a model with the lowest relative value shows the best forecast performance.

Table 7. Measures of Performance of the Model for Forecasting the Daily Sectoral Indices in the Month of July, 1999

	MSE		RMSE		MAD		MAPE	
	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative
Finance								
VAR	19165.63	1.00	138.44	1.00	108.71	1.00	1.68	1.00
VEC	16014.90	0.84	126.55	0.91	101.27	0.93	1.57	0.93
ARMA(2,2)	16778.44	0.88	129.53	0.94	107.21	0.99	1.65	0.98
RW	15433.09	0.81	124.23	0.90	101.52	0.93	1.56	0.93
Industrial								
VAR	897.60	1.00	29.96	1.00	22.96	1.00	1.63	1.00
VEC	736.04	0.82	27.13	0.91	21.21	0.92	1.51	0.93
ARMA(1,1)	744.08	0.83	27.28	0.91	21.27	0.93	1.51	0.93
RW	726.84	0.81	26.96	0.90	21.39	0.93	1.52	0.93
Plantation								
VAR	830.02	1.00	28.81	1.00	21.40	1.00	1.26	1.00
VEC	730.62	0.88	27.03	0.94	19.76	0.92	1.17	0.93
ARMA(1,1)	642.77	0.77	25.35	0.88	18.05	0.84	1.06	0.84
RW	631.52	0.76	25.13	0.87	17.91	0.84	1.06	0.84
Mining								
VAR	166.67	1.00	12.91	1.00	10.03	1.00	3.70	1.00
VEC	136.42	0.82	11.68	0.90	9.72	0.97	3.57	0.96
ARMA(0,1)	138.23	0.83	11.76	0.91	9.79	0.98	3.58	0.97
RW	138.77	0.83	11.78	0.91	9.81	0.98	3.59	0.97
Property								
VAR	2952.84	1.00	54.34	1.00	41.95	1.00	3.48	1.00
VEC	2776.24	0.94	52.69	0.97	41.50	0.99	3.45	0.99
ARMA(1,2)	2797.87	0.95	52.89	0.97	41.10	0.98	3.41	0.98
RW	2707.12	0.92	52.03	0.96	40.97	0.98	3.40	0.98

Notes :

VAR, VEC, ARMA(p,q) and RW refer to the vector autoregression, vector error correction, autoregressive-moving average of order p and q, and random walk, respectively.

MSE, RMSE, MAD and MAPE refer to mean squared error, root mean squared error, mean absolute deviation and mean absolute percent error, respectively.

The relative measure is the ratio of the actual measure for the model of interest to that for the model with the worst performance.

The MAPE measure shows that the forecasting error ranges from 1.0 to 3.7 per cent. On average, th lowest percentage is for the Plantation Index, and highest is for the Mining Index. The minimum MAP for the Finance, Industrial, Plantation, Mining and Property Index is 1.56, 1.51, 1.06, 3.57 and 3.40 pe cent, respectively. The RMSE and MAD measures indicate that the average error of forecast is in th range of 100-140 points for the Finance Index, 18-30 points for the Industrial and Plantation Index, 10- points for the Mining Index, and 40-55 points for the Property Index. The relative ratios consistent suggest that the VAR model has the poorest forecast performance for all the indices regardless of th criterion used. The difference in the forecasting error between this model and the best method range from 0.08 to 0.20 per cent.

For a clearer illustration, the model with the best performance is ranked 1 and the poorest is ranked using the criteria of RMSE, MAD and MAPE (see Table 8). The model with the second and third be forecast accuracy is given a rank of 2 and 3, respectively. The rank across these three criteria is averag for each model. The random walk has performed the best for the Finance, Plantation and Proper Index. The VEC model is the best method for the Industrial and Mining Index. Note that the performan of the VEC model is only marginally behind the random walk for the Finance Index. The Box-Jenkin ARIMA model can sometimes perform as good as the random walk or VEC model, but its foreca accuracy is generally lower than the accuracy of either of these two models. As is shown earlier, th VAR model has the poorest forecast performance in all cases.

Table 8: Rank of Performance of the Model for Forecasting the Daily Sectoral Indices in the Month of July, 1999

	RMSE Rank	MAD Rank	MAPE Rank	Mean Rank
Finance				
VAR	4	4	4	4.00
VEC	2	1	2	1.67
ARMA(2,2)	3	3	3	3.00
RW	1	2	1	1.33
Industrial				
VAR	4	4	4	4.00
VEC	2	1	1	1.33
ARMA(1,1)	3	2	2	2.33
RW	1	3	3	2.33
Plantation				
VAR	4	4	4	4.00
VEC	3	3	3	3.00
ARMA(1,1)	2	2	2	2.00
RW	1	1	1	1.00
Mining				
VAR	4	4	4	4.00
VEC	1	1	1	1.00
ARMA(0,1)	2	2	2	2.00
RW	3	3	3	3.00
Property				
VAR	4	4	4	4.00
VEC	2	3	3	2.67
ARMA(1,2)	3	2	2	2.33
RW	1	1	1	1.00

Notes :

RMSE, MAD and MAPE refer to root mean squared error, mean absolute deviation and mean absolute percent error, respectively.

The model with the best forecast performance is given a rank of 1, and the poorest is given a rank of 4.

The mean rank is further averaged across the five indices and this is reported in Table 9. The number o times where a model is ranked as 1 and 2 in Table 8 is also given. All these indicate that the random wa has the best forecast performance, and this is followed by the VEC model, univariate ARIMA processe and lastly the VAR model.

Table 9: Summary of Forecast Performance Measures

	Mean rank across five sectoral indices	Number of cases where model is best	Number of cases where model is second best
VAR	4.00	0	0
VEC	1.93	6	4
ARMA(p,q)	2.33	0	10
RW	1.73	9	1

Notes :

The mean rank is computed by averaging the figures in the last column of Table 8 across the fi sectoral indices for each model. This average ranges from 1 (best) to 4 (poorest).

The last two columns in this table indicate the number of times where a model is best and second be respectively, out of 15 cases (among the five sectoral indices and each measured by the criterion RMSE, MAD and MAPE).

CONCLUSION AND DISCUSSION

This paper evaluates the usefulness of univariate models and inter-sectoral short- and long-run relationship in forecasting five sectoral indices of the KLSE. The findings concur with existing evidence in that the behaviour of the sectoral indices is predominantly random walk such that it strongly influences the future trajectory of the indices. In this context, it is not surprising that the random walk out-performs the other univariate ARIMA processes and models that capture the short-run (VAR) and long-run (VEC) dynamics of inter-sectoral relationship for forecasting the sectoral indices. When interpreting this finding, an important point to be borne in mind is the limited practical usefulness of the random walk.

This paper attempts to evaluate the next best alternative that is less constrained in its practical usefulness. A few significant findings are noteworthy. First is that the VEC model almost matches the forecast performance of the random walk. This underscores the importance of incorporating information on long-run relationships for forecasting purposes. Second, exploiting univariate movements by itself via ARIMA modeling does not perform too poorly compared to the random walk, despite that this ranks third after the VEC model. The practical advantage of this approach is the easiness in its computation. However, with the availability of application software packages, one can still argue in favour of the use of VEC models. Third, this paper cautions against the use of limited data set and forecasting method based on short-run sectoral relationship alone. All the forecast performance measures employed in this paper indicated that the VAR model is the poorest method.

It must be recognized that the stock market has become increasingly complex and therefore less amenable to forecasting over time. This demands for a greater sophistication in the tools used for forecasting. The scope of this paper can be extended in this spirit to search for a model that can out-perform a random walk in future research. Two explorable aspects are the underlying non-linear dynamics and the volatility of stock price movements.

REFERENCES

Chinn, M.D. and R.A. Meese, 1995. Banking on currency forecasts: How predictable is change in money?, *Journal of International Economics*, 38, 161-78.

Dickey, D.A. and W.A. Fuller, 1979. Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, 427-37.

Gui, H.K., 1999. Sectoral interdependence in the Kuala Lumpur Stock Exchange, Research Paper, Faculty of Economics and Administration, University of Malaya.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica*, 59, 1551-80.

Kok K.L. and Goh K.L., 1994a. Weak form efficiency in the KLSE: New Evidence, *Capital Markets Review*, 2, 45-60.

Kok K.L. and Goh K.L., 1994b. Weak form efficiency and mean reversion in the Malaysian stock market, *Asia-Pacific Development Journal*, 1, 137-52.

Kok K.L. and Goh K.L., 1996. Weak form efficiency and mean reversion in some Asia-Pacific Stock markets, *Malaysian Management Review*, 31, 50-61.

Kok K.L. and Goh K.L., 1997. Inter-sectoral causal relationship in the Malaysian stock market, *Malaysian Management Review*, 32, 53-7.

Laurence, M.M., 1986. Weak-form efficiency in the Kuala Lumpur and Singapore stock markets, *Journal of Banking and Finance*, 10, 431-45.

MacKinnon, J.G., 1991. Critical values for cointegration tests, in: R.F. Engle and C.W.J. Granger (eds.), *Long-run Economic Relationships: Readings in Cointegration*, Oxford University Press.

Mansor, M.I., 1989. Share price behaviour on the Malaysian stock market: Some empirical evidence, *Malaysian Journal of Economic Studies*, 26(1), 1-20.

Meese, R.A. and K. Rogoff, 1983. Empirical exchange rate models of the seventies: Do they fit out of sample?, *Journal of International Economics*, 14, 3-24.

Osterwald-Lenum, M., 1992. A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics, *Oxford Bulletin of Economics and Statistics*, 54, 461-72.

Phillips, P.C.B. and P. Perron, 1988. Testing for a unit root in time series analysis, *Biometrika*, 75, 335-46.

Saw, S.H. and Tan K.C., 1989. Test of random walk hypothesis in the Malaysian stock market, *Securities Industry Review*, 15, 45-50.

Schwarz, G., 1978. Estimating the dimension of a model, *Annals of Statistics*, 6, 461-4.

Yoon, G., 1998. Forecasting with structural change: Why is the random walk model so damned difficult to beat? *Applied Economics Letters*, 5, 41-2.

Figure 1 : Forecast and Actual Values of the Finance Index, July 1999

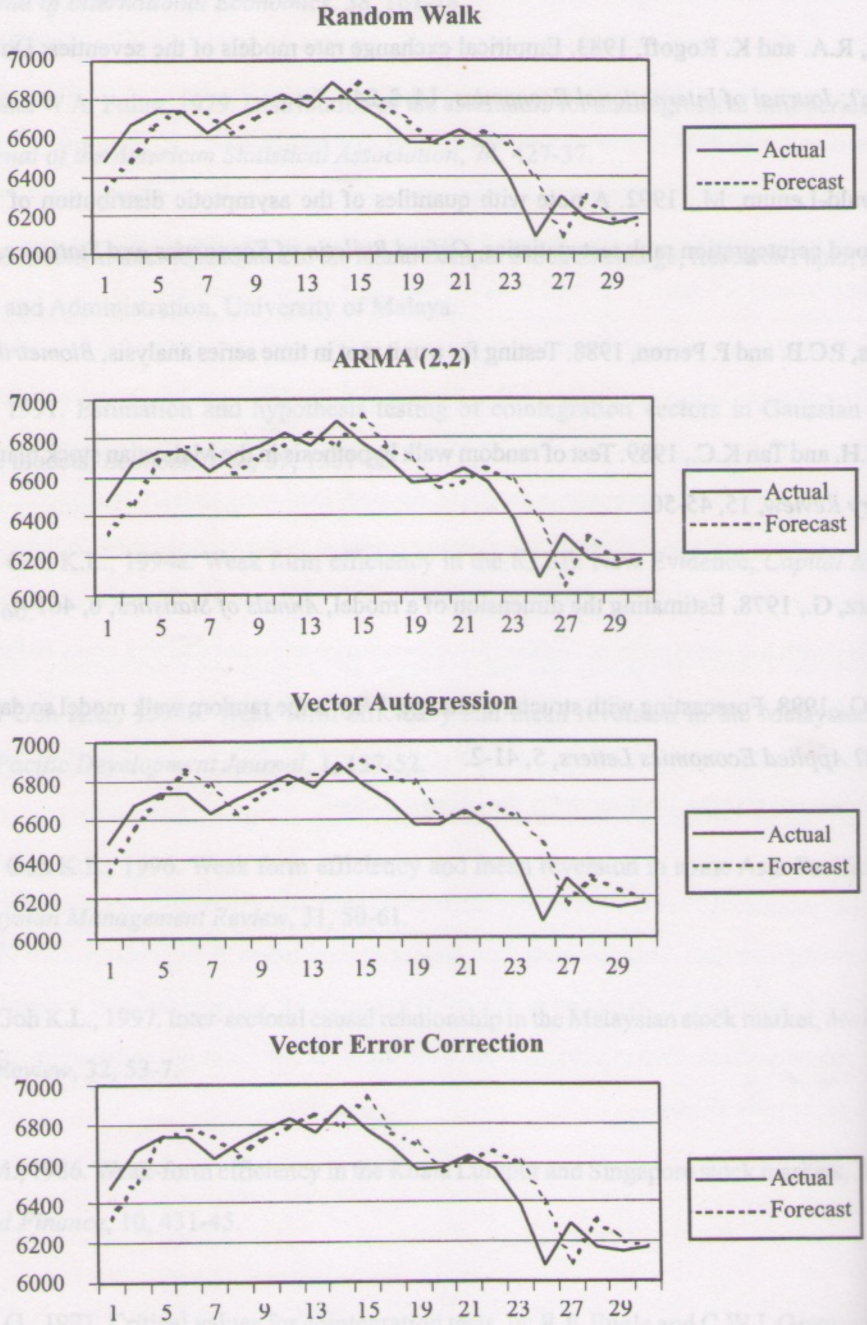

Figure 2 : Forecast and Actual Values of the Industrial Index, July 1999

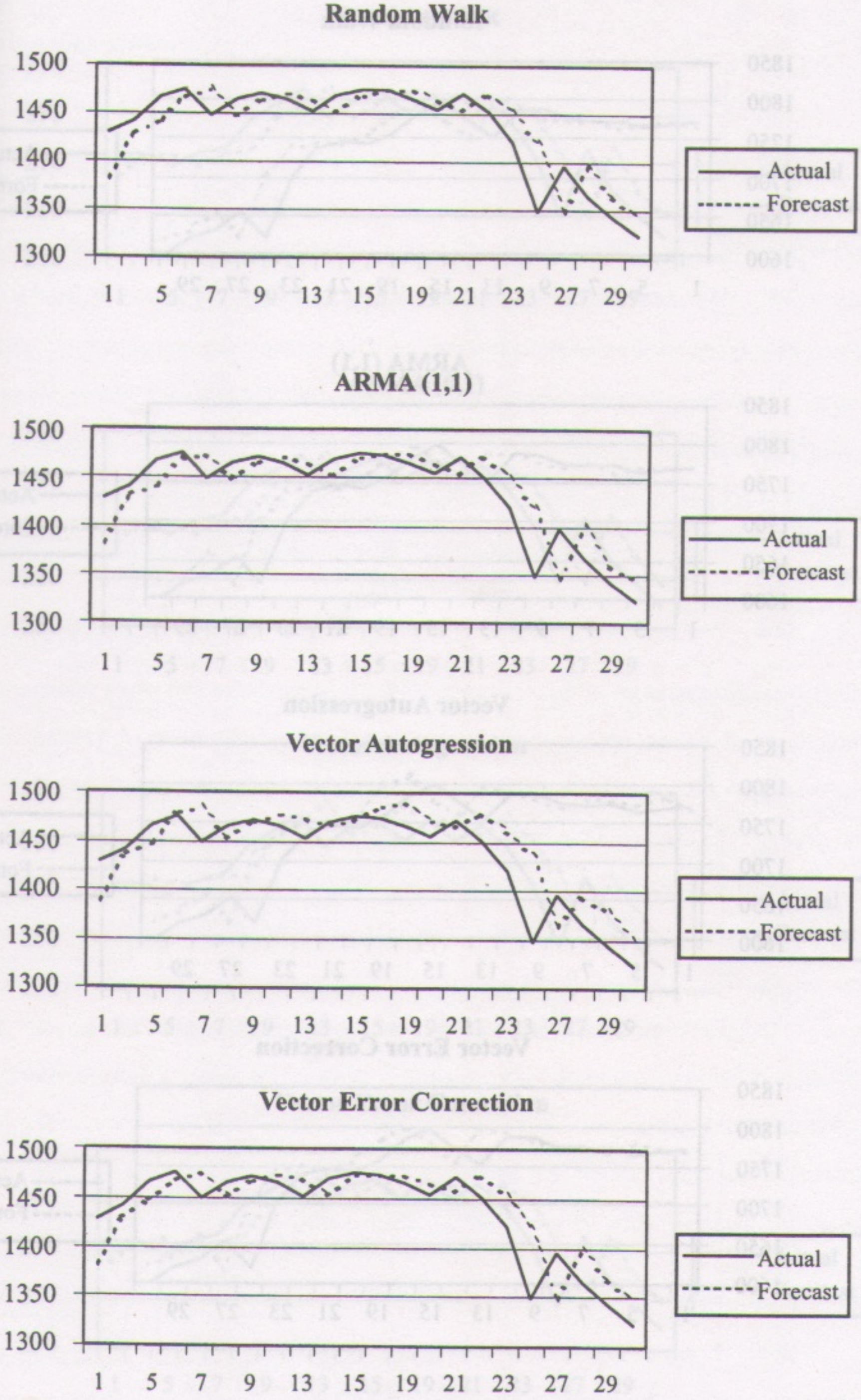

Figure 3 : Forecast and Actual Values of the Plantation Index, July 1999

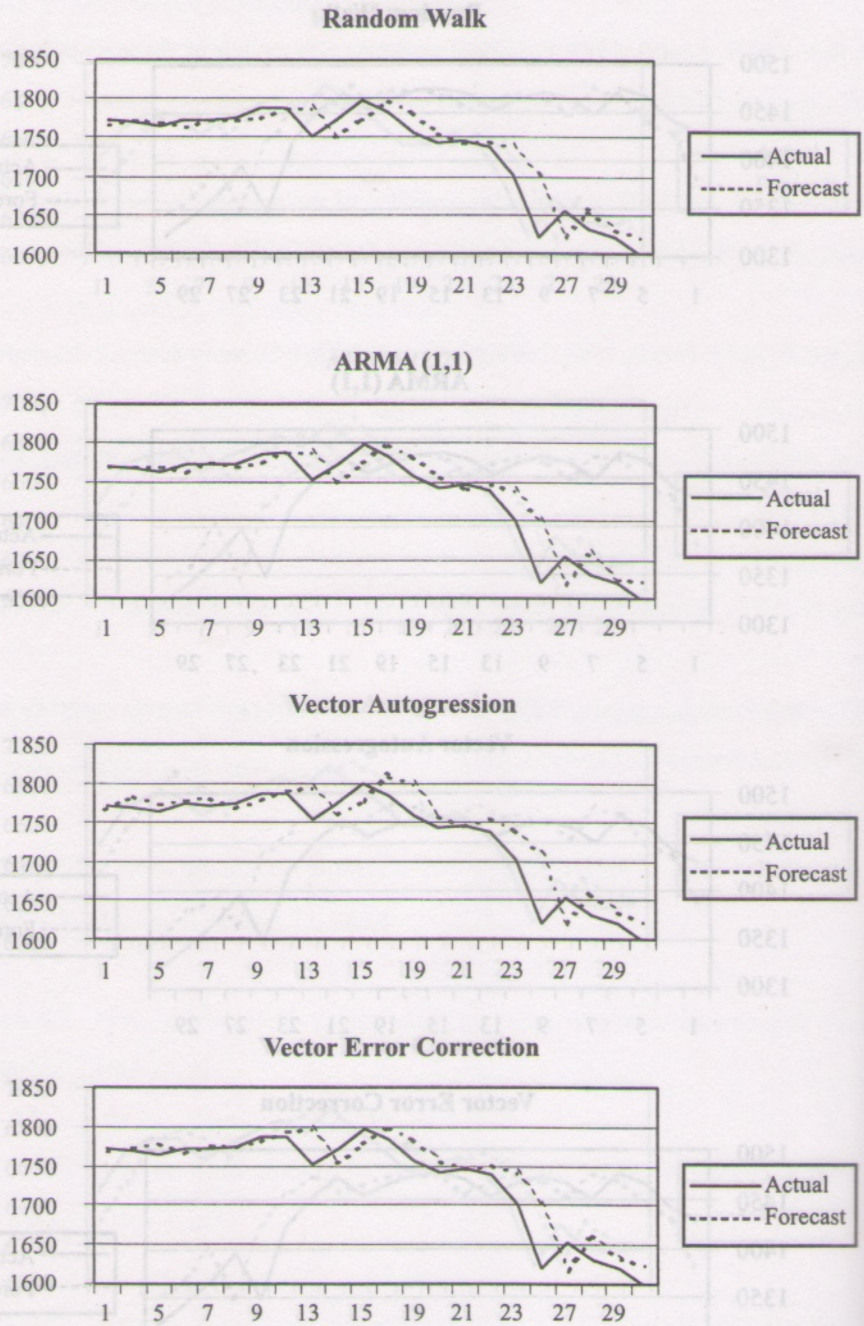

Figure 4 : Forecast and Actual Values of the Mining Index, July 1999

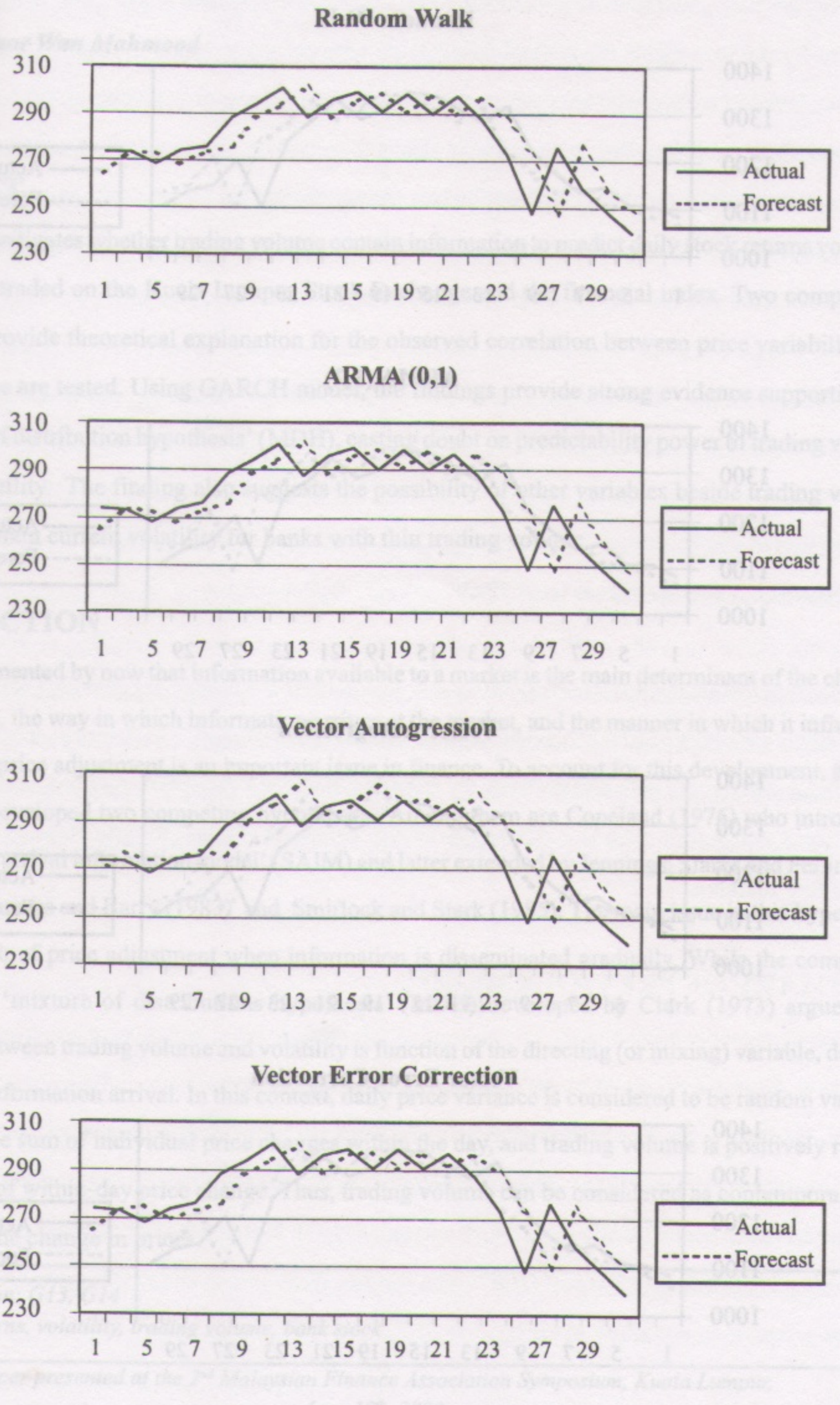

Figure 5 : Forecast and Actual Values of the Property Index, July 1999