NUMBER 1/2019

ISSN 2324-3635

OCCASIONAL & DISCUSSION PAPER SERIES

Models Applied in Stock Market Prediction: A Literature Survey

Wajira Dassanayake,
Chandimal Jayawardena,
Iman Ardekani and
Hamid Sharifzadeh

ePress Unitec

OCCASIONAL AND DISCUSSION
PAPER SERIES 1/2019

Models Applied in Stock Market Prediction:
A Literature Survey

By Wajira Dassanayake, Chandimal Jayawardena,
Iman Ardekani and Hamid Sharifzadeh

This publication may be cited as:

Dassanayake, W., Jayawardena, C., Ardekani. I., & Sharifzadeh, H. (2019). Models Applied in Stock Market Prediction: A Literature Survey. *Unitec ePress Occasional and Discussion Papers Series* (2019:1). Retrieved from http://www.unitec.ac.nz/ epress.

About this series:

Unitec ePress periodically publishes occasional and discussion papers that discuss current and ongoing research authored by members of staff and their research associates. All papers are blind reviewed. For more papers in this series please visit: www.unitec.ac.nz/epress/index.php/category/publications/epress-series/discussion-and-occasionalpapers.

Cover design by Penny Thomson.

Contact:

epress@unitec.ac.nz
www.unitec.ac.nz/epress/

Unitec Institute of Technology
Private Bag 92025, Victoria Street West
Auckland 1142
New Zealand

ISSN
2324-3635

Models Applied in Stock Market Prediction: A Literature Survey

Wajira Dassanayake, Chandimal Jayawardena, Iman Ardekani
and Hamid Sharifzadeh

Abstract

Stock market prices are intrinsically dynamic, volatile, highly sensitive, nonparametric, nonlinear and chaotic in nature, as they are influenced by a myriad of interrelated factors. As such, stock market time series prediction is complex and challenging. Many researchers have been attempting to predict stock market price movements using various techniques and different methodological approaches. Recent literature confirms that hybrid models, integrating linear and non-linear functions or statistical and learning models, are better suited for training, prediction and generalisation performance of stock market prices. The purpose of this review is to investigate different techniques applied in stock market price prediction with special emphasis on hybrid models. This review paper classifies the literature pertaining to hybrid models applied to stock market price prediction in accordance with their input characteristics, makes comparisons between hybridised models, and exhibits the performance evaluation measures used. It summarises the salient characteristics of the contemporary models applied in the stock market price and index prediction. The surveyed papers show that hybrid models are widely used for stock market prediction.

Introduction

The stock market is a forum where market participants can trade stocks, which represent ownership claims on businesses, either through organised exchanges or over-the-counter markets. These ownership entitlements are mostly marketable securities listed publicly. Virtually every country in the world today has its own stock market and trades trillions of dollars' worth of stocks and other marketable securities daily in wider global stock markets. According to the World Federation of Exchanges, in 2016 there were 17 stock exchanges with more than USD 1 trillion market capitalisation.

The stock market is one of the most fascinating, sophisticated and complex financial markets, whose movements are influenced by a multitude of interwoven macroeconomic and non-normal random disturbance factors. The process of globalisation and financial market integration intensifies further complexities. Stock market prices are influenced by the domestic economic environment, government policies, individual and institutional investors' motivation and psychology, global economic conditions, local and international political situations, and the degree of integration with other markets.

Stock markets provide a forum for the market participants to create wealth through investment gains, dividend incomes, diversification benefits, ownership stakes and tax deferrals. The majority of market participants intend to exploit the market situation to make profit through some form of prediction. Generally, stock market prediction is regarded as an extremely challenging and complex problem in time series forecasting more broadly. Thus, an accurate prediction of stock price movement becomes invaluable to private and institutional investors, speculators, arbitrageurs, hedgers, brokers, dealers or government organisations to make informed decisions to generate effective trading strategies through long (buy) or short (sell) positions in the stock markets at the right time to make profits and minimise associated risks.

In the last few decades, many researchers have been attempting to predict the time series of stock market prices and indices. These prediction methodologies can be categorised into two groups: fundamental analysis and technical analysis. Fundamental analysis evaluates stock market prices based on overall macroeconomic conditions, the management strategies of the company and its industry, and political environments. Thus, the fundamentalists employ numeric information pertaining to macroeconomic, financial, and other related factors to predict the intrinsic value of stocks. The technical analysts instead have full reliance on the past market data and typically evaluate the historical price trends to predict stock market price movements. Therefore, the technicians utilise charts and modelling techniques to capture the historical trends to predict stock market price movements. Technical analysts usually employ advanced statistical techniques, soft computing techniques and/or hybrid models. Hybrid models either combine different statistical techniques, or different soft-computing techniques, or statistical techniques with soft-computing techniques to produce better forecasting performances.

The purpose of this review is to investigate different techniques applied in stock market price and index prediction with special emphasis on hybrid models. The paper evaluates statistical techniques, artificial intelligence (AI) techniques and hybrid models applied in stock market predictions. The paper classifies the literature pertaining to hybrid models applied to stock market price prediction in accordance with their input characteristics, compares the invented algorithm with other hybridised models examined and exhibits the performance evaluation measures used in the evaluated models.

This review could contribute to the research on hybrid models applied in stock market price and index prediction in that, firstly, it compares and contrasts the methodologies, techniques and performance measures adopted by hybrid researchers who developed models for stock market-related predictions; secondly, it provides a comparative summary of hybrid algorithms that have been compared, contested, modified and/or invented as superior forecasting models for the tested stock market data sets; and, thirdly, this summary provides a basis to find gaps in the literature and thus enable the synthesis of intelligent, scalable, hybrid models in the future.

The rest of the paper is organised as follows. Section 2 provides an overview of the statistical models applied in stock market prediction. Section 3 provides an overview of the AI models applied in stock market prediction. Section 4 provides an overview of hybrid models applied in stock market prediction. Section 5 lists all the models tested and the evaluation measures applied to ascertain the predictive performance of the formulated hybrid models. Section 6 summarises a range of hybrid models and the combination of statistical and soft-computing techniques used in those evaluated in our review. Concluding remarks related to hybrid research applied in stock market price/index prediction are presented in Section 7.

Statistical models

In the last few decades, many researchers and practitioners have been trying to predict the time series of stock market prices using statistical techniques. These statistical techniques include the exponential smoothing (ES) method, parametric and nonparametric regression models, autoregressive moving average (ARMA) methods, autoregressive integrated moving average (ARIMA) methods, and generalised autoregressive conditionally heteroskedastic (GARCH) methods.

Yule (1927) proposed the notion of stochasticity in time series by hypothesising that every time series could be observed as a realisation of a stochastic process. This conception influenced the work of many subsequent time-series theorists, for example Yule (1927), Slutzky (1937), Walker (1931), and Yaglom (1955), who formulated the notion of auto-regressive (AR) and moving average (MA) models.

The ARMA model, which is a combination of AR and MA models, assumes that the time series is stationary, which is an extremely rare situation in real-world time series due to variations in trends and periodicity. The exponential smoothing (ES) technique in time-series forecasting transpired in the 1950s and 1960s with the contributory work of Brown (1959), Brown

(1963), Holt (1957), Winters (1960), Muth (1960) and Pegels (1969).

Subsequently, Box and Jenkins (1976) integrated moving average and autoregressive models in their influential work, thus, ARIMA models are also known as Box-Jenkins models. The ARIMA model converts non-stationary time series into stationary time series by using differencing (log/difference) before using the ARMA model. Since their publications, the ARIMA models became very popular in time-series forecasting. O'Donovan (1983) showed that these models were better techniques in time series forecasting than the prevailing models at that time.

However, a number of researchers have revealed the existence of serial correlation issues in the residuals of the predictors derived from the ARIMA models. Reid (1969), Chatfield (1978) and Roberts (1982) made significant contribution to time series forecasting. The techniques of Box and Jenkins (1976), Roberts (1982) and Abraham and Ledolter (1986) showed specific linear exponential smoothing predictions appearing as special cases of ARIMA models. However, their results could not be extended to any nonlinear exponential smoothing methods.

Engle (1982) introduced the class of autoregressive conditional heteroscedastic (ARCH) models and Bollerslev (1986) introduced a more parsimonious model than ARCH, called generalised ARCH (GARCH). ARCH/ GARCH models are volatility clustering techniques, and they are applied in financial and econometric analysis. This framework established an insightful structure which led to the development of several other extensions and generalisations such as integration to higher-order moments, generalisation to ultrahigh frequency data, and multivariate extensions.

A number of researchers have used ES technique to predict stock market volatility, stock prices and index prediction. Analysing the Baltic stock market, Girdzijauskas et al. (2009) found that exponential growth models are better suited to modelling long processes. They also uncovered the inability of ES technique to model long-term predictions.

Bley and Olson (2008) evaluated the forecasting performances of ES, linear regression and a number of GARCH models, analysed daily data (S&P 500, S&P 100, NASDAQ indices as well as VIX, VXO and VXN volatility indices) from 1990 to 2004 and found that ES models operated well during the height of stock market volatility.

Maris, Pantou, Nikolopoulos, Pagourtzi and Assimakopoulos (2004) examined the volatility of forecasting accuracies of the random walk (RW) model, mean model, exponential smoothing model (ESM), and four models from the ARCH family by analysing the FTSE/ASE 20 stock index during August 1999 and June 2002. Weekly data was used to avoid over-scribing and under-scribing complications involved in intraday data. They found that FTSE/ASE 20 stock corresponds to the RW model, confirming the presence of chaotic behaviour in the index.

McMillan (2003) examined whether the standard linear model, the logistic smooth-transition threshold (LSTR) model, or the exponential smooth transitions threshold (ESTR) model are more suitable for predicting the quarterly UK FTSE-All stock market index, which exhibits linear and nonlinear dynamics to a certain degree. The study revealed that the ESTR improves both in sample as well as out of sample, confirming its superiority in capturing nonlinear behaviour.

Taylor (2004) examined the weekly volatility forecasting precisions of a number of linear and nonlinear GARCH models, the ES technique, and the adaptive exponential smoothing method. He used eight stock indices to determine the better technique and found that the adaptive exponential smoothing method, which allows smoothing parameters to change over time, provides more satisfactory results.

Balaban, Bayar and Faff (2006) evaluated the out-of-sample forecasting ability of 11 models for monthly volatility in 15 stock market indices. They examined the RW model, historical mean model, moving average (MA) model, weighted moving average (WMA) model, exponentially weighted moving average (EWMA) model, ES model, regression model, ARCH model, GARCH model, GJR GARCH model, and EGARCH model. Symmetric loss functions are used to evaluate the performance of the investigated models whilst asymmetric loss functions are used to penalise under (or over) prediction. They found that the ES model was the best for the symmetric error analysis process, whilst for under (or over) predictions, which are penalised more heavily, ARCH-type (ES) models provided the best forecasts.

Hyung, Poon and Granger (2006) analysed the characteristics of four long-term volatility models, namely fractional integrated (FI), break, component and regime switching, and compared their forecasting performance with the short-term classical memory models (GARCH, Glosten, Jogannathan and Runkle-GARCH [GJR-GARCH], ES and RW). They used daily SP 500 returns to test the models, and established that the FI projected better in volatility forecasts for ten days and beyond, confirming that SP 500 returns exhibit non-stationary processes.

Pereira (2004) examined RW, historical volatility (HIS), MA, weighted moving average (WMA), ES, EWMA, ARCH, GARCH and GJR-GARCH models, applied these to the Portuguese stock market, evaluated symmetric and asymmetric error statistics, and found the ARCH model to be superior.

The ARIMA model is widely used to predict stock market volatility, stock price and index prediction. Kavussanos and Visvikis (2005) evaluated the short-term forecasting performance of univariate (ARIMA) and multivariate (vector autoregressive [VAR], vector error correction [VECM] and restricted VECM) linear models using derivative market time series from the Athens Exchange. Each model was tested using daily closing cash and futures prices for FTSE/ASE-20 and FTSE/ASE-mid40 data over a six-month test period. They found that the prediction power of the univariate ARIMA model was as good as the more complex time-series models tested. Mobarek and Keasey (2002) examined whether the Dhaka Stock Market (DSM) was weak form efficient between 1988 and 1997. They used AR, RW and ARIMA models to analyse the daily price indices of the Dhaka Stock Exchange (DSE) and found that the indices did not follow the RW model, and therefore that the DSM did not exhibit weak form efficiency.

Artificial intelligence (AI) models

The above statistical techniques assume a linear correlation structure and stationarity to exist in the time series. Additionally, conventional statistical methods require data to be normally distributed and also that historical data is available to ameliorate prediction. However, the time series of stock market prices usually exhibit strong nonlinear trends, seasonal, cyclical patterns, random walks and are inherently sensitive to a multitude of interdependent factors. Such non-stationary characteristics of stock market prices make the data chaotic and predictions complex. These factors recently influenced stock market price analysts to utilise artificial intelligence (AI) techniques for more reliable predictions.

Schalkoff (1990) defines AI as a field of study that seeks to explain and emulate intelligent behaviour with the aid of computational processes. AI combines mathematics with engineering to create machines to perform functions which necessitate significant brainpower when they are performed by human beings. In AI, the evolutionary algorithms (EAs) refer to a collection of algorithms which are stimulated from biological and natural evolutionary processes such as selection, mutation and reproduction. EAs work well with situations that have no exact solution due to computational intensity and/or situations where humans do not know how to find solutions. Most optimisation problems fall into this nature of predicament and EAs often identify whether good solutions exist. EAs have been applied to stock market prediction where the solution space is complex and irregular, and it is therefore imprudent to employ conventional optimisation technique/s to explore the global optimum. The genetic algorithm (GA), a search heuristic function, is the most common EA and is often applied to optimisation riddles. For example, Bonde and Khaled (2012) used GA and EA to predict the stock prices of eight companies and found that GA was slightly superior in predicting five of the eight company prices.

Fuzzy logic (FL), which was first introduced by Zadeh (1965), is based on fuzzy set theory, which provides a foundation for the set communion to be in the intermediate values between 'in the set' and 'not in the set'. Thus, FL is established on the norm of the 'degree of truth', that is the truth values of the statements to be any value between 1 (absolute truth) and 0 (absolute false) rather than the 'precise/exact information'. FL enables computers to mimic human cognition effectively by enabling them to make the most logical decisions based on incomplete sets of information. Hence, FL can be applied effectively to find meaningful solutions to noisy, ambiguous, incomplete and chaotic time series such as stock market prices. For example, Govindasamy and Thambidurai (2013) evaluated price prediction of five automobile stocks from July 2012 to April 2013. They used the probabilistic fuzzy logic (PFL) model and evaluated the performance of the model using mean absolute percentage error (MAPE). They found that the PFL model was reliable in the forecasting of all the analysed stocks.

An artificial neural network (ANN) is an information-processing paradigm which has its origin in the biological nervous systems. ANN is also known as neural networking, connectionism, neural computing, adaptive networking, parallel distributed processing and collective computing. The work of McCulloch and Pitts (1943), Widrow and Hoff (1960), and Rosenblatt (1961) are examples of such attempts. Since then ANNs have been used in a wide range of models for pattern/character recognition (Lo et al., 1995), image compression (Watta, Desaie, Dannug, & Hassoun, 1996), the Travelling Salesman's Problem (Ghaziri & Osman, 2003), medical research (Itchhaporia, Snow, Almassy, & Oetgen, 1996; Raji, 2016) and financial market prediction, among others.

Financial market predictions using ANNs have been applied to stock markets (Atiya, Talaat, & Shaheen 1997; Sutheebanjard & Premchaiswadi, 2010; Guresen, Kayakutlu, & Daim, 2011), currency markets (Kondratenko & Kuperin, 2003), futures markets (Kim, 2004), business failure predictions (Tang & Chi, 2005), debt and credit risk assessments (Dissananayake, Hendahewa, & Karunananda, 2007), bank theft detection (Guazzoni & Ronsivalle, 2009), bank failure analysis (Al-Shayea & El-Refae, 2012) and credit card approvals (Ilgun, Mekiš, & Mekiš, 2014).

ANNs have proven applicability to many tasks that are hard to solve using other methods, for example the pronunciation-based NETtalk program (Sejnowski & Rosenberg, 1987), which learns to pronounce written text using neural networks; handwritten digit recognition (Le Cun et al., 1989), which uses neural network methods to read zip codes on handwritten envelopes; or the driving-based Autonomous Land Vehicle In a Neural Network (ALVINN) that controls and guides NavLab vehicles on a highway (Pomerleau, 1989; Pomerleau, 1993).

Compared to statistical techniques, AI techniques such as ANNs, FL and EAs have certain advantages in handling nonlinearities and nonstationarities inherent in the time series of stock market prices. For example, back propagation neural network (BPNN), which is one of the commonly used neural networks in financial time-series forecasting, trains a given feed-forward multilayer neural network to a set of input patterns with known classifications. Sutheebanjard and Premchaiswadi (2010) evaluated the BPNN technique and the adaptive evolution strategy (ES) to predict the movement of the Stock Exchange of Thailand Index (SET Index). They found that BPNN predicted the SET Index with substantially lower prediction error in comparison to the evolutionary strategy (ES).

Guresen et al. (2011) evaluated the effectiveness of neural network models, namely the multi-layer perceptron (MLP), the dynamic artificial neural network (DAN2), and hybrid neural networks which incorporate GARCH to extract new input variables. In their study, they used mean square error (MSE) and mean absolute deviation (MAD) to compare the models, and found MLP to be the best.

Conversely, Govindasamy and Thambidurai (2013) used a probabilistic fuzzy logic (PFL) model to evaluate price prediction of five automobile stocks. They found that the PFL model was suitable for forecasting in all the analysed stocks.

Accommodating a technical analyst approach, Kamijo and Tanigawa (1990) used recurrent neural network approach to analyse stock price patterns of the Tokyo Stock Exchange. They used all corporations listed in the first section of the Tokyo Stock Exchange and 16 triangles were evaluated. Based on their analysis, they found accurate recognition of 15 out of 16 experiments and the average number of mismatching patterns was a 1.06 per name. Wang and Leu (1996) used an ARIMA-based recurrent neural network (RNN) system to forecast the Taiwan Stock Exchange Weighted Stock Index (TSEWSI). They used weekly data from four years (1991-1994) to train the algorithm and found that the algorithm was capable of accurately forecasting for up to four weeks. On the other hand, Saad, Prokhorov and Wunsch (2010) evaluated the prediction accuracy of three neural networks – the time delay neural network (TDNN), the recurrent neural network (RNN) and probabilistic neural networks (PNN). Using daily data relating to selected stocks, they found the RNN to be more powerful in comparison to the other methods evaluated, although implementation complexities existed.

Overview of the hybrid models applied in stock market prediction

The time series of stock market prices exhibit deterministic features combined with varying degrees of stochasticity. These time series mostly exhibit nonlinear characteristics (trends, seasonal changes, cyclical patterns, random walks, high volatility) combined with some degree of linearity. Thus, they are neither purely linear nor purely nonlinear, and it is unwise to implement purely statistical technique/s or purely AI technique/s to forecast the time series of stock market prices. Stated differently, purely statistical techniques and purely AI techniques have their own inherent advantages and disadvantages, however, stock market time series are neither purely linear nor purely nonlinear. As such, the recent literature on stock market price prediction focuses on hybridised methodologies.

Hybridisation refers to either combining alternative statistical techniques, combining statistical techniques with AI techniques, combining short-term models with medium-term models, or combining parametric and nonparametric classifiers, and so on. The primary objective of a hybridised methodology is to blend different techniques to accommodate their strengths, to form a model that performs globally optimal forecasting. The recent literature confirms that hybrid models are better suited to the prediction of stock market prices. Thus, in the last few years, there has been a rapid development of hybrid structures applied to stock market price and index predictions.

The input variables and other relevant characteristics of hybrid models most commonly applied in stock-index prediction are presented in the next sections.

Stock market	Articles
Developed markets (America, Europe, Middle East and Pacific)	
Standard & Poor's 500 (S&P 500) Index	Khadka et al. (2012), Talarposhti et al. (2016), Chen and Chen (2015), Javedani Sadaei and Lee (2014)
National Association of Securities Dealers Automated Quotations (NASDAQ) Index	Khadka et al. (2012), Talarposhti et al. (2016), Javedani Sadaei and Lee (2014)
New York Stock Exchange (NYSE)	Talarposhti et al. (2016)
Dow Jones Industrial Average (DJIA)	Javedani Sadaei et al. (2016), Kao et al. (2013), Chen and Chen (2015), Talarposhti et al. (2016), Javedani Sadaei and Lee (2014), Hsieh et al. (2011)
London FTSE-100 Index (FTSE)	Hsieh et al. (2011)
IBM stock	Bisoi and Dash (2014)
Oracle stock	Bisoi and Dash (2014)
Nikkei 225 Index, Japan	Kao et al. (2013), Hsieh et al. (2011)
Hang Seng Stock Index (HSI)	Wei (2016)
Emerging Markets	
Taiwan Capitalization Weighted Stock Index (TAIEX)	Javedani Sadaei et al. (2016), Talarposhti et al. (2016), Cheng et al. (2010), Chang et al. (2011), Wei (2013), Cheng et al. (2013), Chen and Chen (2015), Chen (2014), Wei (2014), Cheng and Wei (2014), Wei (2016), Javedani Sadaei and Lee (2014), Hsieh et al. (2011)
Shanghai Composite Index of China (SSE)	Kao et al. (2013), Guo et al. (2015), Sun (2014)
Shenzhen Composite Index	Sun (2014), Wang et al. (2012)
HuShen 300 Index	Sun (2014)
Bombay Stock Exchange (BSE)	Bisoi and Dash (2014), Kaur et al. (2016)
Reliance stock market (RIL) stock	Bisoi and Dash (2014)
BOVESPA Stock Index	Kao et al. (2013), Chen and Chen (2015)
Other	
Forex market	Falat and Marcek (2014), Javedani Sadaei et al. (2016),
Enrolment records of the University of Alabama	Chen (2014)

Table 1: List of the surveyed stock markets

SURVEYED STOCK MARKETS AND CHARACTERISTICS OF HYBRID MODELS

Table 1 provides the names of the stock markets where the authors have obtained data for their hybrid model analysis. Each study analysed either a single stock market index or multiple stock market indices. For example, TAIEX data was used by Chang, Wei and Cheng (2011); Cheng, Chen and Wei (2010); Wei (2013); Cheng, Wei, Liu and Chen (2013); Wei, Cheng and Wu (2014); Cheng and Wei (2014); and Hsieh, Hsiao and Yeh (2011), whilst the Bombay Stock Exchange Index (BSE30) was used by Kaur, Dhar and Guha (2016). On the other hand, Sadaei, Enayatifar, Guimarães, Mahmud and Alzamil (2016) used data from the Dow Jones Industrial Average (DJIA), TAIEX and Forex markets, whilst Talarposhti et al. (2016) used data from the TAIEX, National Association of Securities Dealers Automated Quotations (NASDAQ), Standard and Poor's 500 (S&P 500), DJIA and the New York Stock Exchange (NYSE).

Articles in Table 1 can be classified according to the Morgan Stanley Capital International Inc. (MSCI) classification. MSCI classifies the global equity market into developed markets and emerging markets. Developed markets include North American countries (Canada and the US), Europe and the Middle East (Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Israel, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the UK), and the Pacific (Australia, Hong Kong, Japan, New Zealand, and Singapore). Emerging markets include American countries (Brazil, Chile, Colombia, Mexico and Peru), Europe/Middle East/Africa (the Czech Republic, Egypt, Greece, Hungry, Poland, Qatar, Russia, South Africa, Turkey, and the

Articles	Input variables	Time frequency Window	Time period analysed	Duration of training phase	Duration of perdition phase
Wang et al. (2012)	Two indices: Shenzhen Integrated Index (SZII) and DJIA	Monthly	January 1993 to December 2010	168 months	48 months
Sadaei et al. (2016)	Two indices and nine currencies: TAIEX, DJIA, nine main exchange rates versus USD	Daily	TAIEX 1991-1999, DJIA 2001-2009 and exchange rate 2014	First 10 months for training	Last two months prediction purpose
Talarposhti et al. (2016)	46 case studies from five indices: TAIEX, NASDAQ, S&P 500, DJIA and NYSE	Daily	TAIEX & NASDAQ (1991-1999), S&P 500, DJIA (2001-2009) & NYSE (2004-2013)	First 10 months for training	Last two months prediction purpose
Kao et al. (2013)	Four indices: SSE, Bovespa index, Brazil, DJIA, USA and Nikkei	Daily	18/04/2006 to 01/04/2010.	800 trading days (80% of the data)	200 trading days (20% of the data)
Guo et al. (2015)	36 variables from one index: Shanghai stock market index (SSE)	Daily	04/01/2000 to 31/12/2004	957 trading days: 4 Jan 2000 to 31 Dec 2003	243 trading days: set: 1 Jan 2004 to 31 Dec 2004
Cheng et al. (2010)	One index: TAIEX	Daily	04/01/2000 to 30/12/2005	First 10 months for training	Last two months prediction purpose
Chang et al. (2011)	Two case studies from one index: TAIEX	Daily	TAIEX prices from 1997-2003	First 10 months for training	Last two months prediction purpose
Sun (2014)	OP, CP, HP, LP and TV of SSE, Shenzhen Composite Index, and HuShen 300 Index	Daily	05/01/2012 to 17/04/2012.	First 60 days	Last 5 days
Wei (2013)	TAIEX	Daily	Daily 2000-2005	1344 trading days	282 trading days
Falat and Marcek (2014)	USD/CAD exchange rate	Daily closing	31/10/2008 to 31/10/2012	912 trading days	132 trading days
Khadka et al. (2012)	S&P 500 and NASDAQ indices and three stocks	Daily	2/5/2011 to 27/5/2011 and 13/09/2004 to 21/01/2005	Not stated	Not stated
Bisoi and Dash (2014)	BSE, IBM stock market, RIL, and Oracle stock market	Daily	Daily 03/01/2005 to 13/04/2008	500 trading days	400 trading days
Cheng et al. (2013)	TAIEX	Daily	Daily 1997-2003	First 10 months for training	Last two months prediction purpose
Chen and Chen (2015)	TAIEX, DJIA, S&P 500 and IBOVESPA stock indices	Daily	Daily 1998-2003	First 10 months for training	Last two months prediction purpose
Chen (2014)	TAIEX, 2000-2003, and enrolment records of the University of Alabama 1971 to 1992	Daily	Daily TAIEX 2000 to 2003 AND University of Alabama 1971 to 1992	First 10 months for training	Last two months prediction purpose
Wei et al. (2014)	TAIEX	Daily	Daily TAIEX data from 1994-2003	First 10 months for training	Last two months prediction purpose
Cheng and Wei (2014)	TAIEX	Daily	Daily TAIEX data from 2000-2006	First 10 months for training	Last two months prediction purpose
Kaur et al. (2016)	BSE	Daily	Daily Bombay Stock Exchange Index (BSE30) 2006-2012	1400 trading days	350 trading days
Javedani Sadaei and Lee (2014)	TAIEX, NASDAQ (National Association of Securities Dealers Automated Quotations) from 1990-1999, DJIA, and S&P 500, 2000-2009	Daily	TAIEX, NASDAQ from 1990-1999 and DJIA, and S&P 500 from 2000-2009	January-October	November-December
Wei (2016)	TAIEX and Hang Seng Stock Index (HSI)	Daily	TAIEX datasets from 2000-2006 and Hang Seng index (HSI) datasets from 2000-2004	January-October	November-December
Hsieh et al. (2011)	DJIA, FTSE, Nikkei and TAIEX	Daily	DJIA, FTSE, Nikkei and TAIEX 1997-2003 and 2002-2008	1997-2003 first 10 months 2002-2008 first 6 months	1997-2003 last 2 months 2002-2008 last 6 months

Table 2: Input variables, time frequency and duration of training and prediction phases

United Arab Emirates), and Asia (China, India, Indonesia, Korea, Malaysia, the Philippines, Taiwan and Thailand).

For example, as indices in the developed markets, Khadka, George, Park and Kim (2012), Talarposhti et al. (2016), and Chen and Chen (2015) used the S&P 500; Khadka et al. (2012) used the NASDAQ; and Hsieh et al. (2011) used the DJIA, London FTSE-100 Index (FTSE) and the Nikkei-225 Index (Nikkei).

As indices in the emerging markets, Sadaei et al. (2016), Talarposhti et al. (2016), Cheng et al. (2010), Chang et al. (2011), Wei (2013), Cheng (2013), Chen and Chen (2015), Chen (2014), Wei (2014) and Hsieh et al. (2011) used the TAIEX; Kao, Chiu, Lu and Chang (2013), Guo, Wang, Yang and Miller (2015) and Sun (2014) used the Composite Index of China (SSEC); Kaur et al. (2016), and Bisoi and Dash (2014) used the BSE.

VARIABLES, DATA FREQUENCY WINDOW AND DATA SOURCE OF HYBRID MODELS

Table 2 summarises the details of input variable/s, adapted time frequency window/s, analysed time period/s and the duration of training and prediction phases applied in the evaluated hybrid models.

The number of input variables analysed to justify the proposed algorithm varies from one to 46. For example, Cheng et al. (2010), Wei (2013), Cheng et al. (2013), Wei et al. (2014), and Cheng and Wei (2014) used the TAIEX, whilst Kaur et al. (2016) used the BSE30. On the other hand, Talarposhti et al. (2016) used 46 case studies from five indices (TAIEX, NASDAQ, S&P 500, DJIA and NYSE), whilst Guo et al. (2015) used 36 variables from the SSE. Hsieh et al. (2011) used the DJIA, FTSE, Nikkei and TAIEX.

An exception was Wang, Wang, Zhang and Guo (2012) who used monthly closing prices of the Shenzhen Integrated Index (SZII) and monthly opening prices of the DJIA. All the other hybrid researchers used daily time-frequency data in their analysis.

HYBRID MODELS TESTED AND EVALUATION MEASURES APPLIED

Table 3 summarises the details of all the models tested and the measures applied to evaluate the predictive performance of the hybrid research assessed in our review. Applied evaluation methods can be classified as statistical and non-statistical measures.

Statistical measures include mean absolute error (MAE), root mean square error (RMSE), MAPE, Wilcoxon Rank-Sum, Morgan-Granger-Newbold (MGN), Relative Mean Absolute Error (RelMAE), correlation coefficient, symmetric mean absolute percentage error and directional accuracy (DA) among others. The RMSE seems to be the most frequent statistical measure adopted by hybrid researchers.

Wang et al. (2012), Sadaei et al. (2016), Talarposhti et al. (2016), Kao et al. (2013), Chang et al. (2011), Falat and Marcek (2014), Bisoi and Dash (2014), Chen and Chen (2015), Chen (2014), Wei et al. (2014), Cheng and Wei (2014), Kaur et al. (2016), Javedani Sadaei and Lee (2014), Wei (2016) and Hsieh et al. (2011) used purely statistical measures in performance evaluation.

Articles	Models tested	Evaluation measures applied
Wang et al. (2012)	ESM, ARIMA, BPNN, random walk model (RWM), equal weights hybrid model (EWH) and a hybrid approach combining ESM, ARIMA and BPNN	MAE, RMSE, absolute percentage error (APE), mean error (ME) and directional accuracy (DA)
Sadaei et al. (2016)	Chen (1996), Yu (2005)-average, Yu (2005)-distribution, Huarng (2006), Chen (2008), Chen (2011), and Javedani (2014) models, ARFIMA, ARIMA, exponential smoothing state pace (ETS) model and a hybrid method combining ARFIMA models and FTS for the forecast of long memory (long-range) time series	RMSE, MAPE, Wilcoxon Rank-Sum test (WRST)
Talarposhti et al. (2016)	Chen (1996), Yu1 (Yu [2005]-average), Yu2 (Yu [2005]-distribution), Huarng (2006), Cheng (2008), Chao (2011), multi-layer, ARIMA, GARCH, exponential smoothing state space (ETS) models, and a hybrid model based on LAPSO and polynomial fuzzy time series (PFTS) (PFTS–LAPSO)	APE, RMSE, RelMAE
Kao et al. (2013)	Wavelet-support vector regression (SVR), multivariate adaptive regression splines (MARS), single ARIMA, single support vector regression (SVR), a single adaptive neuro fuzzy inference system (ANFIS), and a hybrid model combining wavelet transform, MARS and SVR (Wavelet-MARS-SVR)	RMSE, MAD, MAPE, and root–mean–square percentage error (RMSPE)
Guo et al. (2015)	Principal component analysis (PCA), independent component analysis (ICA) and a hybrid model combining two-directional two-dimensional principal component analysis ([2D] 2PCA) and a radial basis function neural network (RBFNN)	Correlation coefficient actual value and prediction value (r1), non-linear regression multiple correlation coefficient (R2), correlation coefficient between actual return and prediction return (r2), percentage of correct direction, symmetric mean absolute percentage error, MAPE, RMSE, hit rate, total return
Cheng et al. (2010)	Rough set theory, genetic algorithms and a hybrid model combining cumulative probability distribution approach (CPDA), minimise entropy principle approach (MEPA), rough sets theory (RST) and Gas (CDPA, MEPA, RST, and GA)	Maximum accuracy (Table 17) and maximum return (Table 18)
Chang et al. (2011)	Chen (1996) model, Yu (2005) model, and a hybrid adaptive network based fuzzy inference system (ANFIS) model	RMSE
Sun (2014)	Grey neural network (GNN) model and a hybrid model based on the BSO approach and the GNN model (BSO-GNN)	MAE, RMSE, and MAPE
Wei (2013)	Chen (1996) model, Yu (2005) model, and Cheng et al. (2013) model, a hybrid model based on an ANFIS, and an adaptive expectation genetic algorithm	Performance evaluations and RMSE
Falat and Marcek (2014)	Standard RBF network as well as the statistical ARIMA and GARCH model, and a hybrid model based on customised radial basis function (RBF) neural network combined with GA	MSE
Khadka et al. (2012)	ARIMA, and a hybrid model based on concordance and genetic programming	MAE, RMSE, and Wilcoxon Rank-Sum test
Bisoi and Dash (2014)	Unscented kalman filter (UKF), differential evolution (DE), a real-time recurrent learning (RTRL) algorithm, and a hybrid of DE and unscented kalman filter (UKF) (DEUKF)	RMSE, and MAPE, and a MAPE-variant (AMAPE).
Cheng et al. (2013)	Chen (1996) model, Huarng, Yu (2006) model, and a hybrid model which incorporates (OWA) and ANFIS	RMSE
Chen and Chen (2015)	Three experiments: (experiment 1) six hybrid fuzzy time-series models developed by Chen (1996), Yu (2005), Chang et al. (2011), Hsieh et al. (2011), Cheng et al. (2013), and Chen and Kao (2013), and three different SVM models (SVR-Polynomial, SVR-RBF and SVR-Puk the daily stock index of the TAIEX, 1997-2003), divided year by year; (experiment 2) three hybrid fuzzy time-series models developed by Chen and Chen (2011), Chen et al. (2011), and Hsieh et al. (2011), and three different SVM models (SVR-Polynomial, SVR-RBF and SVR-Puk and the daily stock index of the DJIA, 1997 to 2003), divided year by year; (experiment 3) GARCH by Bollerslev (2003), Fuzzy GJR-GARCH by Hung (2011), and GJR-GARCH by Glosten et al. (1993), daily closing prices of the S&P 500 and IBOVESPA, 2000-2011 (divided year by year), and a hybrid model based on a granular computing approach with a binning-based partition and entropy-based discretisation method	MSE, RMSE, MAE, MPE, and MGN
Chen (2014)	Chen (1996), Yu (2005), Yu and Huarng (2010), and Chang et al. (2011) are compared with Song and Chissom (1994), Sullivan and Woodall (1994), Chen (1996), Cheng et al. (2006), Cheng et al. (2008), and a hybrid model developed from a high-order fuzzy time-series model based on entropy-based partitioning, and an adaptive expectation model for time-series forecasting.	RMSE
Wei et al. (2014)	Chen (1996) model, and Yu (2005) model, and a hybrid model based on AR model, moving average, subtractive clustering, ANFIS, and an adaptive equation to improve forecasting performance for stock markets	RMSE
Cheng and Wei (2014)	An autoregressive (AR1) model, a SVR model, and a hybrid time-series SVR model based on empirical mode decomposition (EMD)	RMSE
Kaur et al. (2016)	Random walk model, Chen (1996), Cheng, C.H. et al. (2013), and a hybrid network model based on a fuzzy inference system combined with fuzzy c-means clustering	RMSE
Javedani Sadaei and Lee (2014)	Chen (2011), Yu (2005 and 2006), and a hybrid model based on a multilayer stock forecasting model using FTS	RMSE
Wei (2016)	Chen (1996) model, Yu (2005) model, an AR model (Engle, 1992), the ANFIS (January 1993) model, a SVR model (Vapnik, 1995), and a hybrid time-series ANFIS model that is based on an EMD	RMSE
Hsieh et al. (2011)	Back propagation neural network, a neural network-based artificial bee colony (ABC), the Chen (1996) and Yu (2005) models, and a hybrid model developed from an RNN based on an ABC algorithm.	RMSE, MAE, MAPE and Theil's inequality coefficient (Theil U)

Table 3: Models tested and evaluation measures applied

Guo et al. (2015) used non-statistical measures such as hit rate and total return in addition to the use of statistical techniques.

COMBINATION OF STATISTICAL AND SOFT-COMPUTING TECHNIQUES USED IN THE HYBRID MODELS

Hybrid models applied in stock market prediction either combine different statistical techniques, or combine different soft-computing techniques, or combine statistical techniques with soft-computing techniques to produce better forecasting performances. In the last few years, there has been a rapid

Articles	Composition of hybrid model	Statistical techniques	Artificial neural networks	Fuzzy logic	Evolutionary algorithms
Wang et al. (2012)	A hybrid approach combining ESM, ARIMA, and BPNN is proposed	§	§		
Sadaei et al. (2016)	A hybrid method combining ARFIMA models and FTS for the forecast of long memory (long-range) time series is proposed	§		§	
Talarposhti et al. (2016)	New hybrid model based on LAPSO and PFTS (PFTS–LAPSO) is proposed		§	§	§
Kao et al. (2013)	A hybrid model combining wavelet transform, MARS, and SVR (Wavelet-MARS-SVR) is proposed	§			
Guo et al. (2015)	A hybrid model combining two-directional two-dimensional principal component analysis ([2D] 2PCA) and a RBFNN is proposed	§	§		
Cheng et al. (2010)	A hybrid model combining CDPA and MEPA approaches, RST, and GAs (CDPA, MEPA, RST and GA) is proposed	§	§	§	§
Chang et al. (2011)	A hybrid ANFIS model is proposed			§	
Sun (2014)	A hybrid model based on the BSO approach and the GNN model (BSO-GNN) is proposed		§		§
Wei (2013)	A hybrid model combining technical indicators, subtractive clustering, ANFIS, and GA is proposed			§	§
Falat and Marcek (2014)	A hybrid model based on customised RBFNN combined with GAs is proposed.		§		§
Khadka et al. (2012)	A hybrid model based on concordance and genetic programming is proposed			§	§
Bisoi and Dash (2014)	A hybrid model based on DE and the UKF algorithm (DEUKF) is proposed	§			§
Cheng et al. (2013)	A hybrid model which incorporates OWA and ANFIS is proposed			§	
Chen and Chen (2015)	A hybrid model based on a granular computing approach with a binning-based partition and entropy-based discretisation method is proposed	§		§	
Chen (2014)	A hybrid model based on a high-order FTS model based on an entropy-based partitioning and adaptive expectation model for time-series forecasting is proposed			§	§
Wei et al. (2014)	A hybrid model based on an AR model, moving average, subtractive clustering, ANFIS, and an adaptive equation to improve forecasting performance for stock markets is proposed	§		§	
Cheng and Wei (2014)	A hybrid time-series SVR model based on an EMD is proposed	§			
Kaur et al. (2016)	A hybrid network model based on fuzzy inference system combined with fuzzy c-means clustering is proposed			§	
Javedani Sadaei and Lee (2014)	A hybrid model based on multilayer stock forecasting model using FTS is proposed			§	
Wei (2016)	A hybrid time series ANFIS model that is based on an EMD is proposed	§		§	
Hsieh et al. (2011)	A hybrid model of wavelet transforms and RNN based ABC (ABC-RNN) is proposed	§	§		§

Table 4: Use of statistical and soft-computing techniques in hybrid models studied

development of hybrid models combining linear and/or non-linear models. These hybrid models can be evaluated in terms of the techniques they adopt: statistical techniques, fuzzy logic (FL), artificial neural networks (ANNs) and evolutionary algorithms (EAs). Table 4 summarises the combinations of statistical and soft-computing techniques used in the evaluated hybrid models.

COMBINATION OF STATISTICAL TECHNIQUES USED IN THE HYBRID MODELS

Wang et al. (2012), Sadaei et al. (2016), Kao et al. (2013), Guo et al. (2015), Cheng et al. (2010), Bisoi and Dash (2014), Chen and Chen (2015), Wei et al. (2014), Cheng and Wei (2014) and Wei (2016) all used some form of statistical technique such as exponential smoothing (ES), ARIMA, auto regressive fractional integrated moving average (ARFIMA), multivariate adaptive regression splines (MARS) and support vector regression (SVR). For example, Wang et al. (2012) used ES and ARIMA, and Sadaei et al. (2016) used ARFIMA. Kao et al. (2013) used MARS and SVM. Hsieh et al. (2011) used a wavelet transform-based approach in a proposed ABC-RNN model.

COMBINATION OF FUZZY LOGIC (FL) MODELLING USED IN THE HYBRID MODELS

Sadaei et al. (2016), Talarposhti et al. (2016), Chen (2014), and Javedani Sadaei and Lee (2014) used fuzzy time series (FTS). Chang et al. (2011), Cheng et al. (2013), Wei et al. (2014) and Wei (2016) used an adaptive network-based fuzzy inference system (ANFIS), which is a combination of statistical and fuzzy logic (FL) techniques. Talarposhti et al. (2016) used the learning automata particle swarm optimisation technique (LAPSO), which is a crossbreed of FL and EAs.

COMBINATION OF ARTIFICIAL NEURAL NETWORKS (ANNS) USED IN THE HYBRID MODELS

Wang et al. (2012) used a BPNN, whereas Guo et al. (2015), and Falat and Marcek (2014) used a radial basis function neural network (RBFNN). Sun (2014) used a grey neural network (GNN) and Hsieh et al. (2011) used a recurrent neural network (RNN).

COMBINATION OF EVOLUTIONARY ALGORITHM (EA) USED IN THE HYBRID MODELS

Cheng et al. (2010), Falat and Marcek (2014), and Khadka et al. (2012) used genetic algorithms (GA), whereas learning automata particle swarm optimisation (LAPSO) was adopted by Talarposhti et al. (2016). Sun (2014) used brain-storm optimisation (BSO), which is a form of the swarm intelligence algorithm and therefore a subset of EA. Hsieh et al. (2011) used the artificial bee colony (ABC) algorithm which is one of the population-based optimisation algorithms.

Conclusion

The time series of stock market prices and indices show both deterministic features and varying degrees of stochasticity. These time series predominantly exhibit nonlinear characteristics (interdependence, high volatility, turbulence, random walks, strong trends, seasonal and cyclical patterns, complexities, and so on) combined with some degree of linearity.

Stated differently, the time series of stock market prices are neither purely linear nor purely nonlinear, and it is therefore nonsensical to implement purely statistical techniques or purely AI techniques when attempting their forecast. As such, hybrid models combining different statistical techniques, or different soft-computing techniques, or a blend of statistical and soft-computing techniques, or short-term models with medium-term models, or parametric with nonparametric classifiers, and so on, are capable of predicting the influence of multidimensional determinants on stock market prices.

Recent literature reinforces the fact that many contemporary stock market researchers prefer to use hybrid forecasting models which are capable of capturing linear and non-linear patterns inherent in stock market prices. The hybrid models are known to be better suited for algorithm, training, generalisation performance, and prediction of stock market prices and indices.

Each hybrid model evaluated in this review managed to invent a slightly variant prediction model with sufficient justification. Based on the literature we have evaluated in our study, we find that there is no clearly identifiable, attested, established scalable multi-scale intelligent hybrid model/s utilised in stock market price and index prediction that stands out from the rest as more successful. However, most hybrid researchers have tested the invented algorithms against the Chen (1996), Yu (2005), Chang et al. (2011), and Cheng et al. (2013) models.

Acknowledgements

The effort and goodwill of Paul Pang has enabled us to complete this paper and we truly appreciate all his support.

References

Abraham, B., & Ledolter, J. (1986). Forecast functions implied by autoregressive integrated moving average models and other related forecast procedures. *International Statistical Review/ Revue Internationale de Statistique*, 51-66.

Al-Shayea, Q. K., & El-Refae, G. A. (2012, February). Evaluation of banks' insolvency using artificial neural networks. In *Proceedings of the 11th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Databases (AIKED'12)*, Cambridge, United Kingdom (pp. 22-24).

Atiya, A., Talaat, N., & Shaheen, S. (1997, June). An efficient stock market forecasting model using neural networks. In *IEEE International Conference on Neural Networks,* Vol. 4, 2112-2115.

Balaban, E., Bayar, A., & Faff, R. W. (2006). Forecasting stock market volatility: Further international evidence. *The European Journal of Finance*, *12*(2), 171-188.

Bisoi, R., & Dash, P. K. (2014). A hybrid evolutionary dynamic neural network for stock market trend analysis and prediction using unscented Kalman filter. *Applied Soft Computing*, *19*, 41-56.

Bley, J., & Olson, D. (2008). Volatility forecasting performance with VIX: Implied volatility versus historical standard deviation and conditional volatility. *Advances in Financial Planning and Forecasting*, (3), 67-91.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307-327.

Bonde, G., & Khaled, R. (2012, January). Stock price prediction using genetic algorithms and evolution strategies. In *Proceedings of the International Conference on Genetic and Evolutionary Methods (GEM).* (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control.* San Francisco, CA: Holden Day.

Brown, R. G. (1959). *Statistical forecasting for inventory control.* New York, NY: McGraw-Hill.

Brown, R. G. (1963). *Smoothing, forecasting and prediction of discrete time series.* Englewood Cliffs, NJ: Prentice-Hall.

Chang, J. R., Wei, L. Y., & Cheng, C. H. (2011). A hybrid ANFIS model based on AR and volatility for TAIEX forecasting. *Applied Soft Computing*, *11*(1), 1388-1395.

Chatfield, C. (1978). The Holt-Winters forecasting procedure. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *27*(3), 264-279.

Chen, M. Y. (2014). A high-order fuzzy time series forecasting model for internet stock trading. *Future Generation Computer Systems*, *37*, 461-467.

Chen, M. Y., & Chen, B. T. (2015). A hybrid fuzzy time series model based on granular computing for stock price forecasting. *Information Sciences*, *294*, 227-241.

Chen, S. M. (1996). Forecasting enrolments based on fuzzy time series. *Fuzzy Sets and Systems*, *81*(3), 311-319.

Cheng, C. H., & Wei, L. Y. (2014). A novel time-series model based on empirical mode decomposition for forecasting TAIEX. *Economic Modelling*, *36*, 136-141.

Cheng, C. H., Chen, T. L., & Wei, L. Y. (2010). A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. *Information Sciences*, *180*(9), 1610-1629.

Cheng, C. H., Wei, L. Y., Liu, J. W., & Chen, T. L. (2013). OWA-based ANFIS model for TAIEX forecasting. *Economic Modelling*, *30*, 442-448.

Dissananayake, B. M. M., Hendahewa, C. H., & Karunananda, A. S. (2007, August). Artificial Neural Network approach to credit risk assessment. In *Proceedings of the 2nd International Conference on Industrial and Information Systems*. (pp. 301-306). Piscataway, NJ: IEEE.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Falat, L., & Marcek, D. (2014). Financial time series modeling with hybrid model based on customized RBF neural network combined with genetic algorithm. *Advances in Electrical and Electronic Engineering*, *12*(4), 307.

Ghaziri, H., & Osman, I. H. (2003). A neural network algorithm for the traveling salesman problem with backhauls. *Computers & Industrial Engineering*, *44*(2), 267-281.

Girdzijauskas, S., Štreimikiene, D., Čepinskis, J., Moskaliova, V., Jurkonyte, E., & Mackevičius, R. (2009). Formation of economic bubbles: Causes and possible preventions. *Technological and Economic Development of Economy*, *15*(2), 267-280.

Govindasamy, V., & Thambidurai, P. (2013). Probabilistic fuzzy logic based stock price prediction. *International Journal of Computer Applications*, *71*(5).

Guazzoni, C., & Ronsivalle, G. B. (2009). An artificial neural network for bank robbery risk management: The OS. SI. F web on-line tool of the ABI Anti-crime Department. In *Proceedings of the International Workshop on Computational Intelligence in Security for Information Systems CISIS '08*. (pp. 1-10). Berlin and Heidelberg: Springer.

Guo, Z., Wang, H., Yang, J., & Miller, D. J. (2015). A stock market forecasting model combining two-directional two-dimensional principal component analysis and radial basis function neural network. *PloS one*, *10*(4), e0122385.

Guresen, E., Kayakutlu, G., & Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, *38*(8), 10389-10397.

Holt, C. C. (1957). *Forecasting seasonals and trends by exponentially weighted moving averages*. Pittsburgh, PA: Carnegie Institute of Technology.

Hsieh, T. J., Hsiao, H. F., & Yeh, W. C. (2011). Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm. *Applied Soft Computing*, *11*(2), 2510-2525.

Hyung, N., Poon, S. H., & Granger, C. (2006). The source of long memory in financial market volatility. In E.E. Rapach and M.E. Wohar (Eds.), *Forecasting in the presence of structural breaks and model uncertainty*. Bingley, UK: Emerald.

Ilgun, E., Mekiš, E., & Mekiš, E. Application of ANN in Australian credit card approval. *European Researcher*, *69(*2-2), 334-342.

Itchhaporia, D., Snow, P. B., Almassy, R. J., & Oetgen, W. J. (1996). Artificial neural networks: Current status in cardiovascular medicine. *Journal of the American College of Cardiology*, *28*(2), 515-521.

Javedani Sadaei, H., & Lee, M. H. (2014). Multilayer stock forecasting model using fuzzy time series. *The Scientific World Journal*, *2014*.

Kamijo, K. I., & Tanigawa, T. (1990, June). Stock price pattern recognition: A recurrent neural network approach. In *International Joint Conference on Neural Networks.* (pp. 215-221). Piscataway, NJ: IEEE.

Kao, L. J., Chiu, C. C., Lu, C. J., & Chang, C. H. (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decision Support Systems*, *54*(3), 1228-1244.

Kaur, G., Dhar, J., & Guha, R. K. (2016). Minimal variability OWA operator combining ANFIS and fuzzy c-means for forecasting BSE index. *Mathematics and Computers in Simulation*, *122*, 69-80.

Kavussanos, M., & Visvikis, I. (2005, June). *The predictability of non-overlapping forecasts: Evidence from the derivatives market in Greece*. Working paper presented at the European Financial Management Association Milan Meetings, Athens University of Economics and Business, Athens, Greece.

Khadka, M. S., George, K. M., Park, N., & Kim, J. B. (2012, August). Performance analysis of hybrid forecasting model in stock market forecasting. *International Journal of Managing Information Technology*, *4*(3), 81-88.

Kim, K. J. (2004). Artificial neural networks with feature transformation based on domain knowledge for the prediction of stock index futures. *International Journal of Intelligent Systems in Accounting, Finance & Management*, *12*(3), 167-176.

Kondratenko, V. V., & Kuperin, Y. A. (2003). Using recurrent neural networks to forecasting of forex. *Statistical Finance*. arXiv:cond-mat/0304469.

Le Cun, Y., Jackel, L. D., Boser, B., Denker, J. S., Graf, H. P., Guyon, I., ... & Hubbard, W. (1989). Handwritten digit recognition: Applications of neural network chips and automatic learning. *IEEE Communications Magazine*, *27*(11), 41-46.

Lo, S. C. B., Chan, H. P., Lin, J. S., Li, H., Freedman, M. T., & Mun, S. K. (1995). Artificial convolution neural network for medical image pattern recognition. *Neural Networks*, *8*(7), 1201-1214.

Maris, K., Pantou, G., Nikolopoulos, K., Pagourtzi, E., & Assimakopoulos, V. (2004). A study of financial volatility forecasting techniques in the FTSE/ASE 20 index. *Applied Economics Letters*, *11*(7), 453-457.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115-133.

McMillan, D. G. (2003). Non-linear predictability of UK stock market returns. *Oxford Bulletin of Economics and Statistics*, *65*(5), 557-573.

Mobarek, A., & Keasey, K. (2002). *Weak-form market efficiency of an emerging market: Evidence from Dhaka Stock Market of Bangladesh*. Paper presented at the ENBS Conference, Oslo, May 2000. Retrieved from https://www.researchgate.net/publication/250141858_Weak-form_market_efficiency_of_an_emerging_Market_Evidence_from_Dhaka_Stock_Market_of_Bangladesh

Morgan Stanley Capital International (MSCI) indexes. (n.d.) Retrieved from https://www.msci.com/market-cap-weighted-indexes

Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association*, 299-306.

O'Donovan, T. M. (1983). *Short term forecasting: An introduction to the Box-Jenkins approach*. New York, NY: Wiley, 256.

Pegels, C. C. (1969). Exponential forecasting: Some new variations. *Management Science*, 311-315.

Pereira, R. (2004). Forecasting Portuguese stock market volatility. *Journal of Forecasting*, *18*, 333-343.

Pomerleau, D. A. (1989). *Alvinn: An autonomous land vehicle in a neural network* (No. AIP-77). Pittsburgh, PA: Artificial Intelligence and Psychology Project, Carnegie–Mellon University.

Pomerleau, D. A. (1993). *Neural network perception for mobile robot guidance*. Dordrecht, The Netherlands: Kluwer.

Raji, C. G. (2016). Artificial neural networks in prediction of patient survival after liver transplantation. *Journal of Health and Medical Informatics*, *2016*.

Reid, D. J. (1969). *A comparative study of time series prediction techniques on economic data*. Nottingham, UK: University of Nottingham.

Roberts, S. A. (1982). A general class of Holt–Winters type forecasting models. *Management Science*, *28*(7), 808-820.

Rosenblatt, F. (1961). *Principles of neurodynamics. Perceptrons and the theory of brain mechanisms* (No. VG-1196-G-8). Buffalo, NY: Cornell Aeronautical Laboratory.

Saad, E. W., Prokhorov, D. V., & Wunsch, D. C. (1998). Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. *IEEE Transactions on Neural Networks*, *9*(6), 1456-1470.

Sadaei, H. J., Enayatifar, R., Guimarães, F. G., Mahmud, M., & Alzamil, Z. A. (2016). Combining ARFIMA models and fuzzy time series for the forecast of long memory time series. *Neurocomputing*, *175*, 782-796.

Schalkoff, R. J. (1990). *Artificial intelligence engine*. New York, NY: McGraw-Hill.

Sejnowski, T. J., & Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English text. *Complex Systems*, *1*(1), 145-168.

Slutzky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica: Journal of the Econometric Society*, 105-146.

Sun, Y. (2014). A hybrid approach by integrating brain storm optimization algorithm with grey neural network for stock index forecasting. *Abstract and Applied Analysis*. http://dx.doi.org/10.1155/2014/759862

Sutheebanjard, P., & Premchaiswadi, W. (2010, April). Stock Exchange of Thailand index prediction using back propagation neural networks. In *Second International Conference on Computer and Network Technology*. (pp. 377-380). Piscataway, NJ: IEEE.

Talarposhti, F. M., Sadaei, H. J., Enayatifar, R., Guimarães, F. G., Mahmud, M., & Eslami, T. (2016). Stock market forecasting by using a hybrid model of exponential fuzzy time series. *International Journal of Approximate Reasoning*, *70*, 79-98.

Tang, T. C., & Chi, L. C. (2005). Neural networks analysis in business failure prediction of Chinese importers: A between-countries approach. *Expert Systems with Applications*, *29*(2), 244-255.

Taylor, J. W. (2004). Volatility forecasting with smooth transition exponential smoothing. *International Journal of Forecasting*, *20*(2), 273-286.

Walker, G. (1931). On periodicity in series of related terms. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *131*(818), 518-532.

Wang, J. H., & Leu, J. Y. (1996, June). Stock market trend prediction using ARIMA-based neural networks. In *IEEE International Conference on Neural Networks*, 4(6), 2160-2165).

Wang, J. J., Wang, J. Z., Zhang, Z. G., & Guo, S. P. (2012). Stock index forecasting based on a hybrid model. *Omega*, *40*(6), 758-766.

Watta, P., Desaie, B., Dannug, N., & Hassoun, M. (1996). *Image compression using backprop*. Computation and Neural Networks Laboratory. Retrieved from http://neuron.eng.wayne.edu/prop.html

Wei, L. Y. (2013). A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX. *Economic Modelling*, *33*, 893-899.

Wei, L. Y. (2016). A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting. *Applied Soft Computing*, *42*, 368-376.

Wei, L. Y., Cheng, C. H., & Wu, H. H. (2014). A hybrid ANFIS based on n-period moving average model to forecast TAIEX stock. *Applied Soft Computing*, *19*, 86-92.

Widrow, B., & Hoff, M. E. (1960, August). Adaptive switching circuits. In *IRE WESCON Convention Record*, *4*(1), 96-104.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, *6*(3), 324-342.

World Federation of Exchanges. (2016). Retrieved from https://www.world-exchanges.org

Yaglom, A. M. (1955). The Correlation Theory of Processes whose nth difference constitutes a stationary process. *Matem Sbornik*, *37*, 141-96.

Yu, H. K. (2005). Weighted fuzzy time series models for TAIEX forecasting. *Physica A: Statistical Mechanics and its Applications*, *349*(3-4), 609-624.

Yule, G. U. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *226*, 267-298.

Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, *8*(3), 338-353.

AUTHORS

Wajira Dassanayake, Business Practice Pathway, Unitec Institute of Technology, Auckland, New Zealand.

Iman Ardekani, Computer Science Pathway, Unitec Institute of Technology, Auckland, New Zealand.

Chandimal Jayawardena, Department of Information, Systems Engineering, Sri Lanka Institute of Information Technology, Sri Lanka.

Hamid Sharifzadeh, Computer Science, Pathway, Unitec Institute of Technology, Auckland, New Zealand.