Proceedings of the 2nd International Conference on Management Research and Economic Development
DOI: 10.54254/2754-1169/86/20240887

Research and Forecasting of the FTSE100 Index over Long Time Series

Yangyuhui Zhao[1,a,*]

[1]*Mathematics department, Imperial College London, Exhibition Road, London, United Kingdom*
a. steve.zhao22@imperial.ac.uk
**corresponding author*

Abstract: In order to research the UK and Global stock market, FTSE100 index is one of the most important values to study. This paper uses several models to forecast the future curve of the FTSE100 and compare these models to find a best way on forecasting. Then by analyzing the timeseries, the author shows several factors that might affect the timeseries which will be useful on the further forecast of the market. By using the simple forecast method, including Mean, Naïve, Snaive and Drift, Holt and Holt-winter model, and ARIMA model on the closing price of FTSE100 from Feb.2004 to Feb.2024. The result show that Holt-winter is the best model, and the mean is the worst. Also, by considering the 100 companies of the FTSE100 the factors that will affect the time series are Energy Boom, Global Economy, and Foreign Exchange Market. Therefore, the article might give investors some idea about the UK stock market.

Keywords: Futures, forecasting, Volatility, Statistics-model

1. Introduction

1.1. Research Background and Motivation

The goal of this research's in-depth examination of the FTSE 100 Index, also referred to as the Financial Times Stock Exchange 100 Index, is to find how cyclical and seasonal patterns performed in the long-term time series data. The FTSE 100 Index, introduced on January 3, 1984, serves as a performance benchmark for most investors. An index weighted by market capitalization is the FTSE 100, unlike FT30. The top 100 qualified UK corporations make up this list and the index are determined by multiplying the share price by the total number of shares issued to arrive at their complete market value. In FTSE indices, share prices are adjusted for free-float capitalization, giving greater weight to larger companies with more freely traded stock. The basic formula for these indices is:

$$Index\ level = \frac{\sum_i Price\ of\ sotck \times Number\ of\ shares \times Free\ float\ adjustment\ factor}{Index\ divisor} \tag{1}$$

What represents the proportion of issued shares available for trading is the correction factor for free float, rounded to the next multiple of five percent. When figuring out the free-float capitalization of an organization, multiply its market capitalization by this adjustment factor. Notably, restricted stocks held by company insiders are excluded from this calculation. According to Figure 1, in the last five years, from 2019 to early 2020, the FTSE 100 performed well during this period, reaching historical highs [1]. However, it was severely impacted by the COVID-19 pandemic, leading to a rapid decline. From early 2020 to year-end 2020, The global market turmoil caused by COVID-19 resulted in a significant drop in the FTSE100. Despite government measures to mitigate economic effects, the index suffered. The at the year-end of 2020 to early 2021, with covid vaccines being rolled out and economic recovery underway, the FTSE100 restarted to recover. Between the period between 2021 and 2022, the index remained relatively stable, buoyed by global economic recovery and market confidence. From 2022 to present, the FTSE100 continued to be influenced by global economic and political factors. While there were fluctuations, the overall trend remained relatively steady. Therefore, this article is going to forecast the FTSE100, analyze the factor of the curve fluctuation and provide some views on the basic framework of the UK stock market to help economists, investors, and policymakers make sound decisions.

Figure 1: FTSE100 index from Feb.2019 to Feb.2024

1.2. Literature Review

In the field of time series analysis, various research endeavors focus on forecasting. This includes several significant tasks: There was a study show that, precisely predicting stock prices in the context of stock market research is a crucial problem facing the capital investment sector. One approach explored in this paper is the use of Support Vector Machines (SVMs), a cutting-edge neural network approach for regression. Examining the viability of SVMs regression in stock price prediction is the aim [2]. Again, there was a study focused on forecasting Uniter Kingdom (UK) stock market volatility by using GARCH model.

This groundbreaking study by David McMillan (2000) focuses on FTSE100 index-based stock market volatility forecasting in the context of the United Kingdom. The paper explores the complexities of volatility prediction, which is an essential tool for financial market regulation, option pricing, and investment decisions [3]. Besides, the author David Mendonca Pinho1 undertakes a comprehensive comparison of various volatility models including Naïve, ARIMA, and Univariate GARCH model and their combinations for forecasting volatility in the FTSE 100 index. The research focuses on day-ahead forecasts, which are crucial for risk management, option pricing, and investment decisions [4].

1.3. Research Contents

In order to accurately predict future market trends, this research combines the basic and advanced forecasting techniques to conduct analysis, including mean, naive, seasonal naïve (Snaive), Holt, Holt-Winters, and AutoRegressive Integrated Moving Average (ARIMA) models. Further depth to

DOI: 10.54254/2754-1169/86/20240887

the research is also conducted by analyzing the advantages and disadvantages of several method and how it affects market investment opportunities, including the market Sentiment and Confidence, Sector Exposure, Global Economic Indicators and Currency Exchange Rates.

2. Methodology

To find the best forecasting method, this article will focus on the closing price which founded on the website of London stock exchange [1]. By using the autoplot, the author draw a graph of the FTSE 100 Index (see figure 2).

Figure 2: FTSE100 index from Feb. 2004 to Feb. 2024

Here is the description of the method that will be included in this article.

Autocorrelation Function (ACF): Calculating a time series' linear connection between its lagged values [5].

Mean: Set all the forecasts of all future values by using the mean value of historical data

Naïve: Using the last observation value on all the forecasts.

Snaive: Using the value last seen in the same season on each forecast.

Drift: Drift is defined as the mean variation observed in the past data.

Holt: To handle data with a trend, Holt (1957) extended the idea of simple exponential smoothing. Two smoothing equations—one for the level and another for the trend—as well as a forecast equation are introduced by this method [6].

Holt-Winter: To take seasonality into consideration, Holt (1957) and Winters (1960) expanded Holt's methodology. A forecast equation and three smoothing equations—one for the level, one for the trend, and one for the seasonal component—each with matching smoothing parameters are included in the Holt-Winters seasonal technique [7].

Autoregressive model (AR model): To forecast the variable of interest, a linear combination of the variable's historical observations is utilized.

ARIMA model: Integrate a moving average model and autoregression with differencing.

3. Analysis of the FTSE100 by Using Statistic Models

The linear relationship between lagged values in a time series is quantified by the Autocorrelation Function (ACF). Meanwhile, accounting for the effects of all smaller lags, the Partial Autocorrelation Function precisely assesses the link between a stationary time series and its own lag values. Notably,

DOI: 10.54254/2754-1169/86/20240887

the first partial autocorrelation is equivalent to the initial autocorrelation because there are no intermediate lags to eliminate. In practical terms, each partial autocorrelation can be estimated as the final coefficient in an autoregressive model [8]. Then by using ACF test and PACF test, getting the Figure 3 and Figure4 respectively:

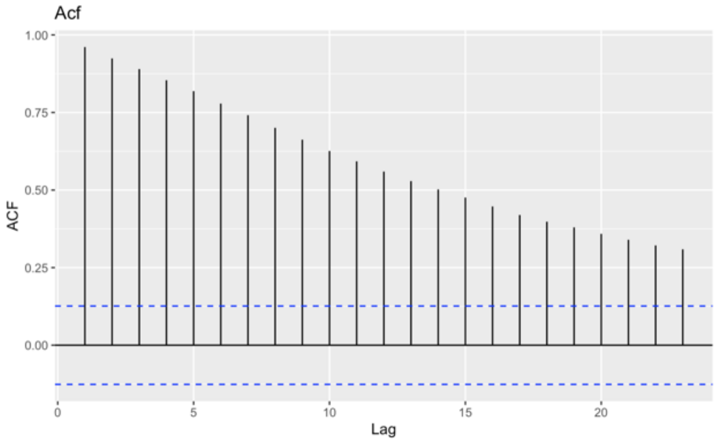

Figure 3: ACF test on FTSE100

According to the figure 3, it shows that the index follows the trend pattern as the ACF of trended time series tend to have positive values that is inversely proportional to lags' value.

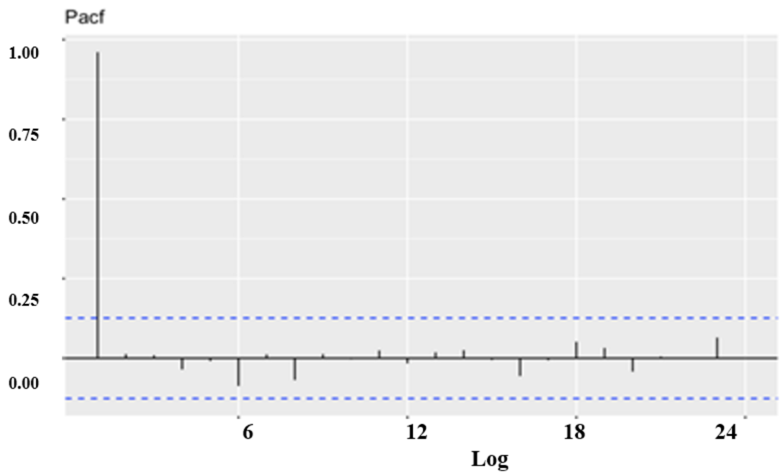

Figure 4: PACF test on FTSE100

According to the figure 4, it shows that there is a meaning correlation at the initial lag, followed by insignificance in later lags, which suggests the presence of AutoRegressive (AR) terms. The number of significant correlations indicates the order of AR terms. In this Figure4, lag 1 shows significant correlation, while subsequent lags do not, indicating an AR order of 1.

After that the author must test the time series is stationary or not, as the author need to use the ARIMA model in the following research. By testing the stationary, it shows the value is far bigger than the critical value, the test fails to reject the null hypothesis of Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, the time series is trend stationary [9].

After that, the author used the stl. function which is used for Seasonal Decomposition of Time Series by Loess (STL). It decomposes a time series into three components: Seasonal: Extracts the seasonal pattern. Trend: Captures the underlying trend. Irregular (Remainder): Represents the residuals after removing the seasonal and trend components.

DOI: 10.54254/2754-1169/86/20240887

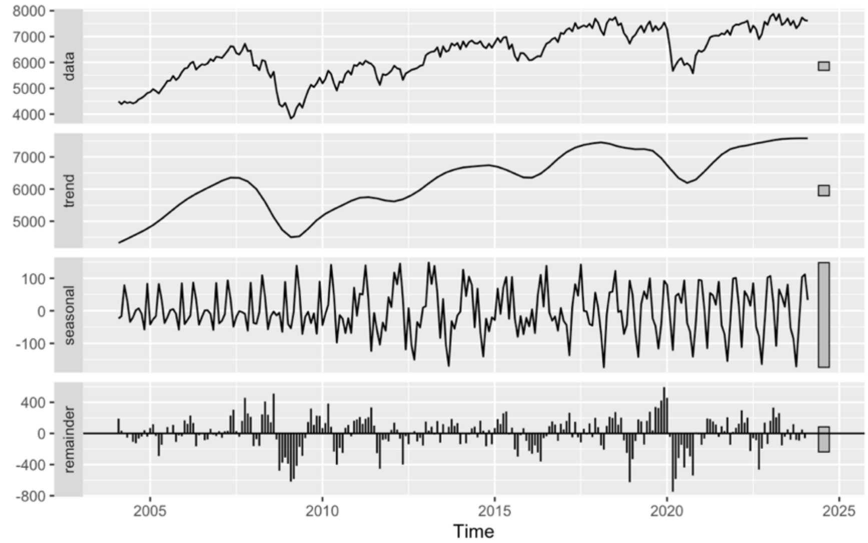

Figure 5: STL graph for FTSE100

From the seasonal part of figure 5, the author can preliminarily say that the time series also follow a seasonal pattern. The chart clearly exhibits periodic oscillations, suggesting a seasonal pattern. Apart from that in the chart, the test can be observed peaks and troughs in the seasonal pattern. Larger amplitudes indicate more pronounced seasonal changes.

In addition, the author uses some simple forecasting method to forecast the timeseries, including, mean, naïve, Snaive and drift.

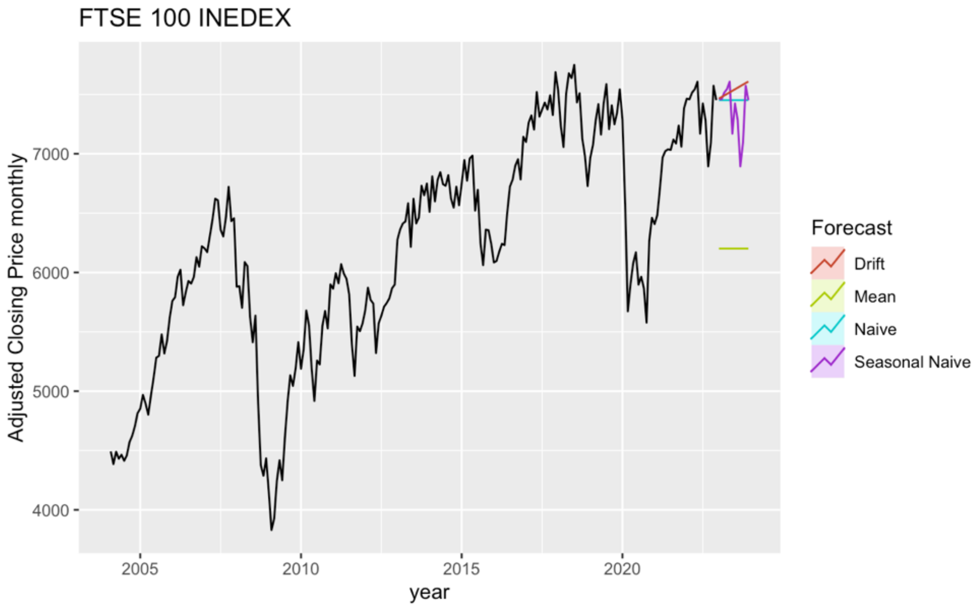

Figure 6: All simple forecast in one graph

Then the author uses the accuracy function to check whether these forecast methods useful or not. And get several values that shown as follow of the predicted value.

ME: The mean error between predicted values and actual values. It indicates the total bias of the model. A value close to zero suggests good performance.

RMSE: The square root of the average of squared errors. It measures the spread of errors and is sensitive to outliers. Smaller RMSE values indicate better fit.

DOI: 10.54254/2754-1169/86/20240887

MAE: The average absolute value of predicted value minus actual value. It provides a robust measure of error, less affected by outliers.

MPE: The mean percentage variation between the expected and observed values. Underestimation is shown by negative values, and overestimation is indicated by positive values.

MAPE: The average absolute percentage difference. It's a relative measure of error, useful for comparing across different scales.

MASE: Compares model performance to a naive model (e.g., persistence model). A smaller MASE indicates better performance.

ACF1 (First-order Autocorrelation Coefficient): Measures the correlation between consecutive residuals in time series data. Close to 1 suggests good autocorrelation.

Theil's U: A statistic that compares forecasted values to actual values. Lower values indicate better fit.

Then by using Holt and Holt-winter method on the timeseries.

Holt: The forecast function no longer remains constant; instead, it exhibits a trend.

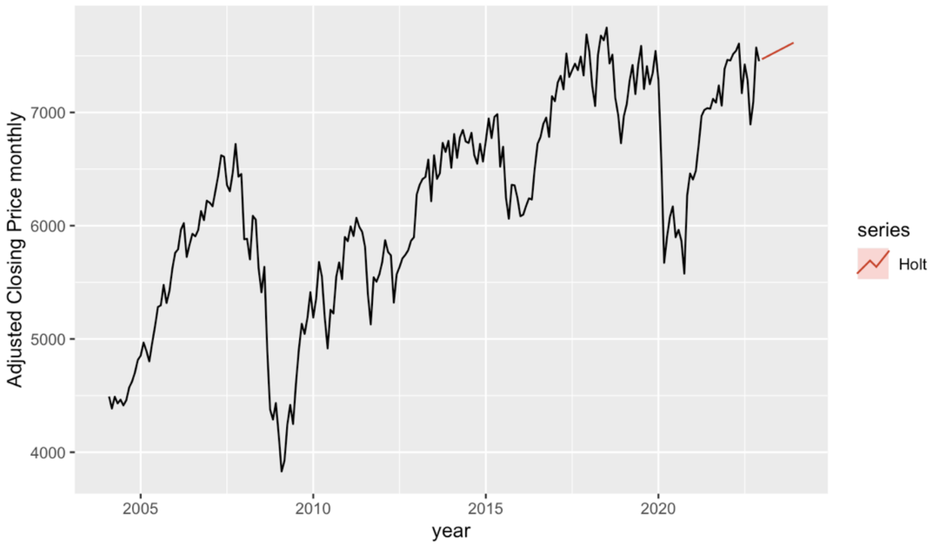

Figure 7: Hole Forecast for FTSE100

Besides, the author also uses Holt-winter's (HW) method (both additive and multiplicative) to capture seasonality. There are two distinct methods for this process, and they vary according to the seasonal component. The multiplicative technique is more acceptable when seasonal variations fluctuate in proportion to the overall level of the series, but the additive approach is advised when seasonal variations are mostly stable throughout the series.

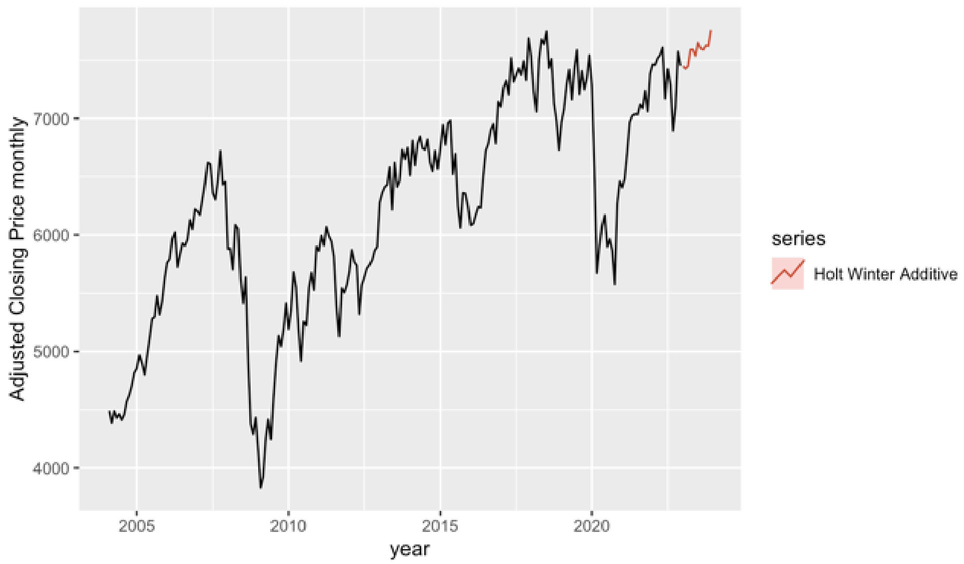

Figure 8: Holt Winter Additive Forecast for FTSE100

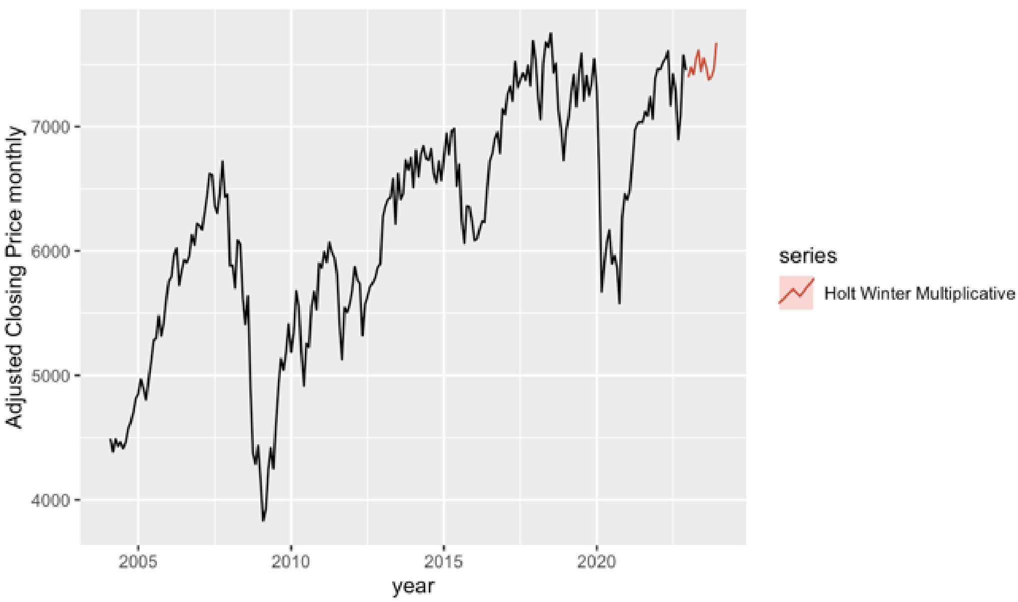

Figure 9: Holt Winter Multiplicative Forecast for FTSE100

The paper will then go into detail on how the ARIMA model is used. First, the autoregressive model is used. By linearly mixing previous values of the same variable, we may forecast the variable of interest in an autoregression model. The term ‘autoregression’ signifies that it involves regressing the variable against its own historical data. Autoregressive models exhibit great versatility in handling various time series patterns [10]. And the author need to decide the order of the Auto regressive model. AIC value can be used by changing the order of the AR model and find the lowest AIC value which will be the one with the best performance.

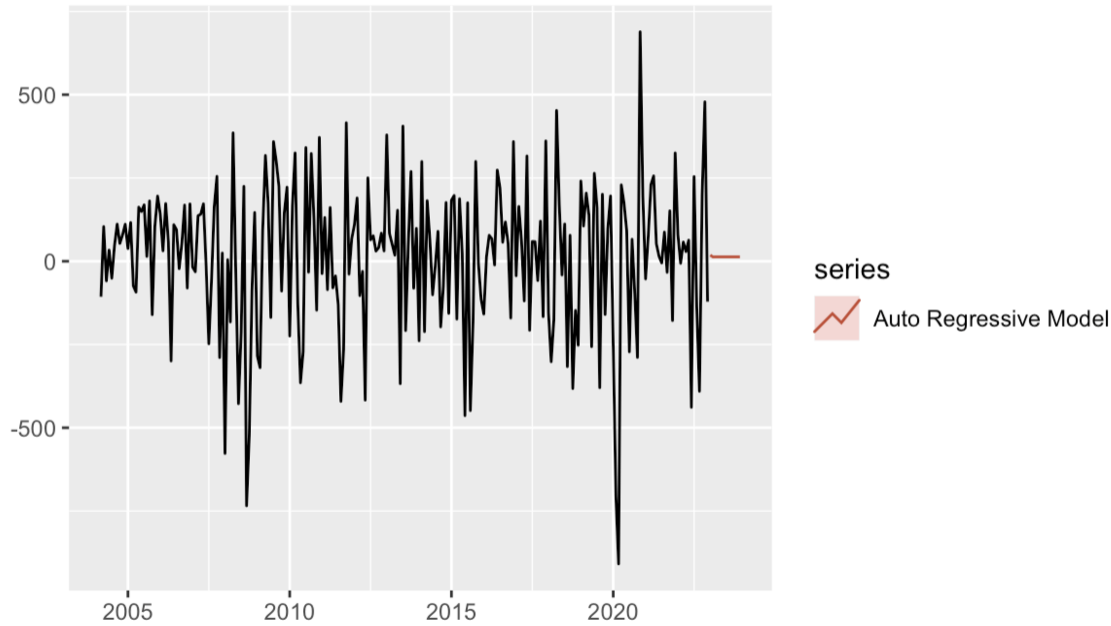

Figure 10: AR forecast for FTSE100

By using the auto.arima function, the author notice that (0,1,0) is the best model. ARIMA (0,1,0) represents a first-order differenced random walk. It is a cumulative sum of an i.i.d. process. The original series is differenced once to obtain this model. Also, the (0,1,0) means that it captures the behavior of a series that exhibits infinitely slow mean reversion.

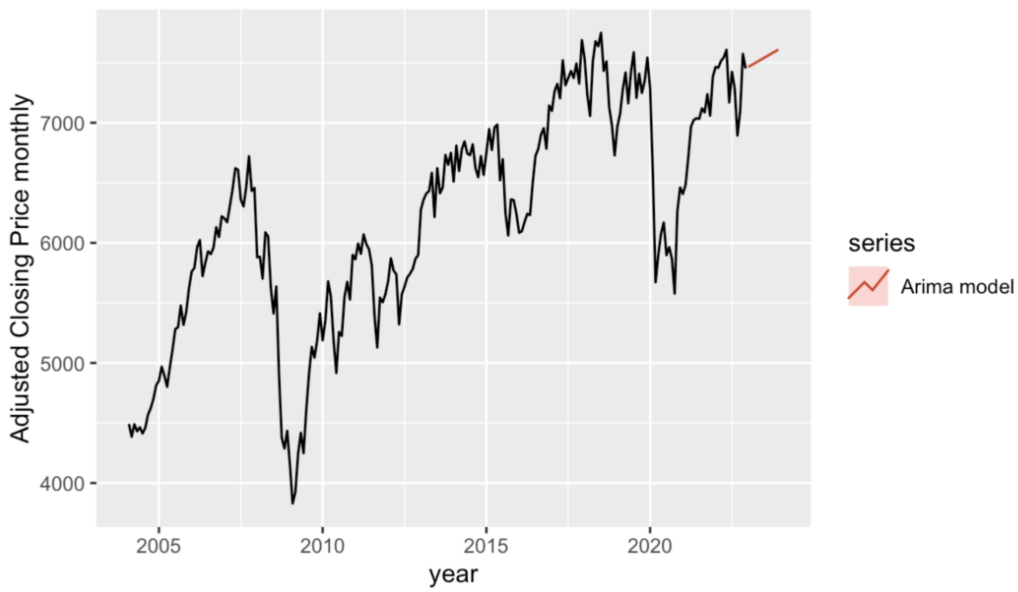

Figure 11: ARIMA Forecast for FTSE100

Finally, by comparing their accuracy, the error function may be used to determine which forecasting approach has the greatest performance and which has the worst. Compare the accuracy metrics of various models to determine the most effective forecasting technique. Better performance is indicated by lower RMSE, MAE, MPE, MAPE, and MASE values. Conversely, the worst forecast method will have higher error metrics. If a model consistently performs poorly across multiple metrics, it may be the worst choice. Therefore, by comparing the accuracy, the author found that the best model is Holt-model and the worst is Mean (see Table1).

Table 1: RMSE, MAE,MPE,MAPE,MASE value of the difference methods.

	Mean	Naive	Snaive	Drift	Holt	HW (additive)	HW (Multiplicative)	ARIMA
RMSE	1424.44	237.86	329.36	215.12	213.36	221.64	218.37	226.15
MAE	1413.93	188.24	288.70	177.00	176.00	176.24	177.58	173.97
MPE	18.52	2.10	3.15	0.97	0.90	0.49	1.64	-0.06
MAPE	18.52	2.43	3.78	2.31	2.30	2.30	2.30	2.88
MASE	2.16	0.29	0.44	0.27	0.27	0.27	0.27	0.27

4. Discussion

For the mean model, it is simple and straightforward, based on historical average, which is suitable for stable time series data without trends or seasonality. While for Naïve model which forecast equals the most recent observation it is suitable for data without clear trends or seasonality. On the perspective of Snaive and Drift model, they are suitable for data with seasonality or linear trends respectively. For the Holt Model, it is suitable when the time series exhibits a consistent upward or downward movement while it does not explicitly address seasonality. On the other hand, the Holt-winter model incorporates both trend and seasonality components. It is well-suited for data with both linear trends and seasonal patterns, but the model's performance depends on the initial parameter values. Choosing appropriate initial values is crucial. For ARIMA model it also suitable for data with trends and seasonality and it is good at forecasting based on past circumstances, but it is better suited for short- to medium-term predictions. From the present time series, we can predict that the value will increase in the future market, but it might also decrease when the global market encounters some shocks.

5. Conclusion

From this article, the author found that by comparing the accuracy of different model, including Mean, Naïve, Snaive, Drift, Holt, HW, and ARIMA, the Holt model is the best on forecasting the future value of FTSE100. Conversely, mean model is the worst. By consider the companies of the FTSE 100 and recent performance of the curve, it is easy to find that there are several factors that will affect the index. Firstly, Energy Boom: The two oil giants in the FTSE 100, Shell and BP, have seen a considerable increase in profits and share prices as a result of the spike in oil and gas prices that followed Putin's invasion of Ukraine. Secondly, Global Economy: The fortunes of FTSE 100 companies largely depend on the global economy, as approximately 75% of their revenues come from overseas. Multinationals like HSBC and Unilever are major players in the index, and their performance is tied to global economic conditions rather than just the UK's. Then, Foreign Exchange Market: Changes in the foreign exchange market impact FTSE earnings. The performance of the index has been impacted by the pound's depreciation versus the US currency and the euro over the past year. Last but not the least, Geo-Political Developments: Geo-political events can create uncertainty and affect investor sentiment, which in turn impacts the index. Keeping an eye on global political developments is crucial for understanding FTSE 100 movements. However, there are also some limitations of this paper, such that the model this article used are not enough. GARCH model can also be used on forecasting the time series. Besides, choosing the best model by comparing the accuracy may not be accurate, using cross-validation or other evaluation can help reader to make

DOI: 10.54254/2754-1169/86/20240887

better decision on choosing model. Overall, this article gives readers a relatively accurate way on the forecast of FTSE100 and might help them on predict the future market.

References

[1] *London Stock Exchange, FTSE100 index overview. (2024). Retrieved from https://www.londonstockexchange.com/indices/ftse-100.*

[2] *Bao Y, Lu Y, Zhang J (2004) Forecasting stock price by svms regression. In: International conference onartificial intelligence: methodology, Systems, and applications, 295–303.*

[3] *David McMillan, Alan Speight & Owain Apgwilym. (2000). Forecasting UK stock market volatility, Applied Financial Economics, 10(4), 435-448.*

[4] *Pinho, David Mendonça. (2020). Forecast Comparison of Volatility Models and Their Combinations For (FTSE100): a Tied Race, Universidade do Minho (Portugal) ProQuest Dissertations Publishing.*

[5] *Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts.*

[6] *Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted averages (O.N.R. Memorandum No. 52). Carnegie Institute of Technology, Pittsburgh USA.*

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

[8] *Jim Forest, Autocorrelation and Partial Autocorrelation in Time Series Data, Retrieved from: https://statisticsbyjim.com/time-series/autocorrelation-partial-autocorrelation/*

[9] *Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics, 54(1-3), 159–178*

[10] *The Guardian. (2023). Five factors to explain FTSE 100's record high despite recession, Publishing 2023, Retrieved from: https://www.theguardian.com/business/2023/feb/16/five-factors-to-explain-ftse-100s-record-high-despite-recession*