PACIFIC RIM PROPERTY RESEARCH JOURNAL, 2017
VOL. 23, NO. 2, 123–160
https://doi.org/10.1080/14445921.2016.1225149

Routledge
Taylor & Francis Group

Check for updates

Forecasting the REITs and stock indices: Group Method of Data Handling Neural Network approach

Rita Yi Man Li[a], Simon Fong[b] and Kyle Weng Sang Chong[b]

[a]Sustainable Real Estate Research Center, Hong Kong Shue Yan University, Hong Kong, China; [b]Department of Computer and Information Science, University of Macau, Macau, China

ABSTRACT

If there is long-term memory in property stocks and REITs prices, historical data is relevant for future prices prediction. Despite previous research adopted various different methods to forecast future asset prices by using historical data; we attempted to forecast the REITs and stock indices by Group Method of Data Handling (GMDH) neural network method with Hurst which is the first of its kind. Our results showed that GMDH neural network performed better than the classical forecasting algorithms such as Single Exponential Smooth, Double Exponential Smooth, ARIMA and back-propagation neural network. The research results also provide useful information for investors when they make investment decisions.

ARTICLE HISTORY
Received 19 October 2015
Accepted 13 August 2016

KEYWORDS
Forecast; Group Method of Data Handling Neural Network; REITs; stocks

1. Introduction

LeRoy and Porter (1981)'s variance-bounds test results supported the existence of random walk and we cannot forecast stock prices accurately. Kendall (1953) suggested that there was no regularity or periodicity in stock prices movements. Fama (1965) indicated that changes in stock price are random and confirmed random walk hypothesis. Stock prices prediction based on historical information was impossible. The Efficient Market Hypothesis (EMH) proposed that if prices reflect all the information in stock market, we cannot use information to increase profit in such an efficient market. Nevertheless, Mandelbrot speculated that capital market price changes were fat-tailed with spikes (LeRoy & Porter, 1981).

Many of the previous research suggested that securitized and direct real estate markets are correlated. For example, Assaf (2006) confirmed that there were long-term co-memories between stocks and securitized property markets in Canada, overthrew EMH theory. Another school of thought, however, illustrate the REITs are unique with some special characteristics such as tax transparency (Newell & Osmadi, 2009; Newell & Peng, 2012). Kuhle (1987)'s research, however, found that risk reduction in common stock was greater than REITs and REITs are more efficient. Besides, Wang, Erickson, and Chen (1995) as compared to the general securities exchange, REITs have less turnovers. All these lead us to study whether the most accurate forecasting methods for REITs and stocks are different

CONTACT Rita Yi Man Li ymli@hksyu.edu

© 2016 Pacific Rim Real Estate Society

as they have different characteristics according to Kuhle (1987) and Wang et al. or they have similar forecasting results due to long-term co-memories between them based on Assaf (2006)'s study.

Previous research used different methods to forecast stocks and REITs prices. For example, Li and Chau (2016) used state space model to forecast property stock prices. Pavlova, Cho, Parhizgari, and Hardin (2014)'s research results showed that the appearance of long memory in REIT return was due to a lack of adjustment for temporal changes in the unconditional mean of volatility and modified FIGARCH model performed better at daily and weekly forecast horizons.

In this research paper, our objectives are to test the following hypotheses:

H1: Random walk theory is disproved, i.e. we can forecast stock and REITs prices.

H2: Long-term memory exists in stock and REITs prices according to R/S analysis.

H3: GMDH neural network results in better forecasting results than Single Exponential Smooth (SES), Double Exponential Smooth (DES), ARIMA and back-propagation neural network (BPNN).

H4: Long-term memory as reflected by Hurst can affect the forecasting ability of GMDH.

H5: There are differences in forecast performance in stocks and REITs indices.

The paper has academic and practical values. It is the first paper which adopts GMDH neural network with Hurst to perform forecasting. It also compares and contrasts the forecasting performance of GMDH neural network with Hursts on stocks and REITs prices with the traditional method such as BPNN, moving average, SES, DES, ARIMA. The results also show that even if long-term memory is weak, with close to random walk characteristics in Italy stock indices, GMDH neural network still outperforms the other traditional forecasting methods.

2. Financial forecasting methodology

Stock prices can be analyzed by technical and fundamental analysis (Huang & Chiu, 2005). Fundamental analysis includes international and political events, aggregate economics environment, industry conditions, and individual enterprises. International and political events include war, natural and man-made disaster, international sanction, international conference or negotiations, large scale financial institutions bankruptcy, government policy and intervention, vote, mass strike, and so on. Economics environment can be affected by incomes, interest rates, exchange rates, commodities prices, oil prices, and tax rates etc. Industry conditions can be affected by types of productions and life-cycle of industry. Individual enterprises can be affected by enterprises' financial statements and their individual events (Huang & Chiu, 2005). On the other hand, previous research shows that long-memory models consistently outperform their short-memory counterparts over a variety of forecast horizons (Zhou & Kang, 2011).

Previous researchers use different research approach to forecast REITs returns, for example, moving average models, SES, DES, autoregressive moving average (ARMA) model, autoregressive integrated moving average (ARIMA) model etc. Pierdzioch and Hartmann (2013) used the think and thin modeling to forecast the real estate returns. Their research found that excess returns are predictable out-of-sample by using the financial and macroeconomic data in real time. Ling, Naranjo, and Ryngaert (2000) found that excess returns are less predictable in the case of out-of-sample than in-sample. Besides, active-trading tactics according to out-of-sample prediction under zero transaction cost assumption perform better than REIT buy-and-hold strategies. Sah, Zhou, and Das (2015) used a sample of 108 additions to the S&P REIT Index from 2000 to 2011, it was found that responses of the analysts to index announcement depended on the type of revision to dividend forecast. Announcement of positive revised estimates does not add any new information to the analysts. However, negative revised estimates add information to the estimates.

2.1. Traditional forecast method: moving average, single exponential smoothing method and ARIMA

In conventional moving average forecast, consider F_t as the forecast at some point within the time series and Yt is the observation (Cadenas, Jaramillo, & Rivera, 2010). In a bull market the effect of the lag of the moving average will fall below the rising price line, whereas in the bear market it will be above. When the price changes direction, from falling to rising or vice versa, moving average and price lines will cross as the moving average reflects the preceding trend by nature of the lag (Ellis & Parbery, 2005).

Single exponential smoothing method is used to forecast the previous periods and adjusted the results by using the forecast error. The new forecast value is the old forecast with an adjustment for the error which obtained in the last forecast (Cadenas et al., 2010). Double exponential smoothing method applies SES twice, once to the original data and then smoothed data. Holt's (2004) method for double exponential smoothing used two different smoothing parameters:

The level estimate: $L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1})$

The trend estimate: $T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$

The p-step ahead forecast at the time origin: $\hat{Y}_{t+p} = L_t + T_t p$

Initial values: (1) $L_t = Y_1$, $T_1 = 0$; (2) where L_1 = average of the first several original values, T_1 = estimated slope.

Wu, Liu, and Yang (2016)'s results show that Grey double exponential smoothing method performs better than double exponential smoothing method in forecasting. Burger, Dohnal, Kathrada, and Law (2001)'s tourism demand forecasting shows that single exponential smoothing's forecasting performance was worse than neural network method. Gardner, Anderson-Fletcher, and Wicks (2001) recorded previous research that exponential smoothing achieved more accurate forecasting results than Focus Forecasting.

2.2. ARIMA method

Autoregressive integrated moving average (ARIMA) forecasting method includes the following steps: (1) model identification and selection; (2) estimation of autoregressive (AR), integration or differencing (I), and moving average (MA) parameters; (3) model checking. ARIMA can be extended to non-stationary series by differencing the time series. Current data values correlate with past values in the same series to produce the AR component p. Current values of random errors were assumed to be correlated with the values in the past to produce the MA component, q. Current and past data's mean and variance are assumed to remain unchanged and stationary. The I component, represented by d, corrects the stationary problem through differencing. P refers to the order of AR terms, d indicates the order of differences requires to reach the goal of stationary, q is the order of MA terms in a non-seasonal ARIMA (p, d, q) model. The p, d, and q parameters are integers equal to or greater than 0.

Though ARIMA is a very well-known traditional forecasting method, many of the recent research showed that it is not better than some of the recently developed forecasting methods. According to Crawford and Fratantoni (2003), ARIMA models generally perform better in out-of-sample housing prices forecasting as compared to regime switching model. Nevertheless, Makridakis and Hibon (2000) showed that forecasting hotel occupancy by ARIMA was worse than simple methods such as Gardner's Dampen Trend Exponential Smoothing.

2.3. Neural network forecasting method

Neural network is a massive parallel complex nonlinear dynamic system. It has parallel processing mechanism, high-speed computing capability, nonlinear operation which processes self-learning and organization ability. BPNN was created with one hidden layer between input and output units (Figure 1). All nodes of a layer were connected to all the nodes in the adjacent layers. The BPNN had two working phases, learning and recall phase. Known data-sets were used as a training signal in the input and output layers during the learning phase. The first operation in learning phase is feed-forward. During this stage, each of the input neuron receives an input signal and broadcasts it to the connected neurons in hidden layer. Each of the neuron in hidden layer computes activation and directs the result to the output neuron (Dosset et al., 2016).

In 1987, Lapedes and Farber first applied the neural network to forecast (Lapedes & Farber, 1987). White (1988) used the neural networks to forecast the IBM daily stock returns. The result was not satisfactory after training the network as it is trapped in local minimum. Kimoto and Kazuo (1990) proposed modular neural networks prediction system for the Tokyo Stock Exchange Prices Indices.

Xiao and Huang (2000) used the Group Method of Data Handling (GMDH) neural network to predict Chinese stamp market price and water level of a river. GMDH provided highly satisfactory forecasting result. Dase and Pawar (2010) indicated that neural network can successfully extract useful information from big data.

2.4. FMH with R/S in financial markets analysis

Based on fractal theory, Peters (1994) presented the new Fractal Market Hypothesis (FMH). The Rescaled Range (R/S) analysis pinpointed that the fractional Brownian motion accurately portrays the volatility of financial markets. Zuang, Zhuang, and Tian (2003) applied R/S analysis of the Hurst exponent to study the volatile Shanghai Stock Exchange (SSE) Composite Index and Shenzhen Stock Exchange (SZSE) Component Index. The results showed that there was autocorrelation between stock price indices and these markets displayed fractal structure feature. Qian and Rasheed (2004) suggested that the BPNNs predicted Dow-Jones index more accurately in times of high Hurst Exponent. The Hurst

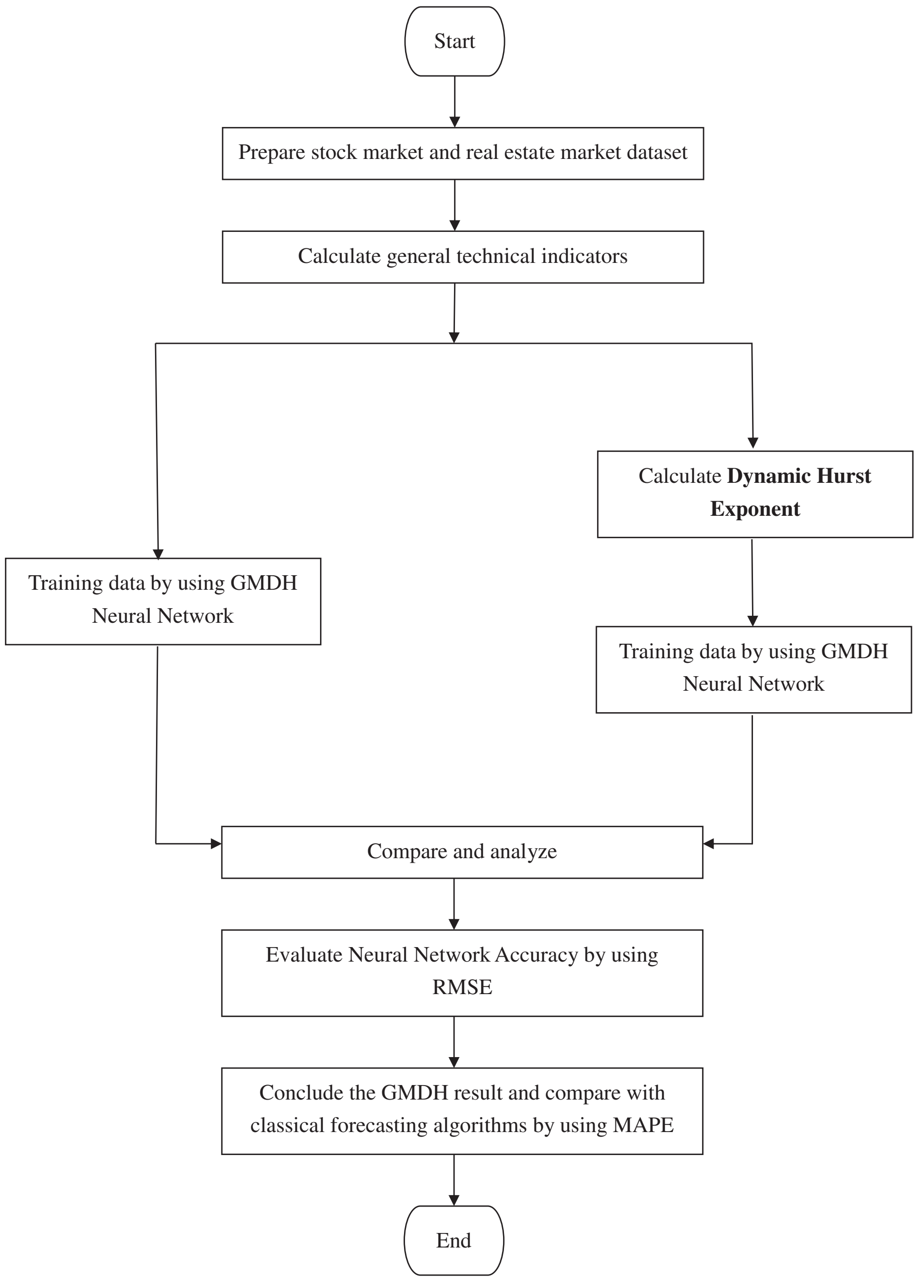

Figure 1. The procedure of training neural network.

calculation showed that stock markets were not random in all periods. Some periods had stronger trend.

Huang (2011) adopted correlation analysis, Brock, Dechert and Scheinkman test and Rescaled Range to study the futures fund's chaos effect. With regard to price forecast, there are back-propagation network and Adaptive network-based fuzzy inference system. By adding five input variables LIBOR, M2, R/J CRB index, Put/call ratio and MSCI global index in the forecast model predict the price of futures fund. The results indicate that futures prices can be predicted. Furthermore, the back-propagation network has better price forecast result.

Zhao and Xu (2011) calculated the Hurst exponent based on the Shanghai Stock Exchange (SSE) Composite Index and found that the dynamic Hurst exponent is able to predict the long-term stock price's trend. Faggini and Parziale (2012) indicated that nowadays the traditional economic theory is outdated. Many of the typical economic theory have become obsolete. As the mainstream economics are totally changing, the new elements are emerging constantly. One typical approach is to throw light on current phenomena with the chaotic theory. The author described different tools to test whether the economic and financial time series is chaotic. The chaotic behavior increases the difficulty in forecasting and introduces the new concept in today's economics and finance.

Mitra (2012) calculated the Hurst exponent of twelve stock index series from financial markets around the world and found that when the data were split into smaller series, there was a strong relationship between Hurst exponent value and shorter series of 60 contiguous trading periods. A high value of Hurst exponent implied there was long memory in time series data. Besides, it showed that Hurst and return from a trading rule are correlated. Therefore, the Hurst exponent is important.

3. GMDH neural network

We developed and used the GMDH neural network to forecast REITs and stock indices. Daily data include Australia All Ordinaries Index (AORD), Hong Kong Hang Seng Index (HSI), Italy FTSE Italia All Share Index (ITLMS) and Turkey Borsa Istanbul 100 Index (XU100) from 1 September 2010 to 2 September 2013. Another REITs daily indices data came from Australia, Hong Kong, Italy, and Turkey from March 2010 to March 2014. Among these stock markets, Australia recorded the largest market capitalization with USD$1,286 billion, followed by Hong Kong 1,108. Italy and Turkey was a small market with USD$480 billion and USD$309 billion market capitalization only.

During this period, Australia REITs market was the largest. The S&P/ASX 300 Property Index included 24 A-REITs with office, retail and industrial sectors (17 REITs were excluded in this REITs index) (Australia Shareholders Associations, 2014). In Turkey, number of REITs in Turkey grows rapidly with 23 public REITs. Its market capitalization reached around USD$3.9 billion & assets of about USD$7.7 billion (Real Estate Easy Property Info, 2016). On June 2012, 7 HK-REITs were listed in Hong Kong Stock Exchange with market capitalization of about US$15 billion (Real Estate Easy Property Info, 2016). Italy REITs is and was the smallest with 2 REITs in this period only (DataStream, 2016).

The step of GMDH neural network was illustrated in Figure 1:

Step 1. Gather stock and real estate indices data;

Step 2. Calculate the general technical indicator (KD, MA, Bias, RSI, W%R, MTM, MACD);

Step 2a. Calculate the dynamic Hurst exponent;

Step 3. Train the data by (1) general technical indicator, and (2) both general technical indicator and dynamic Hurst exponent;

Step 4. Compare and analyze the forecast results;

Step 5. Evaluate the neural network by using RMSE;

Step 6. Summarize and compare the results with classical forecasting algorithms with MAPE.

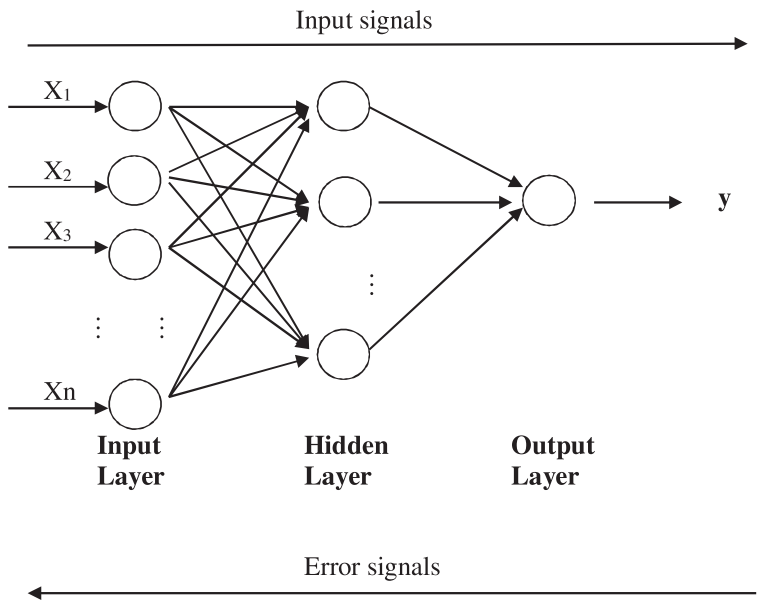

Figure 2. Structure of BPNN with one hidden layer.

3.1. Neural network

As the feed-forward neural networks structure is fixed, its forecasting performance depends on the predetermined structure. BPNN is a widely used neural network to forecast financial data (Kimoto & Kazuo, 1990; Yao, Tan, & Poh, 1999). Typical BPNN usually has one input layer, some hidden layers (at least one) and an output layer. Figure 2 shows a one-hidden layer BPNN.

Each layer contains one or more neurons with mesh connection between layers. Neurons are connected from input to output layer in a feed-forward manner. Weights of the connections are given. To reduce the error between target output and actual value, it is back-propagated via the network with weights updated. The supervised learning procedure minimizes the error between target output and actual value to raise the forecasting accuracy. Sigmoid activation function was used to calculate the outputs of neurons in the hidden layer. BPNN training is then carried by the following steps: (1) initialize the random weights; (2) use the current weight to calculate the output for activation; (3) calculate the error between target output and actual output; (4) adjust the weights to decrease error; (5) repeat step (2)–(5) until the error criterion is satisfied.

The Group Method of Data Handling (GMDH) improves forecasting performance of feed-forward neural network (Pan & Gu, 2007) by simulating nonlinear model. Non-linear heuristic self-organization method is valid for identification of high-order non-linear system (Ivakhnenko, 1970). It is also known as polynomial network model. The network changes continually in the training process. In short, GMDH algorithm displays the following characteristics:

(1) Self-organizing control in the modeling process without initial assumptions;
(2) Optimal complexity model and high precision in forecasting;

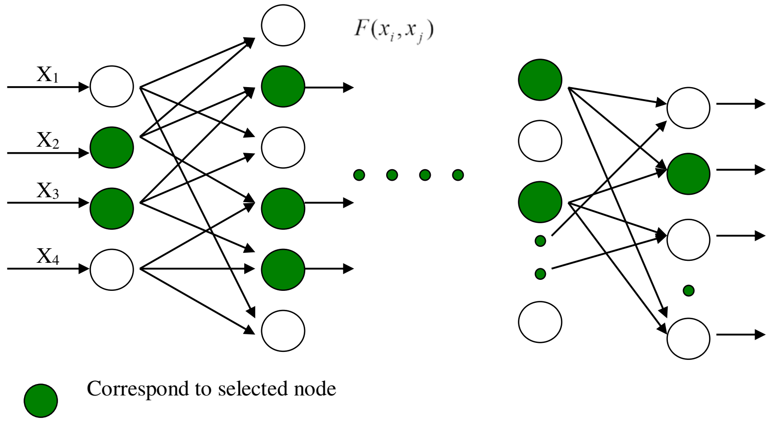

Figure 3. Structure of GMDH network.

(3) It is capable of self-organizing the best structure of each layer in multilayer neural network and it is able to keep the useful variables and remove the redundant variables automatically;
(4) It selects the best number of network layers and optimizes the number of neurons in each layer.

The forecasting performance is enhanced by discarding bad neurons through training. We use the GDMH shell software[1] to develop the neural network model with two learning algorithms: (1) GMDH-type neural networks and (2) Combinatorial GMDH (Fong, Nannan, Wong, & Yang, 2012).

3.2. GMDH-type neural network

GMDH-type neural network is multilayer iterative neural network with the variation of the traditional Multilayer Iterative GMDH algorithm (MIA). It is implemented by a feed-forward neural with multi-layered bi-input neurons. The GMDH algorithm generates a series of neuron by cross combination of each input unit where each neuron has optimal deterministic transfer function. We then select the number of neurons. Selected neurons are combined and generate new neurons. By repeating this genetic, survival competition and evolutionary process, the process continues until a new generation of neuron does not perform better than the previous ones, the optimal model is selected. The structure is shown in Figure 3.

Similar to other neural network, the GMDH algorithm is:

(1) A combination of the black box concept which study the relationship between input and output variables;
(2) The neural approach utilizes the threshold logic and network connectionism.

The GMDH neural network constantly generates neurons, filtered by the external rules, combines the good neurons after filtration and generates the next layer neurons, until the best model is selected. Multiplicative-additive algorithms expand the functional space of GMDH algorithms but face the problem of unstable coefficients due to least square method. Least square method is applied to the logarithms of their values after taking the powers.

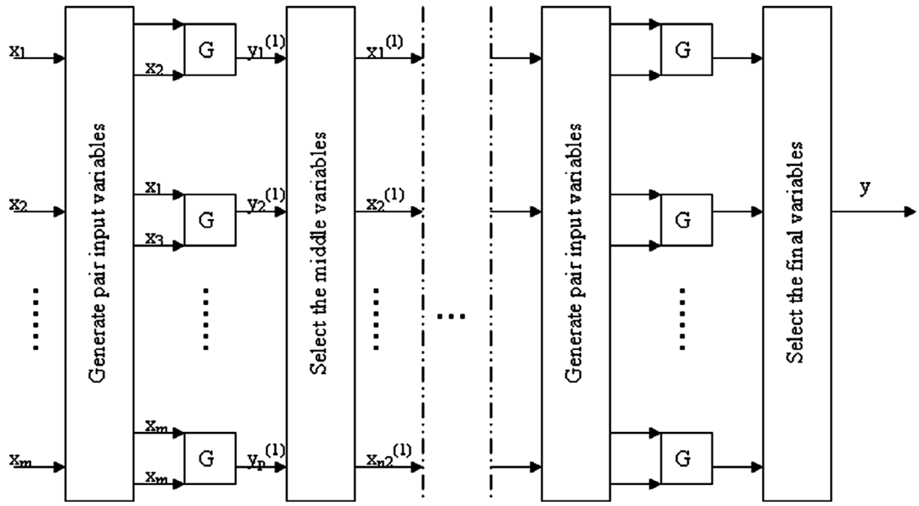

Figure 4. GMDH algorithm Diagram.

Consider a non-linear system, x_1, x_2, …, x_n is the input variables, y is the output variable. Their relationship will be:

$$y = f(x_1, x_2, \cdots, x_n) \tag{1}$$

The function f connects inputs and output variables. It can be expressed by discrete form of the Kolmogorov-Gabor function series:

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} x_i x_j x_k + \cdots \tag{2}$$

The basic GMDH algorithm shown in Figure 4, x_i is the initial input variable. G is the partial polynomials with quadratic polynomial for each two input variables. $y_i^{(k)}$ is the output of the partial polynomials. The partial polynomials are obtained by fitting the measured data. $x_i^{(k)}$ is the mediator, filtered by each layer criteria from $y_i^{(k)}$, used as the input variable in next layer Figure 4.

Training a GMDH network includes an input layer to build the network and weights adjustments of neuron. The network layers increase until they meet the requirement of mapping accuracy. The number of the first layer neurons depends on the number of the input variables, each pair of input variables are linked with one neuron and the output:

$$\hat{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \tag{3}$$

The a_i (i = 0, 1, 2, …, n) is the weight of the neurons, x_i and x_j refer to the pair of input. Therefore, $\hat{y}$ is the weight of input variables in quadratic polynomials. In basic GMDH network, two independent variables out of n input variables are used to construct the quadratic polynomial in Equation (3) that best fit the dependent observations (y_i, i–1, 2, …, M). Each network layer doubles the orders in the polynomial; the output results will be high order (2p orders) polynomial, where p is the number of layer except the input layer.

3.3. Combinatorial GMDH

Combinatorial GMDH is another algorithm in GDMH shell that increases polynomial function's length and power. Data sample is divided into training and test subsample. The training subsample estimates the coefficients of the model, and the test subsample chooses the optimal model structure. In the first layer, information contains every column of the sample as follows:

$$y'_j = a_0 + a_1 x_i, \quad i = 1, 2, \ldots m \tag{4}$$

m is the maximum number of x variables in $\bar{x}$, and y'_j is the jth predicted outcome by this simple linear regression. The least squares method solves the coefficients a_0 and a_1 for each of the m models by treating the model as a system of Gauss's normal equation:

$$\begin{bmatrix} nt & \sum_{j=1}^{nt} x_{j,i} \\ \sum_{j=1}^{nt} x_{j,i} & \sum_{j=1}^{nt} (x_{j,i})^2 \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{nt} y_j \\ \sum_{j=1}^{nt} x_{j,i} y_j \end{bmatrix} \tag{5}$$

As y_j's value is known from the training samples, Equation (5) estimates two coefficient values by a combinatorial search. When all the possible polynomial models have been calculated, the forecast values are checked against the real data. The model ζ with the lowest regularity criterion $\rho(\zeta)$ is retained. The regularity criterion that is used in fitness function is as follows:

$$\rho(\zeta) = \frac{\sum_{j=nt+1}^{n} (y_j - y'_j)^2}{n - nt} \tag{6}$$

$\rho(\zeta)$ is the average error in terms of the squared difference between the predicted and observed outcomes of model ζ. The model contains the variables from the initial layer that yield the lowest error. It is scaled up in the polynomial series and generated the new models in the subsequent layer. The polynomial in the second layer will be:

$$y_j = a_0 + a_1 x_i + a_2 x_k \tag{7}$$

$$y_j = a_0 + a_1 x_i + a_2 x_k + \cdots + a_m x_l \tag{7a}$$

where i and k = 1, 2, …, m. The models in the second layers are checked for compliance with the regularity criterion. The best model is used to generate another model in the third layer. This is repeated in the upper layers until the regularity criterion no longer decrease in value. The output model is selected among the fittest models in each layer. The selection criterion is the forecasting variance criterion δ:

$$\delta(\zeta) = \frac{1}{n} \times \frac{\sum_{j=1}^{n} (y_j - y'_j)^2}{\sum_{j=1}^{n} (y_j - \tilde{y}'_j)^2} \tag{8}$$

where n is the number of samples in learning window. If the predicted outcome y'_j pertaining to model ζ is applied on jth observation, mean value of the model outcome $\tilde{y}'_j$, is obtained. Feed-forward networks structure includes the number of layers and the neurons in each layer are fixed. Performance of the network is affected by the predefined network structure. During the training process, each layer of neuron network increases with the number of new neurons. Poor neurons are discarded.

3.4. Rescaled range analysis (R/S analysis)

R/S Analysis was used to calculate the Hurst exponent, to verify whether the system is a fractal approach. It was firstly used to study the changes in the flow of River Nile (Hurst, 1951) and then used to study the existence of long-term memory. R_n is an accumulated value of the difference between data and its average in a time series. It is called n data's difference. It represents the maximum of the changes in time series data. S_n is the standard deviation of the time series, and represents the degree of deviation from the average. It is a measure of the degree of dispersion. Rescaled range analysis R_n/S_n refers to the difference which is rescaled by S_n. In general, R/S Analysis's assumption is simple, applicable in any time series analysis, can distinguish whether the time series is in random walk and determine whether the time series is persistent. Hurst Exponent is estimated by R/S analysis according to the log return of data (Wang, Song, & Wu, 2004). To eliminate autocorrelation, take log of P_{i+1} over time series P_i will generate new logarithmic series L_i with length $N = M-1$ as follows:

$$L_i = \mathrm{Log}\left(\frac{P_{i+1}}{P_i}\right), \quad i = 1, 2, 3, \cdots, M-1$$

Step 1. A time series Li with time N is cut into shorter time intervals n, i.e. $A \cdot n = N$. The partial time series is $N(k,a)$, $k = 1, 2, 3, \ldots, A$, $a = 1, 2, 3, \ldots, n$. The mean of each of the partial time series is:

$$e_a = \frac{1}{A}\sum_{k=1}^{A} N_{k,a}$$

Step 2. Calculate the cumulative deviate ($X_{k,a}$) from the mean of partial time series:

$$X_{k,a} = \sum_{i=1}^{k} (N_{i,a} - e_a) \quad k = 1, 2, 3, \cdots, n$$

Step 3. Compute R_I which is the difference between the maximum value ($X_{k,a}$) and the minimum value of ($X_{k,a}$). ($X_{k,a}$) is the sum of deviation from the mean.

$$R_{I_a} = \max(X_{k,a}) - \min(X_{k,a}) \quad 1 \le k \le n$$

Step 4. Calculate the standard deviation over the range 1 to A:

$$S_{I_a} = \sqrt{\frac{\sum_{k=1}^{A} (N_{k,a-e_a})^2}{A}}$$

Step 5. R_I divide by S_I, and average all the partial time series of length n:

$$(R/S)_n = \frac{1}{n}\sum_{a=1}^{n}\frac{R_{I_a}}{S_{I_a}}$$

Step 6. Increase the length of A, repeat the Step 1 to 6, until $A = (M-1)/2$. According to Hurst, we establish the following relationship:

$$(R/S)_n = c \times n^H$$

where c is constant.

The Hurst exponent is estimated by least square linear regression log (R/S) = log (c) + $H \cdot \log(n) + \varepsilon$, and the Hurst exponent (H) is the slope of the line. To test the H whether it is equal to .5, and determine its persistency, the t-test statistics will be:

$$H_0{:}H = 0.5 \qquad H_1{:}H \neq 0.5$$

$$T = \frac{\hat{H} - 0.5}{S(\hat{H})}$$

where $\hat{H}$ is the estimate of H, $S(\hat{H})$ is standard deviation of $\hat{H}$ (Wang et al., 2004). In log (R/S) = log (c) + $H \cdot$ log $(n) + \varepsilon$, the slope of log(R/S) to log(n) is estimated as the Hurst exponent. When value of n is very small, ranges between 0 and 1, value of A is very large. Hurst and Feller proved that: $(R/S)_n = (n^*\pi/2)^{1/2}$, when H = .5, the correlation coefficient of time series is $C(t) = 2^{2H-1} - 1$, and the fractional dimension is $D = 2 - H$. the correlation coefficient function $C(t)$ can test the independence of the time series, and the fractional dimension tests whether the time series is random (Fong et al., 2012). Depend on the value of the Hurst Exponent between 0 and 1, there are three types of time series:

(1) When $H = .5$, $C(t) = 0$, $D = 1.5$, time series is random walk, correlation coefficient function is equal to zero, past increment is not associated with a future increment.
(2) When $.5 < H \leq 1$, $0 < C(t) \leq 1$, $1 \leq D < 1.5$, the time series has persistent behavior with long memory. Correlation coefficient falls between 0 and 1, meaning that the past increment has positive correlation with the future increment. If there is an increase from the last period of time, it may increase in the future. When there is a decrease from a last period of time, there will probably be a decrease in future. The time series is no longer a random walk, it is a biased random process.
(3) When $0 \leq H \leq .5$, $-1 \leq C(t) < 0$, $1.5 < D \leq 2$, the time series has anti-persistence behavior. The correlation coefficient function falls between −1 and 0, i.e. past increment has negative correlation with future increment. If the Hurst value is greater than .5, fractal feature exists.

According to Wang et al. (2004), V statistics' formula is:

$$V_n = \frac{(R/S)_n}{\sqrt{n}}$$

V Statistics was originally proposed by Hurst to check the stability of R/S analysis in 1951, but now is used to estimate the length of nonlinear long-term memory process.

To put $(R/S)_n = c \times n^H$ into V_n, then V_n can be: $V_n = \frac{(R/S)_n}{\sqrt{n}} = c \times n^{H-(1/2)}$. Based on the Hurst value, it will be mapped to different V statistics; there are three different types as follows:

(1) If $H = .5$, the time series is the independent random process. V statistics will be equal to constant c and will appear as a horizontal line.
(2) If $H > .5$ persistently, V statistics displays an upward trend.
(3) If $H < .5$ persistently, V statistics shows a downward trend.

If mutation is found in V statistics analysis chart, graphically original from the upward trend will be turned to horizontal or downward trend. This turning point is the length of the aperiodic cycle; n period memory will be vanished at this point.

3.4.1. Dynamic Hurst exponent

If dynamic Hurst exponent exists in period t, $t + n-1$, we select the sub period $(t, t + k-1)$, $k \leq n$, and then use the R/S method to calculate the Hurst exponent in this sub period. t, $t + k-1$ is called the time window of period $(t, t + n-1)$, i.e. the length of $(t, t + n-1)$ time series dynamic Hurst exponent. By fixing k periods, a set of dynamic Hurst exponent is then calculated. Figure 5 shows the Dynamic Hurst Exponent experiment:

3.5. Data analysis

We use different countries or cities' stock market and real estate market data in this research: (1) Australia All Ordinaries Index (AORD); (2) Hong Kong Hang Seng index (HSI); (3) Italy FTSE Italia All Share Index (ITLMS); (4) Turkey Borsa Istanbul 100 Index (XU100); (5) Australia REIT index data-set (Aus); (6) Hong Kong REIT index data-set (HK); (7) Italy REIT index data-set (Italy); (8) Turkey REIT index data-set (Turkey). All of the stock market data are obtained from the Yahoo finance[2] and Google finance.[3]

Testing for stationarity is a necessary and qualifying condition for many of the time series forecasting models such as ARIMA. Presence of stationarity in time series is essential because it affects the behavior of a series where shocks will persist permanently in non-stationary series. Absence of stationarity may give rise to spurious regressions, implying that some statistically significant results may emerge, even when time series are not totally related. When this happens, the standard assumptions for asymptotic analysis (t-ratio) will no longer be valid. The forecast results will not so be credible.

A time series like those stock prices is supposed to be stationary if there is a constant mean and a constant variance without trend in the time series. It looks as if without obvious gradient, with constant variance, with no autocorrelation or periodic fluctuations over time. The basic formulation of stationarity is expressed by considering a simple AR(1) process:

$$y_t = \left(c + t_r\right) + p^n \cdot y_{t-1} + \varepsilon_t \tag{9}$$

where $(c + t_r)$ is a combination of constant and trend in its 1st order, p is a coefficient which can take a nth order power depending on how non-linear the time series is, and ε_t is the white noise. The formal approach to test stationarity of a time series is by the unit root test.

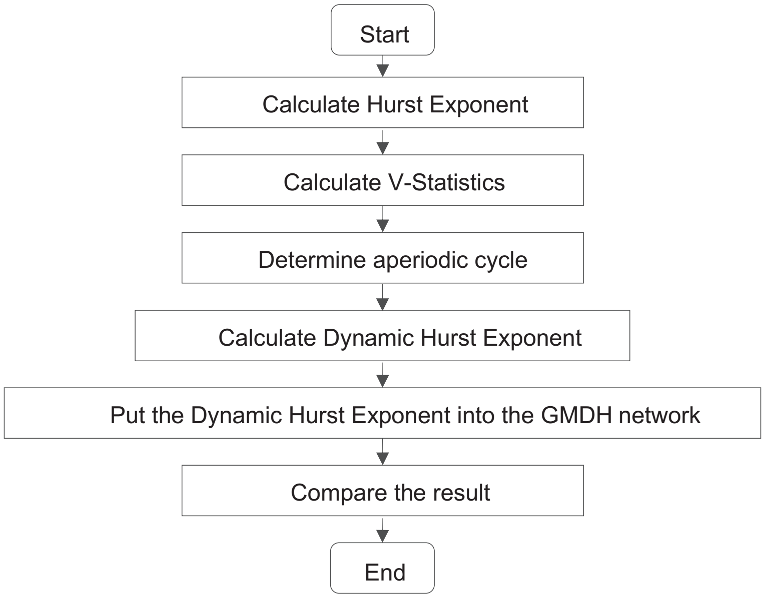

Figure 5. Dynamic Hurst Exponent Experiment Flow Chart.

Depending on the p which is estimated from the stochastic process, two hypotheses are established.

H0: If $|p|$ is greater than or equal to 1, then y is a non-stationary time series as its variance grows over time. i.e. against the alternative H1 when $p \geq 1$.

H1: If $|p|$ is less than 1, then y is a stationary time series.

Two tests are applied for verifying the stationarity of the time series data-sets that we used in this paper. They are Augmented Dickey–Fuller (ADF) (Cheung & Lai, 1995) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) (Kwiatkowski, Phillips, Schmidt, & Shim, 1992). These two popular tests complement each other with respect to setting H0 for non-stationary time series and stationary time series conversely.

3.5.1. Methods (how the test has been handled)

To test whether a given time series is stationary or not, an indirect test is applied checking for the existence of a unit root – the so-called augmented unit root test that generally verifies the hypothesis H0. Specifically two common tests for unit root are applied; they are Augmented Dickey–Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS). The ADF test incorporates a deterministic trend and squared trend, which are the combination of constant and trend ($c + t_r$), and a squared trend p2. respectively. Thus, it embraces tests of trend-stationary process with a moderate to polynomial non-linear trend power to occur. Generally checking over these two types of time series covers most cases of somewhat linear and largely non-linear patterns of time series.

The Augmented Dickey–Fuller test is an extended version of the standard Dickey–Fuller test which is based on Equation (9), and it modifies so by subtracting y_t−1 from both sides of the Equation (9):

$$\Delta y_t = \left(c + t_r\right) + \alpha \cdot y_{t-1} + \varepsilon_t \tag{10}$$

where $p-1 = \alpha$, which is an indicator simplified to represent the null and alternative hypothesis, respectively:

H0: stands true when $\alpha = 0$ that the time series is non-stationary

H1: stands true when $\alpha < 0$ that the time series is stationary

By the statistical properties of ADF, t-ratio is used to assess these hypotheses:

$$\tau_\alpha = \frac{\tilde{\alpha}}{\sigma(\tilde{\alpha})} \tag{11}$$

where σ is the standard error in the estimation process of α, and the estimated α is $\tilde{\alpha}$. It is known that the t-ratio will no longer follow the conventional t-distribution when there exists a unit root. This statistical property is used by DF test; however it is limited to checking only a single lag in the time series. To generalize this checking capability, ADF is extended to an autoregressive model AR(p), which is capable in testing time series with higher order lag. Therefore Equation (10) is extended to (12) as follow:

$$\Delta y_t = \left(c + t_r\right) + \alpha \cdot y_{t-1} + \beta_1 \cdot \Delta y_{t-1} + \beta_2 \cdot \Delta y_{t-2} + \ldots + \beta_n \cdot \Delta y_{t-n} + \varepsilon_t \tag{12}$$

In this way, this ADF formulation can be efficiently used to test the hypotheses H0 and H1 that involve only testing a variable α (equal to or less than 0), by computing the t-ratio of α, as in Equation (11).

ADF by far is effective, except for one shortcoming in determination of the lag order which is a manual step. The choice of the lag length in terms of number of lagged difference terms is influencing the efficacy of the stationarity test. We took the default lag length at 2 as suggested by the software NumXL (URL: http://www.spiderfinancial.com/products/numxl) which was used in the experiment. Since there is no standard lag length that can be proven most effective to remove serial correlation from the residuals, some additional test is required for assuring the reliability of stationary test results.

A popular complementary test to ADF is Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. KPSS tests in econometrics are utilized for testing a null hypothesis which defines the statistical properties of a time series to be stationary. KPSS concerns an observable time series that is trend stationary or stationary around a deterministic trend. The time series is considered as the composite of deterministic trend, stationary error, and random walk. Lagrange multiplier test of the hypothesis is used as the core of the KPSS for verifying if the random walk contains zero variance. The test is based on the residuals from the ordinary least square regression:

$$y_t = (c + t_r) + \varepsilon_t \tag{13}$$

where y_t is the time series, c is a constant that is followed by a trend t_r, ε_t is the standard error of the regression. The hypothesis is tested against the Lagrange multiplier Λ as follow:

$$\Lambda = \sum_t \frac{R\left(t^2\right)}{N^2 r_0} \tag{14}$$

where r_0 at frequency zero is an estimator of the residual spectrum, R is a summing function of the residuals and N is the sample size. KPSS test is usually to pair with ADF test. By double testing with both the stationarity hypothesis and unit root hypothesis, one can distinguish times series that are sufficiently informative to ascertain whether they are stationary or otherwise. It seems that the KPSS test is not a test that the series is stationary but rather that the residual from a deterministic trend is stationary. These two tests complementarily look for different aspects of stationarity.

After applying the ADF and KPSS tests on the data-sets, it is confirmed that all the data-sets used in the experiments are stationary. The test results fall within the acceptable ranges of statistics (all the alpha values are below zero, and *p*-value is lower than 5%) which point to the hypothesis that the time series are stationary. The stronger rejection of the null hypothesis that there is a unit root when the alpha gets more negative implies the time series is stationary. The trend types being tested are (Const + Trend) and (Const + Trend + Trend2), which are corresponding to (c + t_r) and (c + t_r + t_r2), respectively, as in Equations (9) and (13). Significance level of 5% is used, which is a probability threshold below which the alternative hypothesis will be rejected. The *p*-value is the probability, under which the null hypothesis that test statistics are at least as extreme as observed by KPSS. The stationary test results are tabulated as follow:

	Stock index				REITs			
Test	^AORD	^HSI	^ITLMS	^XU100	Aus	HK	Italy	Turkey
Const + Trend								
α by ADF	−2.2	−2.3	−2.1	−2	−3.7	−2.2	−2.3	−2.4
p-value by KPSS	1.3%	1.1%	1.7%	2.4%	.0%	1.2%	1.1%	.9%
Const + Trend + Trend^2								
α by ADF	−2.2	−2	−2.2	−2.7	−6	−2.4	−2.2	−3
p-value by KPSS	1.5%	2.3%	1.5%	.4%	.0%	.8%	1.4%	.1%

After confirming the time series data are stationary, we adopted the following procedures:

(1) *Stochastic, KD*

Step 1. Calculate the Raw stochastic value (RSV):

$$RSV = \frac{C_n - L_n}{H_n - L_n} \times 100\%$$

where C_n is the closing price of day n; H_n is the highest price within day n; L_n is the lowest price within day n.

Step 2. Calculate the K and D value of the current day:

$$K_n = \frac{1}{3}\mathrm{RSV}_n + \frac{2}{3}K_{n-1}$$

$$D_n = \frac{1}{3}K_n + \frac{2}{3}D_{n-1}$$

If there is no value of the day before the K and D, then K and D will be 50 for the first calculation.

(2) *Moving Average, MA*

$$MA_n = \frac{P_1 + P_2 + \cdots + P_n}{n}$$

where P_n is the previous n days' closing prices.

(3) *Bias*

$$\mathrm{Bias}_n = \frac{C_n - A_n}{A_n}$$

where C_n is the closing price of day n; A_n is the average price for n days.

(4) *Relative Strength Index, RSI*

Step 1. Calculate the UP_n and the DN_n:

$$\mathrm{UP}_n = \frac{\mathrm{SUP}_n}{n}$$

$$\mathrm{DN}_n = \frac{\mathrm{SDN}_n}{n}$$

Step 2. Calculate the RSI of current day:

$$\mathrm{RSI}_n = \frac{\mathrm{UP}_n}{\mathrm{UP}_n + \mathrm{DN}_n} \times 100$$

where SUP_n is the sum of the upward change for previous n days; SDN_n is the sum of the downward change for previous n days.

(5) *Williams' Oscillator, W%R*

$$W\%R = \frac{H_n - C_n}{H_n - L_n} \times 100$$

where C_n is the closing price of day n; H_n is the highest price within day n; L_n is the lowest price within day n.

(6) *Momentum Index, MTM*

$$MTM_n = C_n - PC_n$$

where: C_n is the closing price of day n. PC_n is the closing price for previous n days

(7) *Moving Average Convergence and Divergence, MACD*

Step 1. Calculate the Demand Index (DI):

$$\mathrm{DI} = \frac{H + 2 \times C + L}{4}$$

where H is the highest price of current day; C is the closing price of current day; L is the lowest price of current day.

Step 2. Calculate the different days of exponential moving average

$$\mathrm{EMA(n)} = \mathrm{EMA}(n-1) + \alpha \times (\mathrm{DI}(n) - \mathrm{EMA}(n-1))$$

where $\alpha = 2/(1 + \text{no of days of moving average})$

For example, for the 12 days EMA, the α is equal to $2/(1 + 12)=2/13$

Step 3. Calculate the DIF

$$\mathrm{DIF} = 12 \text{ days EMA} - 26 \text{ days EMA}$$

Step 4. Calculate the MACD

$$\mathrm{MACD}(n) = \mathrm{MACD}(n-1) + \alpha \times (\mathrm{DIF}(n) - \mathrm{MACD}(n-1))$$

where $\alpha = 2/(1 + \text{no of days of moving average})$

3.6. Dynamic Hurst exponent (R/S and V statistics analysis procedure)

With regard to R/S and V statistics calculation procedure, we shall illustrate by using HSI as an example. After we take log of the time series prices data, there are 758 observations. Then, we use the R/S analysis to calculate $\log(R/S)$ of $\log(n)$ and calculate Hurst exponent. The V statistics is the average of R/S divided by the square root of n. We used V_n and $\log(n)$ to plot a V statistics analysis chart, it determines whether aperiodic cycles exist. Once the aperiodic cycles are determined, the dynamic Hurst exponents can be calculated. We then use the dynamic Hurst exponents to do the GMDH network to do forecast. There are 18 input variables in the neural network (Tables 1 and 2).

Table 1. Stock market data-set input variables.

Variables (X_k)		Explanation
X_1–X_5	$P_t, P_{t-1}, P_{t-2}, P_{t-3}, P_{t-4}$	Past 5 days closing price
X_6, X_7	%K, %D	9 days Stochastic Oscillator, KD
X_8, X_9	MA6, MA12	6 days and 12 days Moving Average
X_{10}, X_{11}	6Bias 3Bias	6 days and 3 day Bias Ratio
X_{12}	RSI6	6 days Relative Strength Index
X_{13}	W%R12	12 days Williams %R
X_{14}, X_{15}	MTM6 MTM(MA6)	6 days Momentum Index and its moving average
X_{16}, X_{17}	MACD9 DIF9	9 days Moving Average Convergence / Divergence and difference
X_{18}	Dynamic Hurst Exponent	X days Dynamic Hurst exponent (refer the aperiodic cycle from each data-set)

Table 2. Real estate market data-set input variables.

Variables (X_k)		Explanation
X_1–X_5	$P_t, P_{t-1}, P_{t-2}, P_{t-3}, P_{t-4}$	Past 5 days REITs
X_6, X_7	MA6, MA12	6 days and 12 days Moving Average
X_8, X_9	6Bias 3Bias	6 days and 3 days Bias Ratio
X_{10}	RSI6	6 days Relative Strength Index
X_{11}, X_{12}	MTM6 MTM(MA6)	6 days Momentum Index and moving average
X_{13}	Dynamic Hurst Exponent	X days Dynamic Hurst exponent (refer the aperiodic cycle from each data-set)

Note: Output variables are the next day REITs.

With regard to the input variables of the real estate market data-set, since the real estate market data do not have the open, highest and lowest price, some technical indicators such as Stochastic, W%R, MACD cannot be calculated. Finally there are only 13 input variables presented in the neural network as below (Table 3):

The GMDH shell software neural network retains 80% of the data as the training set with remaining 20% used to test the validity of GMDH neural network model. For example, there are 759 data points in HSI index from September 2010 to September 2013. After excluding the first 68 data for dynamic Hurst Exponent calculation, 691 data points remained. The last 20% (138 data points) of HSI index is then used for forecast validation.

4. Experiments and evaluation

4.1. The dynamic Hurst exponent analysis

Cotter and Stevenson (2008) studied REITs' long memory properties by observing the autocorrelation function (ACF) with plots over 100 lags for the returns series and discovered that there was a lack of dependence which was in line with previous findings of financial returns that were believed to be independently distributed with nearly white noise. The absence of long-range correlations in the walk profile displays properties of random walk with $H = .5$. In contrast, values of $H > .5$ indicate a decrement will be followed by a decrement and an increment is likely to be followed by an increment. If $H < .5$, an increment is very likely to be followed by a decrement, showing the existence of oscillatory behavior (Méndez-Acosta, Hernandez-Martinez, Jáuregui-Jáuregui, Alvarez-Ramirez, & Puebla, 2013).

Granero, Segovia, and Pérez (2008) suggested that market long memory implies the failure of EMH. The Hurst exponent is the most important indicator to confirm the existence of market memory. It can be calculated by *R*/*S* analysis. Shyu, Ke, and Tai (2009) used the *R*/*S* method to calculate the Hurst exponent so that different levels of long memory and aperiodic cycle in each country were found. The empirical results indicated that developing countries' stock markets had larger Hurst exponent and longer aperiodic cycle than the developed stock market. Unlike this paper, we used the aperiodic cycle to calculate the dynamic Hurst exponent, and put back the results to the GMDH network for forecasting.

When Hurst exponent declines suddenly, direction of market prices reverses. The last lowest point of Hurst exponent changes the stock market signal. The dynamic Hurst exponent records the memory of earlier market trend, higher value indicates stronger memory. Nevertheless, when the moving Hurst exponent issued a signal, market price may not reach the lowest and there is high chance to reverse. Investor may "forget" the early market signal

Table 3. Data description.

Data Type	Name	Sample period	No. of data	No. of data actually used	No. of training	No. of testing
Stock Index	^AORD	Sep 2010–Sep 2013	760	692	554	138
	^HIS	Sep 2010–Sep 2013	759	691	553	138
	^ITLMS	Sep 2010–Sep 2013	766	697	558	139
	^XU100	Sep 2010–Sep 2013	753	660	528	132
REITs	Aus	Mar 2010–Mar 2014	1045	941	753	188
	HK	Oct 2011–Mar 2014	588	523	418	105
	Italy	Jul 2012–Mar 2014	421	362	290	72
	Turkey	May 2010–Mar 2014	962	856	685	171

when market is reversed. The corresponding moving Hurst exponent reaches a low value. After the aperiodic cycle was calculated, the dynamic Hurst exponent is shown in the following Figures with dynamic Hurst and ^AORD.

4.2. GMDH neural network simulation result

Neural network with and without dynamic Hurst are trained. Several classical forecasting algorithms (SES, DES and ARIMA) and BPNN compare with the GMDH method. Powered by Oracle Crystal Ball (Release 11, 32-bits) and GMDH shell (3.5.7) was used as an optimized solver. Weka machine learning software (3.7.11) was used to install the time series forecasting package for back-propagation in neural network[4]. With regard to SES, DES, ARIMA, we only used the output variable for forecasting. BPNN used 80% of the data for training and 20% for testing, 500 epochs were used to train through one hidden layer as only one variable was included in our study. The GMDH model includes the technical indicators to the neural network for time series prediction. There are 17 technical indicators (X_1–X_{17}), X days dynamic Hurst exponent (X_{18}) for the stock market data-set, 12 technical indicators (X_1–X_{12}) and X days dynamic Hurst exponent (X_{13}) in real estate market data (Table 4).

We evaluate the forecasting accuracy by using Mean Absolute Percentage Error (MAPE). Similar to most of the neural network forecasting method, data in machine learning like GMDH neural network divide the data into training and validation set for forecasting. Training aimed to reduce the error on the training set, it was terminated when the validation set's error began to rise (Hong & Fan, 2016). GMDH shell uses 80% of data for training and the remaining 20% to validate the GMDH neural network model. Mean Absolute Percentage Error (MAPE) calculates the average absolute error between actual and forecast values and it is valid measure to study the forecast accuracy of the forecasting algorithm in percentage (Ghasemi, Shayeghi, Moradzadeh, & Nooshyar, 2016). The equation of MAPE is:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_{t\_}F_t}{A_t}\right| \times 100\%$$

where A_t is the actual value and F_t is the forecast value.

After training neural network by GDMH shell, results are compared with traditional forecasting algorithms as shown in Table 9. GMDH neural network recorded the smallest MAPE from the last 20% testing data among all the REITs and stock indices. GMDH neural network with dynamic Hurst exponent improved the forecasting results in Australia, Hong Kong, and Turkey stock indices since it had the lowest MAPE (error). In REITs, GMDH

Table 4. Descriptions of various forecasting method.

Abbreviation	Descriptions of various forecasting method
SES	Single Exponential Smooth, optimized by Crystal Ball
DES	Double Exponential Smooth, optimized by Crystal Ball
ARIMA	ARIMA model, optimized by Crystal Ball
BPNN	Back-propagation neural network, experimented by Weka time series analysis
GMDH with Dynamic Hurst	Use Table3.1 X_1–X_{17} or Table3.2 X_1–X_{12} and X days/periods dynamic Hurst exponent in GMDH Neural Network, optimized by GMDH shell
GMDH without Dynamic Hurst	Use Table3.1 X_1–X_{17} or Table3.2 X_1–X_{12} in GMDH Neural Network, optimized by GMDH shell

neural network with dynamic Hurst exponent improves the forecasting results of Hong Kong, Italy, and Turkey REITs. One interesting result showed that despite Italy stock index was closest to random walk according to Table 6 (Hurst exponent at .5028), GMDH with Hurst forecasting method still outperformed the other methods.

Figures 6–13 show the forecast time series plot according to classical model and GMDH neural network model. The red line (GMDH with dynamic Hurst) and green dashed line (GMDH without dynamic Hurst) show predicted values.

4.3. Brief comments and overall performance

In Hong Kong Heng Seng index, Hurst exponents are significantly greater than .5, i.e. their volatility was persistent and could be predicted by historical data. In Australia and Turkey, the Hurst exponent was significantly less than .5, i.e. their volatility has anti-persistent and also can be predicted by historical data. Nevertheless, Italy stock index is not significantly different from .5, i.e. the dynamic Hurst exponent does not affect the GMDH neural network.

Besides, Hong Kong and Turkey REITs' Hurst exponents are significantly higher than .5 which means their volatility can be predicted by historical data. The Australia and Italy Hurst exponent are significantly less than .5, which means that their volatility has anti-persistent characteristic which can be predicted by historical data. The GMDH neural network has the best result with the lowest MAPE value in all REITs data. Besides, the GMDH neural network with dynamic Hurst exponent is better than pure GMDH neural network. GMDH neural network without dynamic Hurst exponent achieves the lowest MAPE in Australia real estate market. The Hurst exponent analysis has extremely important research value.

Table 6 shows the best method of forecasting according to the lowest MAPE. GMDH neural network is the best with the lowest MAPE for all the time series data. Therefore, GMDH neural network model helps analyze the stock and real estate market. Moreover, dynamic Hurst exponent improves the forecasting accuracy in most of the research models, such as Australia, Hong Kong, and Turkey stock indices time series and Hong Kong, Italy and Turkey REITs indices time series. If the Hurst exponent is not significantly different from .5, such as Italy stock market, then the dynamic Hurst has no effect. Furthermore, if the Hurst exponent is significantly different from .5, but the difference of Hurst between V_n is too small, such as Australia REIT data, using dynamic Hurst exponent cannot improve the GMDH neural network, other than enhancing the GMDH neural network result.

Type A: Hurst exponent >.5 indicates a decrement will be followed by a decrement and an increment is likely to be followed by an increment in the future.

Type B: If Hurst exponent <.5, an increment is very likely to be followed by a decrement, showing an oscillatory behavior.

Type C: close to random walk

5. Implication of the study

Our results prove that GMDH neural network achieve better forecasts results than the traditional DES, SES and ARIMA methods among all the stock and REITs indices. Moreover, dynamic Hurst exponent in GMDH neural network has smaller MAPE then we only use GMDH neural network method in Australia, Hong Kong, and Turkey stock indices prediction validation. The results suggest that the investors can use the GDMH neural network to increase forecasting accuracy of REITs and stocks prices. As investors can use this information to forecast the changes in stock and REITs prices that can help investors reap more profit. It is also of academic value as this method is the first of its kind which applied GMDH with Hurst in REITs and stock prices' forecasting.

6. Long-term memory of stocks and REITs

Second, with regard to the long-term memory in stocks and REITs, the past and future data volatility is irrelevant for future prediction if the Hurst exponent is .5. The time series is in random walk. If the value ranges from .5 to 1, previous high price often lead to high price in the future. Nevertheless, if the Hurst exponent falls below .5, that implies there is a switch in high and low price in the long run.

In our research study, the Hurst exponent of Hong Kong Hang Seng index is .5314: if the last day's stock price rose, the probability of this stock price rising in the future will be 53.14%. Similarly, the Hurst exponent of Australia All Ordinaries Index is .4805, which means that if the last index rises, probability of this index dropping in the future is 51.95% (= 1−.4805). The Hurst exponent of Turkey Borsa Istanbul 100 Index is .4708, which means that if the last index record rises, probability of this index dropping in the future is 52.92% (= 1−.4708).

The Hurst exponent of Hong Kong REIT index is .5948, which means that if the previous day's Hong Kong REIT rises, the probability of Hong Kong REIT continuing to rise in future is 59.48%. The Hurst exponent of Turkey REIT index is .5183, which means that if the previous day's Turkey REIT rises, the probability of Turkey REIT continuing to rise in future is 51.83%. The Hurst exponent of Australia REIT is .4395, which means that if the last REIT rises, the probability of REIT value dropping in the future is 56.05% (= 1−.4395). And the Hurst exponent of Italy REIT is .4243, which means that if the last REIT rises, the probability of REIT value dropping in the future is 57.57% (= 1−.4243).

The existence of fractional Brownian suggests that changes in the present stock / REITs prices depend on the previous records in prices. Our *R*/*S* analysis method shows that majority of the stocks and REITs prices prove the existence of fractional Brownian motion which reject EMH except Italy stock index. We expect that the difficulty in forecasting the Italy stock prices can be explained by the characteristics of stock prices changes in the Hurst exponent results. The Hurst exponent of Italy stock is far below .5 that means the previous drop often leads to an increase in prices and vice versa. The ever-changing prices imply that (1) it is closer to random walk without a clear trend, (2) long-term memory does not exist. All these increase the difficulty in forecasting. That also coincides with our GMDH neural network simulation results: if the times series is less volatile, the dynamic Hurst exponent has the best forecasting results. Besides, whether the forecasting series are stocks or REITs indices however, may not be an important factor that affects the forecasting accuracy.

7. Conclusion

In this paper, we aim to study the existence of long-term memory in stocks and test if GMDH neural network with Hurst have better forecasting ability as compared to traditional forecasting methods. We also compare and contrast the differences between REITs and stock indices. The results showed that GMDH neural network with Hurst provided the lowest forecasting errors among the majority of REITs and stocks time series except REITs indices in Australia. Our research also confirmed that long-term memory existed for most of the indices except Italy stock index which was close to random walk with Hurst exponent record of .5028. While Australia and Hong Kong's previous increment's impact on the next increment were the same in REITs and stocks indices, Italy and Turkey was different according to Hurst exponent's results. It implies that the long-term co-memory in REITs and stocks may not be the same in every single circumstance despite previous research in Canada was the same for REITs and stocks. Despite Italy stock index showed that it was closest to random walk (least deviation from .5 in Hurst exponent), GMDH with Hurst forecasting method still outperformed the other methods. The results provide an interesting future study direction on whether GMDH with Hurst can be used to forecast the time series in the so-called random walk nature.

Notes

1. http://www.gmdhshell.com/
2. https://finance.yahoo.com/
3. https://www.google.com/finance
4. http://wiki.pentaho.com/display/DATAMINING/Time+Series+Analysis+and+Forecasting+with+Weka

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Assaf, A. (2006). Canadian REITs and stock prices: Fractional cointegration and long memory. *Review of Pacific Basin Financial Markets and Policies, 9*, 441–462.

Australia Shareholders Associations. (2014). https://www.australianshareholders.com.au/resources/areit-sector-view-and-outlook-2014

Burger, C. J. S. C., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide to time-series methods for tourism demand forecasting — A case study of Durban, South Africa. *Tourism Management, 22*, 403–409.

Cadenas, E., Jaramillo, O. A., & Rivera, W. (2010). Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. *Renewable Energy, 30*, 925–930.

Cheung, Y. W. & Lai, K. S. (1995). A search for long memory in international stock market returns. *Journal of International Money and Finance, 14*, 597–615.

Cotter, J., & Stevenson, S. (2008). Modeling long memory in REITs. *Real Estate Economics, 36*, 533–554.

Crawford, G. W., & Fratantoni, M. C. (2003). Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices. *Real Estate Economics, 31*, 223–243.

Dase, R. K., & Pawar, D. D. (2010). Application of Artificial Neural Network for stock market predictions: A review of literature. *International Journal of Machine Intelligence, 2*, 14–17.

DataStream. (2016). http://financial.thomsonreuters.com/en/products/tools-applications/trading-investment-tools/datastream-macroeconomic-analysis.html

Dosset, P., Rassam, P., Fernandez, L., Espenel, C., Rubinstein, E., Margeat, E., & Milhiet, P. E. (2016). Automatic detection of diffusion modes within biological membranes using backpropagation neural network. *BMC Bioinformatics, 17*, 111–208.

Ellis, C. A., & Parbery, S. A. (2005). Is smarter better? A comparison of adaptive, and simple moving average trading strategies. *Research in International Business and Finance, 19*, 399–411.

Faggini, M., & Parziale, A. (2012). The failure of economic theory. Lessons from chaos theory. *Modern Economy, 3*(1), 1–10.

Fama, Eugene F. (1965). Random walks in stock market prices. *Financial Analysts Journa, 21*, 55–59.

Fong, S., Nannan, Z., Wong, R. K., & Yang, X. S. (2012). *Rare events forecasting using a residual-feedback GMDH neural network*. Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on Data Mining Workshops (pp. 464–473), IEEE Computer Society Washington, DC, USA.

Gardner, E. S., Anderson-Fletcher, E. A., & Wicks, A. M. (2001). Further results on focus forecasting vs. exponential smoothing. *International Journal of Forecasting, 17*, 287–293.

Ghasemi, A., Shayeghi, H., Moradzadeh, M., & Nooshyar, M. (2016). A novel hybrid algorithm for electricity price and load forecasting in smart grids with demand-side management. *Applied Energy, 177*, 40–59.

Granero, M. A. S., Segovia, J. E. T., & Pérez, J. G. (2008). Some comments on Hurst exponent and the long memory processes on capital markets. *Physica A, 387*, 5543–5551.

Holt, C. C. (2004). Forecasting trends and seasonal by exponentially weighted averages. *International Journal of Forecasting, 20*, 5–10.

Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting, 32*, 914–938.

Huang, J. H. (2011). *A study of commodity trading advisor performance prediction-an analysis of chaos and artificial neural network*. Taoyuan: Department of Bussiness Administrator, Chung Yuan Christian University, Taiwan.

Huang, H.-S., & Chiu, I.-H. (2005). *The study of neural network to predict Taiwan ETF-50 stock index price*. Changhua: Department of Information Management, National Changhua University of Education.

Hurst, H. E. (1951). Long term storage capacities of reservoirs. *Transactions of the American Society of Civil Engineers, 116*, 776–808.

Ivakhnenko, A. G. (1970). Heuristic self-organization on problems of engineering cybernetics. *Automatica, 6*, 207–219.

Kendall, Maurice. (1953). *The analytics of economic time series, part 1: Prices*. London: London School of Economics, Division of Research Techniques.

Kimoto, T., & Kazuo, A. (1990). Stock market prediction system with modular neural networks. *IEEE International Joint Conference on Neural Networks,* 11–16.

Kuhle, J. L. (1987). Portfolio diversification and return benefits – Common stocks vs. real estate investment trusts. *Journal of Real Estate Research, 2*(2), 1–9.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shim, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics, 54*, 159–178.

Lapedes and Farber. (1987). *Advances in neural information processing systems*. How Neural Nets Work (pp. 442–456). New York, NY: American Institute of Physics.

LeRoy, S. F., & Porter, R. D. (1981). The present-value relation: Tests based on implied variance bounds. *Econometrica, 49*, 555–574.

Li, R. Y. M. & Chau, K. W. (2016). *Econometric analyses of international housing markets*. Routledge.

Ling, D. C., Naranjo, A., & Ryngaert, M. D. (2000). The predictability of equity REIT returns: Time variation and economic significance. *The Journal of Real Estate Finance and Economics, 20*, 117–136.

Makridakis, S., & Hibon, M. (2000). The M3-competition: Results, conclusions and implications. *International Journal of Forecasting, 16*, 451–476.

Méndez-Acosta, H. O., Hernandez-Martinez, E., Jáuregui-Jáuregui, J. A., Alvarez-Ramirez, J., & Puebla, H. (2013). Monitoring anaerobic sequential batch reactors via fractal analysis of pH time series. *Biotechnology and Bioengineering, 110*, 2131–2139.

Mitra, S. K. (2012). *Is Hurst exponent value useful in forecasting financial time series*. Nagpur: Institute of Management Technology.

Newell, G., & Osmadi, A. (2009). The development and preliminary performance analysis of Islamic REITs in Malaysia. *Journal of Property Research, 26*, 329–347.

Newell, G., & Peng, H. W. (2012). The significance and performance of Japan REITs in a mixed-asset portfolio. *Pacific Rim Property Research Journal, 18*, 21–34.
Pan, G. R., & Gu, C. (2007). New method for deformaion prediction. *Journal of Guilin University of Technology, 27*, 529–532.
Pavlova, I., Cho, J. H., Parhizgari, A. M., & Hardin, W. G., III (2014). Long memory in REIT volatility and changes in the unconditional mean: A modified FIGARCH approach. *Journal of Property Research, 31*, 315–332.
Peters, Edgar E. (1994). *Fractal market analysis: Applying chaos theory to investment and economics, [M]*. New York, NY: Wiley.
Pierdzioch, C., & Hartmann, D. (2013). Forecasting Eurozone real-estate returns. *Applied Financial Economics, 23*, 1185–1196.
Qian, B., & Rasheed, K. (2004). *Hurst exponent and financial market predictability*. Athens: Department of Computer Science University of Georgia.
Real Estate Easy Property Info. (2016). http://reitinfo.com/
Sah, V., Zhou, X., & Das, P. K. (2015). Does index addition add any new information? Evidence from REIT dividend forecasts. *Journal of Property Research, 32*, 33–49.
Shyu, S. D., Ke, S. C., & Tai, M. L. (2009). A study of long memory in international stock price indices. *Bank of Taiwan Quarterly, 60*, 277–299.
Wang, K., Erickson, J., & Chen, S. H. (1995). Does the REIT stock market resemble the general stock market? *Journal of Real Estate Research, 10*, 445–460.
Wang, X. Y., Song, X. F., & Wu, R. M. (2004). Fractal analysis of China stock markets. *Journal of Management sciences in China, 7*, 1–8.
White, Halbert. (1988). *Economic prediction using neural networks: The case of IBM daily stock returns*. San Diego: Department of Economics University of California.
Wu, L., Liu, S., & Yang, Y. (2016). Grey double exponential smoothing model and its application on pig price forecasting in China. *Applied Soft Computing, 39*, 117–123.
Xiao, T., & Huang, G. J. (2000). GMDH neural network application to prediction. *Journal of East China Shipbuilding Institute, 14*, 72–76.
Yao, J., Tan, C. L., & Poh, H. L. (1999). Neural networks for technical analysis: A study on klci. *International Journal of Theoretical and Applied Finance, 2*, 221–241.
Zhao, S. J., & Xu, B. Z. (2011). Trend prediction for stock price using dynamic hurst index. *Journal of Ningbo University (Natural Science and Engineering Edition), 24*(4), 79–82.
Zhou, J., & Kang, Z. (2011). A comparison of alternative forecast models of REIT volatility. *Journal of Real Estate Finance and Economics, 42*, 275–294.
Zuang, X. T., Zhuang, X. l., & Tian, Y. (2003). Hurst index and problem of fractal structure in stock market. *Journal of Northeastern University (Natural Science), 24*(9), 79–82.

Appendix

The Hurst exponent analysis

After calculating R/S, all data-set's Hurst exponents are estimated. According to these Hurst values, the correlation function $C(t) = 2^{2H-1}-1$ and fractal dimension $D = 2-H$ (Table 5).

In Table 5, the Hurst exponents of Hong Kong Hang Seng Index is significantly greater than .5. The Italy FTSE Italia All Share Index is not significantly different from .5. The Hurst exponents of Australia All Ordinaries Index and Turkey Borsa Istanbul 100 Index are significantly less than .5, i.e. the Australia and Turkey stock market index exhibits anti-persistence fractional Brownian motion. The upward trend in Australia and Turkey stock market may have decreasing trend in future. All data except the Italy FTSE Italia All Share Index are significantly different from .5 according to the T statistic, i.e. they do not perform according to EMH .

Regression between $\log(R/S)$ and $\log(n)$'s R-square and adjusted R-square achieve over than 90% explanatory power. The correlation function $C(t)$ tests the independence in time series. The fractal dimension D examines whether the time series is random. When $H = .5$, correlation function $C(t)$ is equal to zero, fractal dimension D is equal to 1.5. The Hurst of Australia All Ordinaries Indices and Turkey Borsa Istanbul 100 Indices are less than .5, i.e. the correlation function $C(t)$ is negative, The fractal dimension D is greater than 1.5. Other correlation function $C(t)$ are positive, the fractal dimension D is less than 1.5. Italy stock market's data are random in nature, because their correlation coefficient is .00387, Australia, Hong Kong ,and Turkey stock market are less random with correlation coefficients at −.02674, .04455, and −.03968. In REITs, Turkey REITs data are more random in nature, because their correlation coefficient is −.02568, Australia, Hong Kong, and Turkey REITs data are less independent and random, because their correlation coefficients are −.08040, .14037, and −.09957.

Aperiodic cycle analysis

All stock market and REITs' results are shown in Figures 14–21. The turning point of V_n can be found from the chart. When V statistic mutates, the V_n line has become horizontal or downward trend from upward trend. It means that the original memory at this point disappeared. By checking each interval V_n's turning point, the length of aperiodic cycle can be estimated.

^AORD data-set displays a significant drop of V_n at $n = 69$. The V statistics increases as n increases. Let $n = 68$ as the demarcation, we calculate the Hurst exponent for each segment. The Hurst exponent before the 69 days is .57094 and after the 69 days is .46641. Therefore, $n = 68$ denotes long-term memory vanishing point of the time series data. After 68 days, the Hurst exponent is smaller (Figure 22). In HSI, there is a significant drop of V_n at $n = 69$. The Hurst exponent before 69 days is .60679 and after 69 days is .50089 (Figure 23). In ITLMS, there is significant drop of V_n at $n = 70$. The Hurst exponent before 70 days is .57916 and after 70 days is .46417 (Figure 24). In XU100 data-set, there is a significant drop of V_n at $n = 94$. The Hurst exponent before 94 days is .61906 and after 94 days is .34100 (Figure 25).

The Australia REITs experiences a significant drop of V_n at $n = 105$. The Hurst exponent before and after 105 days is .49232 and .49041, respectively (Figure 26). Hong Kong REITs' V statistics experiences a significant drop of V_n at $n = 66$ as n decrease. The Hurst exponent before and after 66 days is .66780 and .57428 (Figure 27). Italy REITs experiences a significant drop phenomenon of V_n at $n = 60$. The Hurst exponent before and after 60 days is .58255 and .24587 (Figure 28). Turkey REITs recorded

Table 5. All methods' MAPE result of last 20% forecast validation.

	SES	DES	ARIMA	BPNN	GMDH with Hurst	GMDH without Hurst
^AORD MAPE	2.2674	7.0280	2.2721	8.1226	**.7161**	.7381
^HSI MAPE	6.4349	6.8164	6.4351	3.0215	**.8585**	.8706
^ITLMS MAPE	3.9009	3.8642	3.9084	6.0959	**1.0581**	**1.0581**
^XU100 MAPE	7.2180	7.6464	7.2201	8.8756	**1.3970**	1.6292
Aus MAPE	3.1314	4.4299	3.2208	1.8204	.5343	**.5336**
HK MAPE	2.2060	2.0387	2.2134	1.1579	**.3078**	.3495
Italy MAPE	9.9282	9.8251	9.9246	7.4970	**1.5836**	1.7162
Turkey MAPE	8.6897	8.9522	8.6886	6.0488	**1.1468**	1.2454

Note: Figures in bold refers to forecasting methods with the lowest errors.

Table 6. Summary result of all data-set.

Types of data	Data	Hurst exponent	Implications	Aperiodic cycle	Best methods
Stock indices	^AORD (Australia)	.4805	B	68	GMDH with Hurst
	^HIS (Hong Kong)	.5314	A	68	GMDH with Hurst
	^ITLMS (Italy)	.5028	C	69	GMDH with / without Hurst
	^XU100 (Turkey)	.4708	B	83	GMDH with Hurst
REITs Indices	Aus	.4395	B	104	GMDH without Hurst
	HK	.5948	A	65	GMDH with Hurst
	Italy	.4243	B	59	GMDH with Hurst
	Turkey	.5183	A	106	GMDH with Hurst

Table 7. The Hurst exponent results.

	Intercept	Hurst exponent	R-square	Adjusted R-square	Correlation coefficient	Fractal dimension
^AORD	.067668	.480450	.918065	.917845	−.02674	1.51955
^HIS	.06601	.531442	.925218	.925017	.04455	1.46856
^ITLMS	.005630	.502788*	.960819	.960714	.00387	1.49721
^XU100	.194728	.470794	.897971	.897695	−.03968	1.52921
Aus	.10798	.439538	.826926	.826590	−.08040	1.56046
HK	−.22276	.594751	.965539	.965419	.14037	1.40525
Italy	.21645	.424341	.800477	.799498	−.09957	1.57566
Turkey	.03358	.518287	.959115	.959029	.02568	1.48171

*Cannot reject the null hypothesis.

Table 8. The aperiodic cycle of different data-set.

Data	Hurst before V_n	Hurst after V_n	Difference
AORD	.57094	.46641	.10453
HSI	.60679	.50089	.10590
ITLMS	.57916	.46417	.11500
XU100	.61906	.34100	.27806
Aus	.49232	.49041	.00191
HK	.66780	.57428	.09352
Italy	.58255	.24587	.33668
Turkey	.61559	.48264	.13295

Table 9. The aperiodic cycle of different data-set.

Stock			Real estate		
Index	Hurst exponent	Aperiodic cycle	REITs	Hurst exponent	Aperiodic cycle
AORD	.480450	68	Aus	.439538	104
HSI	.531442	68	HK	.594751	65
ITLMS	.502788	69	Italy	.424341	59
XU100	.470794	93	Turkey	.518287	106
Average	.506326	74.5	Average	.597016	83.50

a significant dropped of V_n at $n = 107$. The Hurst exponent before and after 107 days is .61559 and .48264. After 107 days, the Hurst exponent is smaller (Figure 29 and Table 6).

In summary, there is aperiodic cycle from the stock market and real estate market data. The average aperiodic cycle of stock market is 63.75 days and the average aperiodic cycle of REITs market is 83.5 days (Table 7–9).

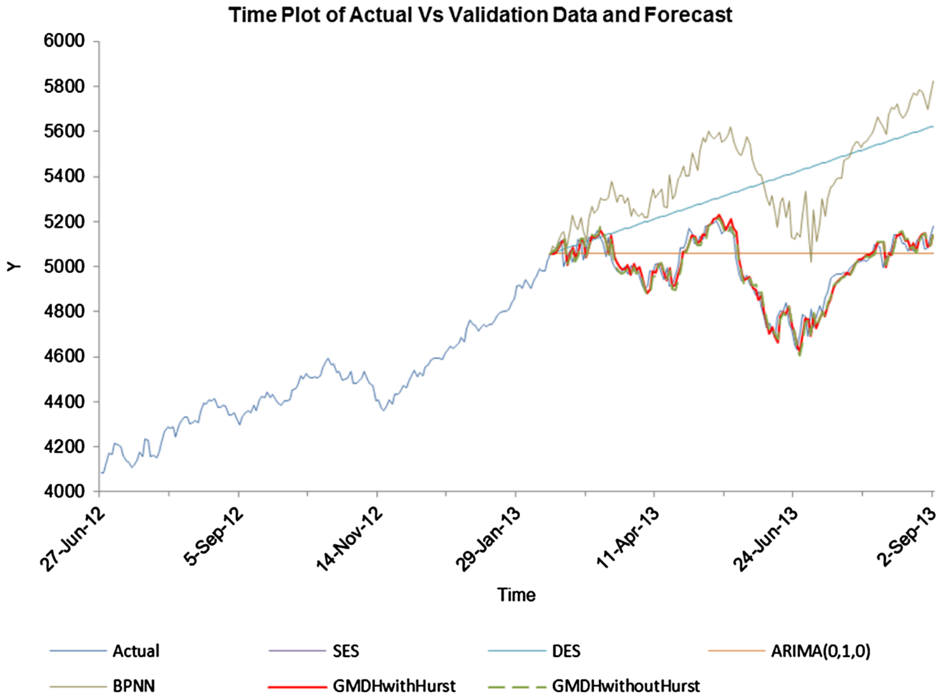

Figure 6. ^AORD prediction comparison.

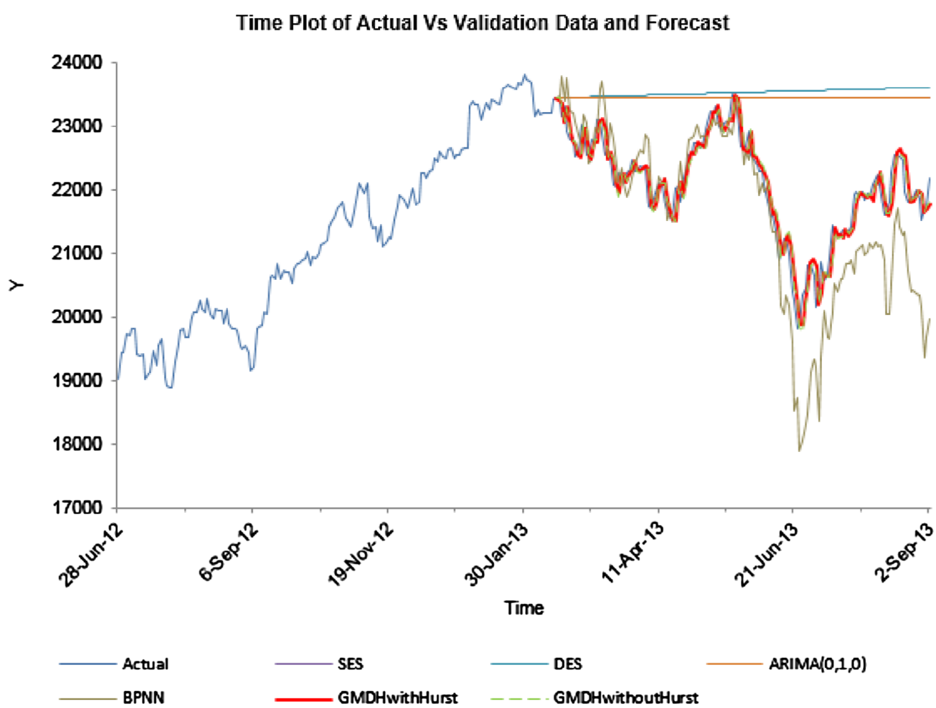

Figure 7. ^HSI prediction comparison.

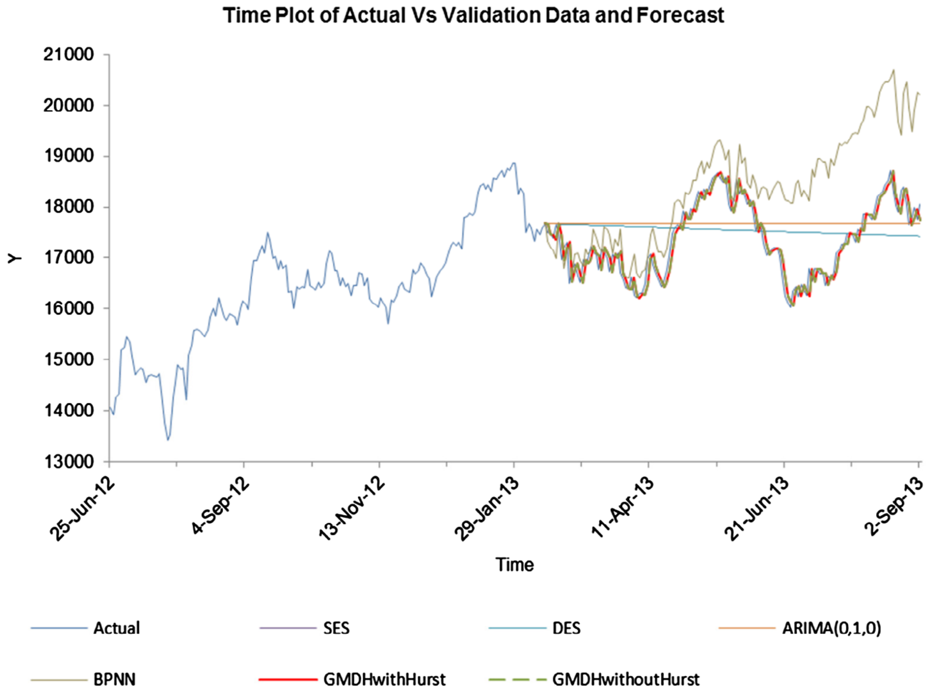

Figure 8. ^ITLMS prediction comparison.

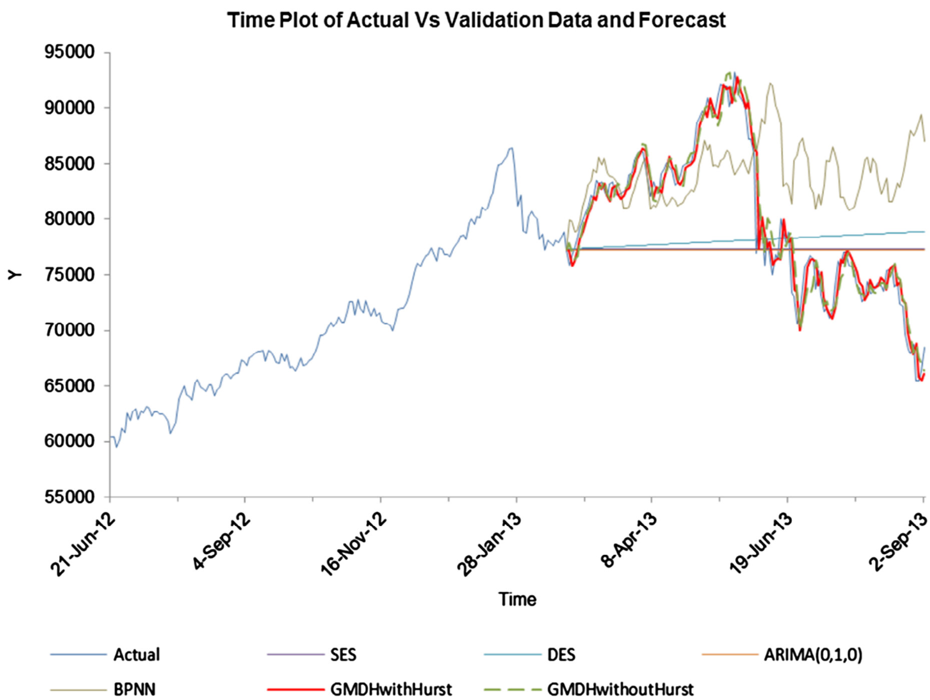

Figure 9. ^XU100 prediction.

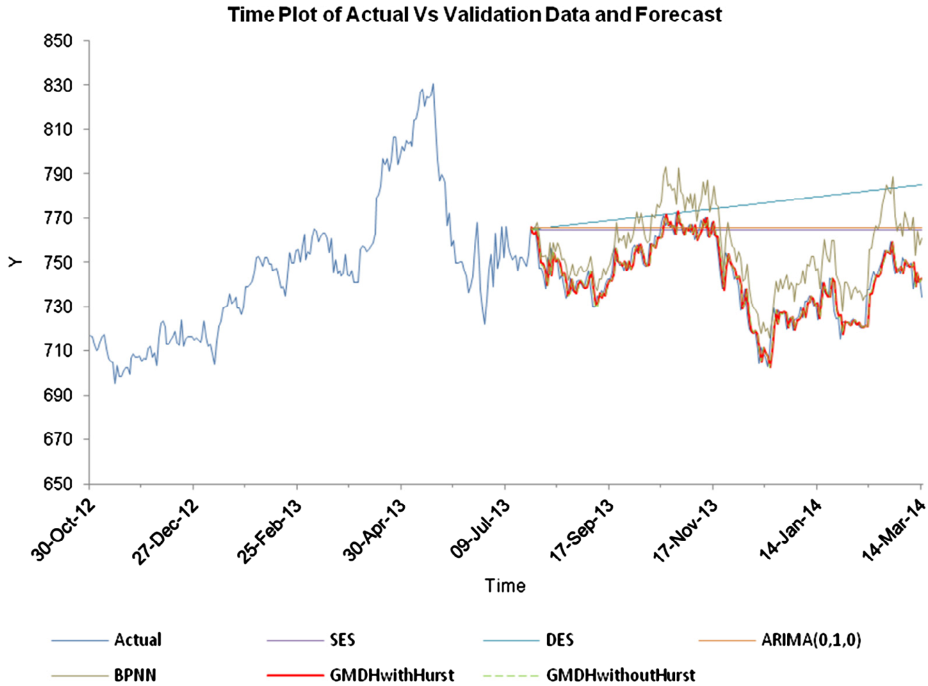

Figure 10. Australia REIT index prediction comparison.

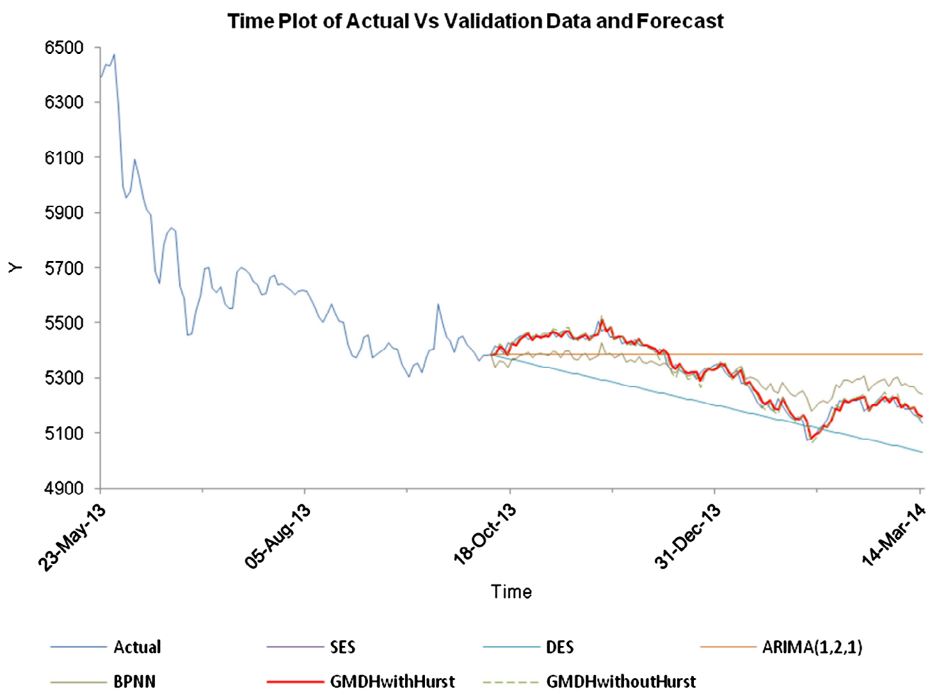

Figure 11. Hong Kong REIT index prediction comparison.

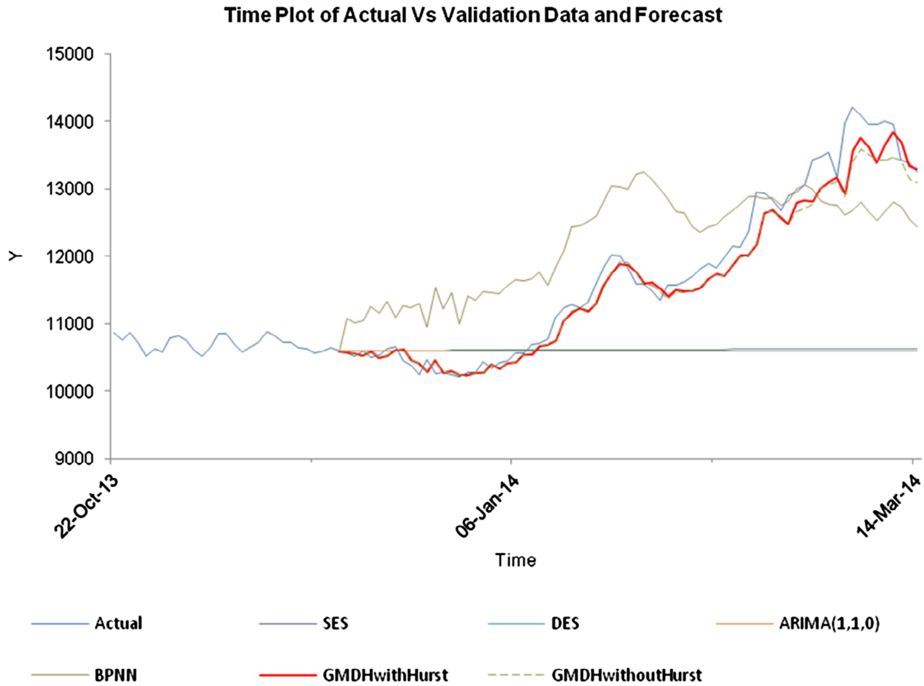

Figure 12. Italy REIT index prediction comparison.

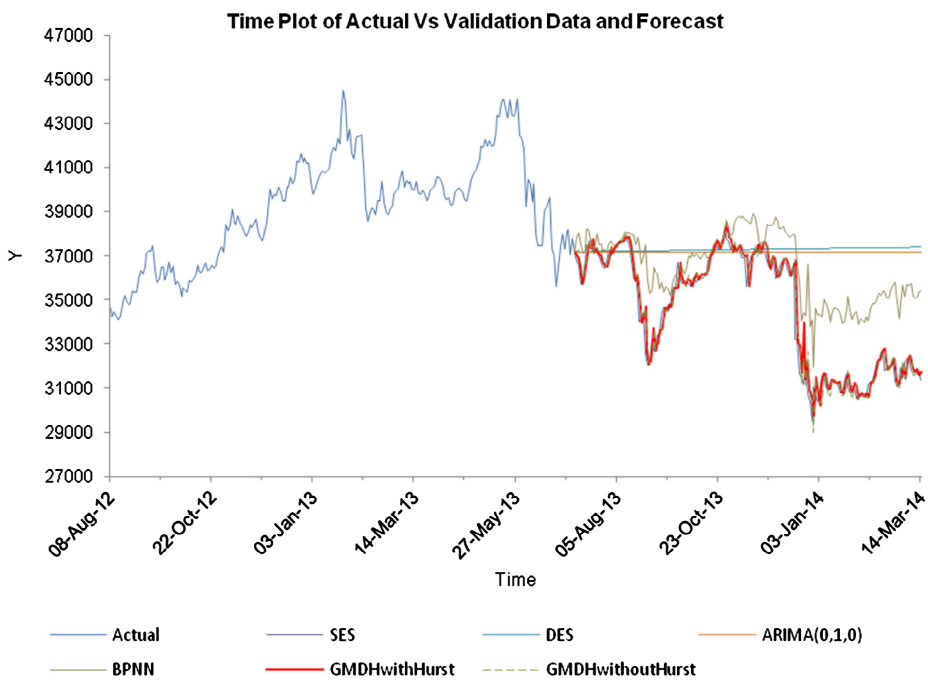

Figure 13. Turkey REIT index prediction comparison.

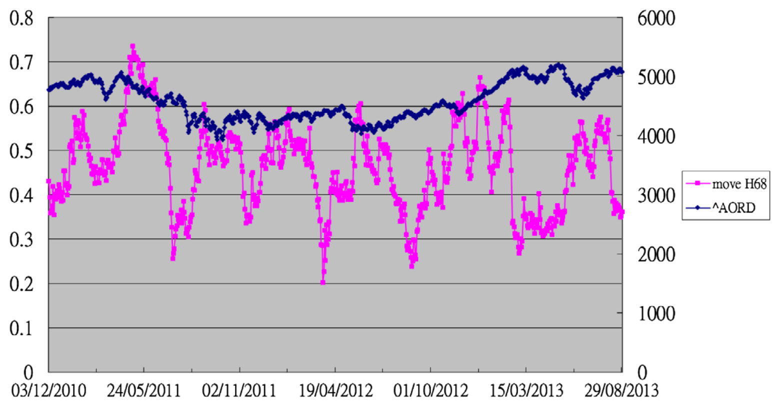

Figure 14. The ^AORD time series with the 68 days dynamic Hurst.

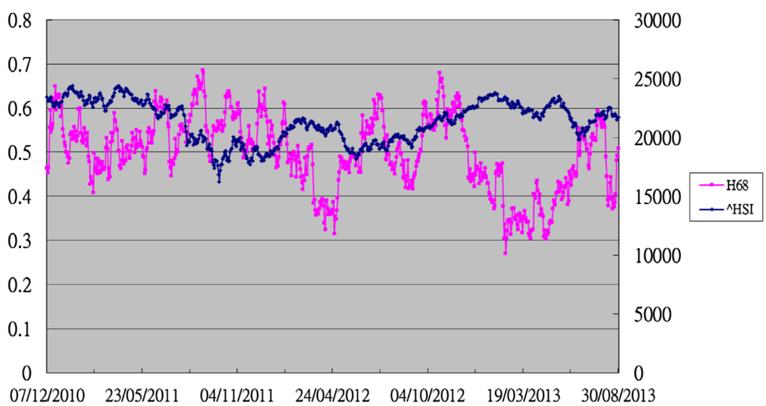

Figure 15. The ^HSI time series with the 68 days dynamic Hurst.
Note: Fixed $n = 68$, the dynamic Hurst and ^HSI are plotted.

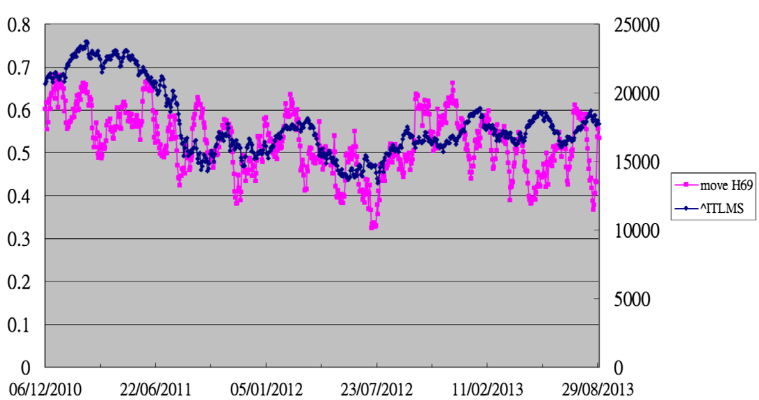

Figure 16. The ^ITLMS time series with the 69 days dynamic Hurst.
Note: Fixed $n = 69$, the dynamic Hurst and ^ITLMS are plotted.

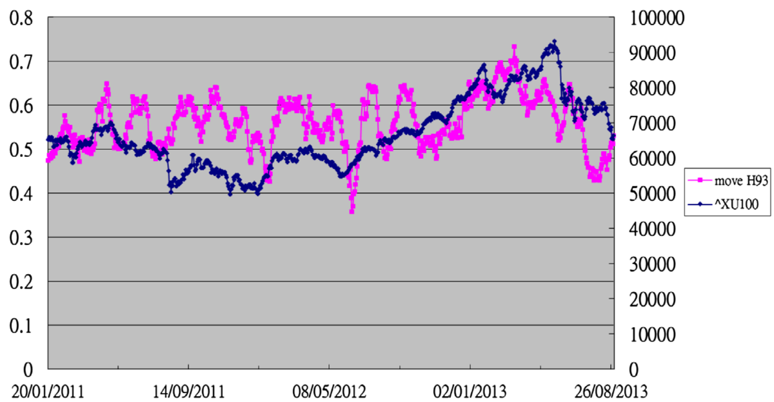

Figure 17. The ^XU100 time series with the 93 days dynamic Hurst.
Note: Fixed $n = 93$, the dynamic Hurst and ^XU100 are plotted.

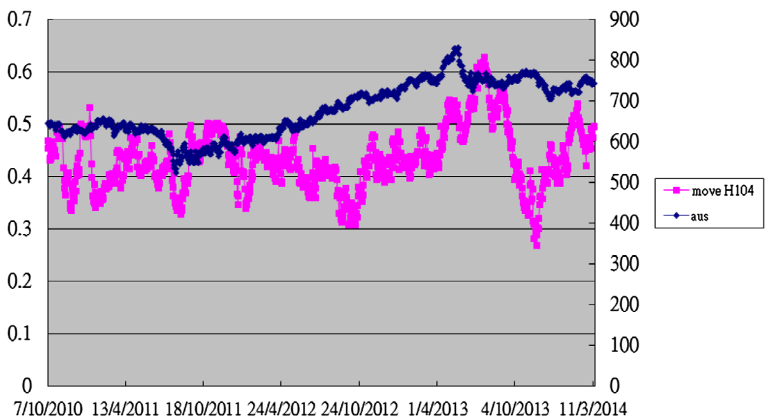

Figure 18. Australia REIT index time series with the 104 days dynamic Hurst.
Note: The corresponding dynamic Hurst with REITs dataset are plotted.

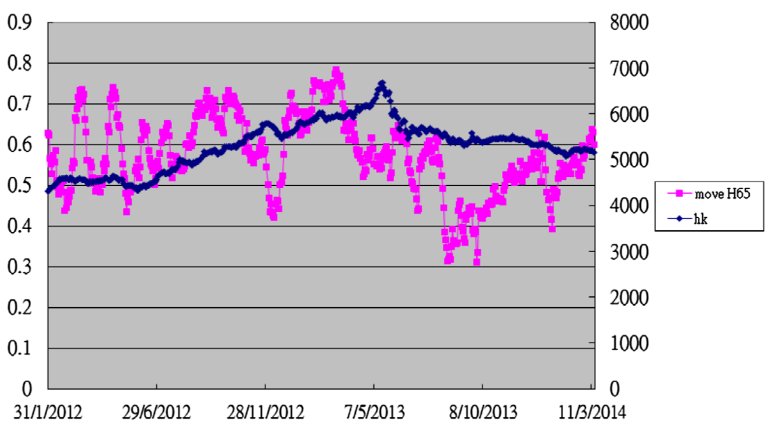

Figure 19. Hong Kong REIT index time series with the 65 days dynamic Hurst.

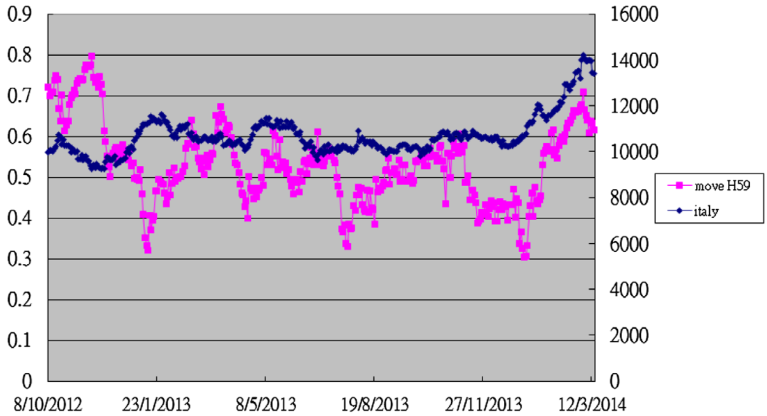

Figure 20. Italy REIT index time series with the 59 days dynamic Hurst.

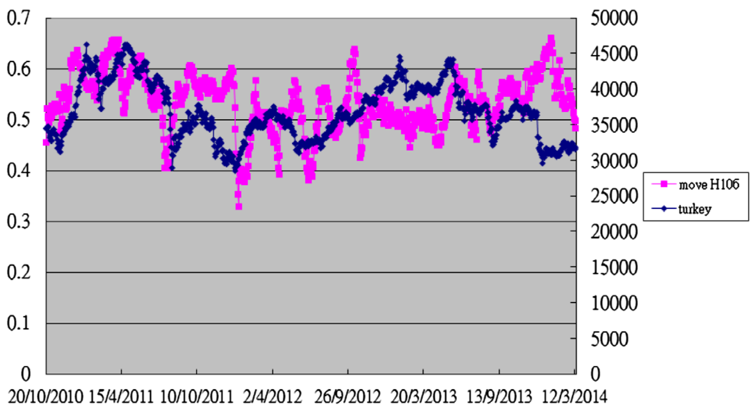

Figure 21. Turkey REIT index time series with the 106 days dynamic Hurst.

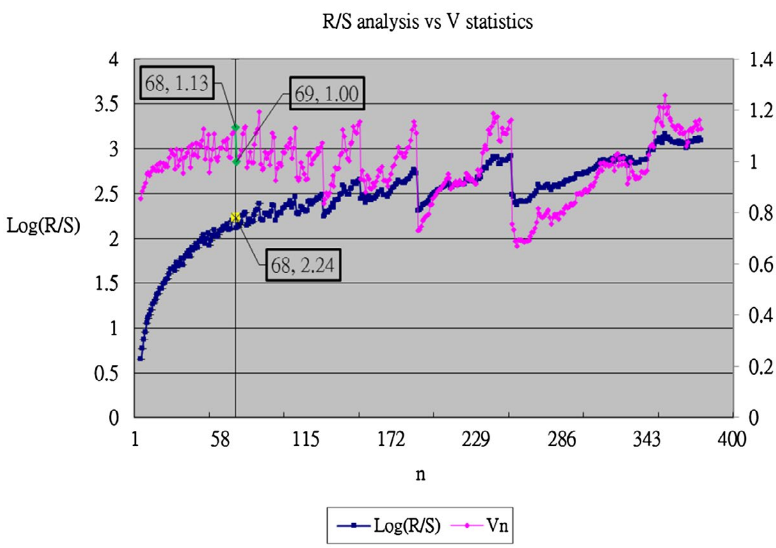

Figure 22. The Australia All Ordinaries Index *R/S* analysis vs. *V* statistics Chart.

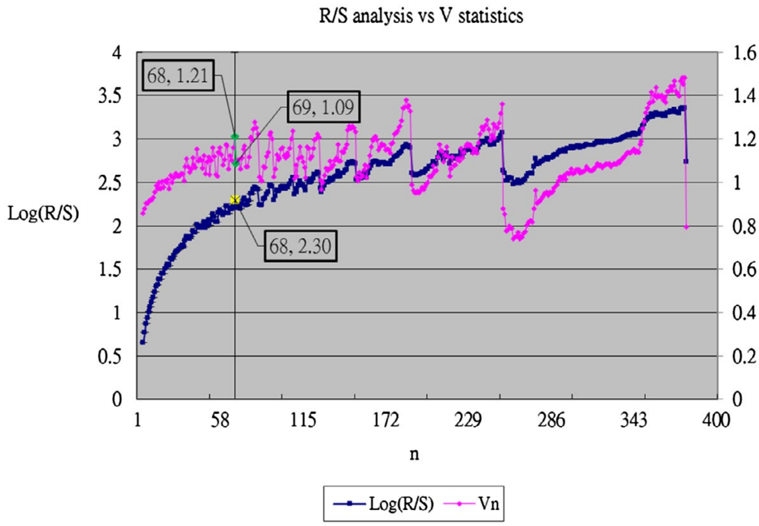

Figure 23. The Hong Kong Hang Seng Index *R/S* analysis vs. *V* statistics Chart.

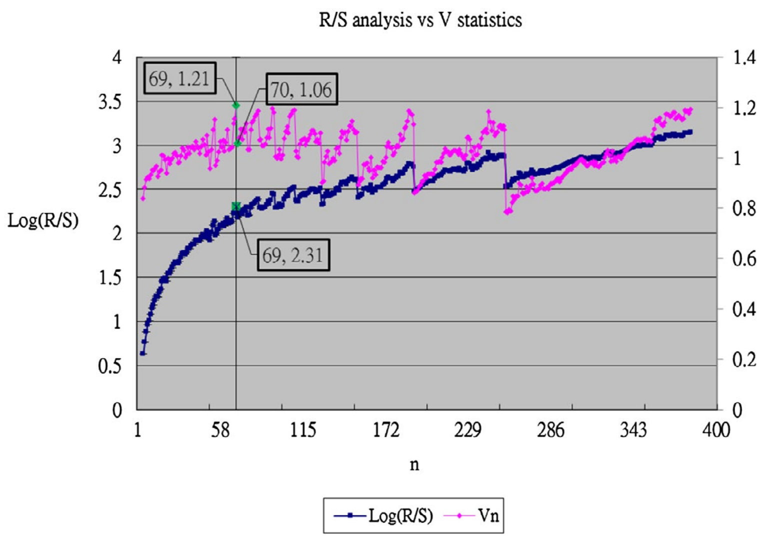

Figure 24. The Italy FTSE Italia All Share Index *R/S* analysis vs. *V* statistics Chart.

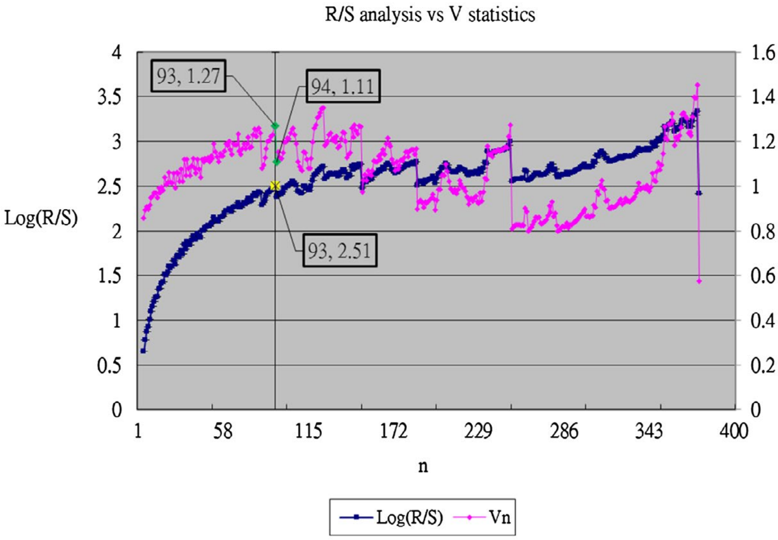

Figure 25. The Turkey Borsa Istanbul 100 Index *R/S* analysis vs. *V* statistics Chart.

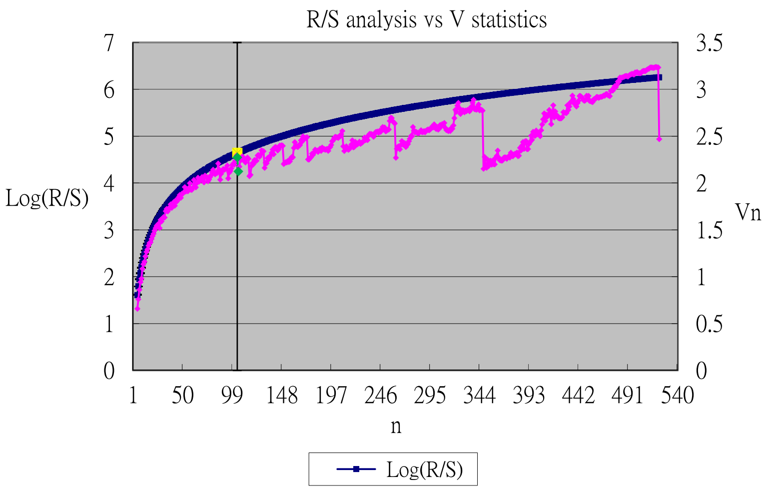

Figure 26. Australia REIT index *R/S* analysis vs. *V* statistics Chart.

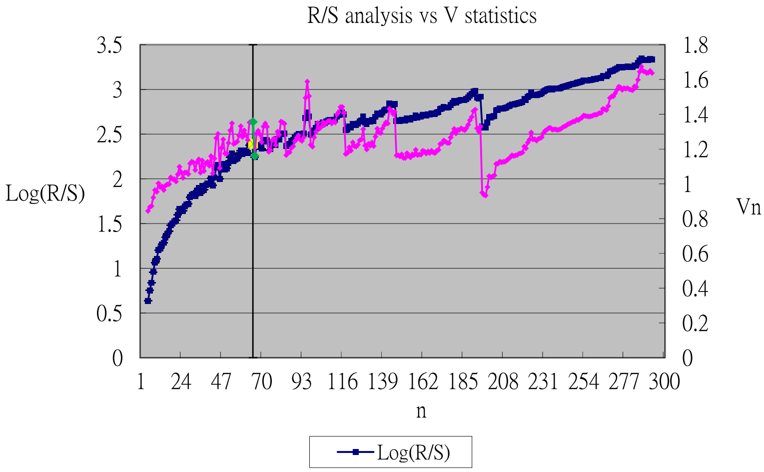

Figure 27. Hong Kong REIT index *R/S* analysis vs. *V* statistics Chart.

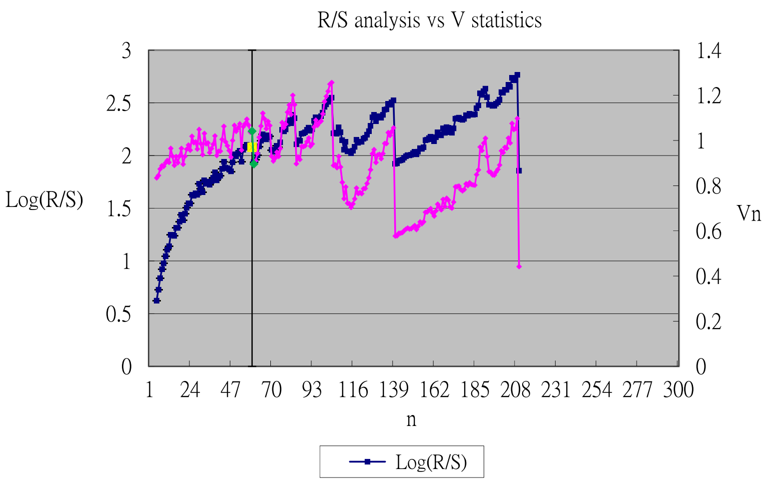

Figure 28. Italy REIT index *R/S* analysis vs. *V* statistics Chart.

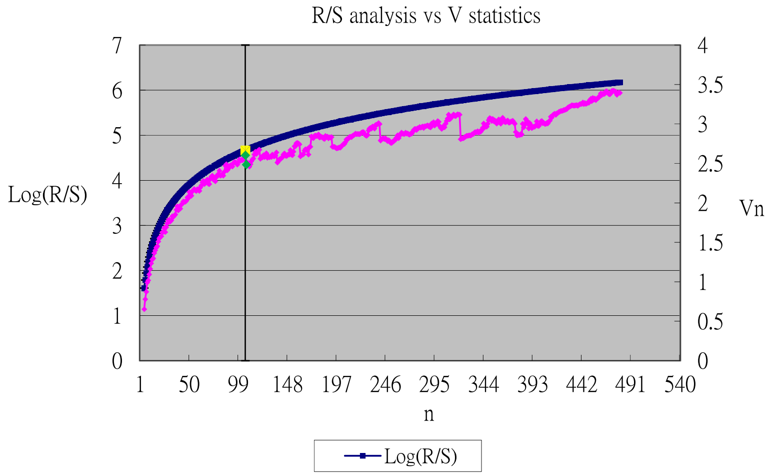

Figure 29. Turkey REIT index *R*/*S* analysis vs. *V* statistics Chart.