


RESEARCH

Open Access

Evaluation of forecasting methods from selected stock market returns



M. Mallikarjuna*  and R. Prabhakara Rao

* Correspondence: mmallikarjuna@sssihl.edu.in

Doctoral Research Scholar,
Department of Economics, Sri
Sathya Sai Institute of Higher
Learning, Vidyagiri, Prasanthi
Nilayam, AP, India

Abstract

Forecasting stock market returns is one of the most effective tools for risk management and portfolio diversification. There are several forecasting techniques in the literature for obtaining accurate forecasts for investment decision making. Numerous empirical studies have employed such methods to investigate the returns of different individual stock indices. However, there have been very few studies of groups of stock markets or indices. The findings of previous studies indicate that there is no single method that can be applied uniformly to all markets. In this context, this study aimed to examine the predictive performance of linear, nonlinear, artificial intelligence, frequency domain, and hybrid models to find an appropriate model to forecast the stock returns of developed, emerging, and frontier markets. We considered the daily stock market returns of selected indices from developed, emerging, and frontier markets for the period 2000–2018 to evaluate the predictive performance of the above models. The results showed that no single model out of the five models could be applied uniformly to all markets. However, traditional linear and nonlinear models outperformed artificial intelligence and frequency domain models in providing accurate forecasts.

Keywords: Financial markets, Stock returns, Linear and nonlinear, Forecasting techniques, Root mean square error

JEL codes: C22, C53, G15, G17

Introduction

Theoretical and empirical studies have shown that a positive relationship exists between financial markets and economic growth (e.g., Levine, 1997; Rajan and Zingales, 1998; Rousseau and Watchel, 2000; Beck et al., 2003; Guptha and Rao, 2018). Given the significance of financial markets, forecasting financial returns occupies a paramount position in investment decision making. However, stock markets are characterized by high volatility, dynamism, and complexity (Johnson et al., 2003; Cristelli, 2014; Wieland, 2015). Movements in stock markets are influenced by several factors, such as macro-economic factors, international events, and human behavior. Hence, forecasting stock returns can become a challenging task. The profitability of investments in stock markets highly depends on the predictability of stock movements. If a forecasting model or technique can precisely predict the direction of the market, investment risk and uncertainty can be minimized. It would enhance investment flows into stock

markets and also be useful for policymakers and regulators in making appropriate decisions and taking corrective measures.

There are two distinct schools of thought—namely, fundamental analysis and technical analysis—for predicting stock price movements. Fundamentalists forecast stock prices on the basis of financial analyses of companies or industries. Technical analysts, meanwhile, use historical securities data and predict future prices on the assumption that stock prices are determined by market forces and that history tends to repeat itself (Levy, 1967). These theories coexisted for several decades as strategies for investment decision making. These approaches were challenged in the 1960s by random walk theory, popularly known as the efficient market hypothesis (Fama, 1970), which proposes that future changes in stock prices cannot be predicted from past price changes. Some empirical studies have shown the presence of ‘random walk’ in stock prices (e.g., Tong et al., 2014; Konak and Seker, 2014; Erdem and Ulucak, 2016). However, most empirical studies have found that stock prices are predictable (Darrat and Zhong, 2000; Lo and MacKinlay, 2002; Harrison and Moore, 2012; Owido et al., 2013; Radikoko, 2014; Said, 2015; Almudhaf, 2018).

Various forecasting techniques are available for time series forecasting. Autoregressive integrated moving average (ARIMA) models were proposed by Box and Jenkins (1970) for time series analysis and forecasting. Some studies have been conducted by employing ARIMA models to forecast stock market returns (Al-Shaib, 2006; Ojo and Olatayo, 2009; Adebisi and Oluinka, 2014; Mondal et al., 2014). Quite a few studies found that ARIMA models produced inferior forecasts for financial time series data (Zhang, 2003; Adebisi and Oluinka, 2014; Khandelwal et al., 2015). To account for nonlinearities resulting from regime changes in economies, some researchers have used Markov regime-switching models and threshold autoregressive (TAR) models assuming nonlinear stationary processes to predict stock prices (Hamilton, 1989; Tong, 1990). Tasy (1989) proposed a simple yet widely applicable model-building procedure for threshold autoregressive models as well as a test for threshold nonlinearity. Gooijer (1998) considered regime switching in a moving average (MA) model and used validation criteria for self-exciting threshold autoregressive (SETAR) model selection. Some empirical studies comparing different methods with SETAR found that this method produced superior results to linear models (e.g., Clements and Smith, 1999; Boero and Marrocu, 2002; Boero, 2003; Firat, 2017).

In the late 1980s, a class of artificial intelligence (AI) models—such as feedforward, backpropagation, and recurrent neural network models—were introduced for forecasting purposes. The distinguishing features of artificial neural networks (ANN) are that they are data-driven, nonlinear, and self-adaptive, and they have very few apriori assumptions. This makes ANNs valuable and attractive for forecasting financial time series. Among ANN models, the feed-forward neural network with a single hidden layer has become the most popular for forecasting stock market returns (Zhang, 2003). Many studies have shown that these models yield more accurate forecasts compared to naïve and linear models (e.g., Ghiassi et al., 2005; Mostafa, 2010; Qiu et al., 2016; Aras and Kocakoc, 2016).

In addition, there are various neural network models for forecasting stock returns. Lu and Wu (2011) used the cerebellar model articulation controller neural network (CAMC NN) model to forecast the stock market indices of the Nikkei 225 and the

Taiwan Stock Exchange. The results showed that CAMC NN made more accurate forecasts than support vector regression and back-propagation neural network (BPNN) models. Guresen et al. (2011) observed that classical ANN models and multilayer perceptron (MLP) outperformed GARCH-class models for the NASDAQ index. Lahmiri (2016) employed variational mode decomposition (VMD) based general regression neural networks (GRNN) for four economic and financial data sets and found that VMD-GRNN models outperformed the ARIMA model and other neural network models. Nayak and Misra's (2018) genetic algorithm-based condensed polynomial neural network (GA-CPNN) improved the accuracy of forecasting stock indices compared to radial basis function neural network (RBFNN) and multilayer perceptron and genetic algorithm (MLP-GA) models. Zhong and Enke (2019) observed that techniques such as deep neural networks using principal component analysis (PCA) and artificial neural networks performed better than traditional models. However, most studies have found that traditional ANN models, as well as ANN models combined with linear models, produce more accurate forecasts than other models (e.g., Asadi et al., 2010; Wang et al., 2011; Khandelwal et al., 2015; Mallikarjuna et al., 2018).

Recently, frequency-domain models, such as spectral analysis, wavelets, and Fourier transformations, have been proposed to improve the forecasting accuracy of financial time series. One widely used technique is singular spectrum analysis (SSA), which is a robust nonparametric method with no prior assumptions about the data (Golyandina et al., 2001; Hassani et al., 2013a). SSA decomposes a time series data into its components and then reconstructs the series by leaving the random noise component before using the reconstructed series to forecast the future points in the series (Hassani, 2007; Ghodsi and Omer, 2014). Since most financial time series data sets exhibit neither purely linear nor purely nonlinear patterns, the combination of linear and nonlinear, i.e., hybrid techniques to model complex data structures for improved accuracy has been proposed (Asadi et al., 2010; Khashei and Bijari, 2010; Khashei and Bijari, 2012; Khandelwal et al., 2015; Ince and Trafalis, 2017). Khashei and Hajirahimi (2017) compared linear and nonlinear models with hybrid models (HM) and concluded that hybrid models perform better than individual models.

Only a few studies have aimed to find a suitable method for forecasting the stock returns of a group of markets. Guidolin et al. (2009) evaluated the performance of linear and nonlinear models for forecasting the financial asset returns of G7 countries. They found that nonlinear models, such as threshold autoregressive (TAR) and smooth transition autoregressive (STAR) models, performed better than linear models in the case of US and UK asset returns. Meanwhile, simple linear models such as random walk and autoregressive models were better for French, German, and Italian asset returns. This suggests that no single model is suitable for forecasting the returns of all stock markets. Awajan et al. (2018) compared the performance of several forecasting methods by applying them to six stock markets and found that the empirical mode decomposition Holt–Winters method (EMD-HW) provided more accurate forecasts than other models.

Though there are various techniques for forecasting stock market returns, no single method can be employed uniformly for the returns of all stock markets. The literature indicates that there is no consensus among researchers regarding the techniques for forecasting stock market returns. The present study, therefore, aimed to evaluate

different forecasting techniques—namely, ARIMA, SETAR, ANN, SSA, and HM models, representing linear, nonlinear, artificial intelligence (AI), frequency domain, and hybrid methods, respectively—as applied to individual stock markets. This study also examined the suitability of different forecasting methods for each category of the world stock markets—namely, developed, emerging, and frontier. Finding a single method that can produce optimal forecasts for all markets could help investors save time and resources and make better decisions. This study is mainly useful for international investors and foreign institutional investors who wish to minimize risks and diversify their portfolios, with the aim of maximizing profits. The objectives of the present study are outlined below.

Objectives

1. To forecast stock market returns using linear, nonlinear, artificial intelligence, frequency domain, and hybrid methods.
2. To find the most appropriate forecasting techniques among the five above-mentioned techniques for developed, emerging, and frontier markets.
3. To check whether any single technique can be applied to all markets to obtain optimal forecasts.

The rest of this paper is organized as follows. Section 2 describes the data and methods employed in the study. Section 3 presents the empirical results. Finally, the conclusions are given in section 4.

Data and methodology

In accordance with the objectives of this study, we considered three types of markets—developed, emerging, and frontier—based on the Morgan Stanley Capital International classification (MSCI, 2018). The market indices taken for the developed category are Australia (ASX 200), Canada (TSX Composite), France (CAC 40), Germany (DAX), Japan (NIKKEI 225), South Korea (KOSPI), Switzerland (SMI), United Kingdom (FTSE 100), and the United States (S&P 500). Those for emerging markets are Brazil (BOVESPA), China (SSEC), Egypt (EGX 30), India (SENSEX), Indonesia (IDX), Mexico (BMV IPC), Russia (MOEX), South Africa (JSE 40), Thailand (SET), and Turkey (BIST 100). Lastly, those in the frontier category are Argentina (S&P Merval), Estonia (TSEG), Kenya (NSE 20), Sri Lanka (CSE AS), and Tunisia (TUNINDEX). The daily closing prices of these indices for the period 1 January 2000 to 30 December 2018 were obtained from the website www.investing.com.

Asset returns (R_t) were calculated from the closing prices of all indices using the formula:

$$R_t = \frac{(P_t - P_{t-1})}{P_{t-1}} * 100 \quad (1)$$

Where, P_t is the price of the asset in the current time period and P_{t-1} is the price of an asset in the previous time period.

Autoregressive integrated moving average (ARIMA)

Proposed by George Box and Gwilym Jenkins in 1970, ARIMA models are among the most popular linear models. In ARIMA models, the future value of a variable is obtained through a linear function of some past observations of the variable and some random errors. The process that generates the time series has the form of:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} \dots \dots \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \dots \theta_q \varepsilon_{t-q} + e_t, \tag{2}$$

where y_t is the variable that will be explained at time t ; c is the constant or intercept ; ϕ_i ($i = 1, 2, \dots p$) and θ_j ($j = 1, 2, \dots q$) are the model parameters; p and q are integers and are often referred to as AR and MA orders of the model, respectively; and e_t is the error term. The assumption regarding the random errors ε_t is that they are independently and identically distributed with a mean zero and constant variance of σ^2 . This model involves a three-step iterative process of identification, estimation, and diagnostic checking. The identification step involves specifying a tentative model by deciding the order of the AR (p) and MA (q) terms. Once a tentative model is specified, the parameters of the model must be estimated, in such a way that the overall measure of errors is minimized, which is generally done with a nonlinear optimization procedure. After the estimation of parameters, diagnostic checking for the adequacy of the model must be done, which involves testing whether the model assumptions about the errors ε_t are satisfied. If the model is adequate, one can proceed to forecast; if not, a new tentative model must be identified following the parameter estimation and model verification. This process with three steps must be repeated until a satisfactory model is selected to forecast the data.

Self-exciting threshold autoregressive (SETAR)

The SETAR model, developed by Tong (1983), is a type of autoregressive model that can be applied to time series data. This model has more flexibility in the parameters which have regime-switching behavior (Watier and Richardson, 1995). Regime switching in this model is based on the dependent variable’s self-dynamics, i.e. self-exciting. In other words, the threshold value in the SETAR model is related to the endogenous variable whereas, in the TAR Model, it is related to an exogenous variable. This model assumes a different autoregressive process in accordance with particular threshold values. SETAR models have the advantage of capturing a commonly observed nonlinear phenomenon which cannot be captured by linear models like exponential smoothing and ARIMA models.

A Threshold Autoregressive model can be transformed into a SETAR model if the threshold variable is taken as a lagged value of the time series itself. The SETAR model with two regimes is specified as:

$$y_t = \begin{cases} \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t \text{ if } y_{t-d} \leq \tau \\ \beta_0 + \sum_{i=1}^p \beta_i y_{t-i} + \varepsilon_t \text{ if } y_{t-d} > \tau \end{cases}, \tag{3}$$

where α_i and β_i are autoregressive coefficients, p is the order of the SETAR model, d is the delay parameter, and y_{t-d} is the threshold variable, ε_t is a series of random variables

that are independent and identically distributed with mean 0 and variance σ_ε^2 . τ is the value of the threshold, and if the value of τ is known, the observations can be separated based on their value in comparison to the threshold, i.e. whether y_{t-d} is below or above the threshold. Then, by using the ordinary least squares method, the AR model is estimated (Ismail and Isa, 2006). The threshold value must be determined along with other parameters of the SETAR model, since the threshold value is unknown in general.

Artificial neural networks (ANN)

Artificial Neural Networks are one of the flexible computing frameworks, that can be used for modeling a broad range of nonlinear data. The major advantages of ANN models are that they are data-driven and universal approximators, which can approximate a large class of functions with great accuracy. This model-building process does not require any prior assumptions about the model form since the characteristics of the data determine the network model. A feedforward neural network with a single hidden layer is one of the most widely used method to forecast the time series data (Zhang, 2003). The structure of the model is defined by a network of three layers of simple processing units connected by acyclic links. The mathematical representation of the relationship between output y_t and the inputs ($y_{t-1}, y_{t-2}, \dots, y_{t-p}$) can be defined as:

$$y_t = w_0 + \sum_{j=1}^q w_j \cdot g \left(w_{0,j} + \sum_{i=1}^p w_{ij} \cdot y_{t-i} \right) + \varepsilon_t, \quad (4)$$

where w_j ($j = 0, 1, 2, \dots, q$) and w_{ij} ($i = 0, 1, 2, \dots, p; j = 0, 1, 2, \dots, q$) are the connection weights or the model parameters, p is the number of input nodes, and q is the number of hidden nodes. The transfer function of the hidden layer is given by the logistic function:

$$\text{Sig}(x) = \frac{1}{1 + \exp(-x)}. \quad (5)$$

Hence, the ANN model in eq. 4 performs the nonlinear functional mapping from past observations ($y_{t-1}, y_{t-2}, \dots, y_{t-p}$) to the future value y_t —that is,

$$y_t = f(y_{t-1}, \dots, y_{t-p}, w) + \varepsilon_t, \quad (6)$$

Where f is a function determined by the network structure and connection weights, and w is a vector of all parameters. Thus, this neural network model is similar to an autoregressive model with nonlinear functionality.

The choice of the value of q depends on the data, as there is no standard procedure for determining this particular parameter. Another vital task of modeling ANN is the choice of the input vector's dimension and the number of lagged observations, p . This is perhaps the most crucial parameter that is to be estimated in an artificial neural network model, as the determination of the nonlinear autocorrelation structure of the time series depends on this parameter. However, there is no rule of thumb that can be followed to select the value of p . Therefore, often trials are conducted to select an optimal value of p and q . After specifying the network structure with the parameters p , and

q , it is ready for training. This is done with efficient nonlinear optimization algorithms, such as gradient descent algorithms and conjugate gradient algorithms, other than the basic backpropagation training algorithm (Hung, 1993).

In ANNs, the most widely used activation functions are the sigmoid functions. Recently, in deep learning, several other functions have been suggested as alternatives to the sigmoid function, such as the hyperbolic tangent (tanh) function, rectified linear units (ReLU), softmax, and Gaussian. These functions are given below.

The hyperbolic tangent (tanh) function is one of the alternatives to the sigmoid function. It can be defined as:

$$\tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}. \quad (7)$$

This function is similar to the sigmoid function, however, it compresses real-value number to a range between -1 and 1 ; i.e., $\tanh(x) \in (-1, 1)$.

Rectified linear units (ReLU) are defined as:

$$f(x) = \max(0, x), \quad (8)$$

Where x is the input for a neuron. In other words, the activation is simply set at a threshold of zero. The range of the ReLU is between 0 and ∞ .

The softmax function, also called as normalized exponential function is a generalization of the logistic function that 'compresses' a K -dimensional vector Z from random real values to a K -dimensional vector $\sigma(z)$ of real values in the range $[0, 1]$, which add up to 1 . The function is defined as:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, j = 1, 2, \dots, K. \quad (9)$$

The Gaussian activation functions are bell-shaped curves that are continuous. The node output is interpreted depending on how close the net input is for a chosen value of average, i.e. it is interpreted in terms of class membership (1 or 0). The function is defined as:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (10)$$

Singular Spectrum analysis (SSA)

Some studies have employed the SSA method to forecast financial time series (Hassani et al., 2013b; Ghodsi and Omer, 2014). The SSA method comprises two stages, one is decomposition and the other is reconstruction. In the first stage, the time series is decomposed to separate the signal and noise, then in the second stage, the series with less noise is reconstructed and applied to forecast by using the following steps (Hassani, 2007):

Step 1. Embedding. Embedding can be considered a mapping that transfers a one-dimensional time series $Y_N = (y_1, \dots, y_N)$ to a multi-dimensional series X_1, \dots, X_K with vectors $X_i = (y_i, \dots, y_{i+L-1})^T \in \mathbb{R}^L$, where L ($2 \leq L \leq N - 1$) is the window length, and $K = N - L + 1$. The result of this step is the trajectory matrix.

$$X = [X_1, \dots, X_K] = (X_{ij})_{i,j=1}^{L,K}$$

Step 2. Singular value decomposition (SVD). In this step, the SVD of X is implemented. Denoted by $\lambda_1, \dots, \lambda_L$ the eigenvalues of XX^T arranged in decreasing order ($\lambda_1, \dots, \lambda_L \geq 0$) and by U_1, \dots, U_L the corresponding eigenvectors. The SVD of X can be written as $X = X_1 + \dots + X_L$, where, $X_i = \sqrt{\lambda_i} U_i V_i^T$.

Step 3. Grouping. This step involves splitting the elementary matrices into several groups and then adding the matrices within each group.

Step 4. Diagonal averaging. The main objective of diagonal averaging is to transform a matrix into the Hankel matrix form, which can be later converted into a time series.

Step 5. Forecasting. There exist two forms of SSA forecasting: recurrent singular spectrum analysis (RSSA) and vector singular spectrum analysis (VSSA). In this study, we employed RSSA. Let $V^2 = \pi_1^2 + \dots + \pi_r^2$ where π_i is the last component of the eigenvector $U_i (i = 1, \dots, r)$. Additionally, for any vector $U \in R^L$, denote by $U^\nabla \in R^{L-1}$ the vector comprising of the first $L - 1$ components of the vector U . Let Y_{N+1}, \dots, Y_{N+h} show the h terms of the SSA recurrent forecast. Then, we can obtain the h -step ahead forecasts by using the following formula.

$$y_t = \begin{cases} \tilde{y}_i \text{ for } i = 1, \dots, N \\ \sum_{j=1}^{L-1} \alpha_j y_{t-j} \text{ for } i = N + 1, \dots, N + h \end{cases} \tag{11}$$

where $\tilde{y}_i (i = 1, \dots, N)$ is the reconstructed series, and vector $A = (\alpha_1, \dots, \alpha_{L-1})$ can be computed by

$$A = \frac{1}{1-v^2} \sum_{i=1}^r \pi_i U_i^\nabla. \tag{12}$$

Hybrid model (HM)

Either purely linear or purely nonlinear models might not be adequate for predicting stock returns since the stock returns are complex in nature. Even data-driven ANNs have produced mixed results in forecasting the time series data. For example, Denton (1995) used simulated data and found that when there is multicollinearity or outliers in the data, neural networks can forecast the data better than the linear regression models. The sample size and noise level play a crucial role in determining the performance of ANNs for linear regression problems (Markham and Rakes, 1998). Therefore, it might not be useful to apply ANNs for all types of data.

Given the complexities in the stock market data, a method that can handle both the linear and nonlinear data, i.e., hybrid model might be an alternative for forecasting. Linear and nonlinear aspects of the underlying patterns in the data can be captured by combining different models.

It might be useful to consider time series data consisting of linear autocorrelation structure and a nonlinear component. That is,

$$y_t = L_t + N_t, \tag{13}$$

where L_t represents the linear component and N_t denotes the nonlinear component. Initially, we must apply a linear model for the data, and then the residuals from the

linear model would contain only the nonlinear relationship. These residuals can be defined as: Let e_t denote the residual at time t from the linear model, then:

$$e_t = y_t - \hat{y}_t + \hat{L}_t, \quad (14)$$

Where, \hat{L}_t is the forecast value at time t from the estimated relationship from eq. 13. Residuals are very crucial in diagnosing the adequacy of linear models because the presence of linear correlation in the residuals indicates the inadequacy of the linear model. In addition, any significant nonlinear pattern in the residuals also indicates the limitation in the linear model. Nonlinear relationships can be discovered by modeling residuals using ANNs. The ANN model for residuals with n input nodes will be:

$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-n}) + \varepsilon_t, \quad (15)$$

Where f is a nonlinear function determined by the neural network, and ε_t is the random error. Denoting the forecast from (13) as \hat{N}_t , the combined forecast will be:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t, \quad (16)$$

where \hat{y}_t is the estimated value from the hybrid model, which is a combination of linear and nonlinear models. We used the inverse mean square forecast error (MSFE) ratio to determine the optimal weights for the hybrid models as it is a widely used method with a robust theoretical background (Bates et al., 1969).

For M models, the combined h -step ahead forecast is:

$$\hat{y}_{t+h} = \sum_{m=1}^M w_{m,h,t} \hat{y}_{t+h,m}, \quad (17)$$

$$w_{m,h,t} = \frac{\left(1/msfe_{m,h,t}\right)^k}{\sum_{j=1}^M \left(1/msfe_{j,h,t}\right)^k}, \quad (18)$$

where $\hat{y}_{t+h,m}$ is the point forecast for h steps ahead at time t from model m . In summary, this hybrid method contains two steps. The first step is to employ the ARIMA to model the linear part of the data. The second step is to apply ANN to model the residuals obtained from the ARIMA, these residuals have information about the nonlinearity in the data. The results from the ANN model can be used as forecasts for the error terms for the ARIMA model. In the manner mentioned above, the hybrid model encompasses the characteristics of the ARIMA and ANN models in modeling time series data. Thus, it could be beneficial to employ hybrid models to improve the accuracy of the forecasts.

Forecast performance measures

The accuracy of forecasts indicates how well a forecasting model predicts the chosen variable. Different accuracy measures are used to validate the suitability of a model for a given data set. There are several accuracy measures in the literature, such as mean error (ME), mean absolute error (MAE), mean absolute percentage error (MAPE), mean squared error (MSE), and root mean squared error (RMSE). In this study, we used RMSE because it is one of the most appropriate methods for measuring forecasting accuracy for data on the same scale, and this criterion has been employed in several

previous studies (Lu and Wu, 2011; Wang et al., 2011; Hyndman and Athanasopoulos, 2015; Makridakis et al., 2015). Also, Chai and Draxler (2014) suggested that RMSE is a suitable measure for models with normally distributed errors. The present study found that the errors in most of the models follow the normal distribution.

If Y_t is the actual observation for time period t , and F_t is the forecast for the same period, then the error is defined as:

$$e_t = Y_t - F_t, \quad (19)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t|, \quad (20)$$

$$MPE = \frac{1}{n} \sum_{t=1}^n PE_t, \quad (21)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n |PE_t|, \quad (22)$$

Where,

$$PE_t = \left(\frac{Y_t - F_t}{Y_t} \right) * 100 \quad (23)$$

The mean squared error (MSE) is:

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2, \quad (24)$$

and the root mean square error (RMSE) is:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}. \quad (25)$$

Empirical results

Here, we present the empirical results, comprising descriptive statistics and the performance measures of various forecasting methods for stock returns in developed, emerging, and frontier markets.

Descriptive statistics of stock returns

Tables 1, 2, and 3 present the summary statistics (e.g., mean, standard deviation, skewness, kurtosis, JB statistic) for developed, emerging, and frontier stock market returns, respectively. From these tables, we can see that the mean returns in all markets are positive, indicating overall positive returns on investments during the period considered for this study. The kurtosis values of the return series of all the markets are observed to be greater than 3, indicating that all of the series are leptokurtic—i.e., they have thick tails, which is a common phenomenon in stock returns (Bouchauda and Potters, 2001; Humala, 2013; Mallikarjuna et al., 2017). The Jarque–Bera test showed that the series are non-normally distributed. Another key feature, from the Tsay (1989) test, is that the returns of all the markets are nonlinear.

Table 1 Descriptive Statistics for Developed Markets

Country	Mean	Standard Deviation	Skewness	Kurtosis	Jarque–Bera Statistic	Tsay Test
Australia	0.016900	0.981318	−0.366178	8.451780	5732.678 (0.00000)	Nonlinear
Canada	0.016719	1.043102	−0.466783	13.58932	21,277.95 (0.00000)	Nonlinear
France	0.005790	1.429676	0.141474	8.712746	6275.928 (0.00000)	Nonlinear
Germany	0.021703	1.475373	0.097060	8.111260	4987.257 (0.00000)	Nonlinear
Japan	0.019562	1.501151	−0.211389	9.413287	7642.163 (0.00000)	Nonlinear
Korea	0.040316	1.384506	−0.350392	9.724636	8473.831 (0.00000)	Nonlinear
Switzerland	0.007731	1.174780	−0.012284	10.18851	9766.624 (0.00000)	Nonlinear
UK	0.009243	1.171826	−0.004334	9.918071	9061.453 (0.00000)	Nonlinear
US	0.020694	1.147998	−0.089045	12.10219	15,630.08 (0.00000)	Nonlinear

Results of forecasting methods

Before applying the forecasting methods, we divided the data into the training set and the test set; we used 80% of the data for training the models and the remaining 20% for testing the models. To forecast the returns using the ARIMA (p, d, q) model, it was necessary to check stationarity to have valid inferences. To test the stationarity of the returns series, we employed the augmented Dickey–Fuller (1979) and Phillips–Perron (1988) tests; the results showed that the returns of all of the markets were stationary. We determined the optimal lag length for the autoregressive (p) and moving average (q) components using the Akaike information criterion (AIC). We observed different

Table 2 Descriptive Statistics for Emerging Markets

Country	Mean	Standard Deviation	Skewness	Kurtosis	Jarque–Bera Statistic	Tsay Test
Brazil	0.055547	1.769091	0.075096	7.347967	3439.261 (0.00000)	Nonlinear
China	0.016499	1.585054	−0.218775	7.684998	4025.909 (0.00000)	Nonlinear
Egypt	0.079649	1.658833	−0.121416	13.15053	18,891.65 (0.00000)	Nonlinear
India	0.059777	1.414324	0.120611	12.83913	17,795.20 (0.00000)	Nonlinear
Indonesia	0.070676	1.326156	−0.501572	9.570707	8083.189 (0.00000)	Nonlinear
Mexico	0.051820	1.213204	0.153714	9.356521	7644.352 (0.00000)	Nonlinear
Russia	0.082735	1.965327	0.370012	24.33804	85,492.66 (0.00000)	Nonlinear
South Africa	0.047609	1.298747	0.036830	6.299922	2044.612 (0.00000)	Nonlinear
Thailand	0.047775	1.263998	−0.520070	13.23751	19,417.31 (0.00000)	Nonlinear
Turkey	0.069655	1.973356	−0.037463	9.689365	8426.657 (0.00000)	Nonlinear

Table 3 Descriptive Statistics for Frontier Markets

Country	Mean	Standard Deviation	Skewness	Kurtosis	Jarque–Bera Statistic	Tsay Test
Argentina	0.120985	2.162861	0.018419	7.202958	3230.709 (0.00000)	Nonlinear
Estonia	0.052646	1.036820	0.316893	14.29612	24,406.12 (0.00000)	Nonlinear
Kenya	0.012726	0.831383	0.555709	15.16038	28,057.55 (0.00000)	Nonlinear
Tunisia	0.035152	1.292467	0.603647	20.46153	59,408.84 (0.00000)	Nonlinear
Sri Lanka	0.066494	1.128874	0.993308	43.75095	300,319.2 (0.00000)	Nonlinear

orders of AR and MA for different series and present them along with RMSE values in Tables 4, 5, and 6. In the SETAR model, the series exhibited nonlinear trends, and we identified two regimes by the minimum AIC values. Then, the model was used to forecast the returns of the markets. To forecast stock returns using the ANN model, we employed feedforward neural networks since many studies have shown that they fit well with asset return data (Zhang, 2003; Qiu et al., 2016). We employed a recurrent singular spectrum analysis (RSSA) model to forecast the returns after decomposing and reconstructing the original returns series by following the four steps involved in forecasting with SSA: embedding, reconstructing, grouping, and diagonal averaging. For the hybrid model, which is a combination of ARIMA and ANN, we fit the model by employing the widely used inverse mean square forecast error (MSFE) ratio (Bates et al., 1969; Winkler and Makridakis, 1983) for assigning the optimal weights for the models in forecasting. Tables 4, 5, and 6 present the RMSE values of the test sets of the forecast series for all techniques (i.e., ARIMA, SETAR, SSA, ANN, and HM) for developed, emerging, and frontier markets, respectively. The model with the lowest RMSE was chosen as the most appropriate model. In addition, we tested RMSE significance

Table 4 RMSE Values of the Forecasting Models for Developed Markets

Country	ARIMA	SETAR	ANN	SSA	HM
Australia	ARIMA (1,0,0) 0.839708	0.8371075	0.839961	0.853739	0.8398178
Canada	ARIMA (4,0,4) 0.7160948	0.7235669	0.7279448	0.721191	0.7178712
France	ARIMA (2,0,3) 1.104531	1.104509	1.121098	1.13334	1.107893
Germany	ARIMA (3,0,3) 1.137963	1.140388	1.156347	1.139859	1.138008
Japan	ARIMA (1,0,1) 1.312612	1.307090	1.312699	1.319925	1.312697
South Korea	ARIMA (1,0,2) 0.7815902	0.7845301	0.7816591	0.7975102	0.7815836
Switzerland	ARIMA (3,0,3) 0.9262604	0.9252304	0.9262753	0.935211	0.9262085
UK	ARIMA (3,0,2) 0.9002485	0.9003106	0.9093658	0.9194308	0.9057976
US	ARIMA (2,0,0) 0.8197795	0.8125207	0.8171281	0.8187795	0.8119642

Table 5 RMSE Values of the Forecasting Models for Emerging Markets

Country	ARIMA	SETAR	ANN	SSA	HM
Brazil	ARIMA (2,0,1) 1.441972	1.442167	1.442900	1.470025	1.442262
China	ARIMA (3,0,3) 1.555623	1.570942	1.56378	1.555431	1.554521
Egypt	ARIMA (0,0,1) 1.357198	1.324729	1.368285	1.385264	1.357206
India	ARIMA (3,0,1) 0.838730	0.844534	0.839839	0.851503	0.838335
Indonesia	ARIMA (1,0,0) 0.919355	0.926838	0.918685	0.921310	0.918988
Mexico	ARIMA (2,0,1) 0.850496	0.848814	0.867506	0.851756	0.848878
Russia	ARIMA (3,0,4) 0.996021	0.993842	1.008659	1.006565	1.000470
South Africa	ARIMA (3,0,1) 1.062336	1.064449	1.063028	1.066819	1.061791
Thailand	ARIMA (2,0,2) 0.756856	0.754199	0.770766	0.775997	0.762275
Turkey	ARIMA (2,0,2) 1.280623	1.28606	1.292737	1.291581	1.283447

using the Diebold–Marino test (1995) and found that the RMSE of all of the models was significant, except for Japan, South Africa, and Sri Lanka.

From Tables 4, 5, and 6, we can observe that no single method performed uniformly for all markets. However, the nonlinear model (i.e., SETAR) performed better than the other models, producing optimal forecasts for 10 markets (i.e., four developed, four emerging, and two frontier markets). This result contrasts with Guidolin et al. (2009). In the case of developed markets, the SETAR model produced optimal forecasts for four of the nine markets (Australia, France, Japan, and Switzerland). The ARIMA model was optimal for Canada, Germany, and the UK, and the HM model was optimal for South Korea and the US. Thus, we can say that nonlinear models are more suitable for developed markets. Meanwhile, ANN and SSA models are not at all useful for developed markets since they did not provide any optimal forecasts.

For emerging markets, the SETAR model was found to be appropriate for four markets (Egypt, Mexico, Russia, Thailand). HM models were appropriate for three markets (China, India, and South Africa) and ARIMA models for two (Brazil and Turkey). The ANN model was appropriate for only one market (Indonesia), while the SSA model

Table 6 RMSE Values of the Forecasting Models for Frontier Markets

Country	ARIMA	SETAR	ANN	SSA	HM
Argentina	ARIMA (1,0,0) 2.058513	2.046473	2.059416	2.067785	2.061399
Estonia	ARIMA (1, 0, 2) 0.554262	0.576389	0.593064	0.567850	0.574015
Kenya	ARIMA (2,0,2) 0.679084	0.656696	0.656707	0.680819	0.684070
Sri Lanka	ARIMA (0,0,2) 0.443346	0.449314	0.444426	0.473222	0.445437
Tunisia	ARIMA (1,0,0) 0.460538	0.44934	0.618865	0.443756	0.464829

was not suitable for any emerging market. Though no single model was suitable for all emerging markets, the SETAR and HM models were relatively more useful. Regarding frontier markets, SETAR was suitable for Argentina and Kenya, ARIMA for Estonia and Sri Lanka, and SSA for Tunisia. The ANN and HM models were not appropriate for any market.

Out of twenty-four stock market indices, the SETAR model produced optimal forecasts for ten, ARIMA for seven, HM models for five, and ANN and SSA models for one market each. From these results, we can observe that nonlinear models are more useful for developed, emerging, and frontier markets alike. Another interesting observation is that the AI and frequency domain models were found to be appropriate only for one market each. Thus, we can say that, even with advancements in AI and frequency domain models, traditional statistical models have not become obsolete; they are still useful and in fact better than AI and frequency domain models for forecasting financial time series data.

Summary and conclusions

Over the years, stock markets have become alternative avenues for surplus funds among individual and institutional investors, especially following globalization and the integration of world financial markets. Given the inherent risk, uncertainty, and dynamic nature of stock markets, accurately forecasting stock returns can help to minimize investors' risks. Thus, forecasting techniques can help with better investment decision making.

This study considered daily data for stock market returns during the period 1 January 2000 to 30 December 2018 to compare forecasting techniques (i.e., ARIMA, SETAR, ANN, SSA, and HM models) representing linear, nonlinear, AI, frequency domain, and hybrid methods. We took the stock indices of 24 stock markets in three market categories (nine developed, ten emerging, and five frontier) to find suitable forecasting techniques for each category. The results showed that no single forecasting technique provided uniformly optimal forecasting for all markets. However, SETAR performed better for ten markets, ARIMA for seven, HM for five, and ANN and SSA for one market each. SETAR and ARIMA techniques can thus be considered the clear winners in forecasting stock market returns for developed, emerging, and frontier markets, as these two methods provided optimal forecasts for seventeen of the twenty-four markets.

Abbreviations

AI: artificial intelligence; ANN: artificial neural networks; AR: autoregressive; ARIMA: autoregressive integrated moving average; EMD-HW: empirical mode decomposition Holt-Winters method; HM: hybrid model; MA: moving average; MSCI: Morgan Stanley Capital International; MSE: mean squared error; RMSE: root mean square error; RSSA: recurrent singular spectrum analysis; SETAR: self-exciting threshold autoregressive; SSA: singular spectrum analysis; STAR: smooth transition autoregressive; TAR: threshold autoregressive; VSSA: vector singular spectrum analysis

Acknowledgments

We thank the editor and the anonymous reviewers for their valuable comments and suggestions that greatly improved the paper.

Authors' contributions

Both the authors have made significant contribution jointly and are in agreement with the contents of this paper.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

The data used for the study are available in the internet. The URL is <https://in.investing.com/indices/major-indices>

Competing interests

The authors declare that they have no competing interests.

Received: 26 April 2019 Accepted: 16 October 2019

Published online: 02 December 2019

References

- Adebiyi AA, Oluinka A (2014) Comparison of ARIMA and artificial neural network models for stock market prediction. *Journal of Applied Mathematics*. <https://doi.org/10.1155/2014/614342>
- Almudhaf F (2018) Predictability, Price bubbles, and efficiency in the Indonesian stock-market. *Bull Indones Econ Stud* 54(1): 113–124
- Al-Shaib M (2006) The predictability of the Amman stock exchange using Univariate autoregressive integrated moving average (ARIMA) model. *Journal of Economic and Administrative Sciences* 22(2):17–35
- Aras S, Kocakoc ID (2016) A new model selection strategy in time series forecasting with artificial neural networks. *IHTS Neurocomputing* 174:974–987
- Asadi S, Tavakoli A, Hejazi SR (2010) A new hybrid for improvement of auto-regressive integrated moving average models applying particle swarm optimization. *Expert Syst Appl* 39:5332–5337
- Awajan AM, Ismail MT, Wadi SA (2018) Improving forecasting accuracy for stock market data using EMD-HW bagging. *PLoS One* 13(7):1–20
- Bates JM, Granger CWJ (1969) The combination of forecasts. *Operational Research Society* 20(4):451–468
- Beck T, Levine R (2003) Stock markets, banks and growth: panel evidence. *J Bank Financ* 28:423–442
- Boero G (2003) The performance of SETAR models: a regime conditional evaluation of point, interval and density forecasts. *Int J Forecast* 20:305–320
- Boero G, Marrocu E (2002) The performance of non-linear exchange rate models: a forecasting comparison. *J Forecast* 21(7): 513–542
- Bouchauda JP, Potters M (2001) More stylized facts of financial markets: leverage effect and downside correlations. *Physica A* 299:60–70
- Box GEP, Jenkins GM (1970) *Time series analysis: forecasting and control*. Holden-Day, San Francisco
- Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding RMSE in the literature. *Geo Scientific Model Development* 7:1247–1250
- Clements MP, Smith J (1999) A Monte Carlo study of the forecasting performance of empirical SETAR models. *J Appl Econ* 14:124–141
- Cristelli M (2014) *Complexity in financial markets*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-00723-6>
- Darrat AF, Zhong M (2000) On testing the random walk hypothesis a model Comparison approach. *The Financial Review* 35: 105–124
- Denton JW (1995) How good are neural networks for causal forecasting? *The Journal of Business Forecasting Methods and Systems* 14(2):17–23
- Dickey D, Fuller W (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of American Statistical Association* 74(366):427–431
- Diebold FX, Mariano RS (1995) Comparing predictive accuracy. *J Bus Econ Stat* 13(3):134–144
- Erdem E, Ulucak R (2016) Efficiency of stock exchange markets in G7 countries: bootstrap causality approach. *Economics World* 4(1):17–24
- Fama EF (1970) Efficient capital markets: a review of theory and empirical work. *J Financ* 25(2):383–417
- Firat EH (2017) SETAR (self-exciting threshold autoregressive) non-linear currency Modelling in EUR/USD, EUR/TRY and USD/TRY parities. *Mathematics and Statistics* 5(1):33–55
- Ghiassi M, Saidane H, Zimbra DK (2005) A dynamic artificial neural network model for forecasting series events. *Int J Forecast* 21:341–362
- Ghods Z, Omer HN (2014) Forecasting energy data using singular Spectrum analysis in the presence of outlier(s). *International Journal of Energy and Statistics* 2(2):125–136
- Golyandina N, Nekrutkin V, Zhigljavsky A (2001) *Analysis of time series structure SSA and related techniques*. Chapman and Hall/CRC, Newyork
- Gooijer DJ (1998) On threshold moving-average models. *J Time Ser Anal* 19(1):1–18
- Guidolin M, Hyde S, McMillan D, Ono S (2009). Non-linear predictability in stock and bond returns: when and where is it exploitable. Federal Reserve Bank of St. Louis: working paper series no 2008-010B
- Guptha SK, Rao RP (2018) The causal relationship between financial development and economic growth experience with BRICS economies. *Journal of Social and Economic Development* 20(2):308–326
- Guresen E, Kayakutlu G, Daim TU (2011) Using artificial neural network models in stock market index prediction. *Expert Syst Appl* 38:10389–10397
- Hamilton JD (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57:357–384
- Harrison B, Moore M (2012) Stock market efficiency, non-linearity, thin trading and asymmetric information in MENA stock markets. *Economic Issues* 17(1):77–93
- Hassani H (2007) Singular spectrum analysis: methodology and comparison. *Journal of Data Science* 5(2):239–257
- Hassani H, Soofi A, Zhiglavsky A (2013a) Forecasting UK industrial production with multivariate singular Spectrum analysis. *J Forecast* 32(5):395–408
- Hassani H, Soofi A, Zhiglavsky A (2013b) Predicting inflation dynamics with singular Spectrum analysis. *J R Stat Soc* 176(3): 743–760
- Humala A (2013) Some stylized facts of return in the foreign exchange and stock markets in Peru. *Stud Econ Financ* 30(2):139–158
- Hung SL, Adeli H (1993) Parallel backpropagation algorithms on CRAY Y-MP8/864 supercomputer. *Neurocomputing* 5(6):287–302

- Hyndman R, Athanasopoulos G (2015) Forecasting principles and practice. Otexts, Melbourne. Available at: <https://otexts.com/fpp3/>. Accessed 20 Mar 2019.
- Ince H, Trafalis TB (2017) A hybrid forecasting model for stock market prediction. *Economic Computation and Economic Cybernetics Studies and Research* 21:263–280
- Ismail MT, Isa Z (2006) Modelling exchange rate using regime switching models. *Sains Malaysiana* 35(2):55–62
- Johnson NF, Jefferies P, Hui PM (2003) *Financial market complexity*. Oxford University Press, Oxford
- Khandelwal I, Adhikari R (2015) Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition. *Procedia Computer Science* 48:173–179
- Khashei M, Bijari M (2010) An artificial neural network model for time series forecasting. *Expert Syst Appl* 37:479–489
- Khashei M, Bijari M (2012) A new class of hybrid models for time series forecasting. *Expert Syst Appl* 39:4344–4357
- Khashei M, Hajirahimi Z (2017) Performance evaluation of series and parallel strategies for financial time series forecasting. *Financial Innovation* 3(24):1–24
- Konak F, Seker Y (2014) The efficiency of developed markets: empirical evidence from FTSE 100. *J Adv Manag Sci* 2(1):29–32
- Lahmiri S (2016) A variational mode decomposition approach for analysis and forecasting of economic and financial time series. *Expert Syst Appl* 55:268–273
- Levine R (1997) Financial development and economic growth: views and agenda. *J Econ Lit* 35:688–726
- Levy RA (1967) The theory of random walks: a study of findings. *Am Econ* 11(2):34–48
- Lo AW, MacKinlay AC (2002) *An non-random walk down Wall street*. Princeton University Press, Princeton
- Lu CJ, Wu JY (2011) An efficient CMAC neural network for stock index forecasting. *Expert Syst Appl* 38:15194–15201
- Makridakis S, Wheelwright SC, Hyndman RJ (2015) *Forecasting: methods and applications*. Wiley India, New Delhi
- Mallikarjuna M, Arti G, Rao RP (2018) Forecasting stock returns of selected sectors of Indian capital market. *SS International Journal of Economics and Management* 8(6):111–126
- Mallikarjuna M, Guptha KS, Rao RP (2017) Modelling Sectoral volatility of Indian stock markets. *Wealth International Journal of Money Banking and Finance* 6(2):4–9
- Markham IS, Rakes TR (1998) The effect of sample size and variability of data on the comparative performance of artificial neural networks and regression. *Comput Oper Res* 25:251–263
- Mondal P, Shit L, Goswami S (2014) Study of effectiveness of time series Modelling (ARIMA) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications* 4(2):13–29
- Mostafa MM (2010) Forecasting stock exchange movements using neural networks: empirical evidence from Kuwait. *Expert Syst Appl* 37:6302–6309
- MSCI (2018) MSCI Announces the Results of Its Annual Market Classification Review. Available at: <https://www.msci.com/market-classification>. Accessed 25 Mar 2019
- Nayak SC, Misra BB (2018) Estimating stock closing indices using a GA-weighted condensed polynomial neural network. *Financial Innovation* 4(21):1–22
- Ojo JF, Olatayo TO (2009) ON the estimation and performance of subset of autoregressive integrated moving average models. *Eur J Sci Res* 28:287–293
- Owido PK, Onyuma SO, Owuor G (2013) A GARCH approach to measuring efficiency: a case study of Nairobi securities exchange. *Research Journal of Finance and Accounting* 4(4):1–16
- Phillips PCB, Perron P (1988) Testing for unit roots in time series regression. *Biometrika* 75:335–346
- Qiu M, Song Y, Akagi F (2016) Application of artificial neural network for the prediction of stock market returns the case of the Japanese stock market. *Chaos, Solitons and Fractals* 85:1–7
- Radikoko I (2014) Testing weak-form market efficiency on the TSX. *J Appl Bus Res* 30(3):647–658
- Rajan R, Zingales L (1998) Financial dependence and growth. *Am Econ Rev* 88:559–586
- Rousseau PL, Watchel P (2000) Equity markets and growth: cross-country evidence on timing and outcomes, 1980-1995. *J Bank Financ* 24(12):1933–1957
- Said A (2015) The efficiency of the Russian stock market: a revisit of the random walk hypothesis. *Academy of Accounting and Financial Studies Journal* 19(1):42–48
- Tong H (1983) *Threshold models in non-linear time series analysis*. Springer, Berlin. <https://doi.org/10.1007/978-1-4684-7888-4>
- Tong H (1990) *Non-Linear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford
- Tong T, Li B, Benkato O (2014) Revisiting the weak form efficiency of the Australian stock market. *Corp Ownersh Control* 11(2):21–28
- Tsay R (1989) Testing and modeling threshold autoregressive processes. *Journal of American Statistical Association* 84:231–240
- Wang JZ, Wang JJ, Zhang ZG, Guo SP (2011) Forecasting stock indices with backpropagation neural network. *Expert Syst Appl* 38:14346–14355
- Watier L, Richardson S (1995) Modelling of an epidemiological time series by a threshold autoregressive model. *Journal of Royal Statistical Society* 44(3):353–364
- Wieland OL (2015) Modern financial markets and the complexity of financial innovation. *Universal Journal of Accounting and Finance* 3(3):117–125
- Winkler RL, Makridakis S (1983) The combination of forecasts. *J R Stat Soc* 146(2):150–157
- Zhang GP (2003) Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50:159–175
- Zhong X, Enke D (2019) Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financial Innovation* 5(4):1–20

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.