DOCTOR OF PHILOSOPHY

Essays in forecasting financial markets with predictive analytics techniques

Alroomi, Azzam J.M.A.H.

Award date:
2018

Awarding institution:
Bangor University

ESSAYS IN FORECASTING FINANCIAL MARKETS WITH PREDICTIVE ANALYTICS TECHNIQUES

Azzam Alroomi

Ph.D. Candidate in Financial Studies

Thesis Submitted in Candidature for the Degree of Doctor of Philosophy at Bangor University

Supervisors:

Prof. Konstantinos Nikolopoulos

Dr. Azhdar Karami

Dr. Ru Xie

July 2018

Extended Abstract

This PhD dissertation comprises four essays on forecasting financial markets with unsupervised predictive analytics techniques, most notably time series extrapolation methods and artificial neural networks. Key objectives of the research were reproducibility and replicability, which are fundamental principles in management science and, as such, the implementation of all of the suggested algorithms has been fully automated and completely unsupervised in R.

As with any predictive analytics exercise, computational intensiveness is a significant challenge and criterion of performance and, thus, both forecasting accuracy and uncertainty as well as computational times are reported in all essays. Multiple horizons, multiple methods and benchmarks and multiple metrics are employed as dictated by good practice in empirical forecasting exercises.

The essays evolve in nature as each one is based on the previous one, testing one more condition as the essays progress, outlined in sequence as follows: which method wins overall in a very extensive evaluation over five frequencies (yearly, quarterly, monthly, weekly and daily data) over 18 time series of stocks with the biggest capitalization from the FTSE 100, over the last 20 years (first essay); the impact of horizon in this exercise and how this promotes different winners for different horizons (second essay); the impact of using uncertainty in the form of maximum-minimum values per period, despite still being interested in forecasting the mean expected value over the next period; and introducing a second variable capturing all other aspects of the behavioural nature of the financial environment – the trading volume – and evaluating whether this improves forecasting performance or not.

The whole endeavour required the use of the High Performance Computing Wales (HPC Wales) for a significant amount of time, incurring computational costs that ultimately paid off in terms of increased forecasting accuracy for the AI approaches; the whole exercise for one series can be repeated on a fast laptop device (i7 with 16 GB of memory).

Overall (forecasting) horses for (data) courses were once again proved to perform best, and the fact that one method cannot win under all conditions was once more evidenced. The introduction of uncertainty (in terms of range for every period), as well as volume as a second variable capturing environmental aspects, was beneficial with regard to forecasting accuracy and, overall, the research provided empirical evidence that predictive analytics approaches have a future in such a forecasting context.

Given this was a predictive analytics exercise, focus was placed on forecasting levels (monetary values) and not log-returns; and out-of-sample forecasting accuracy, rather than causality, was a primary objective, thus multiple regression models were not considered as benchmarks.

As in any empirical predicting analytics exercise, more time series, more artificial intelligence methods, more metrics and more data can be employed so as to allow for full generalization of the results, as long as all of these can be fully automated and forecast unsupervised in a freeware environment – in this thesis that being R.

Keywords: Forecasting, Financial Markets, Predictive Analytics, R, Time Series, Neural Networks.

Essays: The four essays that this PhD dissertation consists of are (in order of development):

1. *Forecasting financial markets with unsupervised predictive analytics techniques*
2. *Forecasting financial markets with predictive analytics: the impact of the forecasting horizon*
3. *Forecasting financial markets with predictive analytics: the impact of uncertainty; in the form of the range of intra-period values*
4. *Forecasting financial markets with predictive analytics: the impact of exogenous variables; in the form of the trading volume*

Acknowledgement

I would like to thank my supervisor Professor Kostas Nikolopoulos for his profound support and guidance in the process of researching and writing this paper. I would also like to thank Bangor University's staff who have assisted me in the process of completing my thesis.

Table of Contents

Chapter 1

1.1 Introduction

Is Artificial Intelligence (AI) the future or are we still better off with statistical methods? This is a question that is commonly debated among academics. Some consider AI to be a myth and that soon everyone will realise this, while on the other hand some people have significantly invested in AI's future. *"Buy And Hold Is No Longer The Case"* the concept of buy and hold is no longer present nor is "*Sell High And Buy Low*". Conducting a test on neural network methods vis-à-vis statistical techniques is critical in demonstrating the efficacy of particular techniques. The following paper will outline the comparative results for each of the respective subjects, where an intensive computing test has been developed and implemented to compare the accuracy and results of the four main approaches to forecasting. Thus, the four main approaches here are tested using a combination of models, statistical techniques and AI. The data being used is that of companies in the FTSE 100. The selected companies' share prices have been collected, while the data consists of opening share price, closing share price, high & low price, adjusted closing price and trading volume. The closing share price is the main variable in our test and has been used in all four different approaches. Every approach covers 18 companies and forecasts for every company have been made using 32 different methods. To test the accuracy of the methods, seven error metrics were computed (ME, RMSE, MSE, MAE, MPE, MAPE and MASE), while logarithms were also computed. The companies' accuracy was ranked in accordance with their performance on the error metrics computation.

How does horizon forecasting affect accuracy testing? Many scholars claim that, for accuracy testing, the 80/20 rule is most suitable when determining initial data forecasting . As the 80/20 rule is becoming more widespread, and more forecasters are using it in their models, we come to a junction where accuracy testing and uncertainty meet. Under this dilemma and with all determination we ran four unique forecasting tests. Test 1 involved the implementation of the 80/20 rule. Model 2 looked at the data differently, where we used the first 100 data points as the first training actuals, and we forecast before adding one more data point and do the forecasting again by adding one point each time before repeating the forecast and so on each time adding one more data point. This model of training data was carried over into the remaining two tests. Test 3 used the same rule of actual plus 1 and forecast, however, model 2 introduced variables of the training data that were inputted into the neural network. Test 4 carries over the same method of training from model 2, however, here we have an exogenous variable, where we introduce the volume traded into the neural network.

All handling of data was kept consistent including methods, period-range, software, error matrices and time series, in order to avoid bias. This would also provide an unbiased conclusion for the four models, and with an introduction of horizon determination we were also able to determine which algorithms would work best in the future.

1.1.2 Introduction to Financial Indices

The main index being tested was the Financial Times Stock Exchange 100 Index (FTSE 100). The FTSE 100 is a share index composed of the top 100 companies with the highest market capitalization in the London Stock Exchange. In the FTSE index, companies are relegated and promoted quarterly. Other indices will be tested in the future including the Dow Jones Industrial Average (DJIA), the Standard and Poor's 500 (S&P 500), the Nikkei 225, the German stock index (DAX), and the Hang Seng Index (HSI). This would widen the spectrum of our test, allowing us to validate and test the methods and accuracy tests that were used.

1.1.3 Introduction to R

The implementation software deployed were "R" and "R Studio." These programmes are widely applied among scholars and academics because of their proven effectiveness. These software are also free to download and use, and they offer a larger number of packages that are easily accessible and freely available on the R platform. Methods are developed and tested by academics to be latterly developed and compacted into packages to be used freely on the R platform. They can adapt more easily to larger datasets compared to other software . "R" and "R Studio" software also have some other important features that make them effective as implementation software. For instance, they have features such as syntax highlighting, smart indentation, integrated R documentation and help, extensive package development tools as well as an interactive debugger for diagnosing and fixing errors quickly. Their superior functionalities and ease of access aid in improving the accuracy, reliability, credibility, replicability and significance of the research findings.

1.2 AI Perspective & Implementation in Financial Markets

Trading in financial markets is a topic that continues to garner immense attention from scholars. Such trading occurs in a variety of markets including stock markets, bond markets, and foreign exchange markets. A number of models have also provided for extensive coverage in examining trading in financial markets. Some of these models and algorithms include: the support vector regression; filtered flag pattern recognition; artificial neural networks; GARCH models; soft computing technologies; hybridized market indicators; multivariate adaptive regression; and simple linear regression. This section explores the existing empirical and theoretical studies that have examined issues related to trading and forecasting in financial markets in order to shed light on the efficacy of different models applicable in financial markets trading and forecasting.

Sermpinis, Stasinakis, Theofilatos & Karathanasopoulos (2015) in their investigation sought to examine the efficacy of support vector regression forecast combinations in forecasting, modelling and trading EUR exchange rates. The hybrid model proposed by the researchers was the Rolling Genetic Algorithm – Support Vector Regression (RG-SVR) for feature subset combination and optimal parameter selection. Sermpinis et al. (2015) applied their proposed algorithm to the task of trading and forecasting the EUR/USD, EUR/JPY, and EUR/GBP exchange rates. Their proposed algorithm worked through genetically searching over a feature space and then combining the optimal feature subsets for each of the aforementioned exchange rates. The researchers derived individual forecasts from several nonlinear and linear models. After investigating the trading and statistical performance of RG-SVR against other established models, including genetically optimized SVMs and ARBF-PSO neural network, Sermpinis et al. (2015) established that RG-SVR had the best performance in terms of both trading efficiency and statistical accuracy for all of the exchange rates investigated.

From a different perspective, Sermpinis, Stasinakis, Rosillo & de la Fuente (2017) also generated two different wSVR (Locally Weighted Support Vector Regression) algorithms and applied these to the task of forecasting as well as trading a number of European exchange-traded funds, specifically covering the EU monetary debt crisis. After benchmarking the proposed wSVR models against traditional SVR models, Sermpinis et al. (2017) found that the wSVR models significantly outperformed the traditionally acclaimed SVR models. Jiang & He (2012), similar to the assertions made by Sermpinis et al. (2017), stated that while traditional SVR models may be effective in forecasting and trading exchanges with linearity in nature, financial trading tends to be nonlinear and non-stationary in nature, often characterised by noise. Furthermore, exogenous factors including political decisions and behavioural elements also affect financial trading to a large degree.

Using smooth transition GARCH models, Chen, Wang, Sriboonchitta & Lee (2017) endeavoured to explore the efficacy of pair trading strategies for quantile forecasting. As an illustration, Chen et al. (2017) conducted an empirical analysis and simulation study of daily stock returns from 36 stocks in the U.S. stock market. Similarly, Perlin (2009) stated that pair trading is an effective mean-reverting strategy, which assumes that the spread computed from two stock returns will revert to its historical trend, thereby having the capability of achieving profits from relatively low-risk and simple positions. To update the estimates and quantile forecasts, Chen et al. (2017) employed Bayesian Markov chain Monte Carlo sampling techniques. After a comprehensive analysis, the proposed pair trading strategies yielded annualized returns of at least 18.4% with a transaction cost and at least 35.5% without a transaction cost.

Unlike Chen et al. (2017) and Sermpinis, Stasinakis, Rosillo & de la Fuente (2017) who examined the efficacy of trading strategies using the GARCH model and SVR respectively, Kurek (2014) sought to examine trading in financial markets using multivariate adaptive regression and simple linear regression splines. The focus of his research was on equity block trades carried out in the Warsaw Stock Exchange market. As envisaged by the researcher, equity block trades relate directly to the valuation of a firm’s equity capital. Using the multivariate adaptive regression and simple linear regression splines, Kurek (2014) established that “equity block trade transactions carry an important signal for investors acting on a stock exchange” (quote taken from p.438). Following the execution of the equity block trade, there were significantly abnormal negative and positive returns.

Apart from the GARCH model, SVR, multivariate adaptive regression and simple linear regression, researchers have also examined the utility of neural networks in trading and forecasting financial markets. Sokolov-Mladenovic, Milovancevic, Mladenovic & Alizamir

(2016) conducted one such study in which they sought to examine the effectiveness of artificial neural networks in economic growth forecasting based on trade, export and import parameters. As an economic growth indicator, the researchers used gross domestic product (GDP). As such, the main purpose of their research was to develop and apply the artificial neural network with extreme learning machine and with a back propagation algorithm in order to forecast the GDP growth rate. Based on trade data, Sokolov-Mladenovic et al. (2016) established that, compared with the artificial neural network with back propagation, the neural network with extreme learning machine was more effective in forecasting GDP growth rate, a view also shared by Ravichandran, Thirunavukarasu, Nallaswamy & Babu (2005).

Ayodele, Ayo, Adebiyi & Otokiti (2012) also examined the efficacy of neural networks, but their emphasis was on stock price prediction. Consistent with the assertions made by Sokolov-Mladenovic et al. (2016), Ayodele et al. (2012) also stated that artificial neural networks are data mining techniques that have found extensive use in financial markets to aid investors to make qualitative decisions. These researchers also affirmed that the use of technical analysis is the predominant approach in stock market prediction using artificial neural networks. Ayodele et al. (2012), in their inquiry, proposed a hybridized approach that combined the use of technical analysis and fundamental analysis for predicting future stock prices. Results from the inquiry revealed that the hybridized model obtained remarkable improvements in stock market prediction compared to the use of technical analysis variables only, prompting researchers to propose that the model could be a satisfactory guide for investors and traders in making qualitative forecast decisions.

A number of studies have also examined the effect of investor sentiments in predicting stock trading returns. Hui & Li (2014) conducted such a study, focusing exclusively on evidence from the Chinese stock market. In this inquiry, the researchers examined cross-sectional analyses to examine the lead-lag relationship between HS300 index and proxy variables. Results from the study indicated that closed-end fund discount (CEFD), SSE share turnover (TURN), and net added accounts (NAA) were leading variables in stock market trading. Findings from this study also revealed that the relative degree of active trading in equity market (RDAT) and average first day return of IPOs (RIPO) were contemporary variables whereas the number of IPOs (NIPO) was a lagging variable in stock market trading. The sample tests from this investigation also indicated that the developed sentiment index proposed by Hui & Li (2014) was robust and had good predictive power for the Chinese stock market.

Joseph, Wintoki & Zhang (2011) similarly conducted a study to examine the forecasting trading volume and abnormal stock return using investor sentiment, albeit using an internet search strategy. Unlike Hui & Li (2014), Joseph et al. (2011) specifically argued that an online ticker search acts as a valid proxy for investor sentiment, generally associated with less sophisticated investors. Based on previous research on investor sentiment, Joseph et al. (2011) expected online search intensity to accurately forecast trading volume and stock returns, and that highly volatile stocks that are more difficult to arbitrage would have a higher sensitivity to search intensity compared to less volatile stocks. Over the 2005-2008 period, Joseph et al. (2011) found out that over a weekly horizon in a sample of S&P 500 firms, online search intensity was able to reliably predict trading volume and abnormal stock returns.

Whereas investor sentiments can have strong ramifications for trading in financial markets, empirical evidence also suggests that stock market trading activities can have a direct influence on forecasting recessions. Chatterjee (2016) endeavoured to examine this intricate relationship through examining current recession forecasting models using stock market liquidity as an additional forecasting variable. The three distinct facets of stock market trading activities investigated by the researcher were stock market liquidity, volatility, and returns as predictors of recessions experienced in the United States. Just like Chatterjee (2016), Erdogan, Paul & Ozyildirim (2015) also explored the relationship between stock market macro liquidity and recessions. Both studies showed a positive correlation between the variables under investigation.

In the study by Chatterjee (2016), vector autoregression results (VAR) showed that lower stock market liquidity is usually a signal of recession whereas returns can forecast recessions for two to three quarters in the future. Nevertheless, stock market volatility was not found to have any forecasting power. Moreover, relying upon the Survey of Professional Forecasters (SPF) estimates, both Chatterjee (2016) and Erdogan et al. (2015) established that stock market liquidity-based models significantly outperformed the survey estimates of professional forecasters' recession probabilities, highlighting the need for professional forecasters to incorporate stock market liquidity models in their forecasts. This conclusion is similar to the one arrived at by (Zeng, Zhang, Liu, Liang & Alsaadi, 2017).

From a different dimension, Arévalo, García, Guijarro & Peris (2017) also looked at stock market price forecasting and its relationship with trading efficacy but, unlike Chatterjee (2016), this study focused on a dynamic trading rule based upon filtered flag pattern recognition. Since the flag pattern proposed by the researchers was a trend-following pattern, they added the EMA indicator in order to filter trades. Arévalo et al. (2017) reiterated that the two main approaches used to make accurate decisions in financial markets are technical analysis and fundamental analysis. Cervelló-Royo, Guijarro & Michniuk (2015) also shared similar sentiments, adding that technical analysis rests on the assumption that historical behaviour or past stock prices have a testable effect on future stock-price evolution while fundamental analysis utilises macroeconomic, business and/or industry variables to predict a firm's stock value.

In a study by Arévalo et al. (2017), the researchers pursued the technical analysis approach, proposing a dynamic window scheme that permitted updating of the stop loss and take profit on a quarterly basis. In the technical analysis, the researchers calculated the EMA indicator both for 1-day and 15-minute timeframes, enabling them to consider the short and medium terms simultaneously. After applying their proposed methodology to 91,309 intraday observations of the DJIA index, Arévalo et al. (2017) found that the trading rule based upon filtered flag pattern recognition was significantly able to improve the results obtained by a buy & hold strategy for both risk and profitability. The results were also true after considering transaction costs.

Research also indicates that financial networks can have significant influences on trading performance, particularly in bond markets. Booth, Gurun & Zhang (2013) examined the influence of financial networks on trading performance and asset prices in the Turkish government bond market. The researchers hypothesized that by having more extensive and strategically placed financial networks, global financial institutions can acquire and process information about asset training more efficiently because of their better access to order flows. Hence, their trading performance may be better than that of local financial institutions that have less extensive financial networks.

From the transactional level data in the Turkish government bonds, Booth et al. (2013) found that global financial institutions had a higher tendency to trade in more liquid bonds, consistently trading at rates that were more favourable compared to local financial institutions. Seru, Shumway & Stoffman (2010) found results similar to those of Booth et al.

(2013), although in their inquiry they concluded that while the global financial institutions do enjoy better trading performance, this informational advantage tends to diminish over time, suggesting that local institutions learn as they trade with global financial counterparts.

1.3 Research Question

The research question that this paper covers asks is: "Is There a Perfect Model for Forecasting Financial Markets?" In endeavouring to answer this question, we came across a significant number of different methodologies, however some were more applicable then others in our case. Researching various models provided us with knowledge of how each model would work under the environments of our tests, however due to computational cost and applicability we had to choose methods that were more widespread and would adapt to our situation more easily. The process of method selection was very significant as it would ultimately decide the outcome of our paper. So, we chose as many methods as applicably as possible from both the spectrum of AI and statistical techniques. As 32 methods were selected, including the most used, tested, researched and implemented methods among academicians, we benefitted from a highly diversified method. After the selection of the methods, the procedure of testing and implementing started, will this lead into finding a perfect method, does a perfect method even exist, would it work under only one set of parameters and/or would it work for all frequencies and under a variety of different environments? If so, how would it work in real-life implementation? Of course, this would only happen if the model beats other models on all different frequencies and horizons, but this is merely the start. What we considered to be a significant accuracy test would also be a great determinant of how the model would perform. Many scholars have debated the use of accuracy testing and when, where and how it works. To be consistent, this paper covers five different mainstream frequencies and six different accuracy tests. The paper moves on to cover the obstacles to determining which models work best, and where and how. Thus, we cover all of the most widespread frequencies, horizons, algorithms, external regressors, exogenous variables and the introduction of neural networks.

Many journals claim to have beaten the market. However, upon closer examination of such claims, it usually transpires that they only beat the market under their own rules. By setting the rules you could indeed claim to have beaten the market. However, we do not set our own rules when it comes to financial markets. We are obliged to follow the market. Some claim to have beaten the market using one share price under one market under one horizon with the use of one accuracy test. Such claims do not stand up in reality . Setting one single set of parameters, in one single environment, and setting a single model, would not bring considerable success in the financial markets. Taking this into consideration, the application that this paper discusses and tests is whether models work regardless of any rules and settings.

1.4 Methods Influenced by the Literature

The method selection was heavily influenced by the literature review, including noticeable journals such as: *Cervelló-Royo (the stock market rule); Makirdakis M3-Competition, Spithourakis, G., Petropoulos, F., Nikolopoulos, K., & Assimakopoulos, V. (2015). Amplifying the learning effects via a Forecasting and Foresight Support System, Hyndman, R. J., & Athanasopoulos, G. (2014),. Forecasting: principles and practice.*

The method selection was also restrained be the algorithms computational cost, where the methods were influenced inpart by the mentioned literature other methods were selectede due to availability, consistency, method dynamic and compatibilty with our test.In each paper the method selection differed due to computational cost and compatibility.

1.5 Methods

Table 1: Companies

COMPANY		TICKER	SECTOR	MARKET CAP (£BN)	EMPLOYEES
1.	ROYAL DUTCH SHELL	RDSA	Oil and gas	160.12	90,000
2.	UNILEVER	ULVR	Consumer goods	90.42	171,000
3.	HSBC	HSBA	Banking	88.11	267,000
4.	BRITISH AMERICAN TOBACCO	BATS	Tobacco	71.4	87,813
5.	GLAXOSMITHKLINE	GSK	Pharmaceuticals	67.38	97,389
6.	SABMILLER	SAB	Beverages	67.32	70,000
7.	BP	BP	Oil and gas	63.13	97,700
8.	VODAFONE GROUP	VOD	Telecom	56.55	86,373
9.	ASTRAZENECA	AZN	Pharmaceuticals	51.23	57,200
10.	RECKITT BENCKISER	RB	Consumer goods	46.32	32,000
11.	DIAGEO	DGE	Beverages	46.01	25,000
12.	BT GROUP	BT.A	Telecom	45.61	89,000
13.	LLOYDS BANKING GROUP	LLOY	Banking	44.11	120,449
14.	BHP BILLITON	BLT	Mining	41.88	46,370
15.	NATIONAL GRID PLC	NG	Energy	36.14	27,000
16.	RIO TINTO GROUP	RIO	Mining	34.84	67,930
17.	PRUDENTIAL PLC	PRU	Finance	31.63	25,414
18.	ROYAL BANK OF SCOTLAND GROUP	RBS	Banking	28.6	150,000

1.5.1 Description

The table above shows the companies from which the data was derived. At the start of this project, the data was collected from Yahoo Finance, a widely used source providing freely accessible data. The source also provides different frequencies of the intended gathered data, which made it more convenient as our tests was going to be performed and compared for different frequencies. The data compromised of five different aspects: share price close, high, low, adjusted close and volume.

The ticker indicates how each company's stock is referred to in the financial markets, and in our case the companies are listed in the London Stock Exchange (LSE). The companies are also listed in the FTSE 100, an index in the LSE.

The sector shows the relevant industry in which each company operates. The industries are diverse, ranging from banking and finance to tobacco, mining and telecom.

The market cap indicates the respective companies' market capitalization in the FTSE 100. As part of our data collection criteria, each company's data were collected according to their capitalization. Finally, the employees section shows how many employees each company had on their books at the time the research was conducted.

All of the above data is from November 2015.

Table 2: Classifications of Test

Classifications of Test						
Frequency	Companies	Methods	Error Matrices	Horizons	Logarithms	Index
Daily	18	32	7	28	1	FTSE 100
Weekly	18	32	7	12	1	FTSE 100
Monthly	18	32	7	18	1	FTSE 100
Quarterly	18	32	7	12	1	FTSE 100
Yearly	18	32	7	4	1	FTSE 100

1.5.2 Description

Table 2 shows the classifications for the tests that were undertaken in this paper the same table will be presented on each chapter to distinguish the tests undertaken in that chapter. The total number of methods is shown in the table, and in each chapter a similar table is included to present each chapter's specific tests. Overall, 18 companies were tested in this project with a total of 32 methods and six error matrices and one log return test. The figures that cross over from chapter to chapter are the companies used, the index used, error matrices and the horizon for each frequency. From chapter to chapter, or paper to paper in this case, the methods and functions used change.

The following table indicates all of the methods that were used:

Table 3: Methods

METHODS		Description
1.	AUTOARIMA	autoarima function
2.	AUTOARIMA_SEASDUMMY	with seasonal dummies as external regressors
3.	AUTOARIMA_FOURIER	with fourier transform as external regressors
4.	SES	Simple exponential smoothing
5.	HOLT WINTERS	Double exponential smoothing with alpha and beta
6.	DOUBLE SEASONAL HOLT WINTERS	Holt-Winters with alpha, beta and gamma
7.	BATS	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components
8.	TBATS	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality. (*)
9.	DSHW	Double seasonal holt winters
10.	NAÏVE	Forecasted by the previous observation
11.	SNAIVE	Forecasted by the last observation in previous period/season.
12.	SINDEX	Seasonal index forecast
13.	NNET	Neural Networks
14.	TSLM	Time series linear model, with level and trend as the X vars
15.	SPLINEF	Splines model
16.	THETAF	Theta model
17.	RWF	Random walk model
18.	MEANF	Mean forecast
19.	STL	Seasonality-Trend-Level. The time series is broken down to these components, the remainder is forecasted and the STL components are added back to the forecasted value.
20.	AUTOARIMA_SES	autoarima function & Simple exponential smoothing
21.	AUTOARIMA_TBATS	autoarima function & state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality
22.	AUTOARIMA_NNET	autoarima function & Neural Networks
23.	SES_THETAF	Simple exponential smoothing & Theta model
24.	SES_MEAN	Simple exponential smoothing & Mean forecast Model
25.	TBATS_THETAF	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality & Theta model
26.	NNET_THETAF	Neural Networks & Theta model
27.	NYMPHY_EXOGONOUS_CLOSE	Neural Network with exogenous variables
28.	NYMPHY_CLOSE_HIGH	Neural Network with Close and High share price
29.	NYMPHY_CLOSE_LOW	Neural Network with Close and Low share price
30.	NYMPHY_CLOSE_HIGH_LOW	Neural Network with Close, High and Low share price
31.	NYMPHY_CLOSE_VOLUME	Neural Network with Close and Volume share price

(*1) But if seasonality is not present, both the models will not have seasonal components. So, they end up optimizing for the same model and, therefore, have the same forecasts.

1.5.3 Description

The aforementioned table 2 shows the functions and methods used in this thesis, ranging from simple statistical techniques to AI. We used as many as 32 methods in total. This aim was to test if the more complex, complicated and advanced can outperform the simpler methods. The widely used Autoarima model was tested as it is widely favoured by scholars and practitioners in their respective fields (Finance, Econometrics, etc). On the other side of the spectrum, the Nntar and Nymphy neural networks were tested as the AI models. Each model had its uniqueness in our testing. In addition, the models' un 2 iqueness arose when testing the respective time series, where some models performed contrarily to other models. Furthermore, some models depended on the time series' seasonality and some would depend on the time series' nonlinearity. Where AI would work better when it comes to nonlinear data. And some would be better equipped to adapt to linear data.

1.6 Papers

Paper 1

Paper 1 tests the 80/20 rule without the use of any horizon or exogenous variables, as reported in chapter 2.

Paper 2

Paper 2 reports on the same methods used in paper 1 with the induction of horizons. The use of 1 plus *actuals* is introduced into the training data as a way to individualise and pin-point abnormalities and/or anomalies. This is reported in chapter 3.

Paper 3 reports on the use of multiple inputs. This paper tests whether there is a more accurate result if the uncertainty of share prices is introduced into the neural network, where a test of high, low and high-low share prices are introduced into the network. This is reported in chapter 4.

Paper 4 reports on exogenous variables. Having introduced horizons in paper 2 and uncertainty in paper 3, we then tested if exogenous variables would provide more accurate tests if introduced into the network. This is reported in chapter 5.

We also took the most accurate method from each paper into the next paper and so forth. This is to determine the winning method. To put us in a better stance to compare and contrast our papers.

By doing so we are able to observe and evaluate the introduction of the new systems introduced into each paper.

The four papers together represent a continuation test progressively developing from one paper to the next, where a layer of complexity is added to each paper. The methods in each paper that achieve greater accuracy are applied to the next paper, where an error comparison is made.

Having compared seven different error matrices, the methods are compared differently under the accuracy test to capture the bigger picture and compress our testing. For this purpose we have presented 2 error matrices: the APE (absolute percentage error) and MSE (mean squared error). This is applicable for paper 2 through to paper 4, however in paper 1 we present six error matrices excluding the log returns.

Papers two, three and four all show their winning methods in the results section in bold and the winning methods median in italics.

All errors are available upon request.

1.7 References

Arévalo, R., García, J., Guijarro , F., & Peris, A. (2017). A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting. Expert Systems With Applications, 81, 177–192.

Ayodele, A., Ayo, C. K., Adebiyi, M. O., & Otokiti, S. O. (2012). Stock price prediction using neural network with hybridized market indicators. Journal of Emerging Trends in Computing and Information Sciences, 3(1), 1-9.

Booth, G., Gurun, U., & Zhang, H. (2013). Financial networks and trading in bond markets. Journal of Financial Markets, 62, 1-32.

Cervelló-Royo, R., Guijarro, F., & Michniuk, K. (2015). Stock market trading rule based on pattern recognition and technical analysis: forecasting the DJIA index with intraday data. Expert Systems with Applications, 42, 5963–5975.

Chatterjee, U. (2016). Do stock market trading activities forecast recessions? Economic Modelling, 59, 370–386.

Chen, C. W., Wang, Z., Sriboonchitta, S., & Lee, S. (2017). Pair trading based on quantile forecasting of smooth transition GARCH models. North American Journal of Economics and Finance, 39, 38–55.

Erdogan, O., Paul, B., & Ozyildirim, C. (2015). Predicting U.S. recessions: yield curve and stock market liquidity deviation as leading indicators. European Financial Review, 19(1), 407–422.

Hui, B., & Li, P. (2014). Does investor sentiment predict stock returns? The evidence from Chinese stock market. Journal of Systems Science and Complexity, 27(1), 130–143.

Jiang, H., & He, W. (2012). Grey relational grade in local support vector regres- sion for financial time series prediction. Expert Systems with Applications, 39(3), 2256–2262.

Joseph, K., Wintoki, M., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. International Journal of Forecasting, 27(4), 1116–1127.

Kurek, B. (2014). The Information Content of Equity Block Trades on the Warsaw Stock Exchange: An Estimation of Shares' Returns with the Usage of Simple Linear Regression and Multivariate Adaptive Regression Splines. Journal of Forecasting, 33(6), 433–454.

Perlin, M. (2009). Evaluation of pairs-trading strategy at the Brazilian financial market. Journal of Derivatives and Hedge Funds, 15(1), 122–136.

Ravichandran, K., Thirunavukarasu, P., Nallaswamy, R., & Babu, R. (2005). Estimation of Return on Investment in Share Market through ANN. Journal of Theoretical and Applied Information Technology, 44-54.

Sermpinis, G., Stasinakis, C., Rosillo, R., & de la Fuente, D. (2017). European exchange trading funds trading with locally weighted support vector regression. European Journal of Operational Research, 258, 372–384.

Sermpinis, G., Stasinakis, C., Theofilatos, K., & Karathanasopoulos, A. (2015). Modeling, forecasting and trading the EUR exchange rates with hybrid rolling genetic algorithms—Support vector regression forecast combinations. European Journal of Operational Research, 247, 831–846.

Seru, A., Shumway, T., & Stoffman, N. (2010). Learning by trading. Review of Financial Studies, 23, 705–739.

Sokolov-Mladenovic, S., Milovancevic, M., Mladenovic, I., & Alizamir, M. (2016). Economic growth forecasting by artificial neural network with extreme learning machine based on trade, import and export parameters. Computers in Human Behavior, 65, 43-45.

Zeng, N., Zhang, H., Liu, W., Liang, J., & Alsaadi, F. E. (2017). A switching delayed PSO optimized extreme learning machine for short-term load forecasting. Neurocomputing, 240, 175-182.

Chapter 2

FORECASTING FINANCIAL MARKETS WITH UNSUPERVISED PREDICTIVE ANALYTICS TECHNIQUES

2.1 Abstract

This paper reports the results for two-six methods consisting of 26 statistical techniques and one AI method. It has used the original 80/20 rule where 80% of the data was used as training data, and the remaining 20% of the data was used as tests for the forecast. The remaining 20% was hidden from the times-series training data, and thus when the forecast was computed we compared and tested the accuracy of our tests using the six error matrices. We presented daily and weekly data in our results section. The winning models are shown in the analysis and evaluation section where it shows which method outperformed other methods for each particular error matrix .

2.2 Introduction

The accuracy of the results was tested using six different error matrices that are used by a wide range of scholars to test the accuracy of implemented methods. Accuracy is being tested for each of the following frequency, method, data series, and all 6 error matrices were tested. The table below presents the output from the algorithm that was run. The errors represent how the function performed for that time series, however for each specific error the all-round average was taken to estimate the performance of that method over the total time series collected. As stated below, the error for auto (autoarima) ME (mean error) is 0.0701, which represents an error of 0.0701 from the mean for all the companies averaged out.

As the algorithm was run, the tested error matrices would calculate the error presented according to the difference between the forecast and the actual time series data. Representations are shown below in the graphs; an illustration of forecasts and actuals are shown clearly while some examples are presented. However, the test was an exhaustive test and it was run for all 18 companies and all methods and errors. So, individual analysis could be performed and any anomalies in individuality testing could be investigated thoroughly.

All market and company mispricing was removed prior to testing to overcome any bias within the time series, thereby ensuring a significant and fair test.

A hybrid function was also tested, where two functions were merged in the algorithm. Furthermore, this was implemented to test if a combination of two methods would eventually produce better accuracy. Nevertheless, all the hybrid tests were also tested individually to confirm their significance. Similarly, a test was run to ascertain if the hybrid model compared relatively well with the individual model.

Overall, all of the companies were tested significantly and fairly. All of the according graphs and significant tests were implemented for all companies using the mentioned models. However, some 600 graphs were produced, to sum up, a piece of the pie from the test will be shown in the appendices section.

2.3 Literature

Accurate forecasting of financial market indices is an important undertaking because it can help investors and financial analysts to make informed decisions. Zhang, Cao &

Schniederjans (2004) contended that forecasting is a process that generates a set of output variables from a given set of inputs. The variables are usually historical data. Forecasting generally assumes that past and presently observable events can help to predict future occurrences at least in part. According to Graves & Pedrycz (2009), forecasting financial market indices has always been a challenging feat. The difficulty has often stemmed from the complex interactions between unknown random processes, such as unexpected news and market influencing factors. Ullrich, Seese & Chalup (2007) made a similar observation, adding that stock market forecasting is a rather difficult process because of its many complex features including irregularities, shifting trends, volatility and noise. Nevertheless, there is a consensus among researchers that superior techniques that have the lowest forecasting error are naturally functional and viable (Haykin (2007), Kim & Shin (2007) and Martens (2002)). It is worth mentioning here that scholars have come up with different methods for forecasting financial market indices. As presented in their literature, most of the researchers favour their models as the ideal model for forecasting various indices such as the DJIA and the S&P 500. Some examples of soft computing techniques used for forecasting financial market indices include artificial neural networks (ANNs), generalized autoregressive conditional heteroscedasticity (GARCH), support vector machines (SVMs), backpropagation networks (BNs), particle swarm optimization (PSO) and fuzzy logic (FL). These theoretical and empirical studies have yielded different results in terms of the accuracy of the diverse models employed.

2.3.1 Risk and Trading in Financial Markets

Trading in financial markets comes with a variety of risks that financial analysts, researchers and other practitioners should identify and mitigate in order to realise their financial goals. Indeed, many scholars have undertaken empirical and theoretical studies to identify the issues related to risk in trading in financial markets. According to Pham, Cooper, Cao & Kamei (2014), stock assessment and risk management are the key strategies employed by financial practitioners involved in stock trading.

Pham et al. (2014), in their empirical study, strived to present an innovative stock trading method that would aid in mitigating or minimising the potential for risks. This method utilised the *Kansei* evaluation, integrating it with a self-organising map model in order to improve the stock trading system. The researchers acknowledged that the major aims of their proposed approach included achieving the greatest investment returns, aggregating multiple expert decisions and reducing losses through dealing with the complexities of the dynamic market environment. Williams (2011) asserted that some examples of such complexities included the upward, downward or steady market trends as well as other uncertain market conditions that pose a risk to the stock trading system.

To quantify trader sensibilities related to market conditions, stock trading and stock market factors that have uncertain risks, Pham et al. (2014) applied fuzzy evaluation models and the *Kansei* evaluation. The results of the experiments revealed that the approach of applying the *Kansei* evaluation was able to minimise losses, reduce risks and increase the capability of investment returns (Pham et al., 2014). Indeed, the researchers tested the effectiveness of their approach on daily stock trading in NYSE, NASDAQ, HOSE and HNX (Vietnam) stock markets, and the approach performed well in all of them. Vella & Ng (2016) also conducted a study to ascertain ways of minimising risk and uncertainty when trading in financial markets. Specifically, the researchers endeavoured to investigate the capability of higher order fuzzy systems in handling increased uncertainty induced by market microstructure noise present in high-frequency trading scenarios. The predominant interest of the researchers was risk-adjusted performance.

2.3.1.1 Risk Assessment

Vella & Ng (2016) proposed an ANFIS/T2 model, which is a resourceful approach of designing an interval type-2 model based on a generalization of a type-1 ANFIS model. The primary objective of the proposed model was to improve risk-adjusted performance without increasing computational and design complexities. Overall, the ANFIS/T2 model proposed by the researchers yielded significant performance improvements compared to standard buy-and-hold and standard ANFIS methods. Vella & Ng (2016) concluded that the results of their investigation can aid regulators, researchers and practitioners in designing effective intelligent and expert systems for managing risks in the field of high-frequency trading. In a preceding study, Vella & Ng (Enhancing risk-adjusted performance of stock market intraday trading with Neuro-Fuzzysystems, 2014) had aspired to find ways of improving risk-adjusted performance in trading systems that were controlled by dynamic evolving neuro fuzzy systems (DENFISs), adaptive neuro-fuzzy systems (ANFISs) and ANNs. Similar to the present study, this study also established that using risk-adjusted objective functions and accounting for transaction costs yielded satisfactory results in out-of-sample tests. Furthermore, Vella &

Ng (2014) also found that combining many risk-adjusted objective functions with an ANFIS ensemble generated promising results.

Riedel and Wagner (2015) looked into a different dimension in their investigation. The academicians endeavoured to examine the periods that had higher potential for risk when trading in financial markets. In their study, Riedel & Wagner (2015) were interested in the magnitude of tail risk, specifically lower tail downside risk present in overnight versus intraday market returns. They utilised the GARCH model to ascertain market return components in different countries including the United States, Japan, Germany and France. After testing for tail index equality, the researchers found that the overnight return innovations displayed significant tail risk, whereas the intraday innovations did not exhibit tail risks. Sortino & Satchell (2011) also found a similar result and noted that a risk assessment based entirely on volatility could severely underestimate the overnight downside risk.

Li, Tang, Mei, Li & Zhang (2016) also conducted an inquiry pertaining to risks and trading in financial markets. Specifically, the researchers endeavoured to examine empirically a stock investment's trading time risk via escape time in Hushen300 (CSI300) and the DJIA. To proceed with their inquiry, the researchers observed a two-peak distribution and a peak distribution for short trading days. The researchers found that "There is monotonicity (or non-monotonicity) for the stability of the absolute (or relative) trading time risk", an assertion also echoed by Long, Shleifer, Summers & Waldmann (1990).

Shoji & Kanehiro (2016) relied upon a number of numerical simulations to shed light on the mechanisms responsible for disposition effects. The researchers constructed their computational model from basic ideas of prospect theory. The computational model transformed the objective rewards into subjective rewards through the value function advocated by prospect theory. According to Barberis & Xiong (2009), the value function in prospect theory characterises risk-seeking in losses and risk-aversion in gains. Overall, the results of the numerical simulations by Shoji & Kanehiro (2016) revealed that risk-seeking in losses plays a significant role in driving the disposition effect.

From an exhaustive review of literature, it is apparent that risk is an inherent component when trading in financial markets. Nonetheless, practitioners can initiate various strategies to limit these risks. For example, as a number of the studies have indicated, applying the *Kansei* evaluation and utilising the tenets of prospect theory are some of the strategies that financial practitioners can use to avert risks when trading in financial markets. Regulators, researchers and practitioners can reap tremendous benefits from applying this recommendation, as by doing so they will be in a better position to design effective, intelligent and expert systems for managing risks in the field of high-frequency trading.

2.3.2 Neural Networks versus Statistical Techniques

Various empirical studies have strived to examine the effectiveness of AI models and statistical benchmarks in forecasting financial market indices. Unsurprisingly, the inquiries have yielded mixed results, with some indicating that neural network methods have higher predictive accuracy while other studies have presented evidence supporting the effectiveness of statistical techniques that will serve as benchmarks for the investigation. Thenmozhi (2006) conducted an empirical study that focused on predicting stock index returns by using neural networks. The author began by reiterating that ANNs have found widespread application in forecasting different realms of financial markets. Essentially, the author attributed this to studies indicating that ANNs usually having a higher capacity for learning the financial markets' underlying mechanics. Ord & Fildes (2013) also made a similar observation, stating that some of the typical applications of neural networks or AI tools in finance include index construction, portfolio diversification/selection, risk rating of fixed income investments and mortgages, simulation of market behaviour, identification of financial explanatory variables, as well as economic forecasting.

In his study, Thenmozhi (2006) strived to apply neural network models to predict daily returns of the Bombay Stock Exchange (BSE). To build the model of the daily returns, the researcher utilised a multilayer-perception and trained the network using the error back-propagation algorithm. As the researcher rightly pointed out, neural networks are analogous to nonlinear and non-parametric regression models. The period under investigation by the researcher spanned from 16/01/1980 to 26/09/1997. For this inquiry, the data set comprised 3667 data points. Thenmozhi (2006) obtained the data from Capitaline 2000, a database dedicated to providing daily stock market data of the BSE. This research revealed that although neural network models were effective in forecasting daily returns of the BSE, they nonetheless had some inherent shortcomings that could undermine their reliability.

A similar study by Kutsurelis (1998) endeavoured to dissect the accuracy of different neural network methods in predicting financial markets. Specifically, the researcher strived to examine and analyse the ability of a neural network to predict future stock market index trends. He compared the accuracy of the neural network with multiple linear regression analysis, which is a traditional, statistical forecasting technique. Finally, the researcher used conditional probability to calculate the likelihood of the forecast of the neural network model being correct. In order to minimise the error term between the neural network's output value and the desired actual output value, Kutsurelis (1998) utilised the back-propagation algorithm (BBPN). Moreover, to compare the accuracy of the two models, the researcher looked at various characteristics including the coefficient of multiple determination, error statistics,(particularly the standard deviation and the mean), conditional probabilities, as well as combined actual versus predicted S&P 500 Close Chart.

The results obtained by Kutsurelis (1998) indicated that neural networks have higher accuracy compared to traditional forecasting techniques such as the multiple linear regression analysis. The empirical results of this investigation showed that the developed neural network, which integrated BBPN and hybrid AI, was able to register 93.3% probability in forecasting a market rise in the S&P 500 and 88.07% probability in forecasting a market drop (Kutsurelis, 1998). For the multiple linear regression analysis, the probability in forecasting a market rise was 72.5% whereas the probability of forecasting a market drop was 57.71%. The researcher attributed the efficacy of the neural network models to a variety of factors including the ability of such models to model both curvilinear and linear systems. Walczak (2001) asserted that the governing regression assumptions when using multiple regressions must hold true. However, this is not true in many cases, a factor that limits their forecasting accuracy.

Armstrong (2001) and Brockwell & Davis (2002) have also identified some factors that they believe help neural networks to achieve greater predictive accuracy compared to the statistical benchmarks. For instance, all of these scholars have mentioned that the sensitivity of neural networks to error rate assumptions is low, that they can tolerate chaotic components or noise and that they can tolerate heavy tails better than many other forecasting methods including statistical methods. Other advantages of neural networks over statistical benchmarks as identified by Lawrence (1998) include robustness, greater fault tolerance as well as adaptability compared to other expert systems. In his concluding remarks, Kutsurelis (1998) affirmed that neural networks have a higher capability of forecasting financial markets and, if trained properly, this forecasting tool can provide numerous benefits to the individual investor. The author also mentioned that this forecasting tool has high practicality and feasibility.

Despite the overwhelming evidence providing support for neural networks, some researchers have also provided some practical recommendations to boost forecasting accuracy when developing or deploying these models. For example, Kutsurelis (1998) acknowledged that a variety of factors may influence the effectiveness and predictive accuracy of neural networks. Some of these factors include identifying reliable raw data, training a network and pre-processing the data. For individual investors to reap the potential benefits of the neural network models, they must have a comprehensive understanding of neural networks as well as their limitations, they must understand basic principles of probability and statistical measures, and they must have high proficiency in spreadsheet software programmes in order to pre-process data efficiently. Detienne, Detienne & Joshi (2003) also provided some recommendations to enable investors to benefit from the use of neural network models in making accurate forecasts. Apart from the recommendations provided by Kutsurelis (1998), the researchers also proposed the following: mastering financial markets theory in order to avoid inappropriate or random selection of input data; and thorough testing and evaluating of many neural networks to determine if results are replicable.

Even though his study found neural network architectures to be superior to statistical benchmarks, Thenmozhi (2006) nonetheless cautioned that, basically, neural networks are experimental methods that involve a significant amount of trial and error. For his neural network model, the researcher saw a need for further experimentation in order to produce better stock price predictions. Since the investigator examined the predictive accuracy of the BSE index on a daily basis, he recommended testing it for longer durations, such as monthly or weekly returns, to provide sufficient comparable data that would enhance the accuracy of the returns predictions. In addition, Thenmozhi (2006) acknowledged that incorporating other macro- and micro-economic variables in the form of inputs can be beneficial in terms of increasing the accuracy, significance and reliability of the neural network methods when it comes to predictions. Marquez, Hill, O'Connor & Remus (1992) and Tseng, Kwon & Tjung (2012) also provided similar recommendations, adding that investors, financial analysts, forecasters and other concerned stakeholders should consider the influence of macro-economic variables including interest rates, GDP, stability of government, employment trends and global stock market trends in order to create a An effective network structure. Apart from fundamental data, developers of the network architectures may also consider using technical indicators in order to enhance the accuracy of predictions derived from the neural network models (Hyndman & Athanasopoulos, 2014).

Fok, Tam & Ng (2008) also conducted a study that sought to compare the effectiveness of neural networks with statistical benchmarks. In this seminal paper, the researchers used computational data mining methodology to predict stock price indexes in four major markets, namely the United States, China, Europe and Hong Kong. The researchers tested and compared two learning algorithms, which were the standard back-propagation, a neural network model, and linear regression, a statistical benchmark. The researchers trained the models from the historical data of two years spanning from January 2006 to December 2007. After evaluating the performance of the two models using statistical metrics, Fok et al. (2008) found that the standard back-propagation algorithm, which is a neural network, had better predictive accuracy compared to the linear regression algorithm. The researchers acknowledged that neural networks' methodologies are superior in predicting trends, pattern recognition and generalizations. This is attributable to various reasons including their ability to tolerate imperfect data as well as the fact that they do not require rules or formulas. Similar to the observations made by Adhikari, Agrawal & Kant (2013), Fok et al. (2008) also established that, because of their unique non-assumable, noise-tolerant, non-parametric and adaptive properties, neural networks are more effective in explaining non-stationary time series dynamics.

2.3.2.1 Neural Network Tested

The analysis process followed by Fok, Tam & Ng (2008) consisted of various stages including data collection, data benchmarking, building a neural network model, the generation of a predicted result and, finally, performance metrics evaluation. Data from this inquiry came from Yahoo Finance. The specific indices studied were the DJIA, the Shanghai B-Share Index (SHB), the HSI and the FTSE100. The data under consideration spanned 590 days, with data trained by the Tiberius Data Mining Software. Fok, Tam & Ng (2008) evaluated the predictive power of the neural networks and the regression model against the statistical metrics of mean absolute error (MAE) and normalized mean square error (NMSE).

The deviation between the forecast and actual values was significantly lower for the neural network in all markets, implying that this forecasting method has high predictive power when compared to statistical methods such as regression. After normalization, both values of NMSE and MAE achieved by the neural network were quite low, amounting to only 60% to 80% of the values achieved by the regression model. For instance, after normalization, the MAE for the SH-B, the FTSE 100, the HSI and the DJIA for the regression model were 0.11, 0.03, 0.05 and 0.02 respectively while for the neural network the values were 0.06, 0.02, 0.03 and 0.02 respectively (Fok, Tam & Ng, 2008). Similarly, the neural network model outperformed the regression model in the NMSE. For the regression model, the metrics for the SH-B, the FTSE 100, the HSI and the DJIA respectively were 0.181, 0.039, 0.070 and 0.028 whereas for the neural network the values were 0.102, 0.027, 0.039 and 0.024 respectively.

In line with most of the aforementioned studies, Kuo & Reitsch (1996) also made a comprehensive investigation into the usefulness of neural networks in making accurate forecasts, comparing these AI models with conventional methods of forecasting. The conventional forecasting models that the researchers investigated in this inquiry included regression and time series decomposition. According to the authors, conventional methods, such as the time series decomposition, moving averages and regression, tended to yield inferior results because they, at times, make assumptions about the data distributions in the selected datasets, assumptions that might not be subject to verification in many instances. For instance, the authors mentioned that regression models usually assume that datasets for the variables of interest follow a normal distribution. Green & Armstrong (2015) also acknowledged this shortcoming with conventional methods, adding that statistical methods such as moving averages are only appropriate for very irregular or very short datasets.

In their comprehensive investigation, Kaastra & Boyd (2006) provided deep insights into developing a neural network to accurately forecast economic and financial time series. The researchers begin by acknowledging that artificial neural networks are highly flexible and universal function parameters that have registered tremendous success in terms of their predictive power. In recent years, the application of neural networks in finance for tasks such as pattern recognition, time series forecasting and classification has increased drastically. Nevertheless, the substantial number of parameters that analysts must select in developing neural network models has meant that these models still involve trial and error. The main objective of the study carried out by Kaastra & Boyd (2006) was to provide an introductory guide to designing effective neural networks for economic and financial time series forecasting. The authors presented an eight-step procedure that practitioners can utilise to develop effective forecasting models based on the neural network architecture. They also highlighted some common pitfalls in the design of neural networks, discussed some trade-offs in parameter selection, and identified points of disagreement among forecasting practitioners.

Consistent with assertions by Jarrett & Kyper (2011), Kaastra & Boyd (2006) have postulated that neural networks have faced stiff criticism for various reasons. For instance, the researchers contended that neural network models have received criticism because of excessive training times, tedious software, the danger of over-fitting, difficulty to obtain and later replicate a stable solution, and the large number of parameters required to generate good forecasts. When designing neural network models, researchers must choose various parameters to bolster the model's predictive accuracy. Ortiz (2015) asserted that these parameters fall under three distinct categories: data pre-processing; training; and topology. For data pre-processing, the variables to consider include frequency of data (quarterly, monthly, weekly or daily), the type of data (fundamental or technical), the method of data sampling, and the technique of data scaling. With regard to training, the parameters to consider include training tolerance, momentum term, learning rate per layer, learning rate limit, epoch size, maximum number of runs, and frequency of randomising weights, as well as extent of training, validation and testing sets. For topology, the parameters to choose from include number of input and output neurons, number of hidden layers, hidden neurons in every layer, transfer function for every neuron, and error function. Researchers need to select the right combination of parameters if they are to enhance the accuracy of forecasts derived from neural network models (Toloeiashlaghi & Haghdost, 2004).

Unlike Ortiz (2015), Kaastra & Boyd (2006) advocated for a more comprehensive eight-step procedure. The eight steps in sequential order are: variable selection; data collection; data pre-processing; training, testing and validation sets; neural network paradigms; evaluation criteria; neural network training; and implementation. The researchers cautioned that the procedure is not usually a single-pass one. Rather, it may require revisiting preceding steps, especially between variable selection and training (Kaastra & Boyd, 2006). The scholars further maintained that the success of neural network models depends on a variety of factors. First, the analyst must have necessary resources, time and patience to experiment with the model. Another factor in success is that the neural network architecture needs to allow automated routines including testing of input variable combinations, optimization of hidden neurons and walk-forward testing. The model can accomplish these automated routines through the use of script/batch files or through direct programming.

In a comprehensive literature review, Marquez, Hill, O'Connor & Remus (1992) attempted to analyse existing literature pertaining to the effectiveness of neural networks as published in empirical studies. While also acknowledging the popularity that ANNs have achieved in the recent past, the authors strived to compare this method with classical forecasting techniques including judgmental forecasting, causal forecasting and time series forecasting. Overall, their comprehensive review of literature found that classical techniques were as accurate as neural networks when it comes to forecasting. Nevertheless, the authors cautioned that, despite these preliminary findings, it is important to conduct further investigations into areas such as implementation situations or data conditions favouring neural networks. The authors also called for rigorous studies examining the relative robustness of classical models and neural networks. In addition, the authors affirmed that because most of the neural network models employ back-propagation, it would be beneficial to test the predictive accuracy of these models using other algorithms, an assertion also echoed by Elliott & Timmermann (2013).

Moshiri & Cameron (2000) also endeavoured to conduct a study examining the predictive power of neural network models vis-à-vis statistical methods. Their primary focus was on predicting inflation. The main neural network methods developed by the researchers for testing were back-propagation, hybrid AI and fuzzy time series whereas the statistical methods tested by the researchers included an ARIMA model, a Bayesian vector regression, a vector autoregressive model, and a structural reduced-form model. Just like other scholars such as Önder, Bayır & Hepșen (2013), Moshiri & Cameron (2000) acknowledged that the BPNN (backpropagation neural network) is popular for many neural network architectures because it is static or feed-forward only, it is hetero-associative and it has supervised learning. The researchers compared dynamic forecasts for three distinct horizons encompassing one, three and twelve months in the future. In comparing the forecast quality, Moshiri & Cameron (2000) relied upon MAE and RMSE. Another important neural network model is the support vector machine (SVM). Tay & Cao (2001) examined the predictive accuracy and reliability of this technique in the forecasting of financial time series using Chicago Mercantile Market futures contracts as datasets. The researcher compared the feasibility of the SVM against backpropagation, with results suggesting the SVM was superior as measured by directional symmetry (DS), weighted directional symmetry (WDS), MAE and NMSE. This research indicated that other neural network models such as SVMs are superior to BPNN depending on the selected parameters.

Fok, Tam & Ng (2008) stated that introducing too many hidden units in the model can diminish the accuracy of neural networks. This is because many hidden layers can create additional parameters, thereby introducing redundancy and deteriorating the performance of the neural network model. Fok, Tam & Ng (2008) have also identified some gaps in the literature. These scholars asserted that even though traditional knowledge supports the view that a longer training period characterised by more training data can help to increase the accuracy of the prediction model, this is not always the case. The authors found the neural network to be effective in the short-term prediction of China’s stock index. Their paper revealed that data collected from a shorter and closer period could aid in reducing prediction error for fast-changing and highly-speculated environments such as the China’s stock index. In the Chinese experience, the rapid appreciation of the Yuan coupled with the establishment of the Qualified Foreign Institutional Investor (QFII) scheme and Qualified Domestic Institutional Investor (QDII) scheme led to major changes in the stock market environment.

2.3.3 Big Data Analytics, Nonlinearity and Deep Neural Networks

Deep learning permits the generation of computational models composed of multiple processing layers designed to learn representations of data with multiple levels of abstraction. Chong, Han & Park (2017) acknowledged that deep learning neural networks discover intricate structures in large datasets by using the back-propagation algorithm. It is important to mention that a number of empirical and theoretical studies have examined different problem domains where deep neural networks have found application including stock market analysis, stock market forecasting, demand forecasting, crude oil price forecasting, exchange rates forecasting, and nonlinear time series forecasting. This review synthesizes findings from the existing literature dealing with deep learning neural networks to shed light on their application in real-life situations.

A number of existing empirical and theoretical studies have examined different aspects of deep learning neural networks. Chong, Han & Park (2017) conducted a study to examine the utility of deep learning neural networks in the analysis and prediction of stock market trends. The researchers affirmed that the ability of deep learning networks to extract features from large datasets without relying upon previous knowledge of predictors made these networks potentially lucrative tools for stock market forecasting at high frequencies. Zhao, Li & Yu (2017) also echoed similar sentiments, adding that deep learning neural networks are efficient for forecasting since they rely on multiple hidden layers that can effectively learn complex mappings between labels and features.

In a study by Chong et al. (2017), the scholars strived to examine the effects of unsupervised feature extraction methods on the overall ability of the network to predict future stock market behaviour in the Korea Composite Stock Price Index (KOSPI). The unsupervised feature extraction methods examined were principal component analysis, the restricted Boltzmann machine, and the autoencoder. As input data, Chong et al. (2017) used high frequency intraday stock returns. Results from this study suggested that deep learning neural networks could extract additional data from the residuals of an autoregressive model, thereby improving prediction performance. When tested against a covariance-based market structure analysis, the deep learning neural network was also able to improve covariance estimation compared to a standard autoregressive model and an ANN model. Schmidhuber (2015) also found deep learning algorithms to be superior to an ANN model. The researcher contended that this was because, unlike the ANN model, deep learning neural networks do not require a careful selection of input variables or network parameters including learning rate, number of nodes and number of hidden layers in order to generate satisfactory results.

Furthermore, Zhao et al. (2017) and Chong et al. (2017) concluded that the ability to extract data's abstract features, and to identify hidden nonlinear relationships without necessarily relying on human expertise or econometric assumptions, makes deep learning algorithms an attractive alternative to existing approaches and models of stock market prediction. Looking at a different dimension, Jiang, Chin, Wang, Qu & Tsui (2017) sought to examine the utility of a pre-trained deep neural network combined with a modified version of genetic algorithm (MGA) for forecasting demand in an outpatient department in northeast China. The researchers proposed the MGA for feature selection and introduced a feed-forward deep neural network as the forecast model. Results indicated that, compared to PCA and GA, the MGA significantly improved the efficiency and quality of feature selection. Similarly, the pre-trained deep neural network optimally strengthened the benefits of the MGA in demand forecasting compared with SANN, ARIMAX, and MLR.

Jiang et al. (2017) in their study concluded that their hybrid methodology combining pre-trained deep neural network and the MGA could have crucial implications on surge capacity and resource allocation at the outpatient department. Furthermore, they hypothesized that the hybrid method could find practical applications as a tool for data mining and knowledge discovery in other intelligent systems such as supply chain demand forecasting and load forecasting for power systems operation, as these intelligent systems have similar characteristics to the outpatient department’s forecasting of demand. These findings of Jiang et al. (2017) are consistent with those of Zhao et al. (2017) and Chong et al. (2017) all of whom found deep learning algorithms to be superior to other alternatives in forecasting. Jiang et al. (2017) made an interesting observation through presenting the advantages associated with appropriate feature selection. Some of these advantages include decreasing implementation and training time, reducing storage requirements, and facilitating better understanding of data. Nevertheless, whereas the deep learning neural networks have some promising potential, Zhao et al. (2017) in their inquiry proposed some ideas to improve the utility of these algorithms in the future. The authors proposed quantifying factors such as political risks, extreme climate, and psychological factors in order to enhance their predictive accuracy.

Coelho, Coelho, Luz, Ochi & Guimarães (2017) from another perspective conducted a study to examine the effectiveness of a Graphics Processing Unit (GPU) deep learning model, but with specific emphasis on time series forecasting. The scholars asserted that GPU architecture offers a greener alternative with relatively low energy consumption for big data mining. Computational results from this inquiry indicated that the GPU deep learning model was highly scalable as training rounds increased, emerging as a promising tool that can find meaningful application in smart sensors. Nevertheless, unlike Chong et al. (2017) and Sun et al. (2017) who found principal component analysis, the restricted Boltzmann machine, and deep autoencoder all to be effective in stock market prediction, Coelho et al. (2017) affirmed that these popular deep learning approaches have some shortcomings, most notably the difficulty often encountered with reproducing and interpreting. Another limitation of these popular deep learning methods, as identified by Coelho et al. (2017), is their high memory consumption. Just like Lollia, Gamberinia, Regattierib, Balugania & Gatos (2017), Coelho et al. (2017) recommended embedding their GPU architecture inside SM using supercomputers with accelerators such as Intel MIC architecture, ARM and FPGA to allow more precise and real-time decision-making.

2.3.3.1 Nonlinearity vs Linearity, Deep Neural Networks

Akin to a study by Coelho et al. (2017), Tealab, Hefny & Badr (2017) also focused on forecasting nonlinear time series, albeit using ANNs. The scholars asserted that when forecasting time series, it is vital to classify them based on linearity behaviour. This is because linear time series continue to remain at the forefront of applied and academic research. Nevertheless, simple linear time series methodologies tend to leave various aspects of financial and economic data unexplained. In real life, most time series have dynamic behaviour, with inherent moving average and autoregressive components. In their study, Tealab et al. (2017) demonstrated that common neural networks used in most forecasting models are not efficient for diagnosing the behaviour of dynamic or nonlinear time series that have moving average terms. Hence, these neural networks have low forecasting capabilities. Tealab et al. (2017) therefore proposed formulating and testing new neural network models such as deep learning with or without hybrid methodologies, for example fuzzy logic, in order to improve the effectiveness of these networks in forecasting financial time series.

Deep convolutional neural networks have also received extensive coverage in the existing literature. Hafemann, Sabourin & Oliveira (2017) conducted a seminal study that used deep convolutional neural networks to examine offline handwritten signature verification. As the researchers vividly pointed out, verifying the identity of an individual using a handwritten signature is a challenging endeavour given the existence of skilled forgeries. Using convolutional neural networks, Hafemann et al. (2017) sought to address the challenges of improving system performance and obtaining good features. Particularly, the scholars proposed a novel problem formulation that included knowledge of skilled forgeries from a users' subset in the feature learning process with the aim of capturing visual cues distinguishing forgeries and genuine signatures regardless of the user. After conducting extensive experiments on GPDS-160, CEDAR, the Brazilian PUC-PR and MCYT datasets, Hafemann et al. (2017), just like Zeng, Zhang, Liu, Liang & Alsaadi (2017), established that the convolutional neural networks, using deep learning algorithms, were able to achieve significant improvements in state-of-the-art performance in distinguishing genuine signatures from forgeries.

Even though a big proportion of the existing studies have attempted to improve deep learning performance with multilayer perceptron (MLP), a number of scholars have found this approach to be unproductive. For example, Chandra & Sharma (2016) acknowledged that attempts to improve deep learning performance using MLP was a major challenge because finding an optimum learning rate, or tuning the learning rate, in MLP is problematic. Furthermore, depending on the learning rate value, classification accuracy in MLP can vary drastically. Chandra & Sharma (2016) proposed a new approach in which they combined the concept of Laplacian score with an adaptive learning rate for varying the weights. The scholars used the Laplacian score of the neuron to update the incoming weights while taking the learning rate as a function of parameter, updated based on error gradient. Unlike a study by Ma, Sheridan, Liaw, Dahl & Svetnik (2015), Chandra & Sharma (2016) established that a deep learning method that combines adaptive learning rate with Laplacian score increased classification accuracy when compared with other well-known techniques of deep learning. The Laplacian score played an instrumental role in increasing the classification accuracy through improving the weight updation.

Sun, Zhang, Zhang & Hu (2017) also tested the effectiveness of a new deep learning neural network based upon the extreme learning machine (ELM) autoencoder. As the researchers affirmed, the ELM is an efficient learning algorithm especially for training single-layer feed-forward neural networks. Sun et al. (2017) developed a new variant of the ELM autoencoder (ELM-AE) known as the generalized extreme learning machine autoencoder (GELM-AE). The GELM-AE adds a manifold regularization to the ELM-AE objective. Experiments testing the efficacy of the GELM-AE using IRIS, GLASS, and WINE datasets from the UCI repository revealed that this deep learning algorithm outperformed other unsupervised learning algorithms such as spectral clustering (SC), deep belief network (DBN), Laplacian embedding (LE), k-means, and the ELM-AE. This is contrary to the findings of Chandra & Sharma (2016) who found the Laplacian score to be a good predictor of classification accuracy.

Luo, Wu & Wu (2016) also conducted a comprehensive study to examine the effectiveness of a deep learning algorithm for credit scoring utilising credit default swaps. In this inquiry, the investigators examined the singular of models used in credit scoring as applied to CDS datasets. They then evaluated the classification performance of some renowned deep learning algorithms, including the restricted Boltzmann machine and deep belief networks, compared to some popular models of credit scoring such as SVM, MLP and logistic regression. Consistent with the assertions of Chandra & Sharma (2016), Luo et al. (2016) also asserted that, whereas some shallow architectures such as MLP and SVMs have found widespread applications in credit scoring, these architectures nonetheless have some major shortcomings. For example, even though they may be effective in solving well-constrained and simple problems, these methods primarily focus on outputs of classifiers mainly at the abstract level, thereby neglecting the rich information embedded at the confidence level.

In addition, as Zeng, Zhang, Liu, Liang & Alsaadi (2017) vividly pointed out, shallow architectures such as MLP and SVMs have limited representational and modelling power, a factor that can cause problems when dealing with complicated real-world applications. To address these drawbacks, Luo et al. (2016) contended that it has become necessary to introduce training algorithms for deep architectures. To calculate efficiency in their study, Luo et al. (2016) used the Waikato environment for knowledge analysis (WEKA) tool based on accuracy about correct instances generated with a confusion matrix for SVMs, logistic regression, MLP and statistical software R for the deep belief network. To enhance the readability of the estimates and to minimise the effects of data dependency, the researchers used a 10-fold cross validation to create random petitions of the datasets. The findings indicated that the accuracy rate of the deep belief network was the highest at 100%, followed by MLP at 87.75%, SVMs at 87.4%, and logistics regression at 77.21%.

Kuremoto, Kimura, Kobayashi & Obayashi (2014) similarly acknowledged the shortcomings of MLP. They affirmed that even though MLP and other ANNs have found widespread application in forecasting financial time series since the 1980s, some inherent problems with these networks such as local optima and initialization exist. Therefore, the development of deep neural networks will be essential to guarantee increased accuracy in the future. The development of these neural networks is not only important for forecasting time series but also for other intelligent computing fields. In their study, Kuremoto et al.

(2014) proposed a method for forecasting time series using the deep belief network composed of a restricted Boltzmann machine (RBM).

While the numerous empirical studies and theoretical studies have examined the efficacy of deep learning neural networks, some have looked at the prospects of deep learning neural networks in resolving big data challenges. A study by Najafabadi, Villanustre, Khoshgoftaar, Seliya & Wald (2015) sought to examine deep learning applications in addressing the challenges associated with big data analytics. Big data has become increasingly important, as many public and private organisations have started collecting huge domain-specific information that can contain useful insights about problems such as fraud detection, cyber security, national intelligence, medical informatics, and marketing informatics. Organisations including Microsoft and Google analyse huge volumes of data to conduct business analysis and guide decisions, all of which affects current and future technologies.

A key benefit of deep learning as identified by Najafabadi et al. (2015) is that it aids in the analysis of massive amounts of unsupervised data, while rendering is a valuable tool for big data analytics where raw data is mainly uncategorised and unlabeled. The researchers established that deep learning algorithms could greatly help in addressing challenges that are present in big data analytics. These challenges include semantic indexing, fast information retrieval and data tagging, simplifying discriminative tasks, and extracting complex patterns from massive data volumes. A review by Schmidhuber (2015) similarly highlighted the importance of deep learning algorithms in big data analytics. According to the researchers, deep learning algorithms tend to perform better when extracting global and non-local patterns and relationships in the data compared to other learning architectures that are relatively shallow. Furthermore, since deep learning algorithms deal with data representation and abstractions, they are more suitable for analysing raw data from different sources or raw data presented in diverse formats (i.e. variety in big data).

Najafabadi et al. (2015) and Schmidhuber (2015) both acknowledged that while deep learning helps in providing a relational and semantic understanding of raw data, it also provides a vector representation of data for faster information searching and retrieval. Deep learning algorithms additionally make it possible to learn complicated nonlinear representations between occurrences of words, thereby allowing for the capturing of high-level semantic aspects of a document. Even though deep learning algorithms have contributed greatly to the success of big data analytics, Najafabadi et al. (2015) and Schmidhuber (2015) both identified some areas that may warrant further exploration in order to enhance the utility of deep learning in big data analytics. Some of these areas include learning with streaming data, scalability of models, dealing with high-dimensional data, incremental learning for non-stationary data, and distributed computing.

Paradarami, Bastian & Wightman (2017) on their part endeavoured to propose a system that integrated a deep learning neural network using a combination of content-based features in addition to reviews to generate model-based forecasts for business-user combinations. On their own, content-based filtering and collaborative filtering are popular methods for recommending new products to business users, but they suffer from some inherent limitations and in many situations fail to offer effective recommendations. The amount of input in these systems often exceeds the processing capabilities of the systems, leading to information overload. Due to this information overload, the effectiveness and quality of decisions suffer. An application built to cope with the information overload problem is the recommender system (RS). Apart from resolving the information overload problem, the RS also offers intelligent suggestions about items to business users. Examples of RSs that offer personalised recommendations include product recommendations by Amazon.com, recommendations for financial services, movies by Netflix and twitter.

In their study, Paradarami et al. (2017) showed that a set of collaborative and content features allowed for the development of a deep architecture neural system that was able to minimise rating misclassification era and logloss using a stochastic gradient descent optimization algorithm. This optimization algorithm varied in a number of hyperparameters including learning rate, momentum and decay. Learning rate determined how slow or fast the model iterated towards the optimal weights, momentum added weights in current iterations through increasing the size of steps towards the minimum, and decay was represented as learning rate decreases over iterations. After performing experiments and evaluating performance against a test dataset, Paradarami et al. (2017) established that, when compared to the standalone collaborative filtering method, the integrated deep learning neural network that used a combination of content-based features in addition to reviews proved to be a very promising solution in providing intelligent recommendations for business users.

In yet another intriguing study, Shen, Chao & Zhao (2015) also examined the efficacy of deep belief networks in forecasting exchange rates. Nevertheless, unlike Luo et al. (2016) who compared deep belief against MLP, SVMs and logistics regression, Shen et al. (2015) applied the conjugate gradient method to accelerate learning for the deep belief network. The researchers held that forecasting exchange rates is a critical financial problem because the markets for exchange rates are multivariable nonlinear systems wherein the mutuality of factors is quite complex. The linear unpredictability of the exchange rate markets makes it very challenging to forecast when using shallow architectures. Both Shen et al. (2015) and Spithourakis, Petropoulos, Nikolopoulos & Assimakopoulos (2015) affirmed that neural networks such as deep belief networks can solve this problem of linearity because they possess learning and extensive adaptability, factors that can help in modelling and controlling multivariate nonlinear systems.

Furthermore, consistent with the assertions made by Zakarya, Abbas & Belal (2017), Shen et al. (2015) also contended that neural networks have a number of beneficial features that make them attractive in forecasting exchange rates. For instance, neural networks such as the deep belief network have general nonlinear function mapping capabilities that can approximate any continuous function with the desired accuracy. As such, these networks are capable of solving a variety of complex problems. In addition, as a nonparametric data-driven model, a neural network does not include the restrictive assumption about the underlying processes used for generating data. This makes it less vulnerable to model misspecification problems compared to most parametric nonlinear methods.

In a study by Shen et al. (2015), the researchers used continuous restricted Boltzmann machines to construct the deep belief network to forecast weekly BRL/USD and GBP/USD exchange rate series. They then updated the classical deep belief network to model continuous data while applying the conjugate gradient method to accelerate learning for the system. In the subsequent experiments, Shen et al. (2015) tested three exchange rate series and adopted six evaluation criteria to assess the performance of their proposed methodology. Results from this investigation revealed that the deep belief network accelerated by the conjugate gradient method was superior in forecasting exchange rates when compared to other architectures such as the feed-forward neural networks, random walk (RW) and auto-regressive-moving-average (ARMA).

Consistent with many other studies, Krauss, Do & Huck (2017) in their inquiry analyzed the efficacy of different algorithms including deep neural networks, random forests (RAFs) and gradient-boosted-trees (GBTs) and various ensembles of these methodologies in the context of statistical arbitrage. The researchers trained each model on the lagged returns of all S&P 500 stocks after the elimination of survivor bias. Krauss et al. (2017) then generated daily one-day-ahead signals from 1992 to 2015 based upon the probability of a stock forecast to outperform the general market. The next step was to convert the lowest *k* probabilities into short positions and the highest k probabilities into long positions, thereby censoring the less certain middle section of the ranking. The empirical findings from the study were promising. Krauss et al. (2017) established that a simple, equal weighted ensemble comprising of one deep neural network, one RAF, and one GBT produced out-of-sample returns that exceeded 0.45% per day for $k = 10$ before transaction costs. The equally weighted ensemble also produced economically and statistically significant daily alphas ranging between 0.14% and 0.24%, confirming the results of Chong et al. (2017) that deep learning algorithms were indeed effective in forecasting stock markets.

2.3.4 The Forecasting Dilemma & AI

The debate about single-layer neural networks and multiple-layer neural networks has received extensive coverage in the existing literature. Empirical and theoretical studies have yielded mixed results as to the efficacy of these architectures in forecasting. Some studies that have explored this contentious topic seem to support the notion that multiple-layer neural networks are far superior in their predictive power and forecasting accuracy compared to single-layer neural networks. A major limitation cited by researchers about single-layer neural networks pertains to their inclination to represent only a small set of functions. Other scholars have nonetheless conducted investigations that support the effectiveness of single-layer neural networks. Advantages of these networks cited by scholars supporting them include the ease of set-up and training, as well as their explicit link to statistical models that allow for posterior probability and interpretable representation. This review explores the existing literature with the aim of identifying the utility of single-layer and multiple-layer neural networks as they are applied to the field of forecasting.

A number of studies have endeavoured to examine the efficacy and utility of multiple-layer neural networks in the field of forecasting. Pedzisz & Mandic (2008) in their inquiry sought to examine the effectiveness of a homomorphic neural network for predicting and modelling. The homomorphic neural network developed by the researchers utilised a two-layer feed-forward architecture that had an exponential hidden layer accompanied by a logarithmic pre-processing step. The researchers affirmed that, this way, they were able to see the overall output/input relationship as a bank of homomorphic filters or as a generalized Volterra model. After introducing gradient-based learning and addressing practical issues pertaining to weight initialization and choice of optimal learning parameters, the researchers compared this two-layer neural network to a single-layer sigmoidal feed-forward neural network.

Results obtained by Pedzisz & Mandic (2008) after verifying performance and convergence speed by using extensive simulations revealed that the homomorphic neural network with a two-layer feed-forward architecture had better predictive accuracy compared to the single-layer sigmoidal feed-forward neural network, particularly with regard to medium-scale and small-scale datasets. Schmidhuber (2015) similarly affirmed that the two-layer feed-forward architecture can increase predictive accuracy in forecasting. As the researcher vividly pointed out, adding the number of hidden layers in these models can help in reducing both the computational time as well as the total number of network weights, which subsequently improves the predictive accuracy of the multiple-layer neural networks (Schmidhuber, 2015).

Nguyen & Chan (2004) likewise conducted an inquiry to ascertain the efficacy of multiple neural networks (MNNs), with specific emphasis on long-term time series forecasting. According to the researchers, a multiplelayer neural network model is one in which a group of neural networks work together in solving a pertinent forecasting problem, a view also shared by Sheela & Deepa (2013) and Panchal, Ganatra, Kosta & Panchal (2011). To address the problem of propagation errors that might lead to inaccuracies in time series forecasting, Nguyen & Chan (2004) proposed a multiple neural network model combining long-term and short-term neural networks to accommodate a wide array of prediction terms.

Nguyen & Chan (2004) hypothesized that the multiple-layer neural network model that they developed would be able to address the problem through reducing the number of necessary recursions. The MNN model developed by Nguyen & Chan (2004) showed superior performance compared to a single-layer neural network in long-term prediction.

Nevertheless, the researchers established that predictive accuracy seemed to diminish if the period was too long. Just like Sheela & Deepa (2013), Nguyen & Chan (2004) also stated that a major weakness of the MNN model is that it requires a dataset involving more continuous and longer time series in order to build it.

In an investigation designed to explore the utility of improved neural networks for short-term load forecasting, Lang, Zhang & Yuan (2015) essentially introduced a weighting technique to the multiple-layer neural network inputs in order to ascertain its efficacy in forecasting daily maximum load in an electric power system in China. The inputs selected for this investigation included day of the month, month of the year, day of the week, holiday indicator, week number, and maximum electricity load of the previous days. After conducting simulation experiments and applying the multiple-layer neural network with kernels and random weights to approximate nonlinear function between the daily maximum load and the selected inputs, Lang et al. (2015) found that their proposed method had a superior forecasting accuracy resulting from good generalization performance and fast learning speed. Pellakuri & Rao (2016) made an interesting observation by contending that learning rate can help to ensure faster training and increase momentum to enhance predictive power.

Claveria, Monte & Torra (2015) in their study investigating the effectiveness of different neural network models in forecasting tourism demand also found that an MLP far outperformed an Elman network and a radial basis function. The results were unswerving even after replicating the experiment to account for different topologies regarding the number of concatenation lags. Dhamija & Bhalla (2011), on the contrary, found contradicting results in their study that compared the effectiveness of different neural network architectures in exchange rate forecasting. Unlike Claveria et al. (2015), Dhamija & Bhalla (2011) found that even though forecasters can generally use neural networks to predict exchange rates with better accuracy, the radial basis function (RBF) networks were superior to MLP networks in forecasting British Pound versus US Dollar (GBP-USD), German Mark versus US Dollar (DEM-USD), JPY-USD, IR-USD, and EUR-USD.

Similar to a study by Lang et al. (2015), Liu, Li & Sun (2013) also aimed to develop a model for short-term load forecasting, although they based their model on a combination of multiple-layer neural networks and multi-wavelet transform for extracting training data. As the training network, the researchers adopted a wavelet neural network, a BP network and

RBF network, after which they inputted training data from the three neural networks into a three-layer feed-forward neural network for load forecasting. The results of Liu et al. (2013) supported those of Lang et al. (2015) and Mall & Chakraverty (2015), indicating that the short-term load forecasting accuracy was superior because of good generalization performance as well as fast learning speed.

It is worth mentioning that while the majority of the studies seem to support the effectiveness of multiple-layer neural networks, other studies, on the contrary, seem to suggest that single-layer neural networks can also be effective in forecasting. Guliyev & Ismailov (2016) endeavoured to look at the utility of a feed-forward neural network that had only one neuron in the hidden layer in approximating univariate functions. The researchers algorithmically constructed a sigmoidal, smooth, and almost monotone activation function that provided approximation to arbitrary continuous functions. After implementing the algorithm, Guliyev & Ismailov (2016) found that the single-layer neural network was able to provide reliable approximations within any degree of accuracy.

Akin to the study by Guliyev & Ismailov (2016), Catillo, Fontenla-Romero, Guijarro-Berdinas & Alonso-Betanzos (2002) designed a global optimum approach in which one-layer neural networks managed to minimise either the MAE or the sum of squared errors as measured in the input scale. The scholars established that the global optimum algorithm was able to solve linear programming problems and linear systems of equations using less computational power. When compared with other high-performance learning algorithms, the researchers proved that their global optimum algorithm that utilised only a single-layer neural network was at least 10 times faster while simultaneously allowing for higher computation of a large number of estimates for weights and providing robust median and mean estimates for them as well as their associated standard errors. Cumulatively, consistent with the assertions made by Giusti & Itskov (2014) and Guliyev & Ismailov (2016), Castillo et al. (2002) found that, when applied properly, their algorithm based on a single-layer neural network gave a good measure of the quality of fit.

Wu, Yao, Li & Zhang (2016) correspondingly conducted a study that seemed to support the efficacy of single-layer neural networks especially within the domain of pseudoconvex optimization. In this inquiry, the researchers developed a single-layer recurrent neural network for solving pseudoconvex optimization with box constraints. They also described the developed system with a differential inclusion system. Results from the investigation established that, compared to other existing neural networks designed for solving pseudoconvex optimization, the single-layer recurrent neural network developed by Wu et al. (2016) had a wider implementation domain and a stable sense of Lyapunov when tested against the Lyapunov stable theory. The researchers also applied Clarke's nonsmooth analysis technique, with findings indicating that the proposed single-layer neural network was able to address finite time-state convergence on the constraint conditions.

Based on the methodology adopted by Wu et al. (2016), Liu, Guo & Wang (2012) also tested the efficacy of single-layer neural networks in the domain of pseudoconvex optimization problems, with their results supporting those of Wu et al. (2016). Liu et al. (2012) went further to state that the single-layer neural network is effective in solving nonlinear programming problems because these problems tend to be of a time-varying nature and are easily solvable by the one-layer neural network. Nonetheless, Liu et al. (2012) proceeded to identify some limitation of single-layer neural networks that can undermine their applicability in the field of forecasting. Apart from the limitation of representing only a limited set of functions, other drawbacks of single-layer neural networks as identified by Liu et al. (2012) include the fact that decision planes must be hyperplanes in these models and the models can perfectly separate only linearly separable data.

Abbas, Belkheiri & Zegnini (2015) also conducted an investigation to examine the efficacy of single-layer neural networks. The main objective of the research by Abbas et al. (2015) was to design an adaptive output feedback control using only one single hidden layer neural network in a class of highly uncertain nonlinear systems. The scholars hypothesized that their proposed model would be able to eliminate unstructured uncertainties and thereby help in increasing its predictive accuracy. The approach by Abbas et al. (2015) employed feedback linearization as well as an online neural network designed to compensate for modelling errors. The researchers additionally designed a fixed structure dynamic compensator, which as Lollia, Gamberinia, Regattierib, Balugania & Gatos (2017) asserted, can help in stabilising the linearized system.

In order to adapt the single-layer neural network weights, Abbas et al. (2015) used a signal comprising a linear combination of compensator states and the measured tracking error, deriving the network weight-adaptation rule from the Lyapunov stability analysis. Results obtained after numerical simulations revealed that, compared to other models including the tunnel diode circuit model, the proposed single-layer hidden neural network was superior since it manifested strong robustness in handling modelling inaccuracies, had the ability to handle both arbitrary complexity and complicated nonlinearity, and succeeded in having excellent tracking performance (Abbas et al., 2015).

In another study examining the effectiveness of single-layer neural network models vis-à-vis multiple-layer linear regression models in time series forecasting of total ozone, Bandyopadhyay & Chattopadhyay (2007) found that the single-layer neural network had better predictive accuracy compared to the multiple-layer regression models. In the inquiry, the researchers developed the single-layer hidden neural network with a variable number of nodes and evaluated their performance based on the method of least squares and error estimation. The researchers trained the single-layer neural network with sigmoidal activation function with the aim of minimising the mean squared error (Bandyopadhyay & Chattopadhyay, 2007). The single-hidden layer neural network was able to predict the mean monthly total ozone concentration more efficiently compared to the multiple linear regression model using past data values. Overall, the consulted literature seems to support the utility of both single-layer and multiple-layer neural networks in the field of forecasting, although their efficacy depends on the variables examined.

Strategic trading and forecasting is an issue that has received notable attention in empirical studies. These studies have identified a number of strategies that are applicable to trading as well as forecasting. Some of these strategies include those modelled on artificial neural networks, GARCH models, and SVMs. A consensus exists that market complexity makes the relationship between future and past financial data nonlinear. This implies that linear statistical methods including ARMA and ARIMA are seemingly powerless in predicting stock returns and trading efficiency as compared to nonlinear approaches including GARCH, SVR, and ANNs. Relying upon evidence-based empirical and theoretical studies, this section provides a comprehensive analysis of the utility and efficacy of different strategies utilised in trading in, and forecasting of, financial markets.

2.3.5 Does Forecasting Optimise Trading Returns?

A number of scholars have endeavoured to examine the efficacy of different strategies and models employed in trading and financial forecasting. The studies have yielded mixed results with regard to the utility and effectiveness of the different any models. Choudhry, McGroarty, Peng & Wang (2012) conducted an extensive study to examine the effectiveness of market microstructure variables in forecasting foreign exchange rates at frequencies ranging from one minute to several minutes. In conducting the study, the researchers utilised a unique foreign exchange dataset comprising electronic transactions from a global inter-dealer. The data utilised in this inquiry were dollar-euro (USD-EUR), German mark-dollar (DEM-USD) and yen-dollar (JPY-USD) exchange rate observations.

As the predictor model, Choudhry et al. (2012) applied an ANN. Jiang, Chin, Wang, Qu & Tsui (2017) shared similar views to those of Choudhry et al. (2012), asserting that when using market microstructure variables in forecasting foreign exchange rates, the immediately preceding ask and bid prices are critical factors, which is in line with the market microstructure theory. Out-of-sample results from the study by Choudhry et al. (2012) revealed that high-frequency trading strategies based upon the ANN model were indeed profitable even after the inclusion of transaction costs. The researchers concluded that the practitioners most likely to benefit from the findings of their study were currency traders working for specialist currency funds and those working in commercial banks.

Bekiros & Georgoutsos (2008), from a different perspective, endeavoured to ascertain the profitability of a trading strategy based upon recurrent neural networks (RNNs) in predicting the NASDAQ composite index's direction-of-change. The sample for this investigation extended over the period between February 1991 and April 1998, with the researchers reserving the sub-period between April 1998 and February 2002 for purposes of out-of-sample testing (Bekiros & Georgoutsos, 2008). The results demonstrated that incorporating estimates of conditional volatility changes in the trading rule strongly enhanced the profitability of the trading strategy based upon RNNs compared to a buy-and-hold strategy and a nested model. Just like Choudhry et al. (2012), Bekiros & Georgoutsos (2008) also found that the profitability increased despite the inclusion of transaction costs.

Akin to the studies by Choudhry et al. (2012) and Bekiros & Georgoutsos (2008), Dunis & Huang (2002) also sought to examine the utility of neural network models in forecasting and trading. The researchers examined the utility of non-parametric RNN and neural network regression (NNR) models in trading and forecasting currency vitality within the domain of USD/JPY and GBP/USD exchange rates, benchmarking the results of both RNNs and NNR against a simpler GARCH alternative. Specifically, the scholars sought to identify mispriced options, develop a non-parametric nonlinear approach to forecast foreign exchange volatility, and then come up with a realistic trading strategy based on this process.

After identifying mispriced options, Dunis & Huang (2002) applied a volatility trading strategy that used foreign exchange option straddles in order to account for both forecasting accuracy and trading efficiency. Similar to the inquiries conducted by Tealab, Hefny & Badr (2017) and Bekiros & Georgoutsos (2008), Dunis & Huang (2002) found that the RNN strategies produced positive returns even after allowing for transaction costs. In a study looking at economic growth forecasting via an ANN based on export, import, and trade parameters, Sokolov-Mladenovic, Milovancevic, Mladenovic & Alizamir (2016) similarly established that a strategy relying on ANNs would generate higher forecasting accuracy even after accounting for transaction costs.

Aside from the studies examining the utility of strategies modelled on neural networks, other studies have investigated the efficacy of trading strategies based on SVMs. For example, Dunis, Rosillo, de la Fuente & Pino (2013) sought to investigate the effectiveness of SVMs in predicting weekly change in the IBEX-35 stock index from October 1990 to October 2010. The researchers implemented a trading simulation whereby measures of economic performance would complement statistical efficiency. The inputs retained by the researchers were traditional technical trading rules including moving average convergence divergence (MACD) and relative strength index (RSI) decision rules. In order to determine the ideal situations to sell or in the market, Dunis et al. (2013) used the SVMs with given values of MACD and RSI.

After benchmarking the model with a buy and hold strategy and an MLP neural network, Dunis, Rosillo, de la Fuente & Pino (2013) established that SVMs yielded better results with a 100% hit ratio when the researchers applied a 90% probability of recurrence. However, promising results were only evident when the researchers used shorter training periods compared to longer training periods. Just like Dunis et al. (2013), Sermpinis, Stasinakis, Rosillo & de la Fuente (2017), in their inquiry that examined trading in European exchange trading funds, also found that trading strategies based on SVMs had a propensity to yield better predictive results over shorter durations as opposed to longer durations. Both Dunis et al. (2013) and Sermpinis et al. (2017) attributed the poorer predictive accuracy over longer training periods to changing market conditions and overtraining.

Qu & Zhang (2016), in the same light, endeavoured to test the efficacy of using support vector regression in the prediction of high-frequency stock returns of the Chinese CSI 300 index. Assuming that each return triggers reversal and momentum periodically, the researchers decomposed each return into decaying cosine waves collections that were functions of past returns. They then reached an analytical countenance of the nonlinear relationship existing between future and past returns, after which they introduced a new kernel for predicting future returns accordingly. Qu & Zhang (2016) then used high-frequency prices of the CSI 300 index for the period covering January 2010 to March 2014. A significant observation using the strategy developed by the researchers was that the new kernel significantly outperformed the benchmarked sigmoid function kernel and radial basis function kernel in both directional forecast accuracy rate and mean squared prediction error.

In s study by Qu & Zhang (2016), capital gains of the trading strategy that integrated the new kernel with support vector regression were also higher compared to the sigmoid function kernel and radial basis function kernel. According to Chin & Hong (2008), a capital gain can only occur when the investor sells stocks at a price that is relatively high compared to the purchase price. Capital loss, on the contrary, occurs when the capital asset price declines instead of appreciating. In their study, which investigated the efficacy of financial ratios in predicting the Malaysian stock market, Chin & Hong (2008) found that dividend yield had a higher predictive accuracy compared to earning yield.

These results of Chin & Hong (2008) implied that dividend yield for listed companies in the Kuala Lumpur Stock Exchange have the ability to offer improved portfolio decisions and a better trading strategy than one based upon earning to price ratio alone. Ultimately, the researchers concluded that it is economically and statistically viable to design a new kernel bolstered by support vector regression to predict high-frequency stock returns, a sentiment also echoed by Steinwart (2001) who assessed the influence of kernel on SVM consistency.

Unlike studies that have examined the utility of ANNs and SVMs in trading and forecasting, V’yugina & Trunov (2016) tested the applications of combined financial strategies for universal adaptive forecasting using BATS and MICEX trading platforms. Based on an adaptive forecasting algorithm, the researchers designed a universal strategy that ensured asymptotically maximal profit compared to other trading strategies in which the forecasters make decisions based upon rules that continuously depend on the input information. In order to minimise risk in their universal strategy, V’yugina & Trunov (2016) performed an adaptive redistribution of current capital among the combined financial instruments according to the AdaHedge algorithm. The results from this inquiry indicated that the combined financial strategies for universal adaptive forecasting using BATS and MICEX trading platforms passed the generalized calibration tests as well as tests given by functions from the RKHS functional space, implying that the combined strategies were effective at trading and forecasting (V'yugina & Trunov, 2016).

Relying upon options price data retrieved from the Taiwanese stock market, Sheu & Wei (2011) also aspired to test an ingenious options-trading strategy based upon the forecasting of volatility direction. The researchers constructed a forecasting model by incorporating absolute returns, proxy of investor sentiment, and heterogeneous autoregressive-realized volatility (HAR-RV). They accumulated data from the Taiwanese Futures Exchange (TAIFEX) on a daily basis from 2003 to 2007, amounting to 1,240 trading days. After taking into account margin-based transaction costs, findings from the simulated trading by Sheu & Wei (2011) suggested that a straddle trading strategy, specifically one that incorporates market turnover in the forecasting of volatility direction, attains the best Sharpe ratios. The researchers concluded that their trading paradigm had immense utility in bridging the gap between market volatility, options trading, and the information content generated by investor overreaction. Results from the investigation by Sheu & Wei (2011) were also consistent with the revelations presented by Joseph, Wintoki & Zhang (2011) who acknowledged the need for assigning investor sentiment a prominent role when constructing volatility-forecasting models.

Strategies based on the GARCH model have also found notable attention in the existing literature. For example, Barunik, Krehlik & Vacha (2016) conducted a study to examine the efficacy of a trading strategy based on Realized GARCH and Jump-GARCH models for modelling and forecasting volatility in the exchange rate market, with specific emphasis on the time-frequency domain. The researchers based their methodology on the decomposition of volatility into jumps and several time scales. Chen, Wang, Sriboonchitta & Lee (2017) reiterated that decomposing volatility into jumps and several time scales can help in approximating traders' behaviour at corresponding investment horizons. The trading strategy based on Realized GARCH and Jump-GARCH models developed by Barunik et al. (2016) statistically outperformed conventional models in both multi-period-ahead and one-day forecasting. These results are consistent with the assertions made by Chen et al. (2017)

who acknowledged that nonlinear approaches including GARCH, SVR, and ANNs are more effective at forecasting financial markets compared to linear approaches.

Unlike the majority of other aforementioned inquiries, Miralles-Quirós, Miralles-Quirós & Daza-Izquierdo (2015) sought to examine intraday patterns and come up with trading strategies for predicting the Spanish stock market. Through anticipating greater interest in Spanish stock market behaviour, the researchers were able to show that the best or ideal trading strategy was one in which the investor entered short or long after the New York Stock Exchange (NYSE) opening until the end of the trading day at 17:30 when bear and bull markets coincide (Miralles-Quirós et al., 2015). The researchers also recommended that for no-coincidence cases it is important to complement the strategy with that of entering short or long from the opening of the trading day until the closing price. To analyse the intraday behaviour of the selected stock market, Miralles-Quirós et al. (2015) compiled IBEX-35 intraday data from February 2000 to December 2012 beginning from the opening quote at 9:00 until the end of every session at 17:30. Cumulatively, the reviewed literature seems to support the effectiveness of nonlinear strategies in trading and forecasting of financial markets. As Hu (2007), Kwan, Lam, So & Yu (2000), and Tealab, Hefny & Badr (2017) all affirmed, market complexity makes the relationship between future and past financial data nonlinear. This implies that linear statistical methods, including ARMA and ARIMA, are seemingly powerless in predicting stock returns compared to nonlinear approaches including GARCH, SVR, and ANNs.

2.3.6 Diversified Method Selection to Overcome the Random Walk Dilemma

There have also been some comparative studies conducted to ascertain the best methods for forecasting financial market indices. For instance, Ong, Smola & Williamson (2005) conducted a comprehensive comparative analysis of different models to determine which was most effective. The comparative analysis looked at forecasting time series using multi-layer perception, multiple regression, and adaptive neuro fuzzy inference system and radial basis functional models. As the researchers pointed out, the adaptive neuro-fuzzy inference system (ANFIS) utilises both the reasoning capabilities of fuzzy logic and the learning capabilities of an ANN in order to provide enhanced reasoning capabilities. Although this analytical method for forecasting financial market indices been highly effective, it nonetheless has some shortcomings, such as low convergence (CristianinI & Kandola, 2006).

The random walk dilemma (RWD) is one problem that has often undermined the accuracy of many forecasts. Araújo, Oliveira & Meira (2015) formulated a study geared towards eliminating the random walk dilemma encountered when using sophisticated techniques to forecast financial time series. In their endeavour to overcome this dilemma, the researchers proposed the concept of time phase adjustment. The model proposed by the authors was called increasing decreasing linear neuron (IDLN), particularly suited for high-frequency stock market forecasting. The researchers presented a gradient-based learning process that had automatic time phase adjustment for designing the proposed model using a systematic approach and ideas from the back-propagation algorithm to overcome the problem of non-differentiability.

Upon conducting their investigations, Araújo et al. (2015) collected results from three frequency financial time series (BBAS3, BRML3 and BRFS3) from the Brazilian stock market. Experimental results demonstrated a consistently better performance compared to other statistical, hybrid, or neural prediction models. The researchers attributed the better performance of their model to the nonlinear component and linear component with decreasing and increasing behaviour as well as the integration of the time phase adjustment in the learning process. The proposed model's success was highly dependent on the accurate adjustment of the parameters of the predictive model. All of the other predictive models investigated in this research failed to overcome the RWD for high-frequency time series (THEILP≥ 1). Nevertheless, the proposed IDLN model was able to overcome the RWD in high-frequency financial time series (THEIL~0) (Araújo, Oliveira & Meira, 2015).
Research on the RWD has included extensive exploration in the work carried out by Zhang, Cao & Schniederjans (2004). The authors presented an intelligent hybrid model comprising a modular morphological neural network (MMNN) joined with an MGA that searches for particular time lags, able to achieve a fine-tuned characterisation of time series and estimating the sub-optimal (initial) parameters including weights, number of modules, and architecture of the MMNN. The researchers then used a back-propagation algorithm to train each component of the MGA population. Adjusting the model enabled it to perform behavioural statistical tests and to phase fix procedures to adjust time phase distortions detected in the financial time series. According to Corrado & Truong (2007), morphological neural networks (MNNs) are an important class of nonlinear systems. They differ from classical ANNs in that simple morphological operators carry out the computation in each

node in the context of image algebra. Typically, MNNs are used in image processing applications including image restoration, pattern recognition, and edge extraction.

2.4 Data

The data in this paper has been collected from the from 18 companies selected from the FTSE 100 index . The companies were selected in accordance with their capitalization within their respective index, with the 18 most capitalized selected.

The data was constructed to be consistently significant by making the unit measure consistent with all companies rather than the dates of the share price. Thus, as it was data collected and rather than the period, the data points were daily 4000 units, weekly 825 units, monthly 191 units, quarterly 65 units, and yearly 17 units with all the data units ranging from 2000 to 2016. The data consists of the respective companies share price. Data was collected regarding the shares prices' high, low, close, adjusted close and volume.

2.5 Methodology

The methodology that has been tested in this paper is a quantitative test to compare intensive methods within both the statistical techniques and AI spectrum. There were 26 methods tested in total of which 23 were statistical techniques and one was an AI method. The two remaining methods either did not produce outputs or could not adapt to the time series. The adaptation of some models was dependent on seasonality, and where the time series did not have seasonality the models would not compute the forecast and thus would not produce results. All of the methods were implemented using the R software and Python considered in AI implementation.

In continuation, the 26 methods are widely used in the forecasting field; these methods can sometimes depend on different perspectives of raw data. As mentioned earlier, some methods depend on the data's seasonality, and if the data does not have seasonality the method within its own computation will not perform well and will possibly result in an (NA) output. Thus, in most cases it will produce a naturally aspirated result.

Table 4: Classifications

Classifications of time series and test

	Companies	Methods	Error Matrices	Index
Daily	18	26	6	FTSE 100
Weekly	18	26	6	FTSE 100
Monthly	18	26	6	FTSE 100
Quarterly	18	26	6	FTSE 100
Yearly	18	26	6	FTSE 100

2.5.1 Description

The table above shows details of the test that was undertaken in this paper. There are 26 methods that were implemented and tested in this paper. That is one of the differences from our other tests. The companies, indices, and error matrices all carry over, however horizon was not tested in this paper and will instead be introduced in other testing that we conduct. Furthermore, though we tested six error matrices, we used MASE in this paper but not in the other three papers.

The following table illustrates the methods that were used:

Table 5: Methods

METHODS		Description
1.	AUTOARIMA	autoarima function
2.	AUTOARIMA_SEASDUMMY	with seasonal dummies as external regressors
3.	AUTOARIMA_FOURIER	with fourier transform as external regressors
4.	SES	Simple exponential smoothing
5.	HOLT WINTERS	Double exponential smoothing with alpha and beta
6.	DOUBLE SEASONAL HOLT WINTERS	Holt-Winters with alpha, beta and gamma
7.	BATS	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components
8.	TBATS	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality. (*)
9.	DSHW	Double seasonal holt winters
10.	NAÏVE	Forecasted by the previous observation
11.	SNAIVE	Forecasted by the last observation in previous period/season.
12.	SINDEX	Seasonal index forecast
13.	NNET	Neural Networks
14.	TSLM	Time series linear model, with level and trend as the X vars
15.	SPLINEF	Splines model
16.	THETAF	Theta model
17.	RWF	Random walk model
18.	MEANF	Mean forecast
19.	STL	Seasonality-Trend-Level. The time series is broken down to these components, the remainder is forecasted and the STL components are added back to the forecasted value.
20.	AUTOARIMA_SES	autoarima function & Simple exponential smoothing
21.	AUTOARIMA_TBATS	autoarima function & state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality
22.	AUTOARIMA_NNET	autoarima function & Neural Networks
23.	SES_THETAF	Simple exponential smoothing & Theta model
24.	SES_MEAN	Simple exponential smoothing & Mean forecast Model
25.	TBATS_THETAF	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality & Theta model
26.	NNET_THETAF	Neural Networks & Theta model

(*) indicates, if seasonality is not present, both the models will not have seasonal components. So, they end up optimising on the same model and, therefore, will have the same forecasts.

2.5.2 Description

The methods that are shown above were all tested, however not all produced results due to time series seasonality, as mentioned earlier. The 26 methods were all run in R studio. The packages for the methods are freely available to download in the R software.

2.6 Results

Table 6

METHOD	**Daily**		**Values**				
	ME	RMSE	MAE	MPE	MAPE	MASE	AVE
AUTO	0.070	25.606	15.688	-0.019	1.350	1.007	7.284
SES	0.299	25.609	15.666	-0.015	1.340	1.001	7.317
HOLT	0.058	25.606	15.670	-0.010	1.343	1.002	7.278
BATS & TBATS	0.374	25.600	15.661	-0.012	1.339	1.001	7.327
DSHW	0.262	26.930	17.008	0.014	1.478	1.101	7.799
NAÏVE & SNAIVE	0.267	25.702	15.650	-0.016	1.339	1.000	7.324
NNET	-0.004	25.005	15.748	-0.461	1.753	1.010	7.177
SPLINE	-7.566	140.077	122.477	-2.073	9.413	NA	NA
THETA	0.299	25.609	15.666	-0.015	1.340	1.001	7.317
RWF	0.000	25.698	15.661	-0.022	1.341	1.001	7.280
MEAN	0.000	580.412	497.461	-60.378	83.876	32.086	188.909
AUTO % SES	-67.656	132.715	112.648	-8.224	9.765	5.852	30.850
AUTO % TBATS	-65.080	131.354	111.403	-8.116	9.718	5.805	30.847
AUTO % NNET	-50.687	147.116	125.796	-6.910	10.425	7.126	38.811
SES & THETA	-62.079	127.192	108.092	-6.790	8.453	5.852	30.120
SES % MEAN	243.634	474.152	464.385	-9.615	35.531	27.146	205.872
TBATS & THETA	-59.502	125.888	106.866	-6.682	8.419	4.855	29.974
NNET & THETA	-67.656	132.715	112.648	-8.224	9.765	5.852	30.850

2.6.1 Discussion

The table above presents the results from our test for each method. The methods were tested under the six error matrices above. This shows us which method performed better under which error. A detailed table on the classification of accuracy testing shows an overview of how the methods performed. The table above reveals strength in the random walk forecast (RWF) and the mean in the mean error (ME) test, but this strength diminishes along the error matrices. Specifically, the mean became the least accurate test in all of the other error testing apart from the mentioned ME test. The naïve model showed better accuracy than the other models in three of the six accuracy tests, performing better in MAE, MAPE and MASE. Neural networks and the Holt-Winters performed better in RMSE and MPE, respectively. The Nnet function would take advantage of the RMSE high values and variance thus here being uncertainty and thus meaning nonlinearity of the test it would eventually perform better.

Table 7

Methods	**Weekly**		**Values**				
	ME	RMSE	MAE	MPE	MAPE	MASE	AVE
AUTO	0.577	85.463	57.664	-0.407	5.961	0.963	25.037
SES	0.577	85.463	57.664	-0.407	5.961	0.963	26.547
HOLT	0.267	86.895	58.387	-0.313	6.047	0.983	25.378
BATS & TBATS	-0.308	86.332	58.239	-0.750	5.952	0.980	25.074
DSHW	1.063	95.586	64.763	1.092	7.916	1.091	28.585
NAÏVE & SNAIVE	5.738	88.257	59.434	-0.317	6.042	1.000	26.692
NNET	-0.051	82.118	56.562	-0.739	5.714	0.939	24.091
SPLINE	898.464	1252.456	1073.246	73.509	80.697	NA	NA
THETAF	5.880	87.759	58.980	-0.335	6.004	0.992	26.547
RWF	0.000	87.679	58.843	-0.322	6.074	0.992	25.544
MEAN	0.000	452.212	395.048	-57.127	79.138	6.289	145.927
AUTO % SES	188.543	530.967	438.400	15.892	31.776	6.290	201.978
AUTO % TBATS	94.809	449.495	364.327	12.970	29.857	5.713	159.528
AUTO % NNET	193.753	549.685	459.314	15.377	32.748	6.507	209.564
SES & THETAF	209.285	564.813	470.709	19.446	34.611	6.290	217.526
SES % MEAN	493.431	668.433	599.344	-11.466	57.966	8.685	302.732
TBATS & THETA	115.550	466.547	381.223	16.524	32.233	6.313	169.732
NNET & THETA	188.543	530.967	438.400	15.892	31.776	6.290	201.978

2.6.2 Discussion

In our weekly tests, the neural network model performed significantly better than other models in four out of the six error matrices, beating the naïve model in the daily test where the naïve model had performed better in three out of the six error matrices. The neural network performed better in the following errors: RMSE, MAE, MAPE, and MASE. The high performance in four errors shows that there is more nonlinearity in the weekly test where Nnet recorded better performance under the mentioned conditions. The lowest performing model was produced by the spline model, which had the least accurate results in four out of the six error measurements. RMSE, MAE, MPE, and MAPE were the measurements where the spline model showed its weakness.

2.7 Conclusion

To conclude, the test showed significant diversity within the error measurements, and this greatly enhanced our understanding of whether the random walk and statistical techniques can outperform AI (AI here being the neural network). In our test, we observed that, from a generally, the naïve model performed better than the neural network under our daily frequency testing. However, the neural network performed better in our weekly testing. Furthermore, the results revealed methods that would perform weakly across the board like the mean method and the spline method. To complement our testing, we need future testing to introduce horizon testing to allow for more detailed inner analysis of the methodology. We would then be able to observably understand our error testing on a different platform where horizons are being introduced. This will show which method performed better, and where and when it did so.

2.7.1 Analysis & Evaluation

Table 8

Classifications of Accuracy Testing

	ME	RMSE	MAE	MPE	MAPE	MASE
Daily 1st	!RWF&MEAN	NNET	NAÏVE	HOLT	NAÏVE	NAÏVE
Daily LE	SES%MEAN^	MEAN	MEAN	MEAN	MEAN	MEAN
Weekly 1st	!RWF/MEAN	NNET	NNET	HOLT	NNET	NNET
Weekly LE	SES%MEAN^	SPLINE	SPLINE	SPLINE	SPLINE	SES/MEAN^

* (LE) here meaning least accurate model.

^ Here representing a hybrid model combining two functions.

& here meaning a joint accurate best accurate model. So, both methods have the same result of error.

! Nnet was very close behind with -0.0043

2.7.2 Description

As seen from the tables above, both neural networks and the naïve methods have performed very well in the test run.

As the RWD enforced its will upon other models, Nnet performed better in the weekly frequency and also almost did so in the daily frequency. If an average of the errors was taken for each method, neural networks performed better in both daily and weekly frequencies.

This also takes into consideration that the neural network test was implemented with a single input and no alternative inputs were chosen in order for the network to learn, work, and develop.

2.7.3 Future Research

While using the same methods, future research should add horizon testing to the methodology and should be able to observe how the methods change over time. Horizon will be tested for all of the used frequencies; the same errors will be tested to keep the accuracy testing consistent. Furthermore, the horizons will be developed according to real financial market trading.

2.8 References

Abbas, H., Belkheiri, M., & Zegnini, B. (2015). Feedback linearization control for highly uncertain nonlinear systems augmented by single-hidden-layer neural networks. Journal of Engineering Science and Technology Review, 8(2), 215 - 224.

Adhikari, R., Agrawal, R. K., & Kant, L. (2013). PSO based Neural Networks vs. Traditional Statistical Models for Seasonal Time Series Forecasting. Advance Computing Conference (IACC). Ghaziabad: IEEE

Araújo, R. d., Oliveira, A. L., & Meira, S. (2015). A hybrid model for high-frequency stock market forecasting. Expert Systems with Applications , 42, 4081–4096.

Armstrong, J. S. (2001). Should We Redesign Forecasting Competitions? International Journal of Forecasting , 17, 542-545.

Bandyopadhyay, G., & Chattopadhyay, S. (2007). Single hidden layer artificial neural network models versus multiple linear regression model in forecasting the time series of total ozone. Int. J. Environ. Sci. Tech., 4(1), 141-149.

Barberis, N., & Xiong, W. (2009). What Drives the Disposition Effect? An Analysis. The Journal of Finance , LXIV (2), 751-784.

Barunik, J., Krehlik, T., & Vacha, L. (2016). Modeling and forecasting exchange rate volatility in time-frequency domain. European Journal of Operational Research, 251(1), 329-340.

Bekiros, S., & Georgoutsos, D. A. (2008). Direction-of-Change Forecasting Using a Volatility-Based Recurrent Neural Network. Journal of Forecasting, 27, 407–417.

Brockwell, P. J., & Davis, R. A. (2002). Introduction to time series and forecasting (2 ed.). New York: Springer.

Catillo, E., Fontenla-Romero, O., Guijarro-Berdinas, B., & Alonso-Betanzos. (2002). A global optimum approach for one-layer neural networks. Neural Computation, 14, 1429-1449.

Chandra , B., & Sharma, R. (2016). Deep learning with adaptive learning rate using laplacian score. Expert Systems With Applications, 63, 1-7.

Chen, C. W., Wang, Z., Sriboonchitta, S., & Lee, S. (2017). Pair trading based on quantile forecasting of smooth transition GARCH models. North American Journal of Economics and Finance, 39, 38–55.

Chin, L., & Hong, L. W. (2008). Can financial ratios predict the Malaysian stock return? Integration & Dissemination, 2, 7-9.

Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Systems With Applications, 83, 187-205.

Choudhry, T., McGroarty, F., Peng, K., & Wang, S. (2012). High-frequency exchange-rate prediction with an artificial neural network. Intelligent Systems in Accounting, Finance and Management, 19, 170–178.

Claveria, O., Monte, E., & Torra, S. (2015). Tourism demand forecasting with neural network models: Different ways of treating information. International Journal of Tourism Research, 17, 492-500.

Coelho, I., Coelho, V., Luz, E., Ochi, L., & Guimarães, F. (2017). A GPU deep learning metaheuristic based model for time series. Applied Energy, 201, 412-418.

Corrado, C., & Truong, C. (2007). Forecasting stock index volatility: comparing implied volatility and the intraday high–low price range. Journal of Financial Research , 30 (2), 201-215

CristianinI, N., & Kandola, J. (2006). On kernel target alignment. Innovations in machine learning: theory and applications. Berlin: Springer.

Detienne, K. B., Detienne, D. H., & Joshi, S. A. (2003). Neural Networks as Statistical Tools for Business Researchers. Organizational Research Methods , 6, 236-265.

Dhamija , A., & Bhalla, V. (2011). Exchange rate forecasting: comparison of various architectures of neural networks. Neural Comput & Applic., 20, 355-363.

Dunis, C. L., Rosillo, R., de la Fuente, D., & Pino, R. (2013). Forecasting IBEX-35 moves using support vector machines. Neural Comput & Applic, 23, 229–236.

Dunis, C., & Huang, X. (2002). Forecasting and Trading Currency Volatility: An Application of Recurrent Neural Regression and Model Combination. Journal of Forecasting, 21(5), 317-354.

Elliott, G., & Timmermann, A. (2013). Handbook of economic forecasting Volume 2B. Amsterdam : Elsevier North-Holland.

Fok, W., Tam, V., & Ng, H. (2008). Computational Neural Network for Global Stock Indexes Prediction. Proceedings of the World Congress on Engineering , 2, 1-5.

Giusti, C., & Itskov, V. (2014). A no-go theorem for one-layer feedforward networks. Neural Computation, 26, 2527–2540.

Graves, D., & Pedrycz, W. (2009). Fuzzy prediction architecture using recurrent neural networks. Neurocomputing , 72 (7), 1668–1678

Green, K. C., & Armstrong, J. S. (2015). Simple versus complex forecasting: The evidence. Journal of Business Research , 68 (8), 1678–1685.

Guliyev, N., & Ismailov, V. (2016). A single hidden layer feedforward network with only one neuron in the hidden layer can approximate any univariate function. Neural Computation, 28, 1289–1304.

Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2017). Learning features for offline handwritten signature verification using deep convolutional neural networks. Pattern Recognition, 70, 163-176.

Haykin, S. (2007). Neural networks and learning machines. Canada: McMaster.

Hribar, P., & McInnis, J. (2012). Investor sentiment and analysts' earnings forecast errors,. Management Science, 58, 293–307.

Hu, O. (2007). Applicability of the Fama-French three-factor model in forecasting portfolio returns. Journal of Financial Research, 30(1), 111-127.

Huang, W., Lai, K., Nakamori, Y., Wang, S., & Yu, L. (2007). Neural networks in finance and economics forecasting,. International Journal of Information Technology and Decision Making, 6(1), 113–140.

Hyndman, R. J., & Athanasopoulos, G. (2014). Forecasting : principles and practice. Heathmont, Vic.: OTexts.

Jarrett, J. E., & Kyper, E. (2011). ARIMA Modeling With Intervention to Forecast and Analyze Chinese Stock Prices. Int. J. Eng. Bus. Manag , 3 (3), 53-58.

Jiang, S., Chin, K.-S., Wang, L., Qu, G., & Tsui, K. (2017). Modified genetic algorithm-based feature selection combined with pre-trained deep neural network for demand forecasting in outpatient department. Expert Systems With Applications, 82, 216-230.

Joseph, K., Wintoki, M., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. International Journal of Forecasting, 27(4), 1116–1127.

Kaastra, I., & Boyd, M. (2006). Designing a Neural Network for Forecasting Financial and Economic Time Series. Neurocomputing , 10, 215-236.

Khashei, M., & Bijari, M. (2014). Fuzzy artificial neural network (p, d, q) model for incomplete financial time series forecasting. Journal of Intelligent & Fuzzy Systems, 26(2), 831-845.

Kim, H., & Shin, K. (2007). A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. Applied Soft Computing , 7 (2), 569–576.

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. European Journal of Operational Research, 259, 689-702.

Kuo, C. C., & Reitsch, A. (1996). Neural Networks vs. Conventional Methods of Forecasting. The Journal of Business Forecasting Methods & Systems , 14 (4), 112-131.

Kuremoto, T., Kimura, S., Kobayashi, K., & Obayashi, M. (2014). Time series forecasting using a deep belief network with restricted Boltzmann machines. Neurocomputing, 137, 47-56.

Kutsurelis, J. E. (1998, September). Forecasting Financial Markets Using Neural Networks: An Analysis Of Methods And Accuracy.

Kwan, J., Lam, K., So, M., & Yu, P. (2000). Forecasting and Trading Strategies Based on a Price Trend Model. Journal of Forecasting, 19, 485-498.

Lang, K., Zhang, M., & Yuan, Y. (2015). Improved neural networks with random weights for short-term load forecasting. PLoS ONE, 10(12), 1-14.

Lawrence, R. (1998, January). Using neural networks to forecast stock market prices. Retrieved May 16, 2016, from https://www.researchgate.net/publication/228607056_Using_neural_networks_to_forecast_stock_market_prices

Li, J.-C., Tang, N.-S., Mei, D.-C., Li, Y.-X., & Zhang, W. (2016). The trading time risks of stock investment in stock price drop. Physica A , 1-20.

Liu , Z., Li , W., & Sun, W. (2013). A novel method of short-term load forecasting based on multiwavelet transform and multiple neural networks. Neural Comput & Applic, 22, 271–277.

Liu, Q., Guo, Z., & Wang, J. (2012). A one-layer recurrent neural network for constrained pseudoconvex optimization and its application for dynamic portfolio optimization,". Neural Networks, 26, 99-109.

Lollia, F., Gamberinia, R., Regattierib, A., Balugania, E., & Gatos, T. (2017). Single-hidden layer neural networks for forecasting intermittent demand. Int. J. Production Economics, 183, 116-128.

Lollia, F., Gamberinia, R., Regattierib, A., Balugania, E., & Gatos, T. (2017). Single-hidden layer neural networks for forecasting intermittent demand. Int. J. Production Economics, 183, 116-128.

Lollia, F., Gamberinia, R., Regattierib, A., Balugania, E., & Gatos, T. (2017). Single-hidden layer neural networks for forecasting intermittent demand. Int. J. Production Economics, 183, 116-128.

Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. The Journal of Political Economy , 98 (4), 703-738.

Luo, C., Wu, D., & Wu, D. (2016). A deep learning approach for credit scoring using credit default swaps. Engineering Applications of Artificial Intelligence, 1-6.

Ma , J., Sheridan, R., Liaw , A., Dahl, G., & Svetnik, V. (2015). Deep Neural Nets as a Method for Quantitative Structure–Activity Relationships. Journal of Chemical Information and Modeling, 55(2), 263–274.

Mall, S., & Chakraverty, S. (2015). Multi Layer Versus Functional Link Single Layer Neural Network for Solving Nonlinear Singular Initial Value Problems. Proceedings of the Third International Symposium on Women in Computing and Informatics, 678-683 .

Marquez, L., Hill, T., O'Connor, M., & Remus, W. (1992). Neural network models for forecast: a review.

Martens, M. (2002). Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. Journal of Futures Markets , 22 (6), 497–518.

Miralles-Quirós, J., Miralles-Quirós, M., & Daza-Izquierdo, J. (2015). Intraday patterns and trading strategies in the Spanish stock market. Applied Economics, 47(1), 88–99.

Moshiri, S., & Cameron, N. (2000). Neural Network Versus Econometric Models in Forecasting In¯ation. Journal of Forecasting , 19, 201-217.

Najafabadi, M., Villanustre, F., Khoshgoftaar , T., Seliya, N., & Wald, R. (2015). Deep learning applications and challenges in big data analytics. Journal of Big Data, 2(1), 1-21.

Nguyen, H., & Chan, C. (2004). Multiple neural networks for a long term time series forecast. Neural Comput & Applic, 13, 90–98.

Önder, E., Bayır, F., & Hepşen, A. (2013). Forecasting Macroeconomic Variables using Artificial Neural Network and Traditional Smoothing Techniques. Journal of Applied Finance & Banking , 3 (4), 73-104

Ong, Smola, & Williamson (2005). Learning the Kernel with Hyperkernels, Journal of Machine Learning Research 6 (2005) 1043–1071

Ord, J. K., & Fildes, R. (2013). Principles of business forecasting (1 ed.). Mason, OH: South-Western Cengage Learning.

Panchal, G., Ganatra, A., Kosta, Y. P., & Panchal, D. (2011). Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. International Journal of Computer Theory and Engineering, 3(2), 332–337.

Paradarami, T., Bastian, N., & Wightman, J. (2017). A hybrid recommender system using artificial neural networks. Expert Systems With Applications, 83, 300–313.

Pedzisz, M., & Mandic, D. (2008). A Homomorphic Neural Network for Modeling and Prediction. Neural Computation, 20, 1042-1064.

Pellakuri, V., & Rao, R. (2016). Training and development of artificial neural network models: Single layer feedforward and multi layer feedforward neural network. Journal of Theoretical and Applied Information Technology, 84(2), 150-156.

Pham, H. V., Cooper, E. W., Cao, T., & Kamei, K. (2014). Hybrid Kansei-SOM model using risk management and company assessment for stock trading. Information Sciences , 256, 8–24.

Qu, H., & Zhang, Y. (2016). A New Kernel of Support Vector Regression for Forecasting High-Frequency Stock Returns. Mathematical Problems in Engineering, 22, 1-9.

Riedel, C., & Wagner, N. (2015). Is risk higher during non-trading periods? The risk trade-offfor intraday versus overnight market returns. Int. Fin. Markets, Inst. and Money , 39, 53–64.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117.

Sermpinis, G., Stasinakis, C., Rosillo, R., & de la Fuente, D. (2017). European exchange trading funds trading with locally weighted support vector regression. European Journal of Operational Research, 258, 372–384.

Sheela , K., & Deepa, S. (2013). Review on methods to fix number of hidden neurons in neural networks. Mathematical Problems in Engineering, 55, 1-11.

Shen, F., Chao, J., & Zhao, J. (2015). Forecasting exchange rate using deep belief networks and conjugate gradient method. Neurocomputing, 167, 243–253.

Sheu, H.-J., & Wei, Y.-C. (2011). Options Trading Based on the Forecasting of Volatility Direction with the Incorporation of Investor Sentiment. Emerging Markets Finance & Trade, 47(2), 31–47.

Shoji, I., & Kanehiro, S. (2016). Disposition effect as a behavioral trading activity elicited by investors' different risk preferences. International Review of Financial Analysis , 46, 104–112.

Sokolov-Mladenovic, S., Milovancevic, M., Mladenovic, I., & Alizamir, M. (2016). Economic growth forecasting by artificial neural network with extreme learning machine based on trade, import and export parameters. Computers in Human Behavior, 65, 43-45.

Sortino, F., & Satchell, S. (2011). Managing Downside Risk in Financial Markets. Burlington : Elsevier.

Spithourakis , G., Petropoulos, F., Nikolopoulos, K., & Assimakopoulos, V. (2015). Amplifying the learning effects via a Forecasting and Foresight Support System. International Journal of Forecasting, 31, 20-32.

Steinwart, I. (2001). On the Influence of the Kernel on the Consistency of Support Vector Machines. J. Machine Learning Res., 2, 67–93.

Sun, K., & Zhang, J. (2017). Generalized extreme learning machine autoencoder and a new deep neural network. Neurocomputing, 230, 374-381.

Tay, F. E., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. Omega , 29, 309–317

Tealab, A., Hefny, H., & Badr, A. (2017). Forecasting of nonlinear time series using ANN. Future Computing and Informatics Journal, 20, 1-9.

Tealab, A., Hefny, H., & Badr, A. (2017). Forecasting of nonlinear time series using ANN. Future Computing and Informatics Journal, 20, 1-9.

Thenmozhi, M. (2006). Forecasting Stock Index Returns Using Neural Networks. Delhi Business Review , 7 (2), 59-69.

Thenmozhi, M. (2006). Forecasting Stock Index Returns Using Neural Networks. Delhi Business Review , 7 (2), 59-69.

Toloeiashlaghi, A., & Haghdost, S. (2004). The Models of Stock Price Forecasting by Neural Network Compared with Mathematical Models of Forecasting. J. Econom. Res. , 25 (2), 237-250.

Tseng, K.-C., Kwon, O., & Tjung, L. C. (2012). Time series and neural network forecast of daily stock prices. Investment Management and Financial Innovations , 9 (1), 32-54.

Ullrich, C., Seese, D., & Chalup, S. (2007). Foreign Exchange Trading with Support Vector Machines. Advances in Data Analysis , 8, 539-546.

V'yugina , V., & Trunov, V. (2016). Applications of Combined Financial Strategies Based on Universal Adaptive Forecasting. Automation and Remote Control, 77(8), 1428–1446.

Vella, V., & Ng, W. (2014). Enhancing risk-adjusted performance of stock market intraday trading with Neuro-Fuzzysystems. Neurocomputing , 141, 170–187.

Vella, V., & Ng, W. L. (2016). Improving risk-adjusted performance in high frequency trading using interval type-2 fuzzy logic. Expert Systems With Applications 55 (2016) 70–86 , 55, 70–86.

Walczak, S. (2001). An empirical analysis of data reqirements for financial forecasting with neural networks. Journal of Management Information Systems , 17 (4), 203-222.

Williams, R. (2011). An introduction to trading in the financial markets : trading, markets, instruments, and processes (2 ed.). Amsterdam ; Boston: Academic Press.

Wu, H., Yao, R., Li, R., & Zhang, X. (2014). A one-layer recurrent neural network for solving pseudoconvex optimization with box set constraints. Mathematical Problems in Engineering, 1-8.

Zakarya, S., Abbas, H., & Belal, M. (2017). Long-term deep learning load forecasting based on social and economic factors in the Kuwait region. Journal of Theoretical and Applied Information Technology, 95(7), 1524-1535.

Zeng, N., Zhang, H., Liu, W., Liang, J., & Alsaadi, F. E. (2017). A switching delayed PSO optimized extreme learning machine for short-term load forecasting. Neurocomputing, 240, 175-182.

Zeng, N., Zhang, H., Liu, W., Liang, J., & Alsaadi, F. E. (2017). A switching delayed PSO optimized extreme learning machine for short-term load forecasting. Neurocomputing, 240, 175-182.

Zhang, W., Cao, Q., & Schniederjans, M. (2004). Neural Network Earnings per Share Forecasting Models: A Comparative Analysis of Alternative Methods. Decis. Sci. , 35 (2), 205-237.

Zhao, Y., Li , J., & Yu, L. (2017). A deep learning ensemble approach for crude oil price forecasting. Energy Economics, 66, 9-16.

2.9 Appendices

2.9.1 Graphs

Figure 1

AstraZeneca monthly neural network forecast

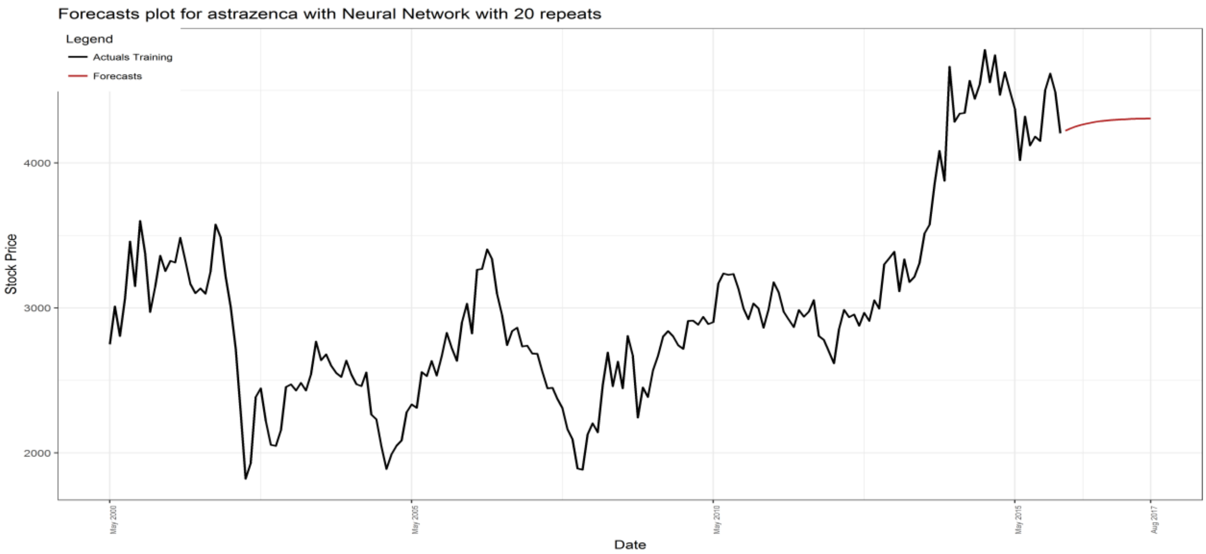
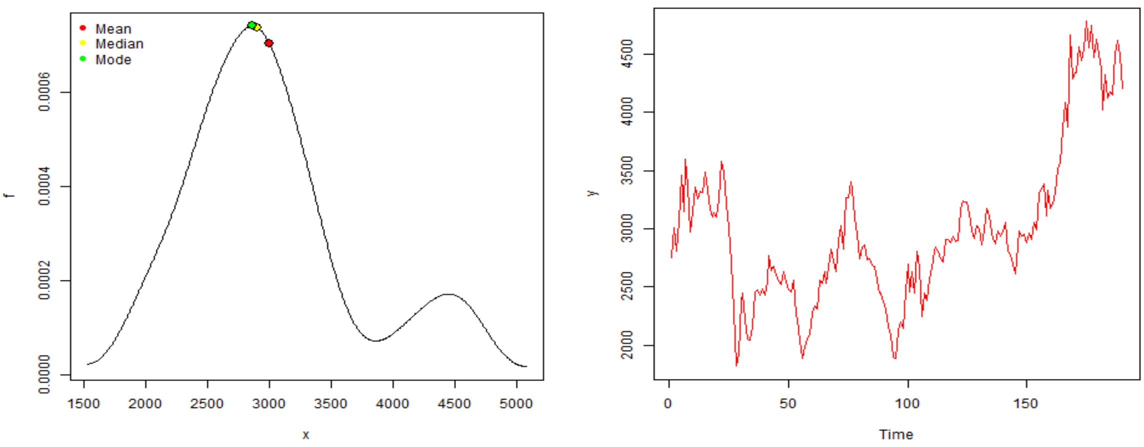


Figure 2



2.9.2 Forecast Accuracy Measures

The forecasting accuracy measures were tested and developed in accordance to the following structures

Error Metrics

Equation 1

1. Symmetric mean absolute percentage error (sMAPE) =

$$\sum \frac{|X-F|}{(X+F)/2} * 100$$

Equation 2

2. Symmetric median absolute percentage error (sMdAPE) =

$$\sum \frac{|X-F|}{(X+F)/2} * 100$$

Equation 3

3. Median relative absolute error (MdRAE_S) =

$median(|r_t|)$ = with $(|r_t|) = \frac{X_t - F_t}{X_t - F_t^*}$

MdRAE recommended by Makridakis (2000) to compare accuracy for alternative models as it controls for outliers.

2.9.3 Error Metrics Tested

Equation 4

Mean Error (ME)

$$ME = \frac{1}{N} \sum_{i=1}^{N} (f_i - k_i)$$

Where:

N = # Forecast / Actuals

F = Forecast

K = Actuals

Equation 5

Mean Absolute Error (MAE)

$$= \text{mean}\ (|e_i|)$$

Equation 6

Root Mean Squared Error (RMSE)

$$= \sqrt{mean\ (e_i^2)}$$

Equation 7

Mean Percentage Error (MPE)

$$MPE = \frac{1}{N}\ \sum_{i=1}^{N} (f_i - k_i) \text{ x } 100$$

Equation 8

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{N}\ \sum_{i=1}^{N} (|f_i| - |k_i|) \text{ x } 100$$

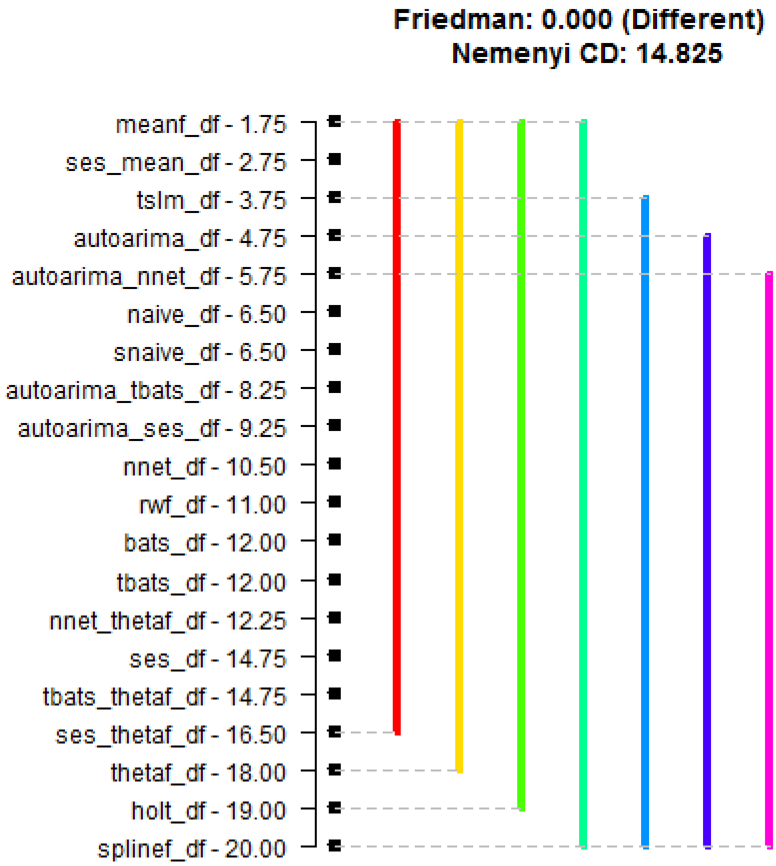
Equation 9

Mean Absolute Scaled Error (MASE)

$$MASE\ q_t = \frac{e_t}{\frac{1}{n-1}\ \sum_{i=2}^{n} (|Y_i - \ Y_{i-2}|)}$$

The aforementioned accuracy tests follow the structure built by Robert Hyndman,

Figure 3



Here the nemenyi test shows no significance.

More significant testing with the nemenyi test is shown below in the appendices.

Figure 4

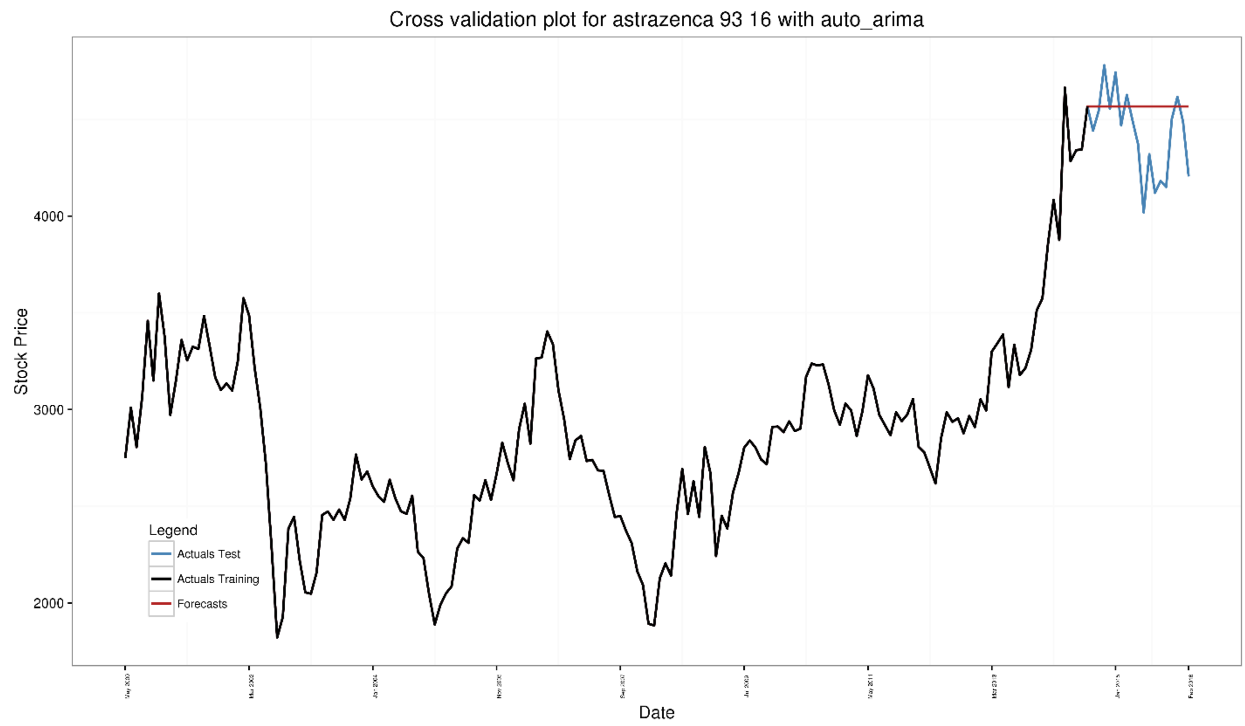
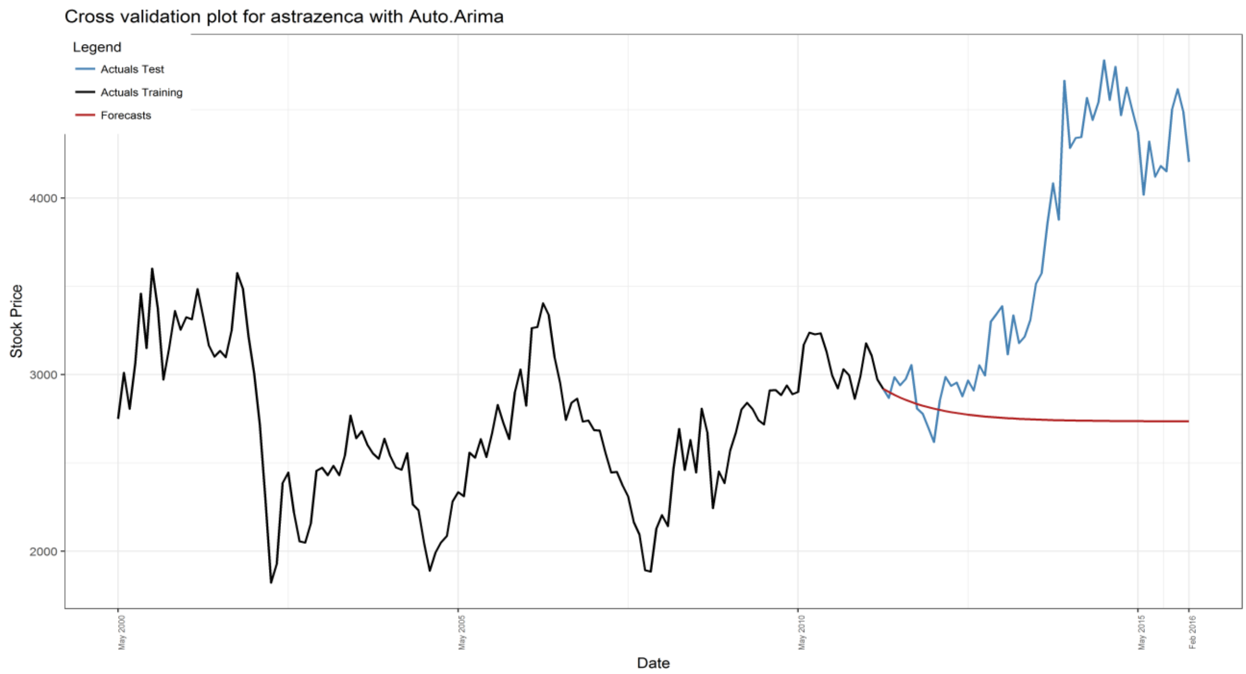


Figure 5



CHAPTER 3

FORECASTING FINANCIAL MARKETS WITH PREDICTIVE ANALYTICS: THE IMPACT OF THE FORECASTING HORIZON

3.0 Abstract

This paper investigates the impact of horizon when introduced into a forecasting accuracy test. It delves further into forecasting, looking at how to train and forecast data. We used a plus 1 training methodology where we used the first 2.5% of the data for training. One more data point was then added to the next training, and so on. Accordingly, the first forecast used 100 points in the daily time series for example and produced 22 horizons ahead. Thereafter, we then used 101 training points to forecast the next 22 horizons. We used 25 methods to forecast 18 different time series, while we also applied six different accuracy tests to determine which methods performed better under different horizons.

3.1 Introduction

This paper introduces horizons to our test, where we compare the difference between five different horizons using five different frequencies. For the daily frequency, 28 horizons were tested, however not all horizons are shown. We chose horizons 1, 2, 3, 4, 5, 10 and 22 to present as those horizons can be developed into a trading strategy as the stocks trade for five days a week and 22 days a month. That concept also applies for the for the other four frequencies frequencies. Horizon is introduced as a factor to better understand the methods applied in this paper. Horizons help us to understand the durability, efficiency and strength of the tested methods. This will, furthermore, enhance our understanding of how a method's accuracy develops.

3.2 Literature

3.2.1 Introduction

Accuracy is an important component of any forecasting undertaking. Horizon accuracy testing and normal accuracy testing have some remarkable differences that may affect the precision of the forecasting endeavour. Empirical studies conducted by various scholars have delved into the issue of whether horizon forecasting and normal forecasting lead to different predictions. Specifically, most of these studies have tested the utility of short-horizon forecasting vis-à-vis long forecasting horizons. Some of the measures for testing the accuracy in different models as identified in the existing literature include mean error, mean absolute error (MAE), mean squared error (MSE), mean absolute percentage error (MAPE), mean percentage error (MPE), and root mean squared error (RMSE). This review explores the literature pertaining to forecasting accuracy with specific emphasis on horizon accuracy testing and normal accuracy testing. Examining the issues related to forecasting accuracy can ultimately help concerned stakeholders to make informed financial decisions that enhance their sustainability in the long term.

Comparing horizons in forecasting is essential when it comes to uncertaitnity and accuracy, as the indiviualisation of singularity in horizons can identify any anomolies and show where the algorithims have worked better then others. Furthermore, this approach would enable us to to identify which method is more accurate under which horizon, also different functions may be more accurate under different horizons.

3.2.2 Literature Review

Moshiri & Cameron (2000) endeavoured to conduct a study examining the predictive power of neural network models vis-à-vis statistical methods. Their primary focus was on predicting inflation. The researchers compared dynamic forecasts for three distinct horizons encompassing one, three and twelve months ahead. Since 2009, the Federal Open Market Committee (FOMC) members, have decided to test and implement long-run forecasting. Thus, through their economic projections, Chen, Wang, Sriboonchitta, & Lee (2017) reiterated that decomposing volatility into jumps and several time scales can help in approximating traders' behaviour at corresponding investment horizons. Based on previous research on investor sentiment, Joseph et al. (2011) expected online search intensity to forecast trading volume and stock returns, and that highly volatile stocks that are more difficult to arbitrage would have a higher sensitivity to search intensity compared to less volatile stocks. Over the 2005-2008 period, Joseph et al. (2011) found that over a weekly horizon in a sample of S&P 500 firms, online search intensity was able to predict trading volume and abnormal stock returns reliably.

A study by Simionescu (2015) sought to explore the accuracy of unemployment rate predictions for the Romanian economy. In this inquiry, three anonymous forecasters with the pseudonyms F1, F2, and F3 provided the forecasts. The researcher applied a multi-criteria ranking to make a hierarchy of forecasters with regard to accuracy, and considered five important accuracy measures at the same time. These accuracy measures were ME, RMSE, MSE, U_1 and U_2 statistics of Theil. F3 provided the most accurate predictions for the horizon 2001-2014, whereas F2 predictions were less accurate as per U_1 Theil's statistic. ME for F3 was -0.7302, MAE was 1.102, RMSE was 1.4028, U_1 was 0.1129, and U_2 was 0.9983 whereas for F2, ME was -0.5833, MAE was 1.6592, RMSE was 1.9037, U_1 was 0.1430, and U_2 was 1.1025. Consistent with the findings of Busemeyer, Wang, Townsend, & Eidels (2015), Simionescu (2015) recommended the use of a combination of various forecasting methods such as optimal combination (OPT), inverse MSE weighting scheme (INV), and equal-weights-scheme (EW), in order to improve the accuracy of unemployment rate predictions.

From a different perspective, Knüppel (2018) endeavoured to examine the relationship between increasing the number of horizon and forecast uncertainty. In carrying out this investigation, the researcher applied the SUR estimator to the forecast errors of the

FOMC, the US Survey of Professional Forecasters, and the Bank of England. The SUR estimator relies on the correlations between forecasting errors from different horizons for the same period, a typical feature of empirical forecast error. After applying the SUR estimator to the increased horizon of the FOMC, the US Survey of Professional Forecasters, and the Bank of England, Knüppel (2018) found the SUR estimator to be robust with respect to non-optimal forecasts and the presence of stationary time-varying uncertainty. The researcher also established that the SUR estimator provided potentially large efficiency gains particularly when introducing new, longer forecasting horizons. This suggests that horizon accuracy testing may be more beneficial in making better predictions when compared to normal accuracy forecasting.

Contrary to the conclusion arrived at by Knüppel (2018), other scholars are of the view that most forecasting models fail to accurately predict the occurrence of failure beyond one year, with their accuracy tending to decline as the forecasting horizon recedes. du Jardin & Séverin (2011) conducted a study in which they aspired to demonstrate a new way of using the Kohonen map to increase the prediction horizon of a financial failure model and improve the model's reliability. Data for this study came from the French database Diane that contains financial data of over 1 million French companies. The findings indicated that, over the duration studied, the generalization error achieved by the self-organizing Kohonen map remained relatively stable, unlike that of other traditional methods such as neural networks, logistic regression, survival analysis, and discriminant analysis used traditionally for financial failure prediction.

For one year before failure, the self-organizing Kohonen map developed and used by Knüppel (2018) had an accuracy rate of 84.09%, compared to 80.00% for Cox's model, 81.59% for discriminant analysis, 81.82% for logistic regression, and 82.05% for the neural network. The gap between the proposed model and the other models grew even larger when the researchers measured accuracy two and three years out, prompting them to conclude that their model would be particularly useful in companies seeking to make financial forecasts over the short- and medium-term. Reikard (2015), from another dimension, aspired to explore the utility of time series models in forecasting geomagnetic activity albeit at annual and monthly horizons. The researcher ran the forecasting tests for the Aa index, beginning in 1868 and providing the longest continuous records of geomagnetic activity. The researcher affirmed that this series is difficult to predict because, while it exhibits cycles at 11-22 years, the period and amplitude of the cycles vary over time.

Solanki & Krivova (2011) echoed similar sentiments to Reikard (2015), adding that evidence of discontinuous trending in which the trend's direction and slope change repeatedly further complicates forecasting geomagnetic activity. In a study by Reikard (2015), the researcher specifically tested a number of models that included regressions, a frequency domain algorithm, neural networks, and combined models, with forecasting tests run in horizons of 1-12 months using the monthly data and 1-11 years using the annual data. Results from this study were intriguing. At the 1-year horizon, the median errors for the models ranged between 10-14% while the mean errors were in the range of 13-17%.

The models' accuracy nonetheless deteriorated at longer horizons, with mean errors ranging between 21-23% at five years and 23-25% at 11 years. At the monthly resolution, median errors in this study by Reikard (2015) ranged between 14-17% whereas the median errors ranged between 17-19%, with the mean error increasing to 23-24% at five months and 25% at 12 months. From these findings, Reikard (2015) concluded that a method that combines frequency and time domain methods was marginally superior to neural networks or regressions alone for short horizons. Solanki & Krivova (2011) supported this view, asserting that the nonlinear variability in monthly series as well as irregular cycles and trends in geomagnetic activity makes forecasting of such activity problematic in long horizons.

Just like Simionescu (2015), Sanders, Manfredo & Boris (2009) also conducted a study to examine the forecasting accuracy at different horizons, although their study focused on the United States Department of Energy's (DOE) energy commodities. Forecast horizons were at the first quarter through to the fourth quarter. In this inquiry, the investigators used a direct test for determining information content albeit at alternative forecast horizons. Results from the study suggested that the DOE's price forecast for electricity and natural gas were informative up until the fourth-quarter horizon whereas forecasts for crude oil, diesel fuel, and gasoline provided incremental information out to three-quarters ahead. For natural gas, the MAE at k=1 was 9.1%, at k=2 it was 24.1%, and at k=4, which was the longest horizon, the MAE was 36.2%. These results seem to offer support for short-term, normal accuracy testing rather than long-term horizon forecasting.

Using a case study from Henan province in China, Reindl, Walsh, Yanqin & Bieri (2017) aspired to evaluate the economic value of forecasting on different horizons in the energy meteorology sector. Their findings suggested that even minor deviations from the requested prediction corridor and the forecasting frequency could result in adverse revenue losses that directly affected the finances of the project. Specifically, the losses pertained to discounted payback period, internal rate of return, and net-present value. The researchers, similar to Diagnea, Mathieu, Lauret, Boland & Schmutz (2013), established that forecasting of irradiance necessitated the use of different techniques for different time horizons.

For long-term forecasts, that is, days ahead, the most appropriate forecasting techniques as identified by Reindl et al. (2017) were complex numerical weather prediction models such as the European Center for Medium Range Weather Forecasts (ECMWF) or Weather Research and Forecasting (WRF). For medium-term forecasting, for example intra-day forecasts, the meteorological models that lead to the best results were those based on combinations of ground observations and satellite-derived data whereas, for short-term forecasting, machine learning time series was found to had greater utility (Reindl et al., 2017).

Consistent with this study by Reindl et al. (2017), Reikard, Haupt & Jensen (2017) similarly initiated a study to examine the effectiveness of forecasting irradiance over short horizons in the meteorological sector. The researchers affirmed that short-range forecasting is one of the enabling technologies to have facilitated the integration of solar energy into the grid. Using data from two locations in the United States, Sacramento and Brookhaven, Reikard et al. (2017) endeavoured to examine a number of forecasting models over horizons ranging from 15 minutes to four hours.

The models tested included the WRF model and the Dynamic Integrated Forecast (DICast) system compared against the primary time series model of ARIMA. Among time series models, Reikard et al. (2017) established that ARIMAs with time-varying coefficients were far superior to fixed coefficient methods. On the contrary, when comparing time series and meteorological models, the results indicated that the ARIMA was more accurate at short horizons, whereas the numerical weather prediction models including the WRF and DICast were more accurate at longer horizons as measured by MAE and RMSE.

To improve the long-horizon forecast accuracy, Snudden (2018) proposed growth rate transformations that have targeted lag selection. In his study, which sought to examine long-horizon crude oil forecasts, the researcher established that targeted growth rates could improve the precision of forecasts significantly at horizons of up to five years. In addition, for the real crude oil price, the targeted lag selection could achieve a higher degree of accuracy up to 10 years ahead that previous standard forecasting methods could only achieve at shorter horizons.

Baumeister & Kilian (2015) echoed similar sentiments. In their study, they found that the application of targeted growth rate transformations to vector autoregressive (VAR) models could consistently outperform VAR models that relied on period-over-period growth rates particularly in out-of-sample forecast performances at longer horizons. Just like Snudden (2018), the results of Baumeister & Kilian (2015) also indicated that when forecasters apply targeted growth rates to autoregressive models, the success ratios are consistently over 0.5 while the MSPE ratios are more often below 1 up until the three-year-ahead horizon.

Interestingly, Cenesizoglu, Ribeiro & Reeves (2017) also found autoregressive models to have superior predictive accuracy for long horizons. The purpose of their investigation was to evaluate a number of beta forecasting techniques for long forecast horizons. The specific models evaluated included the widely used Fama-MacBeth (FM) beta approach based upon five years of monthly returns and an autoregressive model of the realised beta. The researchers calculated realised beta for 15 US companies, with the DJIA index serving as the market portfolio. They sought equity data from for the DJIA via DataStream from 1 January 1952 to 31 December 2011, giving them a time series of 60 years.

Using MSE and MAE as the baseline measures for the models' predictive accuracy, Cenesizoglu et al. (2017) found that the FM beta approach based upon five years of monthly returns was unreliable in terms of statistical bias as well as MSE and MAE. Using an autoregressive model to forecast beta six months ahead yielded a reduction of 26.54% in MAE and 43.55% in MSE (24.87% in RMSE) compared to the FM beta. The analysis of the forecasting bias relying on the Mincer–Zarnowitz regression offered additional support for the autoregressive model being the most effective forecasting model for both yearly and six-month beta forecasts.

In another study comparing multi-horizon forecasts, Capistrán (2006) sought to prove the superiority of applying a Diebold–Mariano test using a multivariate loss function vis-à-vis the common informal method of taking the square of the forecast errors for every horizon and then averaging over the horizons. In order to compare the Diebold–Mariano test to the conventional informal test, the researcher used a Monte Carlo simulation and a statistic based on a squared error loss (SEL) function. Besides having a higher SEL function, the informal conventional model also had a higher MSE loss compared to the Diebold–Mariano test.

In a subsequent research that endeavoured to evaluate multi-horizon evaluation methods for producing monthly inflation predictions for up to 12 months ahead, Capistrán, Constandse & Ramos-Francia (2010) were able to find support for an optimal combination approach and a bottom-up approach. The forecasts in this study relied upon disaggregated Mexican consumer price index (CPI) data as well as individual seasonal time series models, which considered both stochastic and deterministic seasonality. The forecasts produced by an optimal combination approach not only satisfied the hierarchies, but in most instances they had smaller mean squared forecast errors (MSFEs) than the forecasts yielded by the best seasonal model for each series and those produced with the bottom-up approach.

Using a different approach, Degiannakis, Filis & Hassani (2018) strived to examine the forecasting power of parametric versus nonparametric techniques on implied volatility indices over short and long horizons. The volatility indices examined included VIX (S&P 500 Volatility Index — US), VXD (Dow Jones Volatility Index — US), VXN (Nasdaq-100 Volatility Index — US), VFTSE (FTSE 100 Volatility Index — UK), VSTOXX (Euro Stoxx 50 Volatility Index — Europe), VDAX (DAX 30 Volatility Index — Germany), VXJ (Japanese Volatility Index — Japan), and VCAC (CAC 40 Volatility Index — France). The specific nonparametric techniques examined were singular spectrum analysis combined with Holt-Winters (SSA-HW) whereas the parametric technique rested on an autoregressive integrated (ARI) model. Results indicated that the SSA-HW model was significantly superior for the one as well as the ten trading days' ahead forecasting horizon.

In order to assess forecasting accuracy over the horizons, Degiannakis et al. (2018) relied on the MAE and the MSE loss functions, the direction-of-change criterion, and the model confidence set forecasting evaluation procedure. For the SSA-HW model, the MSE for the different indices ranged from 1.0000 to 0.0002 for the one-day ahead forecast and 0.3619 to 0.0000 for the ten-days ahead forecast. MAE on the contrary ranged from 1.0000 to 0.0000 for the one-day ahead forecast and 0.0000 to 1.0000 for the ten-days ahead forecast of the different volatility indices. The SSA-HW combination allowed for a compromise between forecast accuracy and model parsimony, a finding shared by Busemeyer, Wang, Townsend & Eidels (2015) who also established that parsimony additionally allows for better generalizations and predictions of new data as it helps to differentiate the signal from the noise.

In an analysis of integer autoregressive moving average (INARMA) models, Mohammadipour & Boylan (2012) found that a process involving the aggregation of the INARMA process led to increased accuracy in forecasting the conditional mean with minimum mean square error (MMSE) compared to cumulative *h-step* ahead forecast models. For all datasets, the INARMA aggregated model registered an average MSE value of 0.9636 whereas the cumulative *h-step* ahead forecast models had an average MSE value of 1.5557, proving definitively the efficacy of the INARMA aggregated model for both short- and long-term horizons. Motivated by recent successes of integer programming-based procedures used for computing discrete forecast horizons, Dawande, Gavirneni, Naranpanawe & Sethi (2009) instigated a study in which they considered two-product variants of the traditional dynamic lot-size model.

In one of the variants, Dawande et al. (2009) imposed a warehouse capacity restraint on the total ending inventory of two particular products in any period, whereas in the second variant the two products had both joint and individual setup costs for production. Assuming that future demands are discrete, Dawande et al. (2009) characterised the forecast horizon for the two variants as feasibility/optimality questions in 0–1 mixed integer programmes. A detailed computational study affirmed the effectiveness of the researchers' approach and enabled them to gain deep insights into the behaviour exhibited by minimal integer forecast horizons.

Whereas empirical and theoretical studies have highlighted the utility of short horizons vis-à-vis long horizons in determining forecasting accuracy, some studies have also found evidence of bias in predictive accuracy between the developed and the developing countries. For instance, Ince & Molodtsova (2017) examined rationality, economic value, and forecasting accuracy of survey-based exchange rate predictions at 3-, 12-, and 24-month horizons. The countries investigated included 23 developing countries and 10 developed countries, with the scholars utilising data, from two comprehensive surveys for the period between 2004 and 2012. Their analysis found strong evidence that the developing countries' forecasts had bias at all the forecasting horizons, a finding also supported by an empirical study by Blaskowitz & Herwartz (2011).

On the contrary, while strong bias characterised the forecasts of developed countries at the 3-month period, the bias decreased at the 12-month horizon and again increased at the 24-month horizon. Ince & Molodtsova (2017) concluded that, based on the direction of change and magnitude of the forecast errors, long-term forecasting tended to be more accurate compared to short-term forecasting. Economic gains of the developing countries' forecasts also improved with the forecast horizons. For developing nations, the MSPE ratio was less than one for 10 out of 23 developing countries in *FX4Casts* data and for 11 out of 23 currencies in *Consensus Economics*. For developed countries, the MSPE ratio at the 24-month horizon was less than one for 6 out of 10 developed countries (UK, Canada, Euro Area, Norway, Denmark, and Sweden).

In conclusion, it is apparent that normal and horizon forecasting find support for their predictive accuracy under different forecasting models. The utility of these accuracy tests tends to vary depending on the variables considered for forecasting and the duration of the forecasting horizon. However, in general, the forecasting accuracy as evidenced by outcome measures such as ME, MAE, MSE, MAPE, MPE, and RMSE appear to deteriorate over time, providing more support to normal- and short-horizon forecasting.

3.3 Methodology

This paper compares various forecasting horizons. It intends to find the most accurate method under different financial environments and structures. The environments that were pre-set, took into account an all-round base system. Furthermore, the all-round base system behaves structurally to prevent any bias. We set five different frequencies each testing five different horizons. This, eventually, would provide a diverse set of results for further analysis and evaluation where the methods that were tested showed different results and analysis in the different environments and horizons tested.

At first look, all of the algorithms that were tested showed that accuracy decreases as the horizon increases. The functions that were implemented were functions such as Autoarima, double seasonal Holt-Winters, neural networks and more as seen below. A table presenting all the functions is provided later in the section.

As we incorporate our horizon testing in this paper, we faced the issue of which horizons to present, because you can have 100 daily forward horizons and present such horizons, but would that be significant to the trading cycle or financial markets, as to present the best possible solution the best adaption between the trading cycle, frequency and financial markets has been incorporated and presented. Doing so allows us to compare and contrast trading strategies that are being implemented in the financial markets by investors and traders.

The following table will illustrate the methods that were used in this paper:

Table 9

Methods		*Description*
1.	*AUTOARIMA*	autoarima function
2.	*AUTOARIMA_SEASDUMMY*	with seasonal dummies as external regressors
3.	*AUTOARIMA_FOURIER*	with Fourier transform as external regressors
4.	*SES*	Simple exponential smoothing
5.	*HOLT WINTERS*	Double exponential smoothing with alpha and beta
6.	*DOUBLE SEASONAL HOLT WINTERS*	Holt-Winters with alpha, beta and gamma
7.	*BATS*	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components
8.	*TBATS*	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality. (*1)
9.	*DSHW*	Double seasonal holt winters
10.	*NAÏVE*	Forecasted by the previous observation
11.	*SNAIVE*	Forecasted by the last observation in previous period/season.
12.	*SINDEX*	Seasonal index forecast
13.	*NNET*	Neural Networks
14.	*TSLM*	Time series linear model, with level and trend as the X vars
15.	*THETAF*	Theta model
16.	*RWF*	Random walk model
17.	*MEANF*	Mean forecast
18.	*STL*	Seasonality-Trend-Level. The time series is broken down to these components, the remainder is forecasted and the STL components are added back to the forecasted value.
19.	*AUTOARIMA_SES*	autoarima function & Simple exponential smoothing
20.	*AUTOARIMA_TBATS*	autoarima function & state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality
21.	*AUTOARIMA_NNET*	autoarima function & Neural Networks
22.	*SES_THETAF*	Simple exponential smoothing & Theta model
23.	*SES_MEAN*	Simple exponential smoothing & Mean forecast Model
24.	*TBATS_THETAF*	state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components with an inclusion of multiple seasonality & Theta model
25.	*NNET_THETAF*	Neural Networks & Theta model

(*1) But in case seasonality is not present, both the models will not have seasonal components. So, they end up optimising on the same model and therefore will have the same forecasts.

3.3.1 Description

The table above presents descriptions of the methods used in this paper, consisting of 22 statistical techniques, two hybrid models including both statistical techniques and artificial intelligence (AI), and one AI method. The Autoarima function was run in the r-studio platform. Furthermore, all of the methods in this paper was run in the r-studio platform. The packages, including the forecast package, are freely available in the r-studio platform and the packages used to run the mentioned methods are available upon request.

Table 10

Classifications of time series

	Companies	Methods	Error Matrices	Logarithm	Horizons	Index
Daily	18	25	6	1	22	FTSE 100
Weekly	18	25	6	1	12	FTSE 100
Monthly	18	25	6	1	18	FTSE 100
Quarterly	18	25	6	1	12	FTSE 100
Yearly	18	25	6	1	4	FTSE 100

3.3.2 Description

The classifications of time series for this paper is shown above, detailing the methodology that was tested in this paper. The companies are consistent for all of the papers, and the horizon is carried on from paper 2 through to paper 4. However, the log return is tested in this paper only.

3.4 Accuracy Testing

Complete optimization is impossible when it comes to forecasting different academics debate what to optimize and how to implement their models. However, there are more widespread theories out there which many scholars do use and agree on. In our model under error testing, the more commonly used errors matrices like ME, MAPE and MPE were used; we also used MSE, RMSE and MAE. Log returns were also determined by the model, with the same capacity as the error matrices calculation. Furthermore, as six different error matrices were tested, the scale of the test itself increases and develops a more determined and detailed conclusion as we can conclude which function is more accurate under the different mentioned circumstances.

The results that are presented are a mere fraction of the results that were produced. The total output will be freely available upon request. Due to the size of the total output, we were unable to showcase all of the results. Furthermore, the vast output that was computed presented us with the dilemma of how to be fair and consistent when presenting the results from our horizon testing. Therefore, in a commanding rule, the paper presents the absolute percentage error (APE) accuracy test.

The error matrices tested, outputted, are presented below (shown in millions)

Table 11

Classifications of Accuracy Test

	Daily	Weekly	Monthly	Quarterly	Yearly	Total
ME	48.78	4.12	1.199	0.221	0.018	54.338
MAE	48.78	4.12	1.199	0.221	0.018	54.338
MAPE	48.78	4.12	1.199	0.221	0.018	54.338
MPE	48.78	4.12	1.199	0.221	0.018	54.338
MSE	48.78	4.12	1.199	0.221	0.018	54.338
RMSE	48.78	4.12	1.199	0.221	0.018	54.338
TOTAL	292.68	24.72	7.194	1.326	0.108	326.028

3.4.1 Description

The table above shows the errors that were produced, calculated and tested. Due to the significant number of errors that were tested, the implemented test was carried out on the HPC Wales supercomputer (cloud). The implementation of the test on this paper would have taken months on a normal everyday computer. At first, the implementation was carried out on a Dell XPS I7 laptop but after a few weeks of running the test was stopped. At this point, the notion of using a supercomputer arose, while there were also some clouds and clusters to consider. Finally, a decision was taken to implement the test on the HPC Wales supercomputer cloud.

The test took 44.5 hours on 45 cores. This was due to the high load of computation being run. At the start of our code being run on the HPC cloud was going to result in a significant difference time wise to run the code. However, after introducing parameters to the test and implementing dynamic arrays, the time was reduced. however there was more effort taken to reduce the time that it would take to reduce the time the test was taken, in effort to do that, the methods were allocated to individual cores and the more computational intensive the function the more the cores that were allocated to that function.

All of the errors mentioned above are available via a flash drive upon request. The forecasting produced according to the actuals of the company's stock price are also available in the same flash drive.

3.5 Results

Table 12

DAILY	APE								
METHOD	HORIZON								
	1	2	3	4	5	10	22	1-10	1-22
AUTO_ARIMA	0.014	0.020	0.024	0.028	0.031	0.045	0.067	0.031	0.045
AUTOARIMA_FOURIER	**0.013** *0.013*	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
AUTOARIMA_NNET	1.106	1.105	1.103	1.103	1.103	1.102	1.102	1.103	1.102
AUTOARIMA_SEASDUMMY	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
AUTOARIMA_SES	1.106	1.106	1.106	1.106	1.107	1.108	1.110	1.107	1.108
AUTOARIMA_TBATS	1.106	1.107	1.107	1.107	1.107	1.109	1.112	1.108	1.109
BATS	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
DSHW_DAILY	0.016	0.022	0.026	0.030	0.033	0.045	0.068	0.033	0.047
HOLT	0.013	0.019	0.023	0.026	0.030	0.042	0.062	0.030	0.043
HOLT_WINTERS	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
MEANF	0.801	0.802	0.803	0.803	0.804	0.806	0.813	0.804	0.807
naïve	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
NNET	0.017	0.025	0.032	0.038	0.044	0.067	0.104	0.045	0.068
NNET_THETAF	1.105	1.103	1.102	1.101	1.100	1.096	1.088	1.100	1.095
RWF	0.013	0.019	0.023	0.027	0.030	0.042	0.062	0.030	0.043
SES	0.013	**0.019** *0.019*	**0.023** *0.023*	**0.026** *0.026*	**0.029** *0.030*	**0.041** *0.042*	**0.059** *0.062*	**0.029** *0.031*	**0.041** *0.045*
SES_MEAN	0.923	0.923	0.922	0.922	0.922	0.921	0.917	0.922	0.920
SES_THETAF	1.105	1.105	1.105	1.104	1.104	1.101	1.096	1.103	1.101
SINDEX	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
SNAIVE	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
STL	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
TBATS	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041
TBATS_THETAF	1.106	1.106	1.105	1.105	1.104	1.102	1.098	1.104	1.102
THETAF	0.013	0.019	0.023	0.026	0.029	0.041	0.060	0.029	0.042
TSLM	0.013	0.019	0.023	0.026	0.029	0.041	0.059	0.029	0.041

3.5.1 Paper Table Structure & Discussion

The first table of our results section shows the outputted results for our tested APE and the table shows specifically the daily output. The winning method from each horizon is presented in bold and the median for the same horizon is presented in italics (this is carried over through all results).

In our daily APE results, the simple exponential smoothing (SES) takes over the accuracy testing from the second horizon on, becoming more accurate even when the horizons were averaged, demonstrating its near monopoly of the daily APE. However, the Autoarima Fourier had its say when it came to the first horizon. Whereas both SES and Autoarima_Fourier both show *0.13* in the first horizon Autoarima Fourier performed better on the 4th decimal.

Table 13

WEEKLY	APE							
METHOD	HORIZON							
	1	2	3	4	6	12	1-6	1-12
AUTO_ARIMA	0.030	0.042	0.051	0.059	0.072	0.104	0.053	0.073
AUTOARIMA_FOURIER	**0.029** *0.026*	**0.041** *0.037*	0.050	0.057	**0.069** *0.072*	**0.099** *0.104*	**0.051** *0.055*	**0.069** *0.075*
AUTOARIMA_NNET	1.117	1.115	1.108	1.104	1.095	1.072	1.107	1.094
AUTOARIMA_SEASDUMMY	0.029	0.041	0.050	0.057	0.069	0.099	0.051	0.069
AUTOARIMA_SES	1.124	1.122	1.120	1.118	1.114	1.104	1.119	1.114
AUTOARIMA_TBATS	1.124	1.123	1.120	1.119	1.116	1.108	1.120	1.116
BATS	0.029	0.041	0.050	0.057	0.070	0.100	0.052	0.070
DSHW	0.043	0.055	0.067	0.073	0.086	0.118	0.067	0.086
HOLT	0.029	0.042	0.051	0.059	0.073	0.106	0.053	0.073
HOLT_WINTERS	0.029	0.041	0.050	0.057	0.069	0.099	0.051	0.069
MEANF_DAILY	0.828	0.831	0.833	0.836	0.841	0.857	0.835	0.842
NAIVE	0.029	0.041	0.050	0.057	0.069	0.099	0.051	0.069
NNET	0.034	0.050	0.062	0.072	0.089	0.132	0.065	0.090
NNET_THETAF	1.119	1.116	1.110	1.106	1.095	1.069	1.108	1.094
RWF	0.029	0.042	0.051	0.059	0.073	0.106	0.053	0.073
SES	0.029	0.041	**0.050** *0.051*	**0.057** *0.059*	0.069	0.099	0.051	0.069
SES_MEAN	0.936	0.935	0.934	0.932	0.930	0.922	0.933	0.929
SES_THETAF	1.126	1.124	1.122	1.119	1.115	1.101	1.120	1.114
SINDEX	0.029	0.041	0.050	0.057	0.069	0.099	0.051	0.069
SNAIVE	0.029	0.041	0.050	0.057	0.069	0.099	0.051	0.069
STL	0.029	0.041	0.050	0.057	0.069	0.099	0.051	0.069
TBATS	0.029	0.041	0.050	0.057	0.070	0.100	0.052	0.070
TBATS_THETAF	1.127	1.124	1.122	1.120	1.117	1.105	1.122	1.116
THETAF	0.029	0.041	0.050	0.058	0.071	0.102	0.052	0.071
TSLM	0.029	0.041	0.050	0.057	0.069	0.099	0.051	0.069

3.5.2 Discussion

In our second results table, the Autoarima_Fourier carries on from the daily frequency to the weekly frequency from performing best on the first horizon only to winning the 1st, 2nd, 6th, 12th, 1st to 6th and 1st to 12th horizons on the weekly data. The Autoarima_Fourier took over the role of being the dominant method from the SES method on the daily data. A monopoly of sorts is also observable from the table, where SES performed better on two occasions on the 3rd and 4th horizon, narrowly beating the Autoarima_Fourier only in the 4th decimal.

Table 14

MONTHLY	APE							
METHOD	HORIZON							
	1	2	3	4	9	18	1-9	1-18
AUTO_ARIMA	0.060	0.086	0.108	0.129	0.231	0.469	0.148	0.256
AUTOARIMA_FOURIER	**0.057** *0.059*	**0.081** *0.086*	0.101	**0.119** *0.129*	**0.199** *0.231*	**0.350** *0.442*	**0.133** *0.150*	**0.207** *0.242*
AUTOARIMA_NNET	1.130	1.104	1.092	1.071	0.995	0.864	1.060	0.990
AUTOARIMA_SEASDUMMY	0.057	0.081	0.101	0.119	0.199	0.350	0.133	0.207
AUTOARIMA_SES	1.129	1.118	1.108	1.094	1.034	0.929	1.082	1.027
AUTOARIMA_TBATS	1.133	1.125	1.119	1.108	1.062	0.979	1.099	1.056
BATS	0.059	0.085	0.106	0.125	0.216	0.432	0.142	0.238
DSHW	0.084	0.115	0.152	0.158	0.249	0.469	0.177	0.283
HOLT	0.061	0.088	0.112	0.134	0.244	0.447	0.156	0.255
HOLT_WINTERS	0.057	0.081	0.101	0.119	0.199	0.350	0.133	0.207
MEANF	0.919	0.929	0.939	0.949	0.996	1.086	0.958	1.001
NAÏVE	0.057	0.081	**0.101** *0.108*	0.119	0.199	0.350	0.133	0.207
NNET	0.068	0.098	0.125	0.149	0.265	0.495	0.168	0.288
NNET_THETAF	1.124	1.095	1.078	1.055	0.961	0.806	1.040	0.956
RWF	0.059	0.087	0.111	0.135	0.246	0.442	0.157	0.254
SES	0.057	0.082	0.101	0.119	0.200	0.350	0.134	0.207
SES_MEAN	0.943	0.935	0.925	0.916	0.868	0.792	0.906	0.864
SES_THETAF	1.123	1.109	1.094	1.078	0.998	0.865	1.062	0.991
SINDEX	0.057	0.081	0.101	0.119	0.199	0.350	0.133	0.207
SNAIVE	0.057	0.081	0.101	0.119	0.199	0.350	0.133	0.207
STL	0.057	0.081	0.101	0.119	0.199	0.350	0.133	0.207
TBATS	0.059	0.085	0.106	0.125	0.216	0.432	0.142	0.238
TBATS_THETAF	1.127	1.116	1.104	1.092	1.026	0.915	1.078	1.019
THETAF	0.058	0.084	0.106	0.127	0.224	0.396	0.145	0.231
TSLM	0.057	0.081	0.101	0.119	0.199	0.350	0.133	0.207

3.5.3 Discussion

Somewhat close results are carried over from the weekly results, This accounts for all horizons, however the only difference was on the 4th horizon, where on the previous weekly table SES was the dominant factor but the Autoarima_Fourier performed better on the monthly 4th horizon.

Table 15

QUARTERLY	APE							
METHOD				**HORIZON**				
	1	2	3	4	8	12	1-6	1-12
AUTO_ARIMA	0.123	0.193	0.271	0.351	0.700	1.001	0.319	0.565
AUTOARIMA_FOURIER	**0.108** *0.121*	**0.164** *0.191*	**0.213** *0.252*	**0.262** *0.321*	**0.539** *0.628*	0.832	**0.240** *0.287*	**0.441** *0.565*
AUTOARIMA_NNET	0.872	0.823	0.784	0.760	0.664	0.617	0.780	0.712
AUTOARIMA_SEASDUMMY	0.108	0.164	0.213	0.262	0.539	0.832	0.240	0.441
AUTOARIMA_SES	0.895	0.853	0.812	0.783	0.666	0.616	0.802	0.723
AUTOARIMA_TBATS	0.894	0.853	0.814	0.786	0.679	0.643	0.805	0.734
BATS	0.121	0.186	0.248	0.321	1.516	51.236	0.321	6.871
DSHW	0.170	0.234	0.305	0.369	0.784	1.177	0.355	0.643
HOLT	0.129	0.209	0.279	0.351	0.730	1.124	0.318	0.596
HOLT_WINTERS	0.108	0.164	0.213	0.262	0.730	0.832	0.240	0.457
MEANF	1.004	1.037	1.070	1.104	1.234	1.356	1.087	1.183
NAÏVE	0.108	0.164	0.213	0.262	0.539	0.832	0.240	0.441
NNET	0.144	0.214	0.275	0.333	0.628	0.920	0.312	0.529
NNET_THETAF	0.860	0.800	0.753	0.719	0.587	0.514	0.745	0.651
RWF	0.116	0.191	0.252	0.315	0.659	1.036	0.283	0.538
SES	0.112	0.168	0.217	0.269	0.561	0.841	0.247	0.456
SES_MEAN	0.823	0.791	0.763	0.741	0.659	0.634	0.755	0.701
SES_THETAF	0.884	0.830	0.781	0.740	0.586	**0.509** *0.832*	0.765	0.660
SINDEX	0.108	0.164	0.213	0.262	0.539	0.832	0.240	0.441
SNAIVE	0.108	0.164	0.213	0.262	0.539	0.832	0.240	0.441
STL	0.108	0.164	0.213	0.262	0.539	0.832	0.240	0.441
TBATS	0.121	0.186	0.248	0.321	1.516	51.236	0.321	6.871
TBATS_THETAF	0.882	0.829	0.783	0.744	0.601	0.538	0.768	0.671
THETAF	0.117	0.183	0.240	0.298	0.617	0.904	0.272	0.498
TSLM	0.108	0.164	0.213	0.262	0.539	0.832	0.240	0.441

3.5.4 Discussion

On the quarterly data, the dominant factor carries over from the previous frequencies (monthly), where on our previous APE data the results show some sort of a pattern reoccurring from the results. However, on the quarterly results the difference was that the SES_THETAF took the place of the naïve where it performed better on the 12th horizon.

Table 16

Yearly	**APE**					
METHOD	**HORIZON**					
	1	2	3	4	1-2	1-4
AUTO_ARIMA	0.666	0.945	1.442	1.603	0.806	1.164
AUTOARIMA_FOURIER	0.273	0.558	0.937	1.294	0.416	0.766
AUTOARIMA_NNET	0.571	0.503	0.503	0.509	0.537	0.521
AUTOARIMA_SEASDUMMY	0.273	0.558	0.937	1.294	0.416	0.766
AUTOARIMA_SES	0.635	0.544	0.532	0.545	0.589	0.564
AUTOARIMA_TBATS	0.820	0.771	0.824	0.914	0.795	0.832
BATS	0.452	0.820	1.201	1.751	0.636	1.056
DSHW	0.273	0.558	0.937	1.294	0.416	0.766
HOLT	0.273	0.558	0.937	1.294	0.416	0.766
HOLT_WINTERS	0.494	0.954	1.388	1.753	0.724	1.147
MEANF	1.249	1.401	1.551	1.684	1.325	1.471
NAIVE	0.273	0.558	0.937	1.294	0.416	0.766
NNET	0.693	1.239	1.864	2.076	0.966	1.468
NNET_THETAF	0.545	0.448	**0.428** *0.937*	**0.423** *1.294*	0.496	0.461
RWF	0.353	0.713	1.175	1.474	0.533	0.929
SES	0.389	0.734	1.058	1.528	0.561	0.927
SES_MEAN	0.764	0.706	0.718	0.752	0.735	0.735
SES_THETAF	0.613	0.493	0.463	0.460	0.553	0.507
SINDEX	0.273	0.558	0.937	1.294	0.416	0.766
SNAIVE	0.273	0.558	0.937	1.294	0.416	0.766
STL	0.273	0.558	0.937	1.294	0.416	0.766
TBATS	0.452	0.820	1.201	1.751	0.636	1.056
TBATS_THETAF	0.796	0.716	0.746	0.814	0.756	0.768
THETAF	**0.238** *0.452*	**0.325** *0.558*	0.436	0.601	**0.281** *0.505*	**0.400** *0.748*
TSLM	0.273	0.558	0.937	1.294	0.416	0.766

3.5.5 Discussion

On the yearly APE test, THETAF and the NNET_THETAF were the dictating methods, however THETAF dominated four out of six horizons, being stronger in horizons 1, 2, 1-2 and 1-4. In this case, the THETAF function had a stronger performance when taking the overall averaged into account.

Mean Squared Tables

Table 17

DAILY	MSE								
METHOD	HORIZON								
	1	2	3	4	5	10	22	1-10	1-22
AUTO_ARIMA	1080	2074	3084	4111	5185	10158	22342	5652	11626
AUTOARIMA_FOURIER	1059	1984	2847	3681	4536	8535	16887	4892	9421
AUTOARIMA_NNET	2125115	2121173	2117727	2116211	2113960	2107334	2092410	2114488	2105914
AUTOARIMA_SEASDUMMY	1059	1984	2847	3681	4536	8535	16887	4892	9421
AUTOARIMA_SES	2126342	2125317	2124356	2124948	2124465	2124296	2122535	2124766	2124024
AUTOARIMA_TBATS	2127643	2127398	2127059	2127794	2127918	2129763	2132568	2128380	2130010
BATS	1061	1985	2851	3687	4546	8552	16936	4900	9438
DSHW	1311	2284	3178	4089	4976	9138	18666	5332	10279
HOLT	1059	1982	2850	3689	4551	8611	17197	4918	9538
HOLT_WINTERS	1059	1984	2847	3681	4536	8535	16887	4892	9421
MEANF	565658	566499	567356	568228	569096	573552	583844	569561	574753
NAIVE	1059	1984	2847	3681	4536	8535	16887	4892	9421
NNET	1595	3033	4485	5397	6639	12494	24709	7270	13853
NNET_THETAF	2123814	2120237	2117092	2114315	2111883	2102122	2080943	2111783	2099975
RWF	1060	1986	2851	3687	4545	8567	17010	4904	9470
SES_MEAN	931951	931614	931279	930951	930626	929053	924999	930479	928498
SES	**1058** *1061*	**1977** *1985*	**2839** *2851*	**3673** *3687*	**4526** *4545*	**8521** *8566*	**16861** 17010	**4882** *4904*	**9405** *9469*
SES_THETAF	2125039	2124379	2123716	2123049	2122385	2119081	2111042	2122058	2118070
SINDEX	1059	1984	2847	3681	4536	8535	16887	4892	9421
SNAIVE	1059	1984	2847	3681	4536	8535	16887	4892	9421
STL	1059	1984	2847	3681	4536	8535	16887	4892	9421
TBATS	1061	1985	2851	3687	4546	8552	16936	4900	9438
TBATS_THETAF	2126342	2126460	2126416	2125891	2125835	2124536	2121045	2125667	2124040
THETAF	1058	1977	2840	3673	4527	8524	16872	4883	9409
TSLM	1059	1984	2847	3681	4536	8535	16887	4892	9421

3.5.6 Discussion

On our first MSE test for the daily frequency presented above, clear accuracy strength is shown is shown by the simple exponential smoothing (SES). In this case the SES function works best with more data points, where the total number of training data for the daily is 4,000 data points and smaller stock price changes then larger changes resulting in higher volatility changes from day to day, which also affects the results of the functions.

Table 18

WEEKLY	**MSE**							
METHOD					**HORIZON**			
	1	2	3	4	6	12	1-6	1-112
AUTO_ARIMA	4616	8716	12635	16308	23048	44159	14164	24733
AUTOARIMA_FOURIER	4503	8556	12343	15871	22293	42137	13773	23801
AUTOARIMA_NNET	2082904	2078665	2061938	2052111	2021362	1953288	2055809	2018011
AUTOARIMA_SEASDUMMY	4503	8556	12343	15871	22293	42137	13773	23801
AUTOARIMA_SES	2117699	2114996	2112456	2110187	2105830	2093665	2111537	2105102
AUTOARIMA_TBATS	2119433	2120134	2121751	2123507	2126950	2137343	2122847	2127922
BATS	4538	8599	12419	15959	22510	42850	13875	24094
DSHW	6228	10984	16837	19860	26682	49578	17334	28614
HOLT	4573	8718	12686	16375	23183	46447	14217	25327
HOLT_WINTERS	4503	8556	12343	15871	22293	42137	13773	23801
MEANF	588211	592521	596954	601434	609894	636631	599112	612223
NAIVE	**4503** *4538*	**8556** *8599*	**12343** *12419*	15871	22293	**42137** *42850*	13773	23801
NNET	8067	15509	22546	28653	38671	65728	24598	39190
NNET_THETAF	2086889	2081847	2064177	2053206	2019976	1944145	2057293	2015810
RWF	4510	8581	12396	15959	22478	42813	13855	24070
SES_MEAN	910272	908550	906885	905290	901813	892278	906059	901111
SES	4527	8568	12351	**15851** *15959*	**22280** *22510*	42163	**13771** *13875*	**23799** *24094*
SES_THETAF	2121724	2118215	2114702	2111240	2104262	2083670	2112981	2102612
SINDEX	4503	8556	12343	15871	22293	42137	13773	23801
SNAIVE	4503	8556	12343	15871	22293	42137	13773	23801
STL	4503	8556	12343	15871	22293	42137	13773	23801
TBATS	4538	8599	12419	15959	22510	42850	13875	24094
TBATS_THETAF	2123462	2123366	2124016	2124578	2125396	2127254	2124305	2125419
THETAF	4526	8568	12352	15852	22287	42183	13773	23809
TSLM	4503	8556	12343	15871	22293	42137	13773	23801

3.5.7 Discussion

In the weekly MSE table, we can clearly witness the naïve model performing very well for three horizons ahead and the 12th horizon. However, in this case it could be argued that the SES method is the best way forward, if looking from a risk-averse view or looking at the long-term performance and overall performance rather than the individual horizon strength for each of the methods. Furthermore, the two models perform comparatively and are not far apart across the board.

Table 19

MONTHLY **MSE**

METHOD	**HORIZON**							
	1	2	3	4	9	18	1-9	1-18
AUTO_ARIMA	15951	30926	48390	67557	161668	329242	85869	174834
AUTOARIMA_FOURIER	**14628** *15758*	28570	**44750** *49450*	**61296** *68644*	137720	272198	75729	145024
AUTOARIMA_NNET	1534809	1511174	1484878	1461382	1370492	1243036	1446971	1370865
AUTOARIMA_SEASDUMMY	14628	28570	44750	61296	137720	272198	75729	145024
AUTOARIMA_SES	1544165	1531045	1511868	1494254	1419328	1298585	1481083	1414895
AUTOARIMA_TBATS	1559777	1561011	1556239	1553111	1550900	1562698	1554509	1554237
BATS	15758	31659	49450	68644	159636	305858	86487	166475
DSHW	25235	45944	68633	91226	190320	431011	110954	218206
HOLT	16330	35327	60640	91831	293374	875773	136446	372751
HOLT_WINTERS	14628	28570	44750	61296	137720	272198	75729	145024
MEANF	567637	585561	604185	621213	719694	911823	642342	731431
NAÏVE	14628	**28570** *31659*	44750	61296	137720	272198	75729	145024
NNET	38932	68691	89022	114142	198878	358160	124118	211207
NNET_THETAF	1524089	1496532	1470900	1446236	1338856	1190731	1426503	1338886
RWF	14679	28774	45224	62166	141264	280947	77138	149415
SES_MEAN	656860	647914	638518	628434	586239	531608	620718	585652
SES	14788	29003	45220	61743	138144	272408	76109	145377
SES_THETAF	1533569	1516507	1497905	1479016	1387275	1244896	1460499	1382390
SINDEX	14628	28570	44750	61296	137720	272198	75729	145024
SNAIVE	14628	28570	44750	61296	137720	272198	75729	145024
STL	14628	28570	44750	61296	137720	272198	75729	145024
TBATS	15758	31659	49450	68644	159636	305858	86487	166475
TBATS_THETAF	1549167	1546323	1541896	1537189	1515294	1494483	1532554	1516543
THETAF	14768	28933	45052	61439	**136167** *159636*	**265066** *305858*	**75360** *86973*	**142695** *170078*
TSLM	14628	28570	44750	61296	137720	272198	75729	145024

3.5.8 Discussion

On the monthly MSE table, results are rather different compared to the daily and weekly tables. In the daily table, one method was very strong on all horizons, while on the weekly table we saw two methods performing rather well. However, on the monthly results table, three methods performed well on most horizons.

Table 20

QUARTERLY **MSE**

METHOD	**HORIZON**							
	1	2	3	4	8	12	1-6	1-12
AUTO_ARIMA	77589	150585	239579	326865	551615	778992	271532	454996
AUTOARIMA_FOURIER	54706	108625	153372	197683	344889	547139	169837	293670
AUTOARIMA_NNET	1318357	1353830	1394603	1442666	1584396	1652056	1420997	1515169
AUTOARIMA_SEASDUMMY	54706	108625	153372	197683	344889	547139	169837	293670
AUTOARIMA_SES	1259052	1220447	1183941	1154368	1043532	1017900	1172915	1101697
AUTOARIMA_TBATS	1270328	1260494	1253146	1252995	1263195	1377309	1257183	1279833
BATS	61860	140296	336070	1401168	2429894209	4.66E+12	10921227	4.57E+11
DSHW	85657	146859	187281	223531	457496	712351	223975	380397
HOLT	74337	165881	279867	435827	1109181	2102925	375644	919205
HOLT_WINTERS	54706	108625	153372	197683	344889	547139	169837	293670
MEANF	556830	608420	664445	721686	986990	1293118	697935	897328
NAÏVE	**54706** *61860*	**108625** *140296*	153372	197683	**344889** *449253*	547139	169837	293670
NNET	103771	166752	267156	373902	449253	644036	290361	408289
NNET_THETAF	1296778	1301844	1308821	1317800	1284517	1198296	1310784	1280550
RWF	55371	111180	157942	204393	347431	552422	174521	298002
SES_MEAN	493740	461135	433605	410365	345441	**365921** *644036*	427209	387320
SES	57763	110901	155407	199793	346779	549839	171945	295548
SES_THETAF	1239656	1176545	1116319	1060490	845353	733878	1091210	945164
SINDEX	54706	108625	153372	197683	344889	547139	169837	293670
SNAIVE	54706	108625	153372	197683	344889	547139	169837	293670
STL	54706	108625	153372	197683	344889	547139	169837	293670
TBATS	61860	140296	336070	1401168	2429894209	4.66E+12	10921227	4.57E+11
TBATS_THETAF	1251153	1216693	1185075	1157850	1056516	1070901	1174133	1115854
THETAF	57707	110107	**153024** *239579*	**194890** *326865*	325587	519870	**167002** *271532*	**281091** *387320*
TSLM	54706	108625	153372	197683	344889	547139	169837	293670

3.5.9 Discussion

The quarterly MSE results showed the naïve model and the THETAF model to both be dominant. The naïve model proved to be very effective on our quarterly testing for two quarters ahead only. The most consistent model under these circumstances was the THETAF model.

Table 21

YEARLY **MSE**

METHOD	HORIZON					
	1	2	3	4	1-2	1-4
AUTO_ARIMA	292755	418829	600342	942481	355792	563602
AUTOARIMA_FOURIER	251580	369859	591059	933826	310719	536581
AUTOARIMA_NNET	680991	653202	717367	778624	667096	707546
AUTOARIMA_SEASDUMMY	251580	369859	591059	933826	310719	536581
AUTOARIMA_SES	648095	534743	566722	612801	591419	590590
AUTOARIMA_TBATS	1011167	1136177	1560061	2160994	1073672	1467100
BATS	315201	663389	1494019	2928359	489295	1350242
DSHW	**251580** *276596*	369859	591059	933826	310719	536581
HOLT	276596	469424	882639	1524370	373010	788257
HOLT_WINTERS	251580	369859	591059	933826	310719	536581
MEANF	723607	970829	1293465	1636828	847218	1156182
NAÏVE	251580	369859	591059	933826	310719	536581
NNET	1597955	1282929	1487934	1919653	1440442	1572118
NNET_THETAF	672490	601051	645624	698280	636770	654361
RWF	266210	387370	621666	1024460	326790	574927
SES_MEAN	341458	318273	**433568** *591059*	**573922** *933826*	329866	**416805** *563602*
SES	261950	373754	605933	950121	317852	547940
SES_THETAF	651682	518523	553300	609340	585102	583211
SINDEX	251580	369859	591059	933826	310719	536581
SNAIVE	251580	369859	591059	933826	310719	536581
STL	251580	369859	591059	933826	310719	536581
TBATS	315201	663389	1494019	2928359	489295	1350242
TBATS_THETAF	1028782	1154158	1613559	2268105	1091470	1516151
THETAF	258790	**313205** *387370*	579831	971306	**285997** *329866*	530783
TSLM	251580	369859	591059	933826	310719	536581

3.5.10 Discussion

The double seasonal Holt-Winters (DSHW) was the most accurate one year ahead, thus, due to seasonality, this is effective for the first year only. However, it lost strength after the first year, as the THETAF model was the most accurate for both the second horizon and when we calculated the average for the first and second horizons. SES_MEAN was the most accurate after two years ahead being the most accurate for three and four years ahead, thus also becoming the most accurate for the averaged results between years 3 and 4.

3.6 Conclusion

3.6.1 Analysis

After a thorough observation of the results from all of the frequencies and horizons, we can definitely see a pattern emerging. At first we were able to recognise this emerging pattern and observed the same pattern developing as we moved along with the testing. The pattern here is that as the accuracy testing moved further along the horizon spectrum, the accuracy tests became less and less accurate. In the different variations of error matrices that were tested, we specifically presented the APE for the five frequencies that were tested.

Also shown above is the MSE. For the shorter frequencies, the naïve model and the simple exponential smoothing model showed the best accuracy from a general perspective but not from the individual horizon accuracy; the individual horizon is shown under each table. Furthermore, when the frequencies were becoming longer, the testing became more complex and different methods would be more accurate on one horizon but not on the next horizon. And different models were becoming more apparent on their strong accuracy that did not show strength on the lesser frequencies. On the monthly, quarterly and yearly frequencies, the THETAF model was identified as the best way forward and the naïve model was a close second.

3.6.2 Evaluation

From the APE testing, the following is apparent:

Autoarima, simple exponential smoothing (SES) and THETAF performed better than other algorithms. However, frequency does have a significant effect on which models work better and where. Similarly, the horizon also has a noteworthy effect on the models tested, where Autoarima in most cases beat all other models on the first horizon.

The models are competitive. There are somewhat diverse results as we move along the horizon. SES shows complete dominance from the daily APE table where from horizon 2 onwards it performs best in terms of accuracy testing . However, Autoarima shows overall dominance.

Despite Autoarima's overall dominance there are horizons where it showed weakness, especially the third horizon and, in some instances, on the second. It also showed greater strength on medium frequencies and progressively lost strength as the frequencies became greater with less and less training data. Autoarima lost out to other models, which is observable from the yearly accuracy test where it did not win on a single horizon. In the yearly results, other models emerged, for instance THETAF _YEARLY and NNET_THETAF_YEARLY performed significantly better than Autoarima.

While Autoarima showed great dominance overall, SES showed great dominance on the daily accuracy exclusively, In this test, the latter won on all of the horizons except the first. Similar dominance was only shown by Autoarima in the quarterly and monthly APE results. Thus, we can say with great confidence that the best trading model for daily testing is SES and for monthly and quarterly testing it is Autoarima. From these observations we can easily say that, with more training data, SES is the best method, however with medium training data Autoarima is a better option.

Despite constraints in place preventing the collection of larger amounts of data, our test shows that THETAF and NNET_THETAF are the most suitable approaches.

To conclude, for each frequency there is a specific model that prevails. We now take our test to the next level; in the next chapter we introduce neural networks with volatility variables. The aim here is to determine the perfect model, and to test if more complexity is better than less in models. It has been suggested that the simplest models prevail and produce greater accuracy than more complicated models.

3.6.3 Future Research

Future work after testing the horizons will include testing uncertainty and volatility within the data, to further our research and to improve our accuracy and introduce AI to our methodology to widen our method selection. By doing so, we would be able to compare and contrast methods, and test if volatility within the share price can or cannot improve our accuracy testing. The same errors will be chosen for each round of testing to be consistent in our research.

3.7 References

Baumeister, C., & Kilian, L. (2015). Forecasting the real price of oil in a changing world: a forecast combination approach. Journal of Business & Economic Statistics, 33(3), 338–351.

Blaskowitz, O., & Herwartz, H. (2011). On economic evaluation of directional forecasts. International Journal of Forecasting, 27, 1058–1065.

Busemeyer, J., Wang, Z., Townsend, J., & Eidels, A. (2015). 2015. The Oxford Handbook of Computational and Mathematical Psychology. Oxford, NY: Oxford University Press.

Capistra'n, C. (2006). On comparing multi-horizon forecasts. Economics Letters, 93, 176-181.

Capistrán, C., Constandse, C., & Ramos-Francia, M. (2010). Multi-horizon inflation forecasts using disaggregated data. Economic Modelling, 27, 666–677.

Cenesizoglu, T., Ribeiro, F., & Reeves, J. (2017). Beta forecasting at long horizons. International Journal of Forecasting, 33, 936–957.

Chen, C. W., Wang, Z., Sriboonchitta, S., & Lee, S. (2017). Pair trading based on quantile forecasting of smooth transition GARCH models. North American Journal of Economics and Finance, 39, 38–55.

Dawande, M., Gavirneni, S., Naranpanawe, S., & Sethi, S. (2009). Discrete forecast horizons for two-product variants of the dynamic lot-size problem. International Journal of Production Economics, 120, 430–436.

Degiannakis, S., Filis, G., & Hassani, H. (2018). Forecasting global stock market implied volatility indices. Journal of Empirical Finance, 46, 111–129.

Diagnea , M., Mathieu , D., Lauret , P., Boland , J., & Schmutz , N. (2013). Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews, 27, 65-76.

du Jardin, P., & Séverin, E. (2011). Predicting corporate bankruptcy using a self-organizing map: An empirical study to improve the forecasting horizon of a financial failure model. Decision Support Systems, 51, 701–711.

Ince, O., & Molodtsova, T. (2017). Rationality and forecasting accuracy of exchange rate expectations: Evidence from survey-based forecasts. Journal of International Financial Markets, Institutions & Money, 47, 131–151.

Joseph, K., Wintoki, M., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. International Journal of Forecasting, 27(4), 1116–1127.

Knüppel, M. (2018). Forecast-error-based estimation of forecast uncertainty when the horizon is increased. International Journal of Forecasting, 34, 105–116.

Mohammadipour, M., & Boylan, J. (2012). Forecast horizon aggregation in integer autoregressive moving average (INARMA) models. Omega, 40, 703-712.

Moshiri, S., & Cameron, N. (2000). Neural Network Versus Econometric Models in Forecasting In¯ation. Journal of Forecasting , 19, 201-217.

Reikard, G. (2015). Forecasting geomagnetic activity at monthly and annual horizons: Time series models. Journal of Atmospheric and Solar-Terrestrial Physics, 133, 111–120.

Reikard, G., Haupt, S. E., & Jensen, T. (2017). Forecasting ground-level irradiance over short horizons: Time series, meteorological, and time-varying parameter models. Renewable Energy, 112, 474-485.

Reindl, T., Walsh, W., Yanqin, Z., & Bieri, M. (2017). Energy meteorology for accurate forecasting of PV power output on different time horizons. Energy Procedia, 130, 130-138.

Sanders, D., Manfredo, M., & Boris, K. (2009). Evaluating information in multiple horizon forecasts: The DOE's energy price forecasts. Energy Economics, 31, 189–196.

Simionescu, M. (2015). The Improvement of unemployment rate predictions accuracy. Prague Economic Papers, 24(3), 274-286.

Snudden, S. (2018). Targeted growth rates for long-horizon crude oil price forecasts. International Journal of Forecasting, 34, 1–16.

Solanki, S., & Krivova, N. (2011). Analyzing solar cycles. Science, 334, 916-917.

3.8 Appendices

3.8.1 Forecast Accuracy Measures

The forecasting accuracy measures are tested and developed in accordance with the following structures:

Error Metrics Tested

Equation 10

Mean Error (ME)

$$ME = \frac{1}{N} \sum_{i=1}^{N} (f_i - k_i)$$

Where:

N = # Forecast / Actuals

F = Forecast

K = Actuals

Equation 11

Mean Absolute Error (MAE)

$$= \text{mean}\ (|e_i|)$$

Mean Squared Error (MSE)

$$mean\ (e_i^2)$$

Equation 12

Root Mean Squared Error (RMSE)

$$= \sqrt{mean\ (e_i^2)}$$

Equation 13

Mean Percentage Error (MPE)

$$MPE = \frac{1}{N} \sum_{i=1}^{N} (f_i - k_i) \text{ x } 100$$

Equation 14

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} (|f_i| - |k_i|) \text{ x } 100$$

The aforementioned accuracy tests follow the structure built by Robert Hyndman.

CHAPTER 4

FORECASTING FINANCIAL MARKETS WITH PREDICTIVE ANALYTICS: THE IMPACT OF UNCERTAINTY; IN THE FORM OF THE RANGE OF INTRA-PERIOD VALUES

4.1 Abstract

This paper attempts to understand the volatility of movement in share price and whether a model could be developed in an attempt to optimize forecasting share price. Thus, we incorporated high and low share prices into the neural network in an attempt to produce stronger results. The closing share price was also fed into the network as a benchmark with the other benchmark being the winning method or methods from paper 2. This enabled us to compare and contrast whether the complexity of the uncertain data movement can enlighten the test of how the price really moves from horizon to horizon. Furthermore, introducing the uncertainty would enable us to better understand if higher volatility from one horizon can affect the accuracy of the next horizon or the one after that and so on. Does higher volatilely mean more or less forecasting accuracy? That is the main question we endeavour to answer here.

4.2 Introduction

Variables inputted into the neural network can have significant effects on the network and they can greatly influence the output thereof. Numerous studies have investigated this phenomenon with intriguing results. According to Chen, Zhang, Chen & Li (2008), the choice of input variables is a fundamental and critical consideration in identifying the optimal functioning of statistical models, especially those concerned with forecasting. Proper selection and training of the input variables can thus help in increasing the forecasting accuracy of the outcomes under investigation. This review will explore studies that have investigated the input-output relationship in order to ascertain the manner in which input variables affect the network and subsequently influence the output variables.

Uncertainty and volatility are indispensable elements in the field of forecasting. When taken together with the range of data used, they have great potential to influence the forecasting accuracy and the outputs of the forecasts. Numerous empirical and theoretical studies have aspired to look into these three domains and the underlying mechanisms influencing their predictive accuracy. From the literature, it is apparent that some common sources of uncertainty include model uncertainty, parameter uncertainty, risk uncertainty, aggregate uncertainty, and data uncertainty, all of which tend to have a strong link to volatility. The most commonly used and the most important models in studying volatility, as evidenced by the existing literature, are GARCH-type models, with their utility resting mainly on the premise that they can describe time-series fluctuations in volatility more effectively (Petropoulos, Hyndman & Bergmeir, 2018). It is also evident from the consulted literature that global economic policy uncertainties can have significant spillovers that affect volatility at the local level directly. This comprehensive review of literature explores empirical studies that have endeavoured to examine the influence of uncertainty and volatility on forecasting accuracy in addition to highlighting the impact of the different ranges of data used on forecast outputs.

4.3 Literature

A wide array of research exists with regard to the concepts of volatility and uncertainty in the field of forecasting. The studies investigating these important topics, as well as the different ranges of data used in influencing forecasting outputs, have found mixed results. Bonaccolto, Caporin & Gupta (2018) conducted an empirical study to examine the dynamic impact that uncertainty has in causing, as well as forecasting, crude oil returns and risk. The methodological approach utilised by the researchers involved studying causality relationships in quantiles via a nonparametric testing technique and estimating the conditional density of crude oil returns and volatility based upon a collection of quantiles forecasts. They then evaluated the out-of-sample performance using a combination of other suitable tests. After performing a dynamic analysis, Bonaccolto et al. (2018) established that uncertainty indexes are not always significant in causing or forecasting crude oil movements. Colombo (2013) seemed to disagree with the findings of Bonaccolto et al. (2018), arguing that uncertainty indexes tend to be relevant especially during times of market distress, as the role of oil risk becomes the predominant interest. To this researcher, economic policy uncertainty, for example during times of recession, can have adverse effects on business cycles and, in turn, can affect oil price movements.

Consistent with the research conducted by Bonaccolto et al. (2018), Beckmann & Czudaj (2017) similarly conducted an investigation to dissect the impact of uncertainty of forecasting, albeit focusing their study on professional exchange rate forecasts. The scholars relied on survey data furnished by FX4casts to analyse the role that uncertainty plays in both forecast errors and exchange rate professional among professional forecasters for four major currencies, namely the British pound, the Euro, the Japanese yen and the Canadian dollar, all benchmarked against the US dollar. To account for the different uncertainty dimensions, Beckmann & Czudaj (2017) considered economic policy, financial uncertainty, and macroeconomic uncertainty. Based on a Bayesian VAR approach, the researchers established that uncertainty effects on professionals' forecast errors were more significant when compared to the adjustments of their exchange rate expectations. The researchers also found that even though financial and microeconomic uncertainties were mostly even, their effects varied across currencies, a sentiment also echoed by Dick, MacDonald & Menkhoff (2015). For example, whereas financial uncertainty had a greater impact on forecast accuracy and forecast errors of the Japanese yen, macroeconomic uncertainty on the contrary had greater effects on forecast errors of the British pound, particularly over long-forecast horizons.

Nonejada (2017) also concentrated his study on macroeconomic and financial predictors of aggregate stock market volatility in a comprehensive Bayesian model-averaging framework. The specific objective of this scholar was to ascertain the models that give the best forecasts, to highlight when they perform better, and to articulate the reasons behind their superior performance. Candidate models included time-varying and constant-coefficient autoregressive models based on the logarithm of monthly volatility supplemented with exogenous predictors capturing leverage, risk premia, proxies for credit risk, and bond rates. Hence, the ranges of data used by Nonejada (2017) simultaneously addressed model instability and parameter instability that unavoidably tend to affect volatility predictions. After applying the model to monthly S&P 500 volatility from 1926 to 2010, the results suggested that the Bayesian model averaging with time-varying regression coefficients offered modest improvements in point forecasts and very competitive density compared to rival approaches. Akin to the inquiry carried out by Beckmann & Czudaj (2017), Nonejada (2017) also revealed that past volatility and variables that capture proxy for credit risk and time-varying risk premia are important predictors of stock market volatility, particularly over long-term horizons.

From a different dimension, Degiannakis (2018) endeavoured to investigate whether a mixture of prediction models could provide more accurate volatility forecasts compared to the forecasts obtained using a single prediction model. To achieve this goal, the researcher estimated heterogeneous and long-memory autoregressive models under symmetric as well as asymmetric distribution of major indices in the European Union stock market and the Euro exchange rates. Just like Prokopczuk, Symeonidis & Wese-Simen (2015), Degiannakis (2018) acknowledged that ultra-high frequency financial data are increasingly finding utility in estimating and predicting volatility more accurately. In his study, Degiannakis (2018) found that the combined models were more effective in causing and predicting volatility. Specifically, for the forecasting horizon of one week, the heterogeneous autoregressive model proved to be statistically superior to the long-memory autoregressive model. The combination of realized volatility (RV) forecasts increased the predictive accuracy for the two-week forecasting horizon, with volatility asymmetry being an important determinant for the accuracy of the two-week forecasting accuracy (Degiannakis, 2018).

Many other researchers have undertaken comparative studies to evaluate the determinants that are most effective in forecasting market volatility. Wei, Liu, Lai & Hu (2017) conducted an investigation to verify which determinant had the higher predictive power in forecasting crude oil volatility among speculation, fundamental, or uncertainty determinants. These scholars used the dynamic average combination method and a new

GARCH-based model grounded on mixed data sampling regression (GARCH-MIDAS). Wei et al. (2017) and Conrad & Loch (2015) both contended that GARCH-MIDAS is advantageous since it allows for the incorporation of macroeconomic fundamentals in low frequencies to forecast daily asset volatility. Wei et al. (2017) integrated national economic policy uncertainty (EPU) indices and global economic policy uncertainty (GEPU) indices with traditional determinants including global oil demand, supply, and speculation. To evaluate the performance of the GARCH-MIDAS model, the academicians relied on an advanced model confidence set (MCS) test and two loss function tests, which were mean absolute forecast error (MAFE) and MSFE.

According to Wei et al. (2017) and Luo, Sattayatham & Chatpatanasiri (2017), the rationale behind using the MCS test stems from the fact that the loss functions are not able to distinguish whether the loss differences of models such as GARCH-MIDAS are statistically significant. Luo et al. (2017) additionally identified other advantages of MCS tests compared to conventional tests. The researchers posited that, unlike conventional tests such as MAPE, MAE, and RMSE, the MCS test does not necessarily have to specify a benchmark model, which makes it extremely useful in application. Moreover, the test allows for the possibility of more than one “best” model. After performing their analysis on the GARCH-MIDAS model, Wei et al. (2017) established that GEPU indices and the US EPU index had superior predictive power for daily oil volatility, particularly for West Texas Intermediate (WTI) crude oil. The finding underscored the importance of EPU indices, affirming that these indices have higher predictive accuracy across different time horizons compared to speculation or oil demand-supply fundamentals, a sentiment corroborated by Yin & Zhou (2016).

It is clear from the review of Wei et al. (2017) that global economic policy has a strong link to volatility in different sectors of the national economy. In a bid to substantiate this claim, Yu, Fang & Sun (2018) likewise designed a study to scrutinize the impact of GEPU on market volatility, focusing exclusively on the effects in the Chinese stock market. Mirroring the methodological approach used by Wei et al. (2017), Yu et al. (2018) also utilised the GARCH-MIDAS model to evaluate the impact of the monthly GEPU index on the SSE Composite Index (daily). Empirical results from the study suggested that GEPU had a significant and positive influence on the volatility of the SSE Composite Index, prompting Yu et al. (2018) to conclude that the results were a reflection of Chinese stock markets’ successful and gradual integration into the global economy. Furthermore, just like the revelations from the study conducted by Wei et al. (2017), the forecasts generated from the GARCH–MIDAS model with GEPU and the MCS test in the study by Yu et al. (2018) produced substantially smaller errors compared to the GARCH-MIDAS with RV measured by RMSE, RMAD, RMSD, RMAE loss functions, and a DM test.

Liow, Liao & Huang (2018) in their empirical research attempted to examine the volatility spillovers in bond, securitized real estate, currency, and stock as well as economic policy uncertainty spillover across seven different countries: the US, Canada, the UK, Germany, France, Japan, and China. The researchers made an interesting observation by asserting that, during times of global financial crises, international spillovers of financial market risk are important to the degree that the co-movements of asset/market volatilities as well as the risk of contagion increase significantly, thereby exerting considerable financial stress on the recovery and growth of economy overall. Knüppel (2018) and Harvey, Leybourne & Whitehouse (2017) agreed with this sentiment, positing that, in responding to such a crisis, central banks and financial authorities tend to construct several financial stress indices to assess and monitor the current state of financial system stress, after which they can combine the indices into aggregate stress indices. In the seven examined countries, Liow et al. (2018) found that spillovers were extremely important, accounting for approximately 72% of the dynamics of financial market stress and approximately 50% of the economic policy uncertainty. Not surprisingly, the researchers also found that in the multi-country context, policy uncertainty spillovers led directly to financial market stress spillovers.

Notably, existing research suggests that uncertainties can be good or bad. Gong & Lin (2017) initiated a study to explore the impacts of the types of uncertainties mentioned above on the volatility of crude oil prices. As the researchers vividly pointed out, good uncertainty is the volatility associated with positive innovations in asset prices whereas bad uncertainty is volatility associated with negative innovations in asset prices. Gong & Lin (2017), in this empirical study, developed and deployed new heterogeneous autoregressive (HAR) type models for forecasting the good as well as the bad uncertainties of crude oil prices. Using in-sample and out-of-sample analyses, they investigated the effects of these uncertainties on daily negative and positive signed jump variations, with the in-sample results indicating that good and bad uncertainties both have long memory property. Furthermore, Gong & Lin (2017) found that the predictability of long-term bad and good uncertainties tended to be stronger compared to the predictability of mid- and short-term bad and good uncertainties. An inquiry by Harvey et al. (2017) explicity lended support to this proposition, with the researchers pinpointing that good uncertainty is usually beneficial to oil producers and bullish traders but can be detrimental to oil consumers and short-sellers. On the contrary, bad uncertainty benefits oil consumers and short sellers, but can harm oil producers and bullish traders.

A number of studies have also endeavoured to explore strategies for eliminating uncertainty and volatility in order to augment forecasting accuracy. Conrad & Loch (2015), for instance, proposed a new measure of anticipated variance risk premium based on a forecast of conditional variance modelled on a GARCH-MIDAS algorithm with the aim of isolating fundamental uncertainty. In congruence with the findings of Bollerslev, Marrone, Xu & Zhou (2014) and Liu, Guo & Qiao (2015), Conrad & Loch (2015) affirmed that variance risk premium has vast potential for predicting aggregate stock market reports, particularly in the medium-term. The researchers rationalized the predictive ability of variance risk premium according to its close relationship with aggregate risk aversion and economic uncertainty. In a study by Conrad & Loch (2015), which utilised monthly US macroeconomic data from 1970 to 2011 and daily continuously compounded returns from S&P 500, the scholars recorded some intriguing results. They found that their new variance risk premium model was able to isolate uncertainty and had a higher predictive power compared to conventional measures.

Clements & Liao (2017) took a different approach in their research that focused on identifying alternatives to improve forecast gains and to isolate uncertainty when dealing with stock index returns. Their study considered how forecasters could use index-level jumps and cojumps across index constituents to forecast the variance of index-level returns. To accomplish their goals, Clements & Liao (2017) used an array of jump and cojump detection techniques including the sequential BNS test, the LM test, and the ABD test. From conducting an in-depth analysis, the researchers arrived at the conclusion that incorporating the estimated jump intensity from a designated point process model helped in isolating volatility and led to significant forecast accuracy gains. Likewise, in the same light as Lahaye, Laurent & Neely (2011), Clements & Liao (2017) also found cojumps across underlying constituent stocks to be useful determinants for forecasting index-level returns, with forecast performance improvements being particularly strong on trading days when jumps and cojumps occur.

Unlike Conrad & Loch (2015) and Clements & Liao (2017), Prak & Teunter (2018) placed emphasis on addressing financial uncertainty in inventory models. The researchers provided valuable background information by asserting that inventory decisions in practice heavily depended on demand forecasts, but the existing literature seemed to assume that decision makers know the demand distributions. Hillier & Lieberman (2014) shared a similar point of view, reiterating that this misguided assumption implies that decision makers and inventory forecasters substitute estimates directly for unknown parameters, thereby leading to stock-outs, insufficient safety stocks, high costs, and limited service. In their investigation, Prak & Teunter (2018) proposed a framework to address this estimation uncertainty, a framework that is applicable to any inventory model, parameter estimator, and demand distribution. To develop the framework, the researchers modelled estimation errors and obtained a predictive lead-time demand distribution, after which they substituted it into the inventory model. After the deployment of the model, this study by Prak & Teunter (2018) revealed that when basing their estimates on 10 observations, the relative savings were typically between 10% and 30% for mean-stationary demand. The savings became even larger in instances when the researchers based their estimates on fewer observations. Hillier & Lieberman (2014) acknowledged that savings can become larger when the lead-time is longer and when backorders are costlier.

Other researchers have relied on survey data to explain the relationship between uncertainty, volatility, and forecasting accuracy. For instance, Kruger & Nolte (2016) utilised a cross-section of economic survey data to predict the distribution of United States macro variables in real time. The specific aim here was to generalize the existing literature that uses disagreement to predict uncertainty. Results from this study demonstrated that even though cross-sectional information garnered from economic survey forecasts could be helpful for distribution forecasting, the information nonetheless needed remodelling in a statistically efficient manner in order to avoid overfitting. In practice, Kruger & Nolte (2016) were adamant that a simple one-parameter model exploiting time variation inherent in the cross-section of survey point forecasts could yield better performance and mitigate the overfitting problem. Driver, Trapani & Urga (2013), from another angle, conducted an exploration on the use of cross-sectional measures of forecasting uncertainty, focusing directly on the use of these measures among private forecasters in addition to assessing their impact on measurement and their use in forecasting uncertainty. To assess the impact of cross-dependence and other distributional properties, Driver et al. (2013) performed a Monte Carlo exercise. Here, analysis and validation revealed that cross-sectional measures including disagreements among forecasters could play a role in improving predictive ability.

Conclusively, this extensive literature review has demonstrated that indeed a direct relationship does exist between uncertainty and volatility, with the range of data used serving mainly as the mediator in the predictive ability of the relationship between uncertainty and volatility. It has also become apparent that GARCH-type models are the most commonly used models for examining volatility and its influence on forecasting. This is because GARCH-type models, most notably GARCH-MIDAS, do not necessarily have to specify a benchmark model and they can provide room for more than one “best model.” Other emerging themes from the literature review have included the effects of global economic policy uncertainties on volatility at the national level. Results from such studies have demonstrated that economic policy uncertainties can affect different markets in the local sectors, and that the more integrated a country is with the global financial systems, the better the outcomes are for its markets. For example, the consulted literature indicated that volatility in the Chinese stock market has responded positively based on global economic uncertainty indices, a testimony to its seamless integration into the global economy.

From comprehensive research, it is apparent that a number of empirical and theoretical studies have endeavoured to examine the manner in which input variables affect the outputs of interest. In a well-designed study by Chandar, Sumathi & Sivanandam (2015), the researchers sought to explore the utility of neural networks in forecasting foreign currency exchange. The inputs utilised by the researchers were 820 historical exchange rate data of four currencies, namely the British pound, the US dollar, the Japanese yen, and the Euro, compared against the Indian rupee. After the normalization process, the researchers divided the input dataset into a training set that consisted of 70-80% of the entire data. The linear transfer function in the output layer served the purpose of approximating the trend of the exchange rate.

After simulation, output results by Chandar et al. (2015) indicated that the Levenberg-Marquardt based model was able to outperform other algorithms including resilient back-propagation, variable learning rate back-propagation, batch gradient descent, and batch gradient descent with momentum, in forecasting the accuracy of the four currencies vis-à-vis the Indian rupee. Pacelli, Bevilacqua & Azzollini (2011) similarly conducted a study to ascertain the efficacy of an adaptive neural network (ADNN) in the forecasting of three-day EUR/USD foreign exchange rates. As inputs, the researchers utilised market data and macroeconomic variables, assuming the conditionality of the EUR/USD exchange rate. The output variable in this study was pricing, while the frequency of data collection of output and input variables was daily. For each of the input variables, the researchers calculated historical memory according to a polynomial interpolation with coefficient R2 equal to 0.98 for 90% of cases.

Just like Chaudhuri & Ghosh (2016), Pacelli et al. (2011) contended that input data for forecasting models are usually explanatory in terms of the phenomenon under investigation and they greatly influence the output for the model, which in their case was pricing. Both sets of researchers also seem to support the view that the input variables enable forecasting models to learn the phenomenon of interest using historical data. In a study by Pacelli et al. (2011), it was apparent that the input variables and their training algorithms greatly influenced the pricing of the EUR/USD exchange rates, prompting the researchers to conclude that the inputs' adaptive metrics can solve problems of trend determination and amplitude changing, thereby avoiding the over-fitting of networks.

Moghaddam, Moghaddam & Esfandyari (2016) similarly endeavoured to establish the utility of an input selection in an ANN in forecasting the daily NASDAQ stock exchange rate. As input variables, the researchers used the day of the week and short-term historical prices. To develop their robust feedforward ANN trained by the back-propagation algorithm, Moghaddam et al. (2016) used daily NASDAQ stock exchange rates for the period spanning from January to June 2015. The researchers were able to develop and validate networks for NASDAQ index prediction for two input datasets (nine prior days and four prior days). The model outputs showed that there were no distinct differences between the prediction abilities of the nine and four prior working days as input parameters.

In another comparable study, Fadlalla & Amani (2014) endeavoured to predict Qatar Exchange (QE) indexes' next trading day closing price using historical data from January 2010 to December 2012. The input variables in this inquiry consisted of 10 technical market indicators including SMAVG, WMAVG, momentum, Stochastic D%, Stochastic K%, RSI, MACD, WLPR, CCI, and ADO. Experimental results revealed that ANNs based on the aforementioned input variables proved to be an effective modelling technique for forecasting QE indexes with high predictive accuracy, outperforming the well-established

ARIMA models (Fadlalla & Amani, 2014). An ANN using 10 technical inputs revealed that simple and weighted moving averages were the most important technical input indicators for forecasting the QE indices whereas the accumulation/distribution oscillator was the least important in predicting the index.

From a different dimension, Antić, Milovanović, Perić, Nikolić & Milojković (2014) endeavoured to explore neural network input parameter selection and pre-processing. The purpose of the network proposed by the researchers was to forecast foreign exchange rates through using AI. In commencing with their investigation, the scholars formed two datasets for two different economic systems. They then represented each system according to six categories comprising 70 economic parameters used in the analysis. They used a principal component analysis (PCA) method to perform reduction of the parameters within each category. They then used the newly formed parameters to create input networks of the multilayer feedforward neural network trained by batch training techniques (Antić et al., 2014).

After performing simulation, Antic et al. (2014) established that their input data preparation method was an effective way of pre-processing neural network data and improving the predictive accuracy of the AI model. Abdi & Williams (2010) also used PCA and batch training technique for input selection, with their study similarly indicating the effectiveness of this technique in improving forecasting accuracy and ensuring the attainment of the desired output outcomes. Akin to Antic et al. (2014), Abdi & Williams (2010) reiterated that PCA finds meaningful applications in a variety of input selection processes including processing large amounts of data, statistical data analysis, signal processing, face recognition, motion analysis, dimension reduction, dataset identification, and data clustering.

Rasouli, Tabesh & Etminani (2016) similarly conducted a study to investigate input variable selection technique that can optimize output outcomes albeit in a hospital setting using an ANN. The input data in this inquiry included a retrospective number of monthly inpatient flows from 2004 to 2015 for four hospitals. The post-sample analysis performed by the researchers revealed that the ANN model utilising selected input variables based upon the partial autocorrelation function (PACF) of time-series data offered significant improvements in forecasting monthly inpatient flows for the first three months of 2016. The ANN model far outperformed the benchmark model, which was neural network autoregression, with MAPE ranging from 2.91 to 6.67% for the ANN model utilising selected input variables based upon the PACF, prompting the scholars to conclude that this form of input variable selection has great accuracy in forecasting hospital inpatient visits (Rasouli et al., 2016).

Wanous, Boussabaine & Lewis (2013), from another perspective, sought to develop and test a novel bid/no bid model using the ANN technique to test its predictive power among contractors operating in Syria. In developing the model, the researchers used a back-propagation network that consisted of an input buffer with 18 input nodes, one output node, and two hidden layers. The researchers used 157 real-life bidding situations for training. After testing the model on 20 other new projects, the findings from this study suggested that the developed model was a powerful tool for modelling and forecasting the bid/no bid decision process, as it predicted the phenomenon of interest with a 90% accuracy rate. Wanous et al. (2013) concluded that, with the right input variables, their model had the potential of offering an easy-to-use and simple tool to aid contractors to improve consistency in the bid/no bid decision-making process and also help them consider the most influential bidding variables.

Input selection in recurrent neural networks (RNNs) has also received some attention in terms of their utility in predicting accurate outcomes in different architectures. For example, Tenti (1996) affirms that RNNs, in which the activity patterns of the input layer pass through the network more than once prior to generating a new output pattern, can assist in the learning of extremely complex temporal patterns, which can subsequently enhance its forecasting accuracy. Using technical indicators as inputs to three RNN architectures, Tenti (1996) was able to establish high predictive accuracy for the output variable, which was a future forecast of the Deutsche mark of the Deutsche mark and its momentous profitability and riskiness of a trading strategy that focused on when to enter, when to exit, and number of contracts per trade. Liu, Guo & Wang (2012) and Tealab, Hefny & Badr (2017) likewise supported the role of proper input selection to guarantee the accuracy of output variables using an RNN. As the scholars vividly pointed out, the main advantage of RNNs stems from the fact that the short-term memory of RNN architecture retains features of the input series that are relevant to the prediction task while simultaneously capturing the network’s prior activation history. As such, the appropriate response at a specific point in time could not only depend on the current input, but also potentially on all previous inputs.

Some studies have also indicated that a number of difficulties may be inherent in input selection. In a comprehensive review of input variable selection techniques for ANNs, May, Dandy & Maier (2011) established that the difficulty of selecting input variables particularly for ANNs arises because of various factors. Some of these factors, as pinpointed by the researchers, include the number of available variables, variables having low or little predictive power, and correlations between the potential input variables, which may create redundancy. Chen, Zhang, Chen & Li (2008), Zeng, Zhang, Liu, Liang & Alsaadi (2017), and Kwak & Choi (2002) echoed similar sentiments, adding that the inherent complexity, nonlinearity, and nonparametric nature of ANN regression made it extremely difficult to apply many existing analytical input variable selection methods.

4.3 Methodology

Table 22

Methods	*Description*
1. NYMPHY_EXOGONOUS_CLOSE	Neural Network with exogenous variables
2. NYMPHY_CLOSE_HIGH	Neural Network with Close and High share price
3. NYMPHY_CLOSE_LOW	Neural Network with Close and Low share price
4. NYMPHY_CLOSE_HIGH_LOW	Neural Network with Close, High and Low share price

4.3.1 Description

The tested models here are the neural networks testing of uncertainty. Nymphy was the function tested in this paper. The algorithm was constructed to test whether the volatility of the share price could predict the future for that share price.

Table 23

Classifications of time series

	Companies	Methods	Error Matrices	Horizons	Index
Daily	18	4	6	22	FTSE 100
Weekly	18	4	6	12	FTSE 100
Monthly	18	4	6	18	FTSE 100
Quarterly	18	4	6	12	FTSE 100
Yearly	18	4	6	4	FTSE 100

The table above shows the overall structure of the paper's methodology.

4.4.1 Results

Table 24

DAILY	APE								
METHOD	HORIZON								
	1	2	3	4	5	10	22	1-10	1-22
CLOSE	0.024248	0.024247	0.024245	0.024244	0.024243	0.024237	0.024219	0.024242	0.024234
CLOSE_HIGH	0.024895	0.024894	0.024894	0.024894	0.024893	0.024889	0.024873	0.024892	0.024886
CLOSE_HIGH_LOW	**0.016707** *0.021847*	**0.016708** *0.021847*	**0.016708** *0.021847*	**0.016708** *0.021848*	**0.016708** *0.021848*	**0.016708** *0.021847*	**0.016702** *0.021843*	**0.016708** *0.021848*	**0.016706** *0.021847*
CLOSE_LOW	0.019447	0.019448	0.019449	0.019451	0.019452	0.019458	0.019467	0.019453	0.019459
SES	**0.013289** *0.013383* **(AUTOARIMA_ FOURIER)**	**0.018832** *0.018959*	**0.022836** *0.023048*	**0.026175** *0.026486*	**0.029162** *0.029591*	**0.040821** *0.042036*	**0.059134** *0.062275*	**0.029275** *0.031007*	**0.041303** *0.045043*

4.4.2 Discussion

The daily APE results for the Nymphy model and the winning method from paper 2 are shown above, with very interesting results. Produced from 4 thousand data points, the method close_high_low shows very consistent results going from horizon 1 through to horizon 22, however that is not the case for the best method carried over from paper 2, only beating the close_high_low method on the first horizon only where it was not the SES method but rather the Autoarima_Fourier method.

Table 25

WEEKLY	APE							
METHOD	HORIZON							
	1	2	3	4	6	12	1-6	1-12
CLOSE	**0.010104** *0.016593*	**0.010125** *0.016617*	**0.010146** *0.016631*	**0.010168** *0.016643*	**0.010213** *0.016688*	**0.010354** *0.016781*	**0.010157** *0.016639*	**0.010226** *0.016693*
CLOSE_HIGH	0.017385	0.017413	0.017434	0.017449	0.017497	0.017614	0.01744	0.017504
CLOSE_HIGH_LOW	0.01685	0.016875	0.016888	0.016897	0.016941	0.017022	0.016894	0.016944
CLOSE_LOW	0.016336	0.016359	0.016373	0.016389	0.016435	0.016541	0.016384	0.016442
AUTOARIMA_FOURIER	**0.029086** *0.025839*	**0.040883** *0.037228*	**0.049625** *0.051038* **(SES)**	**0.056642** *0.058677* **(SES)**	**0.069163** *0.072259*	**0.098690** *0.104405*	**0.051387** *0.054858*	**0.069353** *0.075354*

4.4.3 Discussion

Compared to the daily APE, the weekly APE showed us that the introduction of uncertainty does not mean volatility can produce better results, however the neural network here did perform stronger than the carried over methods from paper 1.

Table 26

MONTHLY	APE							
METHOD	HORIZON							
	1	2	3	4	9	18	1-9	1-18
CLOSE	0.110653	0.10991	0.108936	0.108032	0.102802	0.095695	0.106916	0.102559
CLOSE_HIGH	0.048861	0.04867	0.048108	0.047595	0.045819	0.044107	0.047275	0.045935
CLOSE_HIGH_LOW	**0.043350** *0.048346*	**0.043355** *0.048161*	**0.043174** *0.047784*	**0.042933** *0.047416*	**0.042019** *0.045705*	**0.041146** *0.043429*	**0.042735** *0.047049*	**0.042062** *0.045658*
CLOSE_LOW	0.047832	0.047652	0.047460	0.047237	0.045591	0.042753	0.046823	0.045381
AUTOARIMA_FOURIER	**0.056734** *0.059492*	**0.081133** *0.086376*	**0.100575** *0.108120* **(NAÏVE)**	**0.118678** *0.129480*	**0.199235** *0.230741*	**0.349945** *0.442083*	**0.133154** *0.150454*	**0.206680** *0.242198*

4.4.4 Discussion

Producing the same outcome as the the daily APE, the weekly results show us how uncertainty can produce better results. Furthermore, where in the daily results the winning method for horizon 1 did perform better than the close_high_low method, here the close_high_low method performed stronger even one horizon ahead.

Table 27

QUARTERLY	APE							
METHOD	HORIZON							
	1	2	3	4	8	12	1-6	1-12
CLOSE	0.496267	0.495893	0.495791	0.497656	0.508113	0.533274	0.497914	0.507351
CLOSE_HIGH	0.265978	0.266474	0.267439	0.269942	0.278473	0.291858	0.269350	0.276379
CLOSE_HIGH_LOW	**0.164642** *0.285519*	**0.165393** *0.286098*	**0.166231** *0.286737*	**0.167947** *0.289040*	**0.174909** *0.298342*	**0.183577** *0.312349*	**0.167552** *0.288737*	**0.172817** *0.296064*
CLOSE_LOW	0.305061	0.305721	0.306035	0.308139	0.318211	0.332839	0.308124	0.315750
AUTOARIMA_FOURIER	**0.108245** *0.120869*	**0.164416** *0.190705*	**0.213086** *0.252326*	**0.262497** *0.321027*	**0.538682** *0.628008*	**0.508976** *0.832367* **(SES_THETAF)**	**0.239539** *0.286677*	**0.441059** *0.565188*

4.4.5 Discussion

The quarterly APE outcome also shows us how the close_high_low method can perform better than all other neural network methods tested here, and can also perform stronger and more consistently than the strongest methods from paper 2. However, in this case, the most accurate method from paper 2 did perform better one horizon and two horizons ahead, however it lost accuracy after that. Meanwhile, to compare and contrast the close_high_low showed that it was consistent moving from one horizon to the next, this is because it did not lose accuracy as quickly as the methods from paper 2.

Table 28

YEARLY	APE					
METHOD	HORIZON					
	1	2	3	4	1-2	1-4
CLOSE	0.647111	0.651285	0.676842	0.715215	0.649198	0.672613
CLOSE_HIGH	0.217690	0.218180	0.229277	0.248240	0.217935	0.228347
CLOSE_HIGH_LOW	**0.164094** *0.252530*	**0.166871** *0.250952*	**0.179122** *0.262187*	**0.196443** *0.282098*	**0.165483** *0.251741*	**0.176632** *0.261942*
CLOSE_LOW	0.287370	0.283723	0.295097	0.315955	0.285546	0.295536
THETAF_YEARLY	**0.238219** *0.451838*	**0.324738** *0.558452*	**0.428134** *0.936784* **(NNET_THETAF)**	**0.422913** *1.294401* **(NNET_THETAF)**	**0.281478** *0.505145*	**0.400210** *0.747618*

4.4.6 Discussion

The same occurs in the yearly APE results, where the close_high_low performs better than all the other methods including the methods from paper 2, and in this case even in horizon 1 and 2 where in the quarterly results the methods from paper 2 performed better.

Mean Squared Tables

Table 29

DAILY	MSE								
METHOD	HORIZON								
	1	2	3	4	5	10	22	1-10	1-22
CLOSE	1825	1828	1832	1836	1840	1866	1914	1844	1871
CLOSE_HIGH	**1788** *2020*	**1791** *2024*	**1794** 2029	**1798** 2035	**1801** 2041	**1821** 2073	**1858** 2130	**1804** 2045	**1825** 2080
CLOSE_HIGH_LOW	2215	2220	2227	2234	2241	2280	2354	2246	2289
CLOSE_LOW	2453	2459	2466	2474	2483	2528	2614	2488	2538
SES	**1058** *1061*	**1977** *1985*	**2839** *2851*	**3673** *3687*	**4526** *4545*	**8521** *8566*	**16861** 17010	**4882** *4904*	**9405** *9469*

4.4.7 Discussion

For the daily MSE results, the close_high method performed better than all other methods, however the SES method from paper 2 was stronger one horizon ahead but then lost strength and lost strength faster than the close_high after horizon 1. Compared to the APE daily results, the close_high_low method was the stronger model but this was not the case for the MSE daily.

Table 30

WEEKLY	MSE							
METHOD	HORIZON							
	1	2	3	4	6	12	1-6	1-12
CLOSE	**1137** *1968*	**1149** *1979*	**1161** *1986*	**1175** *1991*	**1200** *2011*	**1278** *2052*	**1169** *1989*	**1206** *2012*
CLOSE_HIGH	2277	2292	2301	2306	2325	2375	2302	2328
CLOSE_HIGH_LOW	1817	1827	1832	1835	1850	1881	1833	1851
CLOSE_LOW	2119	2130	2139	2147	2172	2223	2144	2174
NAIVE	**4503** *4538*	**8556** *8599*	**12343** *12419*	**15871** *15959*	**22293** *22510*	**42137** *42850*	**13771** *13875* **(SES)**	**23799** *24094* **(SES)**

4.4.8 Discussion

The MSE weekly results compare well with the APE weekly results, showing us that volatility is not a suitable approach when it comes to the weekly frequency.

Table 31

MONTHLY	MSE							
METHOD	HORIZON							
	1	2	3	4	9	18	1-9	1-18
CLOSE	23875	24011	24101	24239	25272	28420	24476	25586
CLOSE_HIGH	8401	8471	8416	8391	**8571** *9209*	**8867** *9813*	8448	**8582** *9195*
CLOSE_HIGH_LOW	**7953** *8748*	**8079** *8831*	**8116** *8850*	**8143** *8873*	8630	9429	**8236** *8943*	8658
CLOSE_LOW	9096	9191	9285	9355	9789	10197	9437	9732
THETAF	**14628** *15758* **(AUTOARIMA_FOURIER)**	**28570** *31659* **(NAÏVE)**	**44750** *49450* **(AUTOARIMA_FOURIER)**	**61296** *68644* **(AUTOARIMA_FOURIER)**	**136167** *159636*	**265066** *305858*	**75360** *86973*	**142695** *170078*

4.4.9 Discussion

The results here are rather interesting where strength is visible from both the close_high and the close_high_low methods, but that is not the interesting result here. The interesting result is that both close_high and close_high_low methods performed stronger than the methods from paper 2 across all horizons.

Table 32

QUARTERLY	MSE							
METHOD	HORIZON							
	1	2	3	4	8	12	1-6	1-12
CLOSE	69867	73008	76459	81190	106650	143338	79899	99877
CLOSE_HIGH	27106	27547	28492	30051	35805	41291	29574	33697
CLOSE_HIGH_LOW	**18465** *25223*	**19647** *25518*	**20943** 26152	**21924** 27359	**27318** 32977	**34928** 38312	**21291** 27237	**25807** 31107
CLOSE_LOW	23340	23489	23813	24666	30149	35332	24900	28517
NAIVE	**54706** *61860*	**108625** *140296*	**153372** *239579*	**197683** *326865*	**344889** *449253*	**547139** *644036*	**169837** *271532*	**293670** *387320*

4.4.10 Discussion

The close_high_low method here is far stronger than all other methods.

Table 33

YEARLY	MSE					
METHOD	HORIZON					
	1	2	3	4	1-2	1-4
CLOSE	104193	118344	150483	198633	111268	142913
CLOSE_HIGH	76843	83645	103515	126997	80244	97750
CLOSE_HIGH_LOW	**69173** *88165*	**80559** *96818*	**101363** *118893*	**121866** *142911*	**74866** *92492*	**93240** *111697*
CLOSE_LOW	99487	109992	134271	158825	104740	125644
DSHW	**251580** *276596*	**369859** *387370*	**591059** *591059*	**933826** *933826*	**285997** *329866* **(THETAF)**	**530783** *563602* **(THETAF)**

4.4.11 Discussion

The same results carry over from the MSE quarterly accuracy test, where close_high_low method performs stronger than the other methods. This shows us a pattern where the larger the frequency, the better the performance of the close_high_low method compared to where the close_high method was performing stronger on the lesser frequencies.

4.5 Conclusion

4.5.1 Analysis & Evaluation

In our third paper, with the introduction of share price volatility we observed dominance of the close_high_low method in all frequencies apart from the weekly frequency which was dominated by the close neural network method. Otherwise the close_high_low method showed greater accuracy than any of the other methods. This method beat all other methods for all tested frequencies and horizons. It was also consistent through all the horizons, providing a great margin of error where it lay within a marginal error through the horizons when compared with the newer methods of paper 3.

However, when we bring in the winning methods from the previous papers the close_high_low method also showed better accuracy in all circumstances apart from quarterly and daily frequencies, nevertheless it was only close behind on the first horizon respectively. Furthermore, to elaborate the close_high_low method only lost to the method that had won the previous test on the first horizon of the daily frequency and the first horizon of the quarterly frequency.

Furthermore, as the winning methods from paper 2 significantly decreased in accuracy as the horizon increased, the close_high_low method showed great consistency. As the horizon got longer, the Autoarima method, which had won in the previous paper, kept losing strength but this was the case for all methods carried over from paper 2. At the same time, the neural networks showed great consistency throughout the horizons.

Furthermore, for the MSE outputs a pattern is easily observable with both the close_high and close_high_low methods performing much better than all other methods.

4.5.2 Future Research

For future research an exogenous variable will be introduced to the test, namely trading volume; the trading volume will be fed into the neural networks and the result will be tested if the volume that is traded on that stock will differ on the accuracy of our test. the volume

4.6 References

Abdi, H., & Williams, J. (2010). Principal Components Analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433-459.

Antić, D., Milovanović, M., Perić, S., Nikolić, S., & Milojković, M. (2014). Input Data Preprocessing Method for Exchange Rate Forecasting via Neural Network. Serbian Journal of Electrical Engineering, 11(4), 597-608.

Beckmann, J., & Czudaj, R. (2017). The impact of uncertainty on professional exchange rate forecasts. Journal of International Money and Finance , 73, 296-316.

Bollerslev, T., Marrone, J., Xu, L., & Zhou, H. (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence. Journal of Financial and Quantitative Analysis , 49, 633-661.

Bonaccolto, G., Caporin, M., & Gupta, R. (2018). The dynamic impact of uncertainty in causing and forecasting the distribution of oil returns and risk. Physica A , 2-53.

Chandar, S., Sumathi, M., & Sivanandam, S. (2015). Forecasting of foreign currency exchange Rate using neural network. International Journal of Engineering and Technology, 7(1), 99-108.

Chaudhuri, T., & Ghosh, I. (2016). Artificial Neural Network and Time Series Modeling Based Approach to Forecasting the Exchange Rate in a Multivariate Framework. Journal of Insurance and Financial Management, 1(5), 92-123.

Chen, T., Zhang, C., Chen , X., & Li, L. (2008). An Input Variable Selection Method for the Artificial Neural Network of Shear Stiffness of Worsted Fabrics. Statistical Analysis and Data Mining, 18, 1-9.

Claveria, O., Monte, E., & Torra, S. (2015). Tourism demand forecasting with neural network models: Different ways of treating information. International Journal of Tourism Research, 17, 492-500.

Clements, A., & Liao, Y. (2017). Forecasting the variance of stock index returns using jumps and cojumps. International Journal of Forecasting , 33, 729-742.

Colombo, V. (2013). Economic policy uncertainty in the US: Does it matter for the euro area? Economics Letters , 121, 39–42.

Conrad, C., & Loch, K. (2015). The variance risk premium and fundamental uncertanty. Economic Letters , 132, 56-60.

Degiannakis, S. (2018). Multiple days ahead realized volatility forecasting: Single, combined and average forecasts. Global Finance Journal , 36, 41-61.

Dick, C., MacDonald, R., & Menkhoff, L. (2015). Exchange rate forecasts and expected fundamentals. Journal of International Money and Finance , 53, 235-256.

Driver, C., Trapani, L., & Urga, G. (2013). On the use of cross-sectional measures of forecast uncertainty. International Journal of Forecasting , 29, 367-377.

Fadlalla, A., & Amani, F. (2014). Predicting next closing day closing price of Qatar Exchange Index using technical indicators and artificial neural networks. Intell. Sys. Acc. Fin. Mgmt., 21, 209-223.

Gong, X., & Lin, B. (2017). Forecasting the good and bad uncertainties of crude oil prices using a HAR framework. Energy Economics , 67, 315-327.

Harvey, D., Leybourne, S., & Whitehouse, E. (2017). Forecast evaluation tests and negative long-run variance estimates in small samples. International Journal of Forecasting , 33, 833-847.

Hillier, F., & Lieberman, G. (2014). Introduction to operations research (10 ed.). Boston, MA: McGraw Hill.

Knüppel, M. (2018). Forecast-error-based estimation of forecast uncertainty when the horizon is increased. International Journal of Forecasting , 34, 105-116.

Kruger, F., & Nolte, I. (2016). Disagreement versus uncertanty: Evidence from distribution forecasts. Journal of Banking & Finance , 72, S172-S186.

Kwak, N., & Choi, C. (2002). Input feature selection for classification problems. IEEE Transactions on Neural Networks, 13(1), 143-159.

Lahaye, I., Laurent, S., & Neely, C. (2011). Jumps, cojums, and macro announcement. Journal of Applied Econometrics , 26 (6), 893-921.

Liow, K. H., Liao, W.-C., & Huang, Y. (2018). Dynamics of international spillovers and interaction: Evidence from financial market stress and economic policy uncertainty. Economic Modelling , 68, 96-116.

Liu, Q., Guo, S., & Qiao, G. (2015). VIX forecasting and variance risk premium: A new GARCH approach. North American Journal of Economics and Finance , 34, 314-322.

Liu, Q., Guo, Z., & Wang, J. (2012). A one-layer recurrent neural network for constrained pseudoconvex optimization and its application for dynamic portfolio optimization,". Neural Networks, 26, 99-109.

Luo, L., Sattayatham, P., & Chatpatanasiri, R. (2017). GARCH-type forecasting models for volatility of stock market and MCS test. Communication in Statistics- Simulation and Computation , 46 (7), 4-19.

May, R., Dandy, G., & Maier, H. (2011). Review of Input Variable Selection Methods for Artificial Neural Networks . Adelaide: University of Adelaide.

Moghaddam, A., Moghaddam, M., & Esfandyari, M. (2016). Stock market index prediction using artificial neural network. Journal of Economics, Finance, and Administrative Science, 21, 89-93.

Nonejada, N. (2017). Forecasting aggregate stock market volatility using financial and macroeconomic predictors: Which models forecast best, when and why? Journal of Empirical Finance , 42, 131-154.

Pacelli, V., Bevilacqua, V., & Azzollini, M. (2011). An artificial neural network model to forecast exchange rates. Journal of Intelligent Learning Systems and Applications, 3, 57-69.

Petropoulos, F., Hyndman, R., & Bergmeir, C. (2018). Exploring the source of uncertainty: Why does bagging for time series forecasting work? European Journal of Operational Research , 268, 545-554.

Prak, D., & Teunter, R. (2018). A general method for addressing forecasting uncertainty in inventory models. International Journal of Forecasting , 1-15.

Prokopczuk, M., Symeonidis, L., & Wese-Simen, C. (2015). (2015). Do jumps matter for volatility forecasting? Evidence from energy markets. Journal of Futures Markets , 1-35.

Rasouli, S., Tabesh, H., & Etminani, K. (2016). A Study of Input Variable Selection to Artificial Neural Network for Predicting Hospital Inpatient Flows. British Journal of Applied Science & Technology, 18(4), 1-8.

Shi, J. (2000). Reducing prediction error by transforming input data for neural networks. Journal of Computing in Civil Engineering, 14(2), 109-116.

Tealab, A., Hefny, H., & Badr, A. (2017). Forecasting of nonlinear time series using ANN. Future Computing and Informatics Journal, 20, 1-9.

Tenti, P. (1996). Forecasting foreign exchange rates using recurrent neural networks. Applied Artificial Intelligence, 10, 567-581.

Wanous, M., Boussabaine, H., & Lewis, J. (2013). A neural network bid/no bid model: the case for contractors in Syria. Construction Management and Economics, 21, 737–744.

Wei, Y., Liu, J., Lai, X., & Hu, Y. (2017). Which determinant is the most informative in forecasting crude oil market volatility: Fundamenta, speculation, or uncertainty? Energy Economics , 68, 141-150.

Yin, I., & Zhou, Y. (2016). What drives long-term oil volatility? Fundamentals versus speculation. Economics , 10, 1-14.

Yu, H., Fang, L., & Sun, W. (2018). Forecasting performance of global economic policy uncertainty for volatility of Chinese stock market. Physica A , 505, 931-940.

Zeng, N., Zhang, H., Liu, W., Liang, J., & Alsaadi, F. E. (2017). A switching delayed PSO optimized extreme learning machine for short-term load forecasting. Neurocomputing, 240, 175-182.

4.7 Appendices

4.7.1 Forecast Accuracy Measures

The forecasting accuracy measures are tested and developed in accordance with the following structures:

4.7.2 Error Metrics Tested

Equation 15

Mean Error (ME)

$$ME = \frac{1}{N} \sum_{i=1}^{N} (f_i - k_i)$$

Where:

N = # Forecast / Actuals

F = Forecast

K = Actuals

Equation 16

Mean Absolute Error (MAE)

$$= \text{mean } (|e_i|)$$

Mean Squared Error (MSE)

$$mean\ (e_i^2)$$

Equation 17

Root Mean Squared Error (RMSE)

$$= \sqrt{mean\ (e_i^2)}$$

Equation 18

Mean Percentage Error (MPE)

$$MPE = \frac{1}{N} \sum_{i=1}^{N} (f_i - k_i) \text{ x } 100$$

Equation 19

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} (|f_i| - |k_i|) \text{ x } 100$$

The aforementioned accuracy tests follow the structure built by Robert Hyndman.

CHAPTER 5

FORECASTING FINANCIAL MARKETS WITH PREDICTIVE ANALYTICS: THE IMPACT OF EXOGENOUS VARIABLES; IN THE FORM OF THE TRADING VOLUME

5.1 Abstract

This paper introduces the trading volume of the share price into the neural network in an attempt to test if an exogenous variable, in the form of trading volume, can produce more accurate results compared to only having the closing price fed into the neural network. By feeding the volume into the neural network we can understand if human behaviour and action can affect the future of that share price and the significance of its effect on the future price. By feeding the volume into the network, we can also compare how a greater number of trading days, weeks, months, quarters and years affect the future share price. This may clarify whether higher or lower volumes of trading result in better forecasting accuracy. Comparisons can also be drawn for when the trading volume is being fed into the network and when it is not.

5.2 Introduction

The existing literature suggests that the relationship between returns and trade volume is of undisputable interest to financial markets. The daily trading volume represents the number of shares traded on any particular day (Tangmongkillert & Suwanna, 2016). Many of the published studies clearly provide support for the existence of a positive relationship between trading volume and stock price changes, a factor that depends primarily on the company's closing share price. A consensus seems to exist supporting the assumption that the relationship between trading activity, which includes trading volume, and closing share price changes is critical because trading activity tends to be thin whereas price volatility is quite high. This blueprint provides a comprehensive review of the issues relating to the relationships between share prices, trading volume, volatility, and returns as documented in the existing empirical studies.

5.3 Literature Review

In today's international financial markets, volatility is a critical risk factor, with portfolio allocation methods and asset pricing models relying upon the precision of its estimates. Consistent with a study carried out by Tangmongkillert & Suwanna (2016), Sabbaghi (2011) sought to examine the relationship between asymmetric volatility and trading volume, based on their evidence from the G5 countries. Specifically, the researcher used national equity indices and performed an EGARCH analysis to explicate the asymmetric volatility-trading volume relationship. Findings from this study proved to be quite captivating. After performing the EGARCH analysis, the researcher established that trading volume is a critical variable in explaining conditional volatility. Just like Tangmongkillert & Suwanna (2016), Sabbaghi (2011) also found that the presence of trading volume did not cause the persistence levels of volatility to decrease. Gebka & Wohar (2013), in a similar light, endeavoured to examine the causality that exists between index returns and trading volume in the Pacific Basin countries. While ordinary least squares regression did not show any causal relationships between trading volume and index returns, applying a quantile regression method did reveal strong nonlinear causality that was negative for low return quantiles and positive for high return quantiles at both long-term and short-term horizons.

Caginalp & DeSantis (2017) endeavoured to examine whether price efficiency increases with trading volume, testing this relationship using 124,236 daily observations of

68 large and liquid exchange traded funds (ETFs) in the US equity markets. The main advantage of ETFs, according to Caginalp & DeSantis (2017) and Samonas (2015), is that the forecasters can measure efficiency in terms of deviation between the trading price and the underlying net asset value reported each day. Data for a study by Caginalp & DeSantis (2017) came from the Bloomberg terminal and the findings from this inquiry suggested that the relationship between efficiency and trading volume is nonlinear, with the efficiency increasing as trading volume increases from low to moderately high levels and, in turn, the closing share price increases, but efficiency decreases subsequently as the volume increases even further. The researchers speculated that, as the volume increases even further, it leads to increased speculation that ignores valuation and thus decreases efficiency. In the long-term, this leads to a decline in the closing share price.

Considerable publications have also delved into the relationship between search intensity and trading volume. One such inquiry was by Takeda & Wakao (2014), who tested the relationship between Google search intensity and returns as well as trading volume in Japanese stocks. The scholars measured the search intensity according to the search volume of names of companies listed in Google. Takeda & Wakao (2014) used a sample consisting of 189 Japanese stocks searched in the period between 2008 and 2011. After performing regression analyses, the researchers established that correlations with search intensity were weakly positive for stock returns and strongly positive for trading volume. The rationale behind this trend was that the increase in search activity had a direct association with increases in trading activity, which subsequently increased the share price at the closing time. In their study, Caginalp & DeSantis (2017) refuted the claims by Takeda & Wakao (2014) through asserting that increases in trading activity had a low probability of leading to increases in stock prices.

Doo, Brooks, Treepongkaruna & Wu (2014) took a different approach in which they examined the effects of trading volume on financial return distributions in 18 international currency and equity markets. Data sources for this investigation came from FX transactions and stock market indices for the 5-minute intraday data from the Thomas Reuters Tick History database. They then used the US dollar as the base currency for the FX market, pricing all other national currencies against it. Do et al. (2015) computed for each market the 5-minute intraday returns in the closing and the mid prices of the FX and stock markets. The volume-volatility analysis in this study provided evidence of lead-lag and positive relationships between trading volume and volatility across FX and stock markets, thereby lending supporting to earlier findings by Sabbaghi (2011). Furthermore, whereas the researchers observed a stronger level of interdependence among higher moments in reaction to significant financial events, trading volume dampened the strength of this association.

The relationship between trading volume and returns in financial markets has continued to attract notable attention from not only academicians but also from practitioners. Chen, Qiu, Jiang, Zhong & Wu (2015), similar to studies conducted by many other researchers, also endeavoured to gauge the manner in which trading volume responds to price returns in financial dynamics. The researchers based their analysis on daily data of the United States and the Chinese stock markets. For the United States, analysed datasets came from the Dow Jones Industrial Average (DJIA) and the S&P 500 whereas datasets for the Chinese stock market came from the Shenzhen Composite Index and the Shanghai Composite Index. After deploying a retarded herding model, results from this inquiry differed in various aspects from those conducted by Boonvorachote & Lakmas (2016) and Samonas (2015). Specifically, Chen et al. (2015) observed a positive correlation between trading volume and returns in the Chinese stock markets, but for the United States stock markets they observed a transition from the positive to the negative correlation. They attributed the changes to the differences in financial dynamics between mature or developed markets (e.g. US) and emerging markets (e.g. China).

Brida, Matesanz & Siejas (2016) similarly based their study on the relationship between returns and trading volume in the Euro Stoxx market, albeit applying a network analysis approach to analyse the structure of the market between 2002 and 2014. Volume trading and asset returns were the main variables in this study. They introduced a multidimensional generalization of the minimal spanning tree (MST) concept by adding the role of trading volume to the traditional methodology that only includes price returns. Both Brida et al. (2016) and Chen (2012) asserted that a common adage is that volume is relatively light in bear markets and heavy in bull markets. In this regard, Chen (2012) conducted a comprehensive study to scrutinise whether the return-volume relation is asymmetrical in bear and bull stock markets using trading volume from 2008 to 2014 and monthly data for the S&P 500. The researcher found that the stock return has a higher capability of forecasting trading volume in both bull and bear stock markets. To study the market structure behaviour in critical and normal situations, Brida et al. (2016) in their inquiry used symbolization techniques for the raw data, deriving the hierarchical organisation of their network from the structural topologies of the MST, with findings mirroring those of Chen (2012).

Following a methodological approach similar to that favoured by Chen (2012), Gupta, Das, Hasim & Tiwari (2018) likewise revisited the dynamic relationship between trading volume and stock returns. Differently, Gupta et al. (2018) relied upon a maximum overlap discrete wavelet transform (MODWT) approach to revisit the relationship in a time-frequency domain. The researchers utilised almost concurrent data of over 15 years from the emerging stock markets of India and China. To examine the dynamic relationship, Gupta et al. (2018) first applied the MODWT for the purposes of decomposing the level series. They then applied VAR to the decomposed data in order to obtain a richer picture of the causality between stock returns and trading volume for different time-scale horizons. According to Samonas (2015), the VAR methodology is beneficial for gauging horizon-based investor behaviour and ascertaining whether the stock returns are responsible for predicting trading volume or vice-versa. Gupta et al. (2018), after deploying the MODWT-VAR approach, found that the examined markets work according to an efficient market hypothesis in the short-term horizon, and in the long-term horizon they reach a stage of market inefficiency, a factor that may lead to low closing share prices.

A number of empirical studies have affirmed the possibility of an additional foreign market presence having a direct effect on trading volume. This is especially true for cross-listed forms. Ghadhab (2016) rolled out a study geared towards addressing the question of the effects that additional foreign market presence can have on the trading volume of cross-listed firms and, consequently, their closing share prices. The researcher relied upon a comprehensive and unique sample of 235 firms from 32 nations with 788 foreign listings over the period spanning from the year 1980 to 2013. Just like Dodd, Louca & Paudyal (2015), Ghadhab (2016) made a clever observation by reiterating that extant literature shows that cross-listing can enhance the value of the firm but the source of such increment remains elusive, although corporate managers often attribute the value addition to increases in stocks' liquidity. In his study, Ghadhab (2016) found that, compared with the decline in trading volume after the first cross-listing/trading, additional foreign listings resulted in more shares being traded on the stock. Furthermore, the researcher found that the effect of additional cross-listing/trading was more important for high orders.

Dodd et al. (2015), in a similar study, seemed to contradict the findings arrived at by Ghadhab (2016). The researchers explored the determinants of the foreign trading volume with a particular emphasis on European stocks listed in multiple markets. Unlike Ghadhab (2016), who found that additional foreign listings resulted in more shares being traded on the stock, the results by Dodd et al. (2015) contradicted these findings. Their results suggested that stocks that cross-list in markets that have greater liquidity and are larger than their home markets, and stocks for which foreign investors can obtain information at a lower cost, tend to experience higher trade volumes in foreign destinations. As Dodd et al. (2015) vividly pointed out, this is the reason why stocks cross-listed in the US stock market are usually more attractive to foreign traders compared to stocks cross-listed in European markets. Among the fundamental motives to trade, the researchers added that stock risk and diversification benefit are more important for investors aspiring to trade in American markets while for investors in European markets, the difference in trading costs is more critical. Samonas (2015) also echoed similar sentiments, adding that informational motives to trade are also significant determinants of trading in US markets but not in European markets.

Contrary to many other studies, Alvez-Albelo (2012) initiated an investigation designed to examine whether importing growth via trade volume or via terms of trade matters when determining the extent of competition in the export sector. In order to illustrate their arguments, the researchers constructed two simple growth models that represented a small open economy using its export revenues to import capital goods and enjoying strong export market power. In the second model that specifically tested whether trade volume increases mattered when determining the degree of export sector competition, Alvez-Albelo (2012) proposed that growth directly depended upon externalities associated with the trade volume or the number of shares traded. Results from this study indicated that importing growth via both trade volume increases and terms of trade was relevant for determining the extent of competition in the export sector. They concluded that when growth relies on externalities associated with the trade volume, “more competition is required in the export sector” (Alvez-Albelo, 2012, p. 8). This supported the empirical evidence of Samonas (2015), which showed that high closing share price is responsible for increasing competition in the export market.

Meanwhile, Magkonis & Tsouknidis (2017) examined spillover effects that were evident across petroleum-based commodities and among trading volume, spot-futures volatilities, and open interest from 2341 observations. To accomplish this goal, they looked at daily time-series of closing spot prices, futures total volume, futures prices, and futures open interest using RVs of spot-futures markets as inputs for estimating a VAR model and distinguishing dynamic spillovers in total as well as net effects. When examined pairwise, the results revealed the existence of time-varying and large spillovers across the petroleum-based commodities and among spot-futures volatilities. Furthermore, the researchers found that hedging pressure, as reflected by open interest, and speculative pressures, as reflected by the futures trading volume, transmitted persistent and large spillovers to the futures and spot volatilities of heating oil gasoline and crude oil markets respectively (Magkonis & Tsouknidis, 2017).

Boonvorachote & Lakmas (2016) similarly based their study on commodity markets, although their specific focus was on futures exchanges, unlike Magkonis & Tsouknidis (2017) who explored petroleum-based commodities. The primary objective in the study by Boonvorachote & Lakmas (2016) was to investigate the impact that trading activity, including open interest and trading volume, had on price volatility on futures exchanges in Asian economies. Their study utilised three different definitions of volatility. The first was daily volatility, which they measured according to close-to-close returns. The other two were trading volatility measured via open-to-close returns and non-trading volatility measured according to close-to-open returns. Boonvorachote & Lakmas (2016) subsequently divided volume and open interest into unexpected and expected components, after which they employed an augmented GARCH model using these components as explanatory variables. Findings were consistent with those of Chan, Fung & Leung (2004), with Boonvorachote & Lakmas (2016) suggesting that while open interest was able to mitigate volatility, a positive contemporaneous relationship on the contrary existed between daily volatility as measured via close-to-close returns and unexpected and expected trading volume. Chan et al. (2004), in their empirical investigation that centred on volatility behaviour in the Chinese market, went on to add that hedging activities, as substituted by open interest, tended to stabilise the market while speculative activities, as substituted by the volumes, tended to increase the futures volatility.

Wang, Qian & Wang (2017) endeavoured to examine the dynamic relationship between trade volume and stock returns. The researchers held the view that the popularisation of high-frequency, high-speed trading over the past two decades, which is a conspicuous aspect of financial markets, has attracted increasing attention with regard to the relationship between trading volume and stock return from both practitioners and academicians. In conducting their study, Wang et al. (2017) looked at the relationship from the perspective of out-of-sample stock return predictability. They believed that for the purposes of risk management and real-time predictions, focusing on the dynamic relationship between returns and volume is perhaps more informative than the often elusive contemporaneous causality. Wang et al. (2017) in their study found that in certain markets including the United States, higher stock returns do indeed follow higher trading volume, whether measured by high volume return premiums or by aggregate time-series of turnover, a sentiment also echoed by Dodd et al. (2015). However, Wang et al. (2017) acknowledged that forecasters and academicians should interpret such predictive power with caution because the associated economic gain is quantitatively minimal for the market as a whole.

It is important to mention that, whereas many of the reviewed studies have explicitly examined the relationship between trading volume and returns, others have gone a step further to explore the mediating role of trading volume on volatility of financial markets. For instance, Shahzad, Hernandez, Hanif & Kayani (2018) aspired to investigate the dynamics of long memory and efficiency, as mediated by trading volume, on the volatilities and efficiency of returns of four major traded global currencies (GBP, EUR, JPY, and CHF). The researchers, in a similar light to Tabak & Cajueiro (2006), affirmed that the issue of efficiency in financial markets is of critical importance since it relates to the absence or existence of arbitrage opportunities that can, in turn, enhance or diminish the probability of earning above-average market returns and thus affect the closing share price positively or negatively. Shahzad et al. (2018) in their investigation used a quantile-on-quantile (QQ) approach while simultaneously implementing full sample and rolling window MF-DFA in order to test the mediating role of trading volume and employed high-frequency data (5-minute interval) spanning from 2007 to 2016. After deploying the QQ approach for analysis, the scholars found evidence of higher levels of efficiency in the CHF and JPY currency markets, with further analysis revealing that the trading volumes' impact on efficiency was only significant in these two currencies. The least efficient currency in the investigation appeared to be the GBP, closely followed by the EUR, both of which experienced substantial declines in their closing share prices over the study duration.

Overall, this review of literature has revealed that the relationship between trade volume and returns, while taking into account factors such as volatility and closing share price, is of significant interest to financial markets. In many of the studies explored, it became apparent that trading volume positively affects returns, with factors such as volatility, uncertainty, market structure, and cross-listing in foreign markets serving as mediators to the degree of the relationship. Gaining a heightened understanding of the relationships between the aforementioned variables can help corporate investors, forecasters, and other concerned stakeholders to make informed decisions that are beneficial for their investment ambitions.

Like many other researchers, Wang (2001) also centred his study on the neural network approach to input-output analysis, albeit focusing exclusively on economic systems. The researcher contended that conventional input-output analytic methods are becoming less attractive for various reasons. The first of these reasons pertains to the assumption of a constant linear relationship between the input and output. For the Chinese economy, Wang (2001) affirmed that this assumption is incorrect because of various factors such as the introduction of modern technologies, the introduction of massive amounts of FDI due to the "open door" policy, fast-growing demand from both the government and consumers, as well as unbalanced development in different regions. Hence, as the scholar averred, in conventional input–output analysis, the linear input coefficient matrix calculated based upon the statistics of the previous years would be unacceptably erroneous for the current year.

Consistent with Wang (2001), Claveria, Monte & Torra (2015) in their systematic review also highlighted another shortcoming of the conventional input-output analysis in neural networks. The researcher affirmed that another major shortcoming stems from the reality that all conventional input-output neural network analysis assumes that exogenous sectors (final demands) are given. The alternative neural network input-output analysis model developed by Wang (2001), which was a layered neural network model, had many advantages over traditional or conventional mathematical models, including high adaptive capacity and the advantage of nonlinearity. Nevertheless, the researcher also affirmed that the model had little capability of modelling oscillatory economic systems such as the stock market.

Shi (2000) offered an interesting perspective on reducing prediction error through transforming input data for neural networks. According to the researcher, the primary goal of data transformation is modifying the distribution of input variables so that they can match the outputs better. The three prevalent methods of data transformation are linear transformation, mathematical functions, and statistical standardization. Shi (2000) presented another method of data transformation using cumulative distribution functions, merely addressed as distribution transformation. The researcher contended that this method has the potential to transform a stream of any data distributed in any range of data points that are uniformly distributed on [0, 1]. Therefore, “all neural networks input variables can be transformed to the same ground-uniform distributions on [0, 1] (Shi, 2000, p.109).

5.4 Methodology

Table 34

Methods	*Description*
1. Nymphy_Close_Volume	Neural Network with Close and Volume share price

Nymphy_Close_Volume is the method used in this paper.

Table 35

Classifications of time series

	Companies	Methods	Error Matrices	Horizons	Index
Daily	18	1	6	22	FTSE 100
Weekly	18	1	6	12	FTSE 100
Monthly	18	1	6	18	FTSE 100
Quarterly	18	1	6	12	FTSE 100
Yearly	18	1	6	4	FTSE 100

The table above shows the structure of the methodology.

5.5 Results

Table 36

DAILY	**APE**								
METHOD	**HORIZON**								
	1	2	3	4	5	10	22	1-10	1-22
CLOSE_VOLUME	**0.024027**	**0.024026**	**0.024025**	**0.024024**	**0.024022**	**0.024015**	**0.024003**	**0.024021**	**0.024015**
CLOSE_HIGH_LOW	**0.016707**	**0.016708**	**0.016708**	**0.016708**	**0.016708**	**0.016708**	**0.016702**	**0.016708**	**0.016706**
	0.021847	*0.021847*	*0.021847*	*0.021848*	*0.021848*	*0.021847*	*0.021843*	*0.021848*	*0.021847*
SES	**0.013289**	**0.018832**	**0.022836**	**0.026175**	**0.029162**	**0.040821**	**0.059134**	**0.029275**	**0.041303**
	0.013383	*0.018959*	*0.023048*	*0.026486*	*0.029591*	*0.042036*	*0.062275*	*0.031007*	*0.045043*
	(AUTOARIMA_FOURIER)								

5.5.1 Discussion

The close_volume here does not perform stronger than method carried over from paper 3 but compares well with the method carried over from paper 2, even performing stronger in the latter stages of the horizon.

Table 37

WEEKLY	APE							
	HORIZON							
	1	2	3	4	6	12	1-6	1-12
CLOSE_VOLUME	**0.01235**	**0.012375**	**0.012398**	**0.012423**	**0.012474**	**0.012642**	**0.012411**	**0.012490**
CLOSE	**0.010104** *0.016593*	**0.010125** *0.016617*	**0.010146** *0.016631*	**0.010168** *0.016643*	**0.010213** *0.016688*	**0.010354** *0.016781*	**0.010157** *0.016639*	**0.010226** *0.016693*
AUTOARIMA_FOURIER	**0.029086** *0.025839*	**0.040883** *0.037228*	**0.049625** *0.051038* **(SES)**	**0.056642** *0.058677* **(SES)**	**0.069163** *0.072259*	**0.098690** *0.104405*	**0.051387** *0.054858*	**0.069353** *0.075354*

5.5.2 Discussion

The method from paper 2 here shows significant strength compared to the other two winning methods.

Table 38

MONTHLY	APE							
	HORIZON							
	1	2	3	4	9	18	1-9	1-18
CLOSE_VOLUME	**0.111967**	**0.111374**	**0.110548**	**0.109767**	**0.105142**	**0.099356**	**0.014732**	**0.015132**
CLOSE_HIGH_LOW	**0.043350** *0.048346*	**0.043355** *0.048161*	**0.043174** *0.047784*	**0.042933** *0.047416*	**0.042019** *0.045705*	**0.041146** *0.043429*	**0.042735** *0.047049*	**0.042062** *0.045658*
AUTOARIMA_FOURIER	**0.056734** *0.059492*	**0.081133** *0.086376*	**0.100575** *0.108120* **(NAÏVE)**	**0.118678** *0.129480*	**0.199235** *0.230741*	**0.349945** *0.442083*	**0.133154** *0.150454*	**0.206680** *0.242198*

5.5.3 Discussion

Akin to the weekly results, the winning method from paper 3 is the most accurate.

Table 39

QUARTERLY	APE							
	HORIZON							
	1	2	3	4	8	12	1-8	1-12
CLOSE_VOLUME	**0.262950**	**0.262968**	**0.263902**	**0.265588**	**0.270850**	**0.289465**	**0.265962**	**0.270890**
CLOSE_HIGH_LOW	**0.164642** *0.285519*	**0.165393** *0.286098*	**0.166231** *0.286737*	**0.167947** *0.289040*	**0.174909** *0.298342*	**0.183577** *0.312349*	**0.167552** *0.288737*	**0.172817** *0.296064*
AUTOARIMA_FOURIER	**0.108245** *0.120869*	**0.164416** *0.190705*	**0.213086** *0.252326*	**0.262497** *0.321027*	**0.538682** *0.628008*	**0.508976** *0.832367* **(SES_THETAF)**	**0.239539** *0.286677*	**0.441059** *0.565188*

5.5.4 Discussion

Akin to the weekly and monthly results, the winning method from paper 3 is the most accurate.

Table 40

YEARLY	APE					
	HORIZON					
	1	2	3	4	1-2	1-4
CLOSE_VOLUME	**0.251166**	**0.240078**	**0.252682**	**0.279575**	**0.245622**	**0.255875**
CLOSE_HIGH_LOW	**0.164094** *0.252530*	**0.166871** *0.250952*	**0.179122** *0.262187*	**0.196443** *0.282098*	**0.165483** *0.251741*	**0.176632** *0.261942*
THETAF_YEARLY	**0.238219** *0.451838*	**0.324738** *0.558452*	**0.428134** *0.936784* **(NNET_THETAF)**	**0.422913** *1.294401* **(NNET_THETAF)**	**0.281478** *0.505145*	**0.400210** *0.747618*

5.5.5 Discussion

Akin to the weekly, monthly and yearly results, the winning method from paper 3 is the most accurate.

Table 41

DAILY	MSE								
METHOD	HORIZON								
	1	2	3	4	5	10	22	1-10	1-22
CLOSE_VOLUME	**856**	**856**	**857**	**857**	**858**	**862**	**869**	**858**	**863**
CLOSE_HIGH	**1788** *2020*	**1791** *2024*	**1794** 2029	**1798** 2035	**1801** 2041	**1821** 2073	**1858** 2130	**1804** 2045	**1825** 2080
SES	**1058** *1061*	**1977** *1985*	**2839** *2851*	**3673** *3687*	**4526** *4545*	**8521** *8566*	**16861** 17010	**4882** *4904*	**9405** *9469*

5.5.6 Discussion

Where on our APE results the introduction of trading volume did not produce more accurate results, on the MSE daily data the volume was very effective in terms of producing the most accurate result.

Table 42

WEEKLY	MSE							
METHOD	HORIZON							
	1	2	3	4	6	12	1-6	1-12
CLOSE_VOLUME	**981**	**991**	**1003**	**1014**	**1035**	**1111**	**1008**	**1043**
CLOSE	**1137** *1968*	**1149** *1979*	**1161** *1986*	**1175** *1991*	**1200** *2011*	**1278** *2052*	**1169** *1989*	**1206** *2012*
NAIVE	**4503** *4538*	**8556** *8599*	**12343** *12419*	**15871** *15959*	**22293** *22510*	**42137** *42850*	**13771** *13875* **(SES)**	**23799** *24094* **(SES)**

5.5.7 Discussion

Akin to the daily results, the winning method from paper 4 is the most accurate.

Table 43

MONTHLY	MSE							
METHOD	HORIZON							
	1	2	3	4	9	18	1-9	1-18
CLOSE_VOLUME	**32593**	**33262**	**33814**	**34445**	**38439**	**48093**	**35309**	**39313**
CLOSE_HIGH_LOW	**7953** *8748*	**8079** *8831*	**8116** *8850*	**8143** *8873*	**8630**	**9429**	**8236** *8943*	**8658**
THETAF	**14628** *15758* **(AUTOARIMA _FOURIER)**	**28570** *31659* **(NAÏVE)**	**44750** *49450* **(AUTOARIMA _FOURIER)**	**61296** *68644* **(AUTOARIMA _FOURIER)**	**136167** *159636*	**265066** *305858*	**75360** *86973*	**142695** *170078*

5.5.8 Discussion

The close_volume does lose pace when we reach the monthly frequency where the method from paper 3 is the most accurate.

Table 44

QUARTERLY	MSE							
METHOD	HORIZON							
	1	2	3	4	8	12	1-6	1-12
CLOSE_VOLUME	**136033**	**145783**	**156900**	**169513**	**220089**	**281896**	**163671**	**203033**
CLOSE_HIGH_LOW	**18465** *25223*	**19647** *25518*	**20943** 26152	**21924** 27359	**27318** 32977	**34928** 38312	**21291** 27237	**25807** 31107
NAIVE	**54706** *61860*	**108625** *140296*	**153372** *239579*	**197683** *326865*	**344889** *449253*	**547139** *644036*	**169837** *271532*	**293670** *387320*

5.5.9 Discussion

Akin to the monthly results, the method from paper 3 is the most accurate.

Table 45

YEARLY **MSE**

METHOD	HORIZON					
	1	2	3	4	1-2	1-4
CLOSE_VOLUME	**175727**	**220196**	**293567**	**398809**	**197961**	**272075**
CLOSE_HIGH_LOW	**69173** *88165*	**80559** *96818*	**101363** *118893*	**121866** *142911*	**74866** *92492*	**93240** *111697*
DSHW	**251580** *276596*	**369859** *387370*	**591059** *591059*	**933826** *933826*	**285997** *329866* **(THETAF)**	**530783** *563602* **(THETAF)**

5.5.10 Discussion

Akin to the monthly and quarterly results, the method from paper 3 is the most accurate.

5.6 Conclusion

5.6.1 Analysis & Evaluation

The introduction of the trading volume into the neural network tested whether trading volume would make the model produce more accurate results. This proved to be consistent with our predictions as some days in the financial market there is high trading volume, meaning high liquidity and some days the market is quiet, due to investors and traders waiting for an event or as a result of their fear of market volatility.

In our test, we were able to observe some significant results where the introduction of the trading volume variable showed great robustness and beat all other methods that had won the previous papers and were carried over to this test. This variable also showed more accuracy as the horizons increased. If we take into consideration the models from paper 2, the close_volume model showed greater accuracy in most frequencies and horizons. Nevertheless, the models from paper 2 did produce better accuracy on the first horizon of some of the frequencies. However, overall the model from paper 4 showed greater power than the model from paper 2. It also beat the close_high_low model from paper 3 in some cases, however the close_high_low model performed better than the close_volume model.

Other important variables that affect the relationship between share price, trading volume, and returns include volatility, uncertainty, market efficiency, and listing of stocks in multiple foreign countries.

The MSE results for close_volume did not show great strength against its counterparts in higher frequencies, however it did show accurate results on lesser frequencies.

5.6.2 Future Research

After testing the trading volume in this paper, future research will test other exogenous variables including changes in inflation, interest rate and consumer price index (CPI) and test how these variables affect the volume of trading in the market and how that eventually affects the share price. The share price of some companies would react differently to different exogenous variables depending on the industry they are in and/or the service they provide.

5.7 References

Alvez-Albelo, C. (2012). Does it matter if growth is imported via terms of trade or trade volume increases? *Economics Letters , 114*, 6-8.

Boonvorachote, T., & Lakmas, K. (2016). Price volatility, trading volume, and market depth in Asian commodity futures exchanges. *Kasetsart Journal of Social Sciences , 37*, 53-58.

Brida, J., Matesanz, D., & Siejas, M. (2016). Network analysis of returns and volume trading in stock markets: The Euro stoxx case. *Physica A , 444*, 751-754.

Caginalp, G., & DeSantis, M. (2017). Does price efficiency increase with trading volume? Evidence of nonlinearity and power laws in ETFs. *Physica A , 467*, 436-452.

Chan, K., Fung, H., & Leung, W. (2004). Daily volatility behavior in Chinese futures markets. *Journal of International Financial Markets, Institution, and Money , 14*, 491-505.

Chen, G., Qiu, T., Jiang, X.-F., Zhong, L., & Wu, X. (2015). How trading volume responds to return in financial dynamics. *Physica A , 424*, 73-81.

Chen, S.-S. (2012). Revisiting the empirical linkages between stock returns and trading volume. *Journal of Banking and Finance , 36*, 1781-1788.

Claveria, O., Monte, E., & Torra, S. (2015). Tourism demand forecasting with neural network models: Different ways of treating information. International Journal of Tourism Research, 17, 492-500.

Do, H., Brooks, R., Treepongkaruna, S., & Wu, E. (2014). How does trading volume affect financial return distributions? *International Review of Financial Analysis , 35*, 190-206.

Dodd, O., Louca, C., & Paudyal, K. (2015). The determinants of foreign trading volumes of stocks listed in multiple markets. *Journal of Economics and Business , 79*, 38-61.

Gebka, B., & Wohar, M. (2013). Causality between trading volume and returns: Evidence from quantile regressions. *International Review of Economics and Finance , 27*, 144-159.

Ghadhab, I. (2016). The effect of additional foreign market presence on the trading volume of cross-listed/traded stocks. *Journal of Multinational Financial Management , 34*, 18-27.

Gupta, S., Das, D., Hasim, H., & Tiwari, A. (2018). The dynamic relationship between stock returns and trading volume revisited: A MODWT-VAR approach. *Finance Research Letters* , 1-8.

Magkonis, G., & Tsouknidis, D. (2017). Dynamic spillover effects across petroleum spot and futures volatilities, trading volume and open interest. *International Review of Financial Analysis , 52*, 104-118.

Sabbaghi, O. (2011). Asymmetric volatility and trading volume: The G5 evidence. *Global Finance Journal , 22*, 169-181.

Samonas, M. (2015). *Financial forecasting, analysis, and modelling : a framework for long-term forecasting* (1 ed.). Chichester, West Sussex: Wiley.

Shahzad, S., Hernandez, J., Hanif, W., & Kayani, G. (2018). Intraday return inefficiency and long memory in the volatilities of forex markets and the role of trading volume. *Physica A , 506*, 433-450.

Tabak, B., & Cajueiro, D. (2006). Assessing inefficiency in Euro bilateral exchange rates. *Physica A , 367*, 319-327.

Takeda, F., & Wakao, T. (2014). Google search intensity and its relationship with returns and trading volume of Japanese stocks. *Pacific-Basin Finance Journal , 27*, 1-18.

Tangmongkillert, K., & Suwanna, S. (2016). Asset price and trade volume relation in artificial market impacted by value investors. *Physica A , 450*, 126-133.

Wang, Z., Qian, Y., & Wang, S. (2017). Dynamic trading volume and stock return relation: Does it hold out of sample? *International Review of Financial Analysis* , 1-16.

Wang, S. (2001). The Neural Network Approach to Input–Output Analysis for Economic Systems. Neural Computing & Applications, 10, 22-28.

6.13 Appendices

6.14 Forecast Accuracy Measures

The forecasting accuracy measures were tested and developed in accordance with the following structures:

Error Metrics Tested

Equation 20

Mean Error (ME)

$$ME = \frac{1}{N} \sum_{i=1}^{N} (f_i - k_i)$$

Where:

N = # Forecast / Actuals

F = Forecast

K = Actuals

Equation 21

Mean Absolute Error (MAE)

$$= \text{mean } (|e_i|)$$

Mean Squared Error (MSE)

$$mean\ (e_i^2)$$

Equation 22

Root Mean Squared Error (RMSE)

$$= \sqrt{mean\ (e_i^2)}$$

Equation 23

Mean Percentage Error (MPE)

$$MPE = \frac{1}{N} \sum_{i=1}^{N} (f_i - k_i) \text{ x } 100$$

Equation 24

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} (|f_i| - |k_i|) \text{ x } 100$$

The aforementioned accuracy tests follow the structure built by Robert Hyndman.

CHAPTER 6

6.0 Conclusion

It is highly observable that having external regressors added to the neural network can provide significantly more accurate results. However, we still encounter an obstacle as seen above in the last table which shows all the winning methods from all the companies. Specifically, close_high_low beats all other methods in all frequencies and horizons apart from the weekly horizon. Putting that into perspective, we are able to pin-point which method is more accurate in each paper. Furthermore, during real-life trading we would be able to pick and choose between methods and when to choose which model and when not to choose that model.

By taking each winning method from each paper and progressing that method into the next paper then adding a variable to the newer paper, this gives us the ability to prepare for any significant changes that the new variable brings. Doing so will determine if more is better and if more complicated models result in a better model and thus more accurate results.

The next step would be to implement and test these models and bring them to life with a trading strategy and then test which strategy works better with which model. As observed, the close_high_low model provides the best result in all frequencies apart from the weekly results. From an analytical point of view, to test the results of the papers, we would use the horizon testing and choose where each model was more accurate and choose the appropriate trading strategy accordingly.

Earlier, where we endeavoured to test if AI could overcome the random walk diemma (RWD), which basically meant that less complexity is better than more.

As a starting point, the 80/20 rule, then introduced horizon testing before introducing the high_low variables and finally introducing the trading volume factor into the equation.

As seen in the last table, the errors mix and match on different horizons but we can definitely see consistency from the _close_high_low method. This revealed what trading strategies work best with`each model. We tried to remove volatility and uncertainty, which are significant obstacles for traders. Indeed, by reducing volatility and uncertainty, we remove risk. Reducing risk while increasing returns is referred to as the miracle of trading.

However, looking at the results for all four paper we can see Autoarima performing well all round, carrying significance over frequencies. Regardless of data complexity, Autoarima performed well. While performing well, Autoarima overcame some of the time series diversity as some of the data provided seasonal data where specific models were introduced to test if models dependent on seasonality would beat Autoarima, only if the data had seasonality. However, Autoarima showed a sign of weakness when it was under close pressure from other models, where it beat other models in some but not all horizons; it was the most accurate model in the first and second horizons but then faded away albeit staying close behind other models sometimes. That shows us that there is an issue when it comes to the consistency of Autoarima. Therefore, at the start of this conclusion we illustrated that the winning model is the close_high_low model as consistency is just as important as forecasting. The winning model should be consistent across all horizons more so than frequencies as we would be able to change models when it comes to frequencies.

In paper 3, the volatility and uncertainty of share price was introduced. We used those variables from the company's share price to account for risk.

To conclude, the close_high_low method is the most suitable option, as it showed greater robustness, consistency and accuracy.

6.1 Strategic Trading & Forecasting

Strategic trading follows the most accurate trading tool. The trading tool in this essence, we mean the most accurate percentage of error thus the least error computed. This will be incorporated into future work on how to merge the forecasting that was tested in this paper and strategic trading.

Here, we measure the forecasting accuracy using two different approaches according to the forecasting period. For example, for monthly data we forecast 18 months ahead using all the methods used. After the forecast, we test our six different error matrices for each method and each frequency. After the error has been developed, we analyse the matrices according to a set of ranking points. Here, fewer points are awarded for fewer errors, thus the method with the least points is the more accurate.

The second approach tests the accuracy between methods and horizons. The forecasting here uses different training data to that used in the first approach where we used all data points as training data from the beginning. However, in the second approach we used progressive continuous addition of training data. In the first approach, we used the first 100 data points to forecast 18 horizons; the next data point and training data was re-computed to produce the sequence of 18 horizons. This procedure applies to all of the remaining dataset.

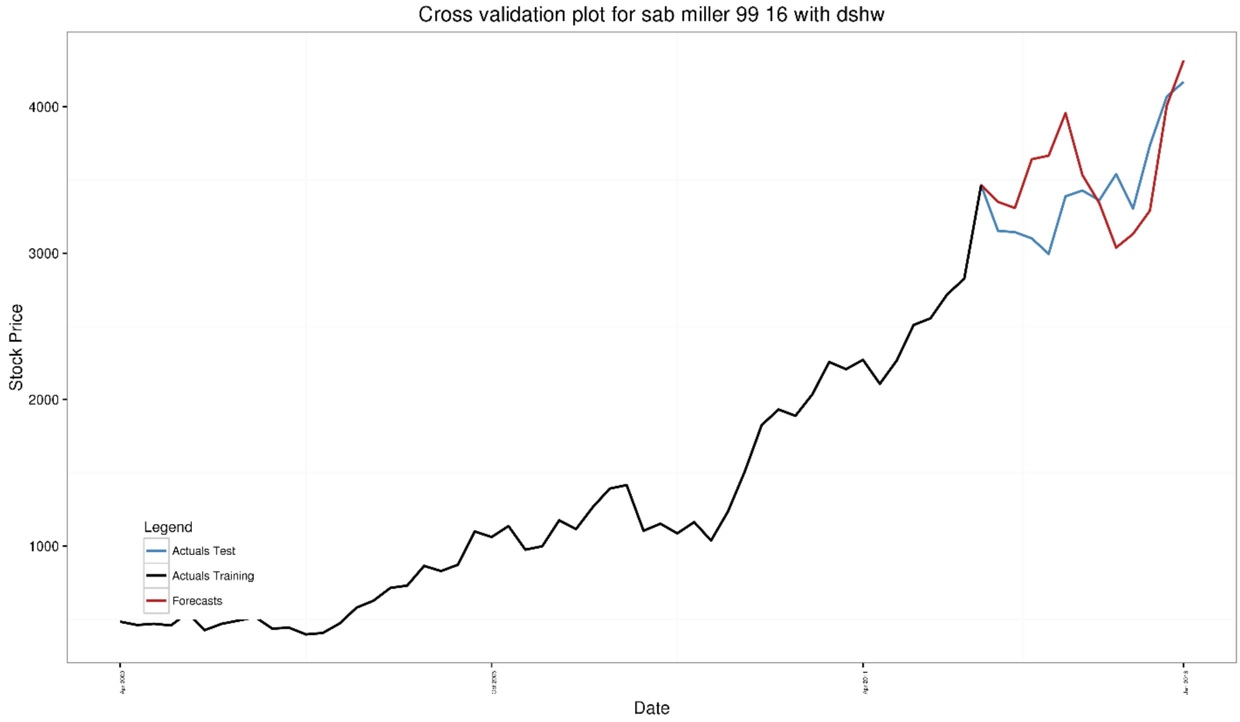
CHAPTER 7

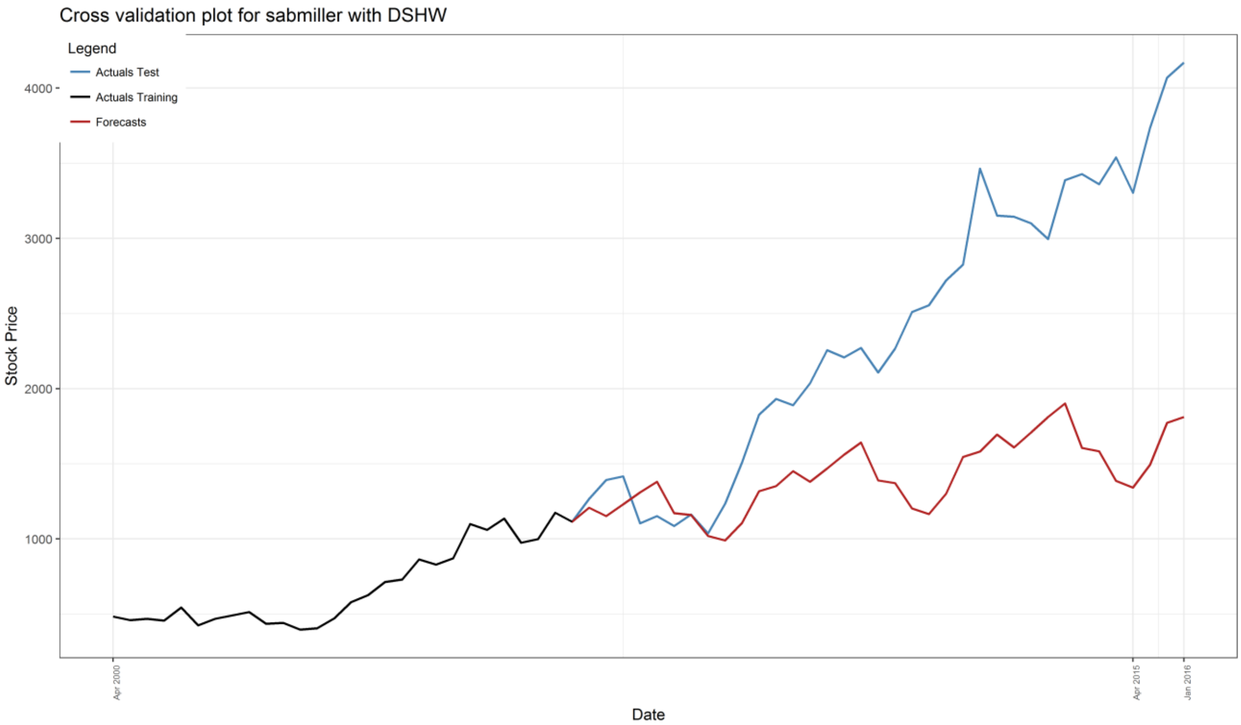
7.0 Appendices

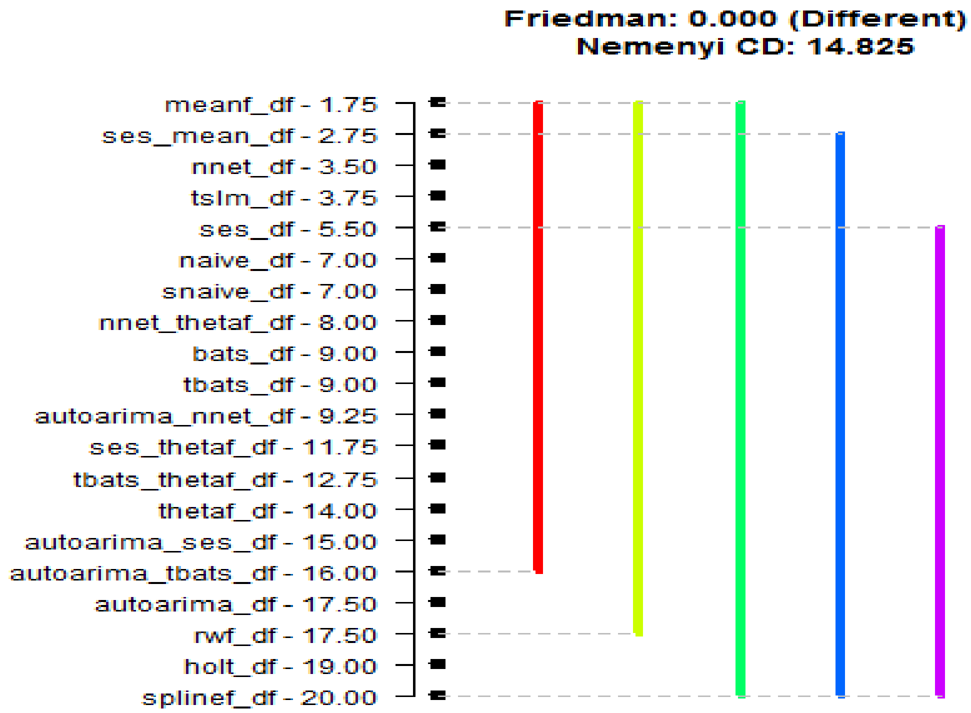
Appendices for Graphs and Significant Tests

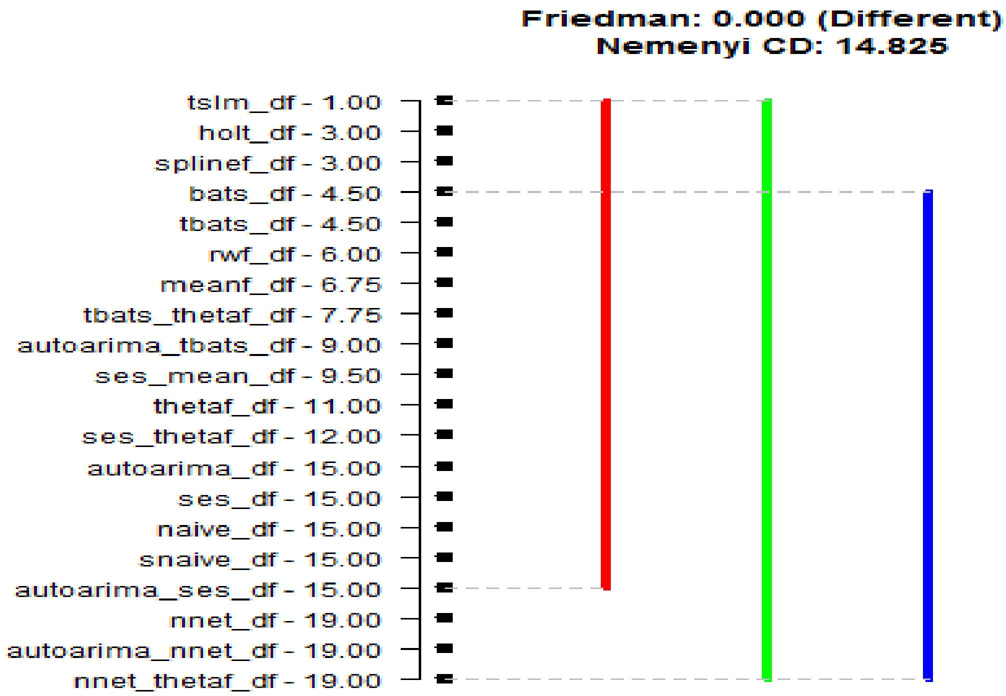
Graph Illustrations

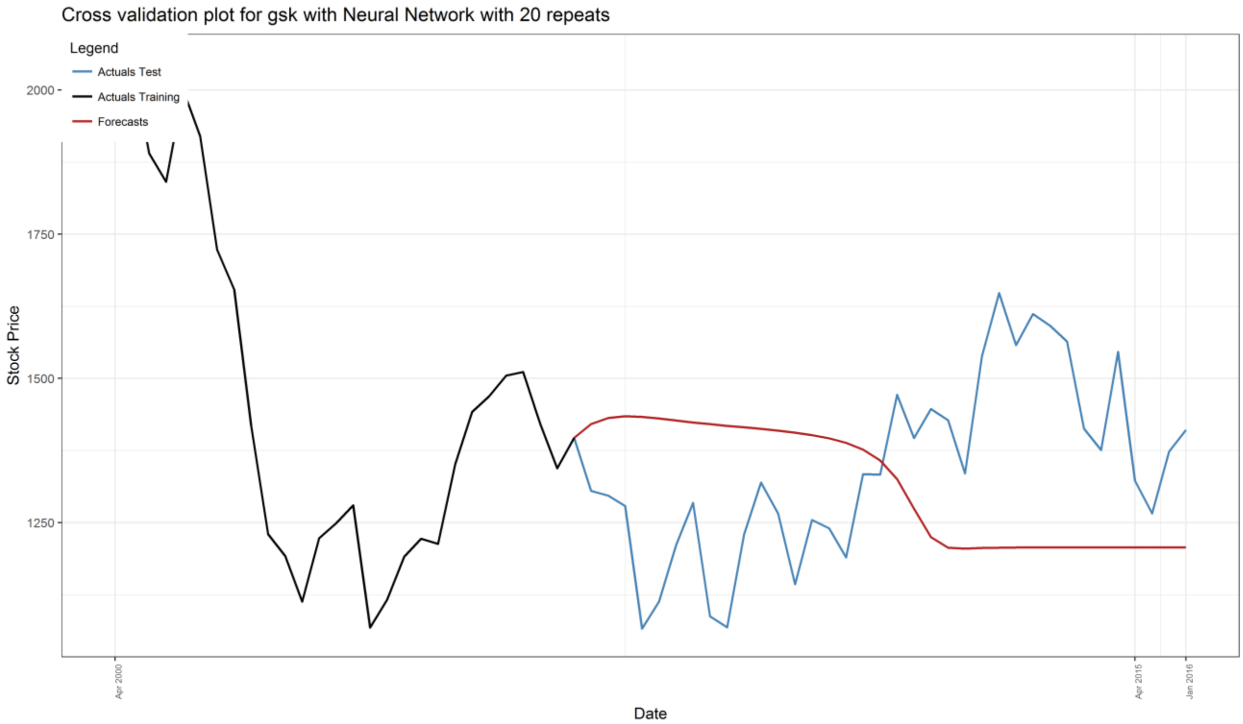
Quarterly

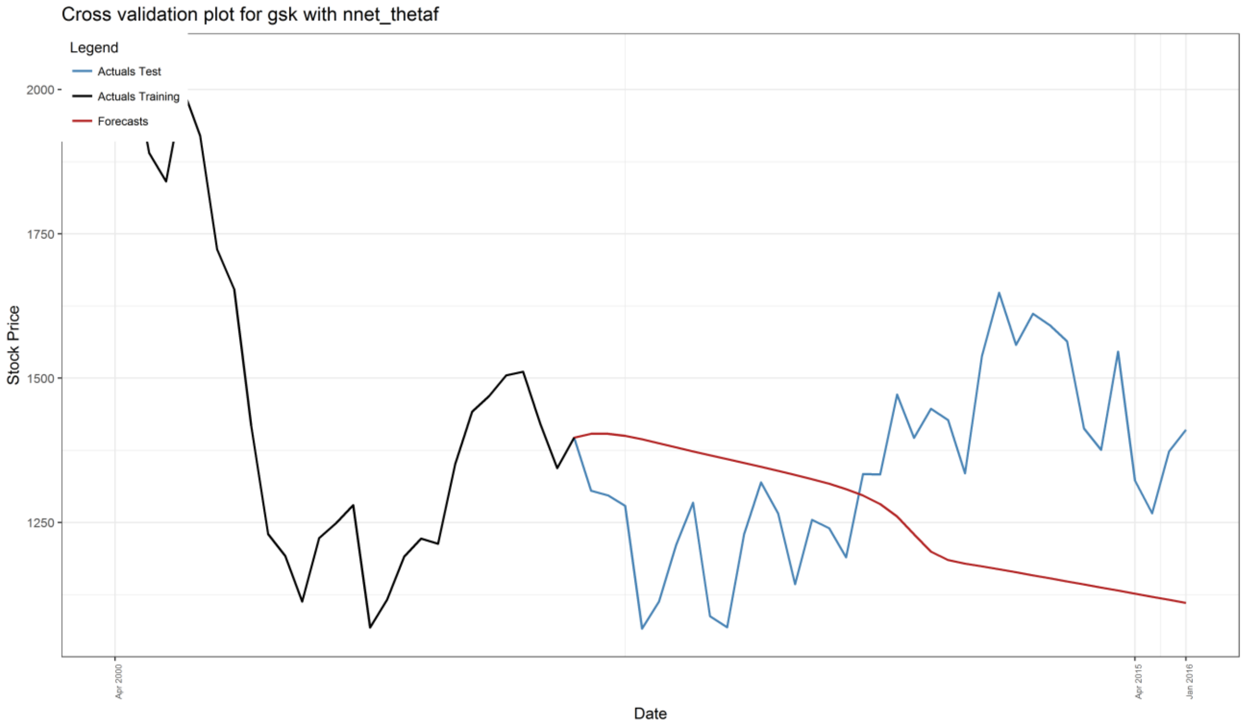


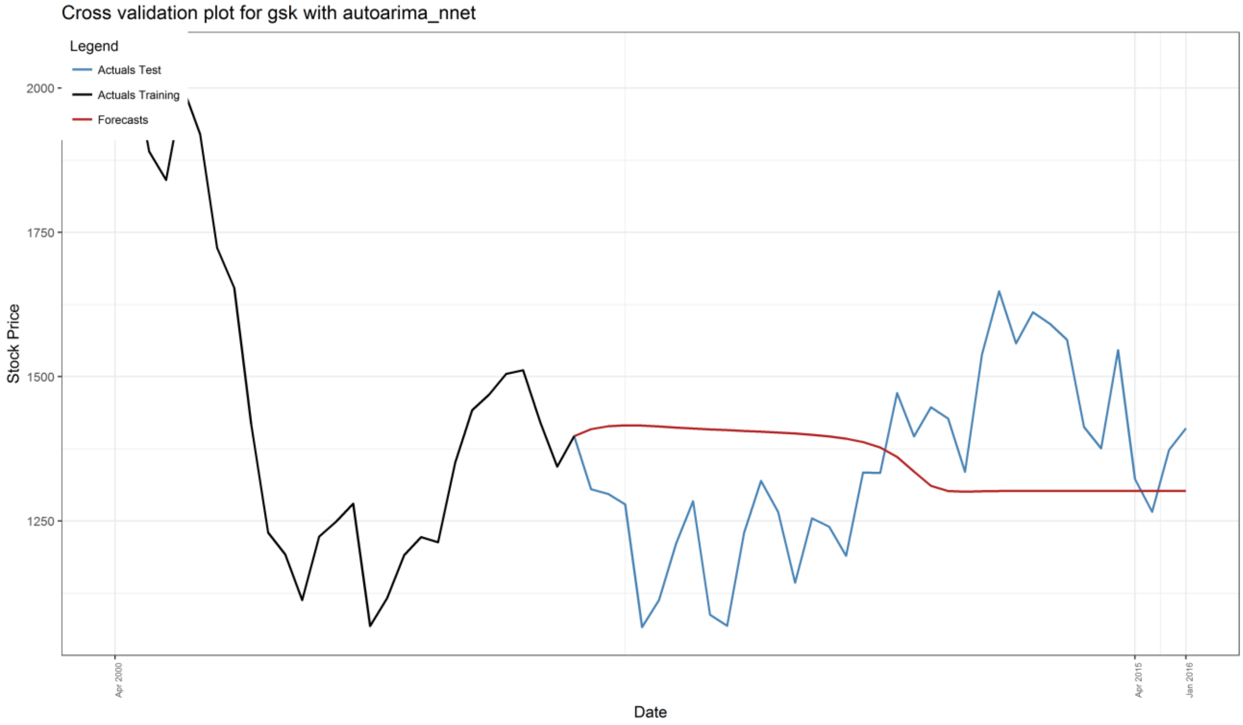



Friedman: 0.000 (Different)
Nemenyi CD: 14.825
meanf_df - 1.75
ses_mean_df - 2.75
nnet_df - 3.50
tslm_df - 3.75
ses_df - 5.50
naive_df - 7.00
snaive_df - 7.00
nnet_thetaf_df - 8.00
bats_df - 9.00
tbats_df - 9.00
autoarima_nnet_df - 9.25
ses_thetaf_df - 11.75
tbats_thetaf_df - 12.75
thetaf_df - 14.00
autoarima_ses_df - 15.00
autoarima_tbats_df - 16.00
autoarima_df - 17.50
rwf_df - 17.50
holt_df - 19.00
splinef_df - 20.00

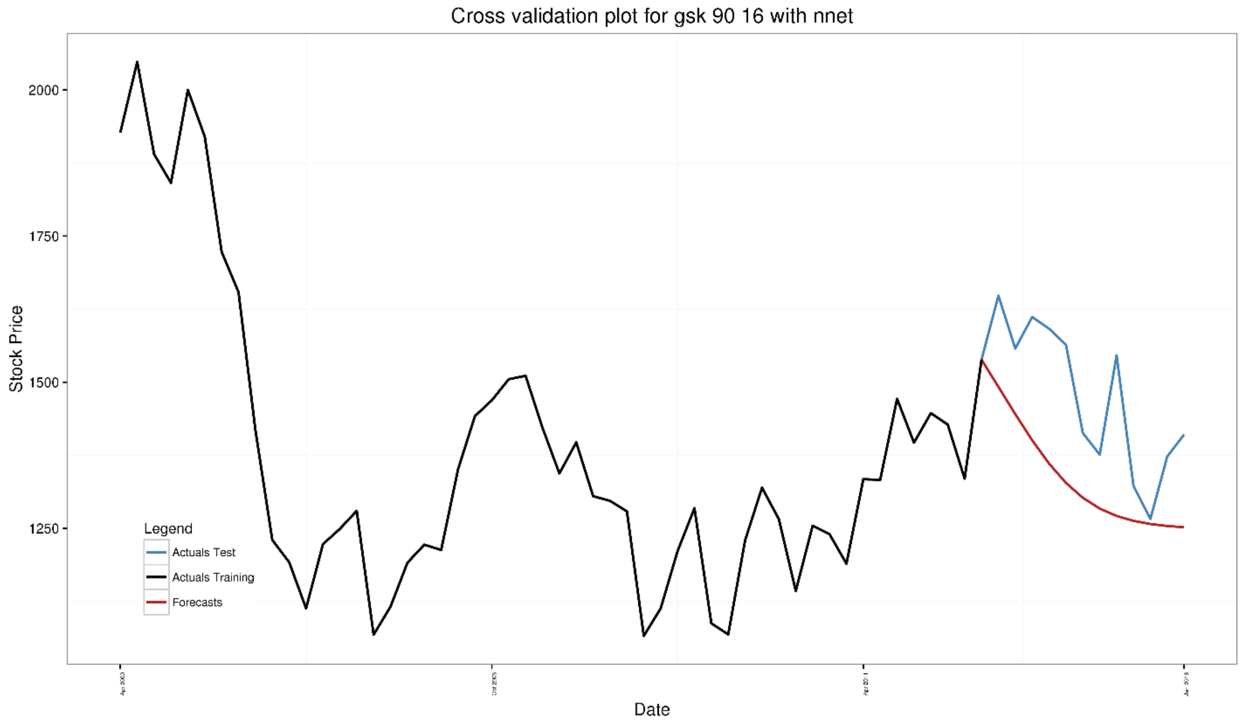

Friedman: 0.000 (Different)
Nemenyi CD: 14.825
tslm_df - 1.00
holt_df - 3.00
splinef_df - 3.00
bats_df - 4.50
tbats_df - 4.50
rwf_df - 6.00
meanf_df - 6.75
tbats_thetaf_df - 7.75
autoarima_tbats_df - 9.00
ses_mean_df - 9.50
thetaf_df - 11.00
ses_thetaf_df - 12.00
autoarima_df - 15.00
ses_df - 15.00
naive_df - 15.00
snaive_df - 15.00
autoarima_ses_df - 15.00
nnet_df - 19.00
autoarima_nnet_df - 19.00
nnet_thetaf_df - 19.00

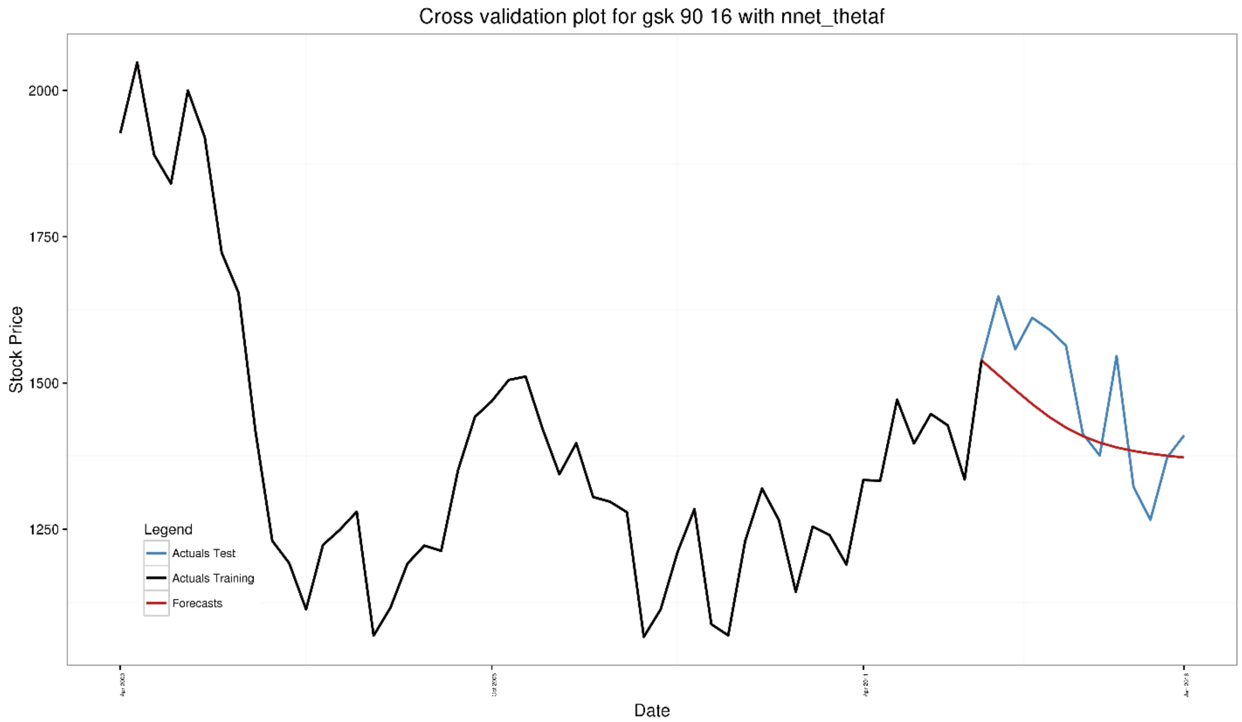

Cross validation plot for gsk with Neural Network with 20 repeats
Legend
Actuals Test
Actuals Training
Forecasts
2000
1750
1500
1250
Stock Price
Apr 2000
Apr 2015
Jan 2016
Date


Cross validation plot for gsk with nnet_thetaf
Legend
Actuals Test
Actuals Training
Forecasts
2000
1750
1500
1250
Stock Price
Apr 2000
Apr 2015
Jan 2016
Date

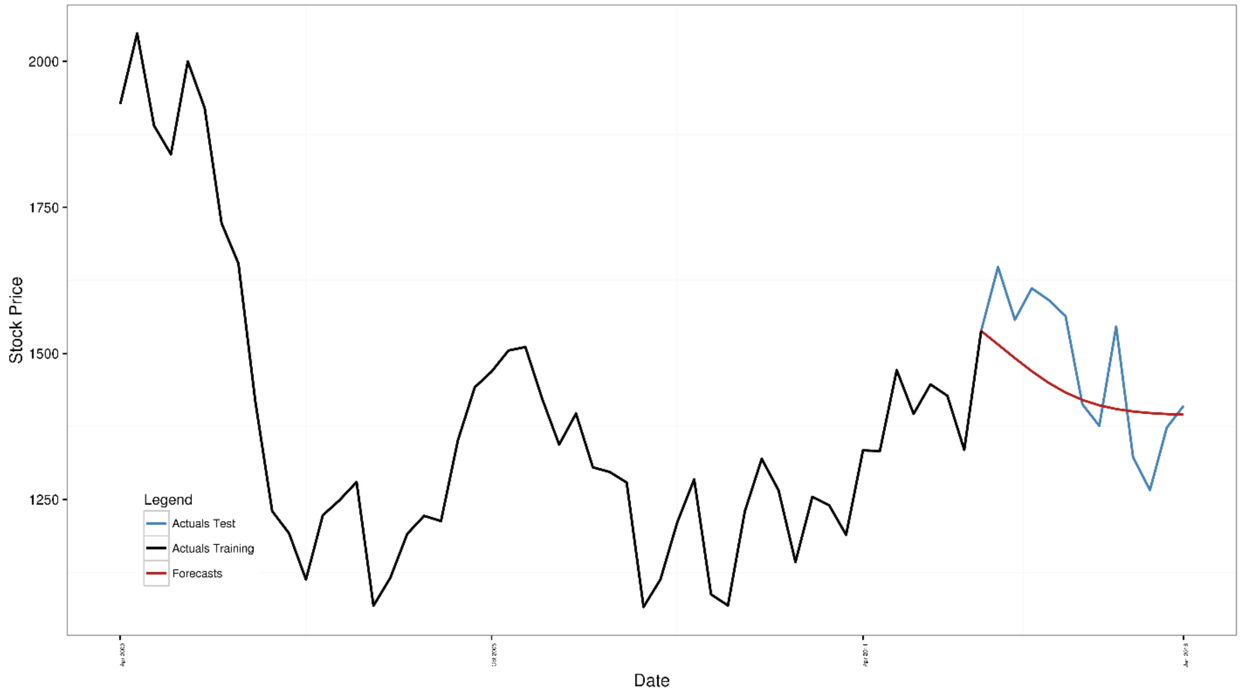

Cross validation plot for gsk with autoarima_nnet
Legend
Actuals Test
Actuals Training
Forecasts
Stock Price
2000
1750
1500
1250
Apr 2000
Apr 2015
Jan 2016
Date


Cross validation plot for gsk with Theta Model
Legend
Actuals Test
Actuals Training
Forecasts
Stock Price
1800
1500
1200
Apr 2000
Apr 2015
Jan 2016
Date

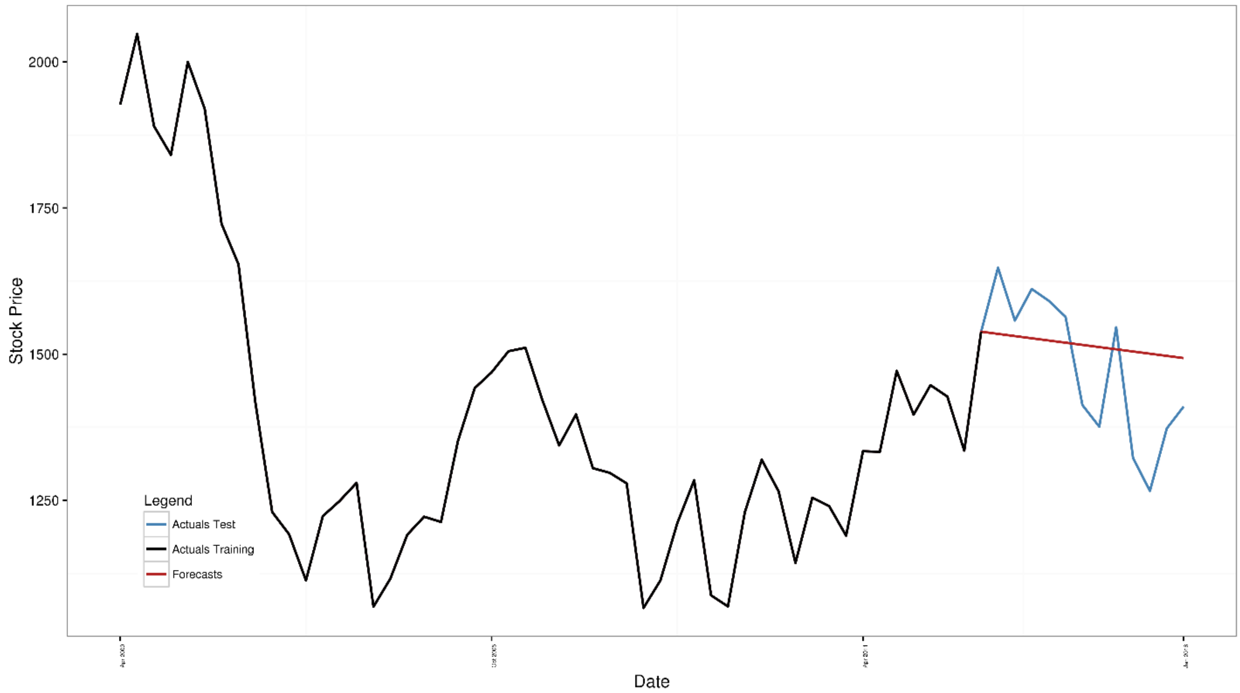

Cross validation plot for gsk 90 16 with nnet
Stock Price
2000
1750
1500
1250
Date
Legend
Actuals Test
Actuals Training
Forecasts


Cross validation plot for gsk 90 16 with nnet_thetaf
Stock Price
2000
1750
1500
1250
Date
Legend
Actuals Test
Actuals Training
Forecasts

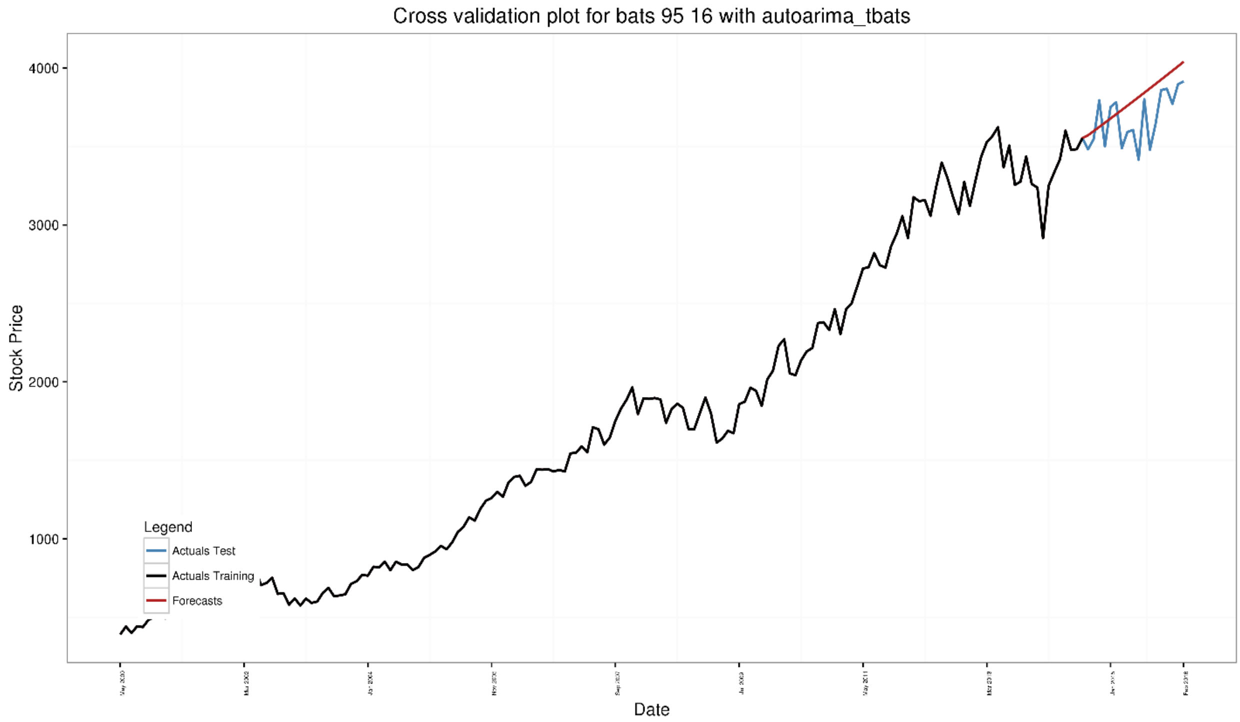
Cross validation plot for gsk 90 16 with autoarima_nnet

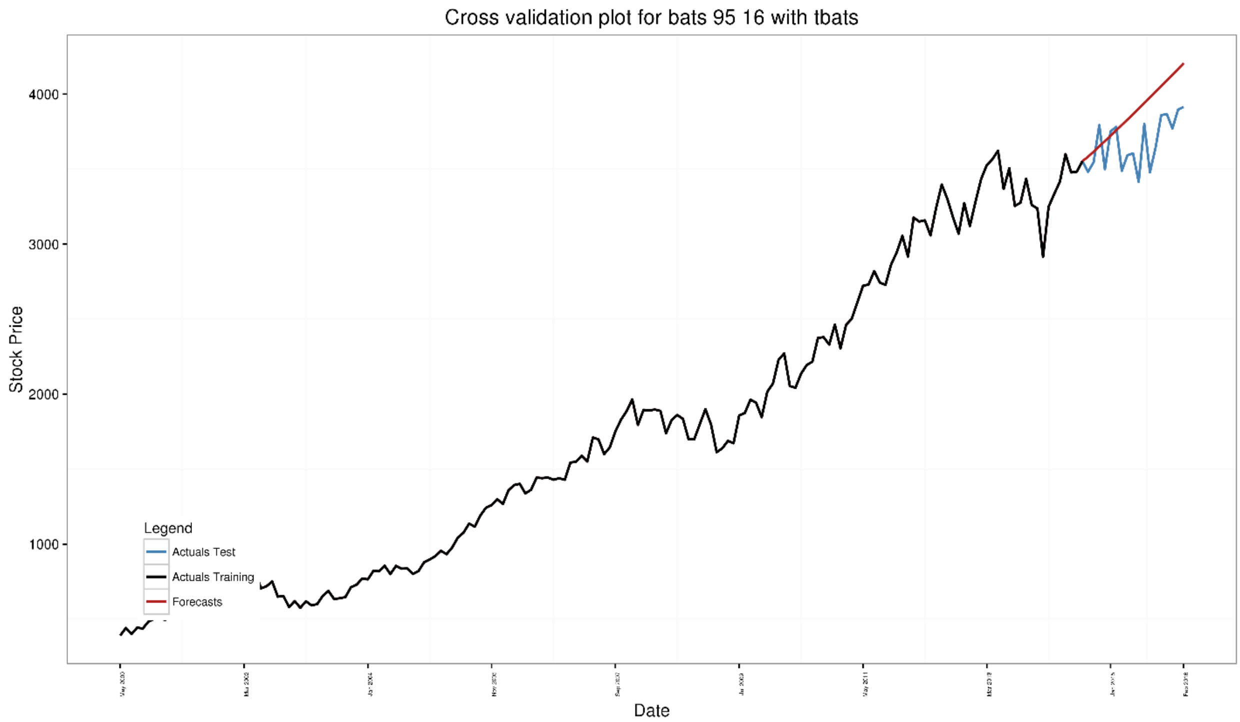


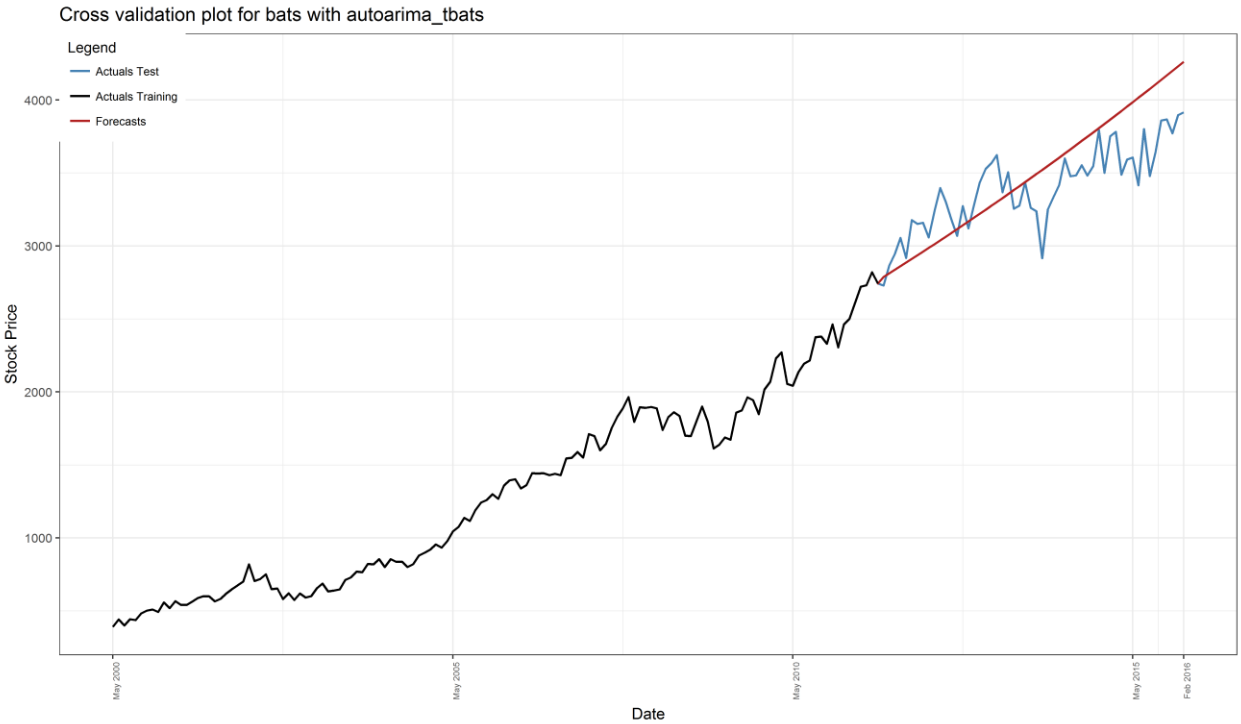
Cross validation plot for gsk 90 16 with thetaf

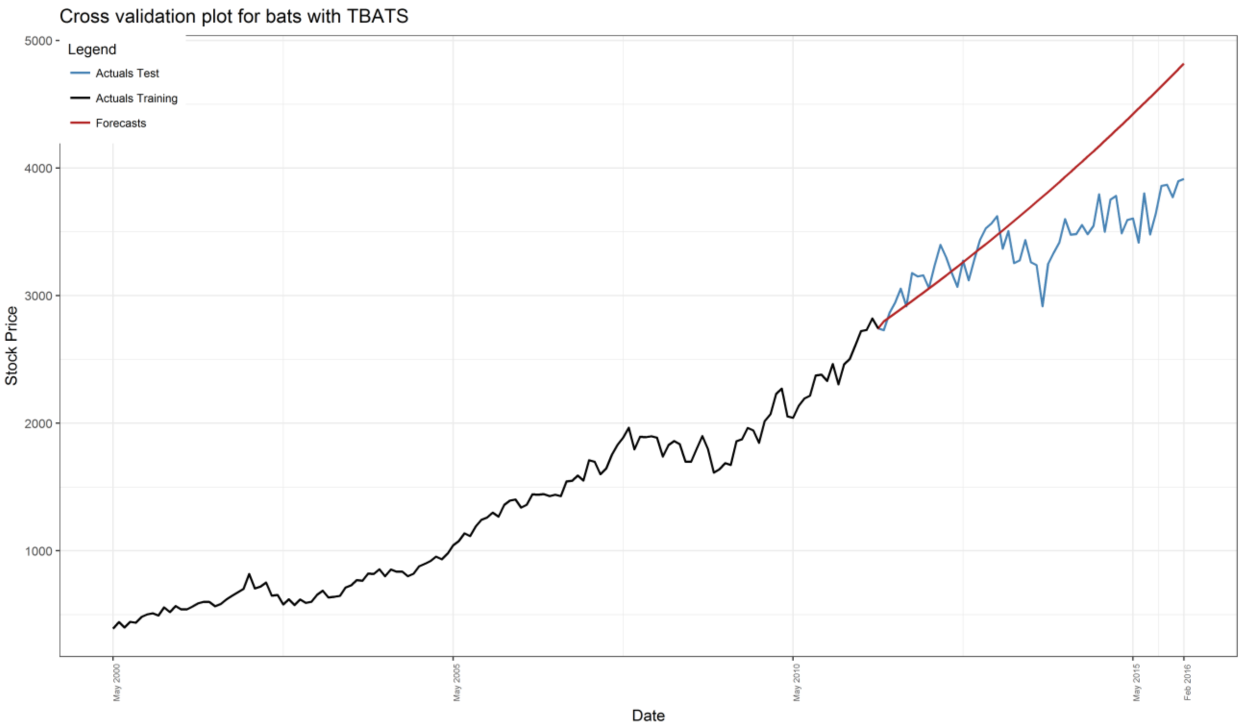


Monthly


Cross validation plot for bats 95 16 with autoarima_tbats
Stock Price
4000
3000
2000
1000
Date
Legend
Actuals Test
Actuals Training
Forecasts


Cross validation plot for bats 95 16 with tbats
Stock Price
4000
3000
2000
1000
Date
Legend
Actuals Test
Actuals Training
Forecasts

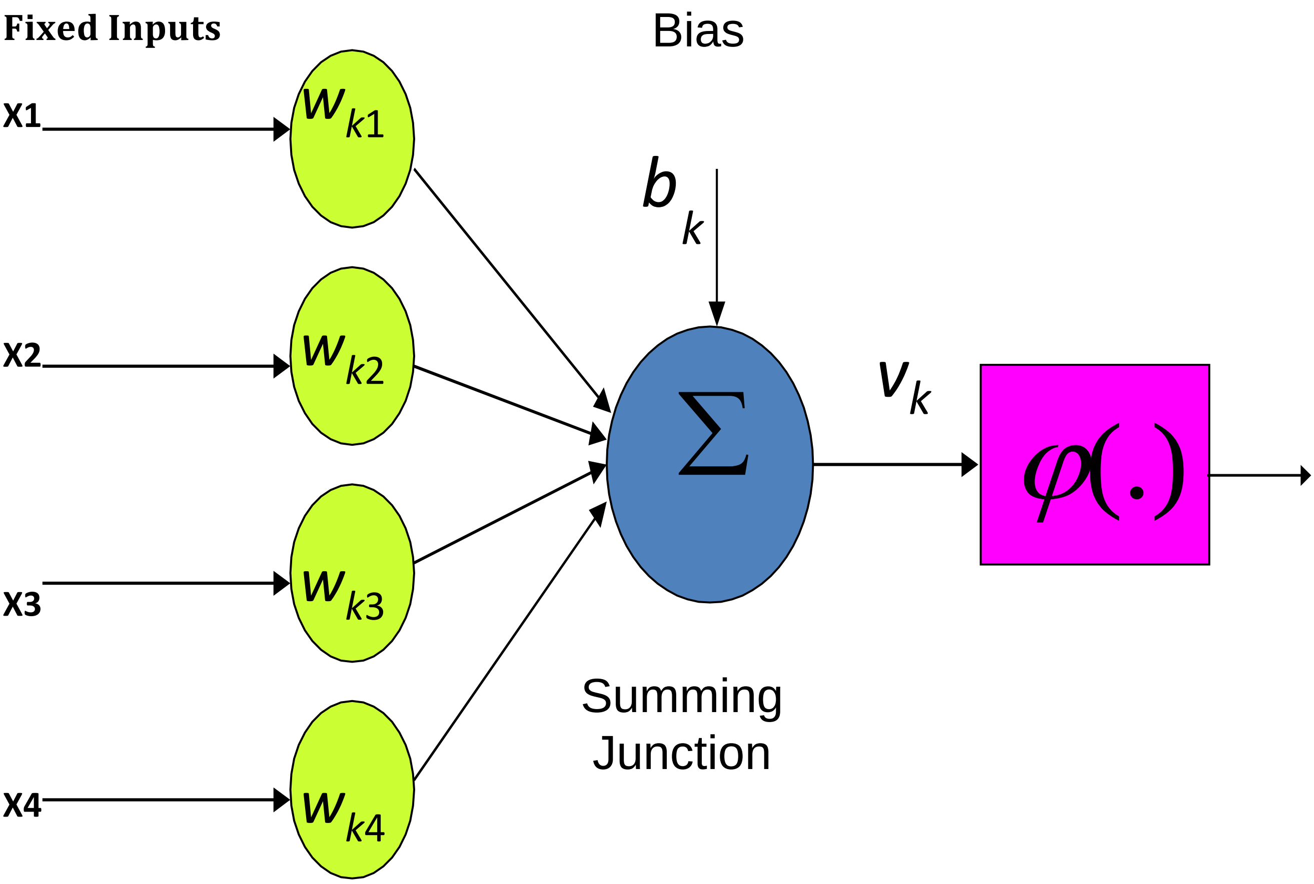

Cross validation plot for bats with autoarima_tbats
Legend
Actuals Test
Actuals Training
Forecasts
Stock Price
4000
3000
2000
1000
May 2000
May 2005
May 2010
May 2015
Feb 2016
Date


Cross validation plot for bats with TBATS
Legend
Actuals Test
Actuals Training
Forecasts
Stock Price
5000
4000
3000
2000
1000
May 2000
May 2005
May 2010
May 2015
Feb 2016
Date

Appendices for Neural Networks

Neural Networks

Single-layer perceptron



The single-layer network above can be represented as:

$$\mu_k = \sum_{j=1}^{m} w_{kj}\, x_j$$

Where:

k = is the neuron

w = is the weight allocated to that neuron

x = the input

$j = 1$ Here as the bias function is not simplified in the equation

And

$$y_k = \varphi(\mu_k + b_k)$$

Where:

y_k = the output

φ = the activation function, for more details see (activation function 1.1)

Where μ_k is the linear combiner output.

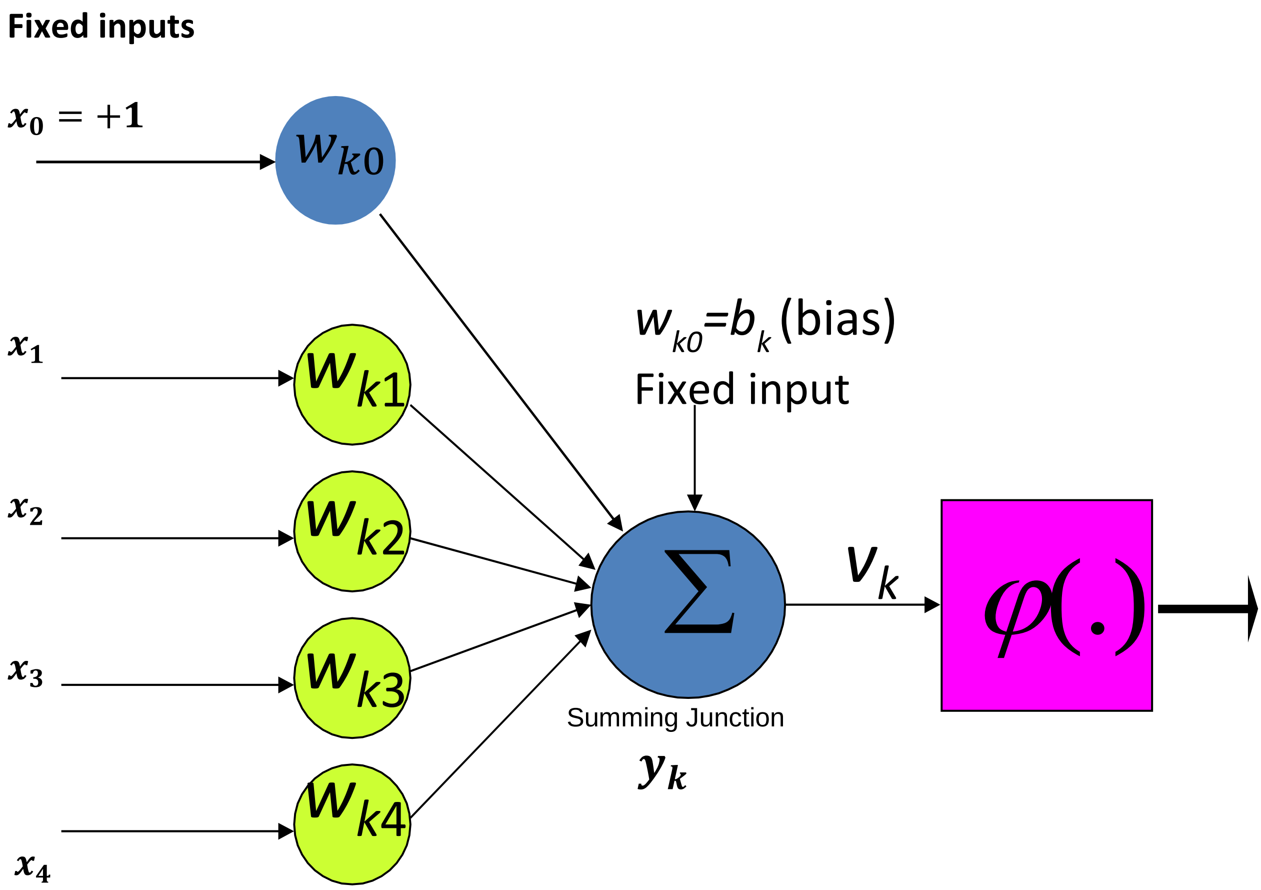
It can be simplified as:

$$\mu_k = \sum_{j=0}^{m} w_{kj}\, x_j$$

Note that the change of limits has changed from 1 to 0

And the output is:

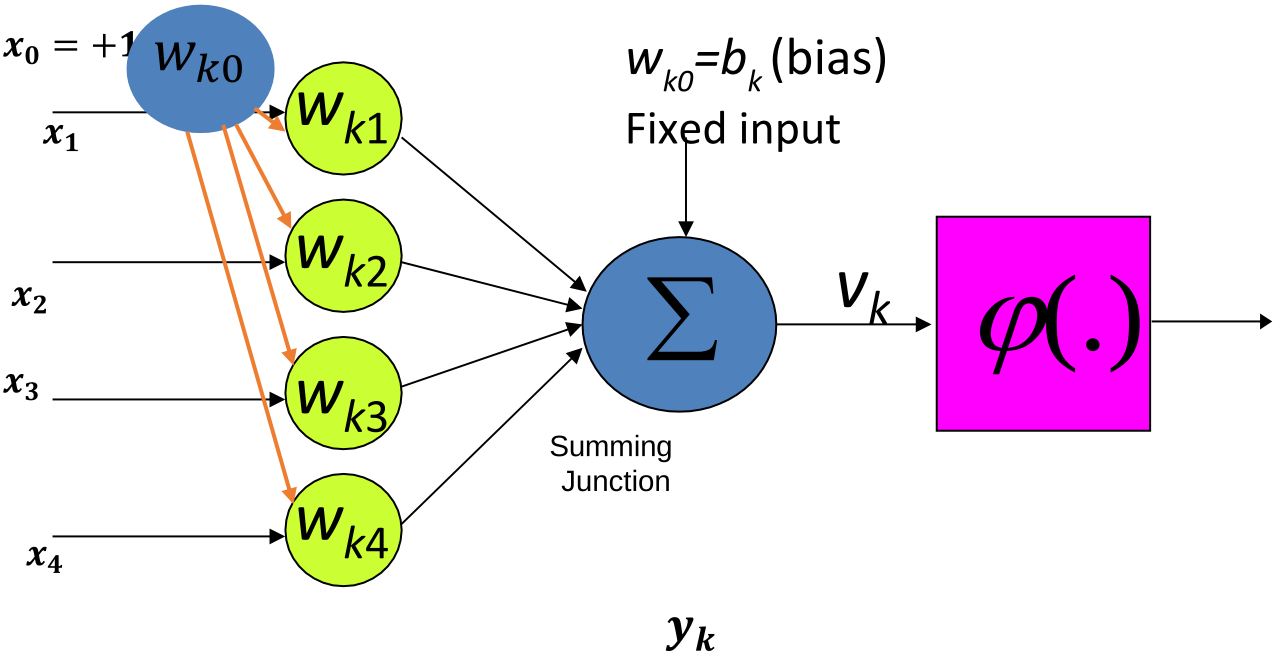
$$y_k = \varphi(v_k)$$


Fixed inputs
$x_0 = +1$
w_{k0}
x_1
w_{k1}
x_2
w_{k2}
x_3
w_{k3}
x_4
w_{k4}
$w_{k0}=b_k$ (bias)
Fixed input
Σ
Summing Junction
y_k
v_k
$\varphi(.)$

Where:

w_{k0} = the bias function

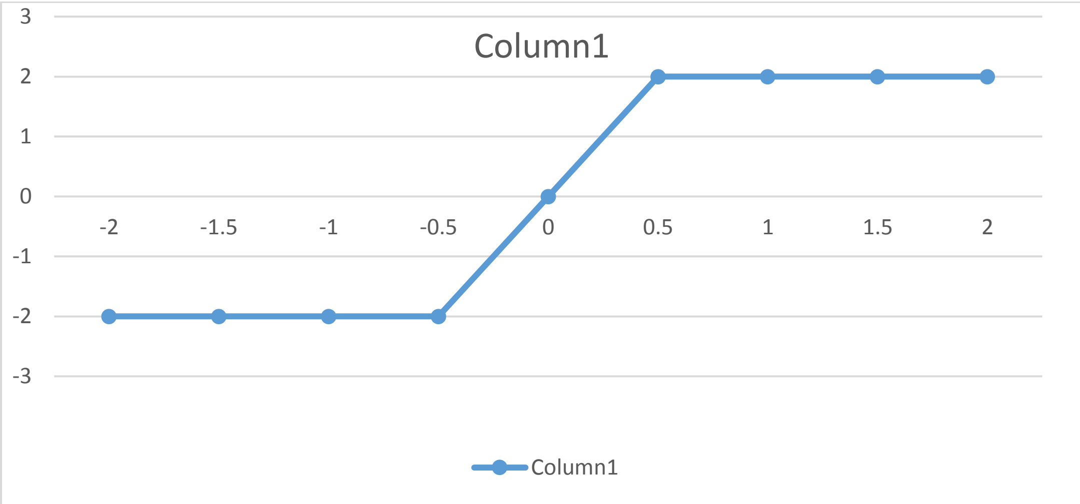
Fixed inputs



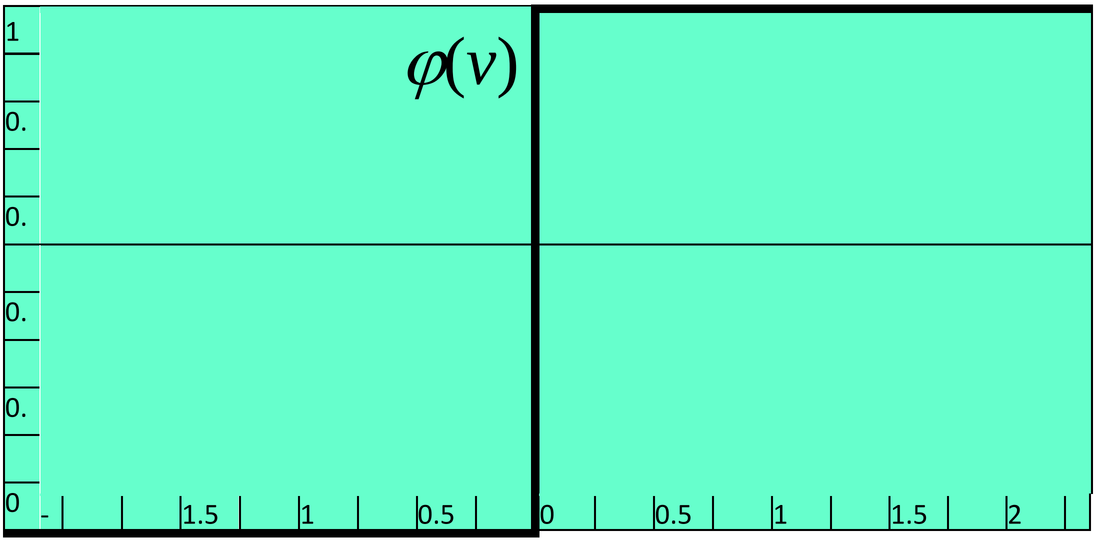
In this bias function, the bias fixed function is calculated after the inputs in the weights and then the bias function is calculated back into all the given weights within the hidden layer. All other equations remain consistent.

Activation function neural networks

Heaviside Step



It can be illustrated as the graph above or the one below



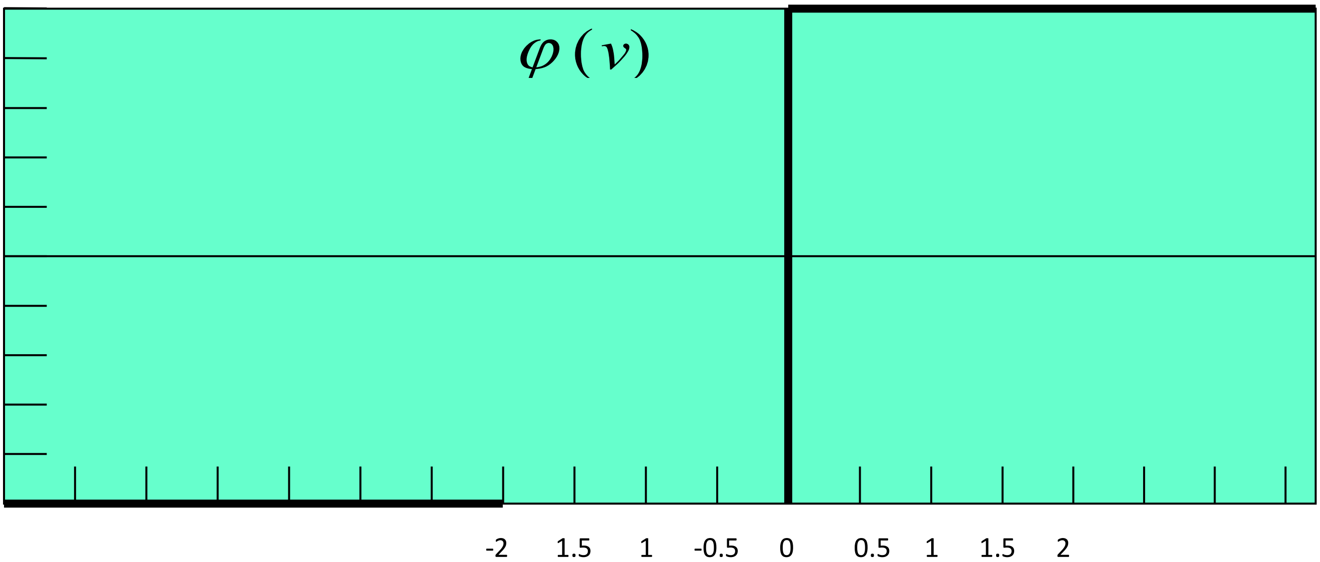
It is defined as:

$$\varphi(v)\{1\ if\ v \succcurlyeq 0$$

$$\varphi(v)\{0\ if\ v \prec 0$$

Output 0 or 1

Piecewise-Linear Function



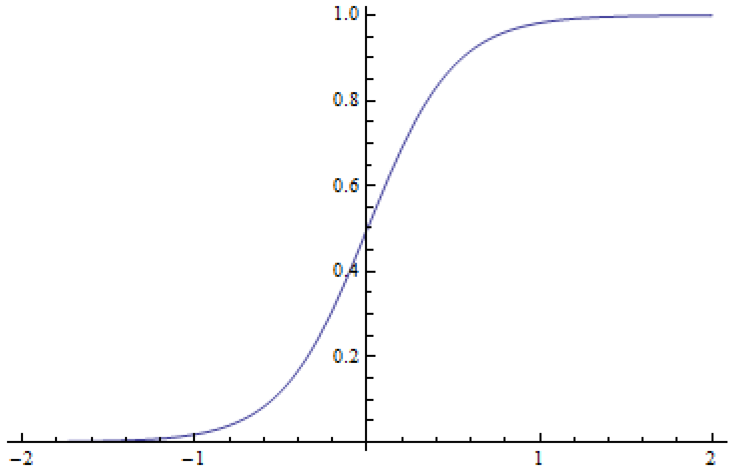
Defined as:

–

$$\varphi(v) = \begin{cases} 1, & v \geq +\frac{1}{2} \\ v, & +\frac{1}{2} > v > -\frac{1}{2} \\ 0, & v \leq -\frac{1}{2} \end{cases}$$

Sigmoid Function:

$$\varphi(v) \frac{1}{1+exp^{-t}}$$



Hyperbolic Tangent

$$y = \tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$$