Wavelet as a Viable Alternative for Time Series Forecasting

Heng Yew Lee
Universiti
Tunku Abdul Rahman

Woan Lin Beh
Universiti
Tunku Abdul Rahman

Kong Hoong Lem
Universiti
Tunku Abdul Rahman

Abstract

Analysis of financial data is always challenging due to the non-linear and non-stationary characteristics of the time series which is further complicated by volatility clustering effect and sudden changes such as jump, steep slopes and valleys. Classical regression based analysis techniques often entail rigorous mathematical treatments albeit with little success in exploiting the differing frequency characteristics to uncover hidden but valuable trending information. Wavelet, on the other hand provides an efficient way to represent time series with such complex dynamics by decomposing it into time-frequency space and at the same time preserve both temporal and spectral information. This property enables analysts to identify the dominant modes (spectral information) of a time series and observe how those modes vary over time (temporal information). Most importantly, wavelet transform is computationally efficient, only a small number of wavelet coefficients are needed to describe complicated signals. This paper seeks to establish cases for the use of wavelets as viable tools in time series forecasting. Two time series, Kijang Emas Daily Index and Bit Coin Daily Price with differing characteristics are used as subjects of study. Out-of-sample dynamic forecasting of 20 points is made using best-fit ARIMA and prior-point imitation follow by wavelet de-noising methods (imitate-wavelet). Comparisons made with MAPE measurements of ARIMA and imitate-wavelet methods indicated comparable forecasting performance between the simpler imitate-wavelet techniques and ARIMA model.

Keywords: financial analysis, forecasting, time series, wavelet.

1. Introduction

Analysis of financial data is always challenging due to the non-linear and non-stationary characteristics of the time series contributed by numerous economic, political and social factors. Classical regression based analysis techniques often fail to exploit the differing frequency characteristics and thus render the estimations of the model parameters less stable. On top of that, the generally unknown trend, volatility clustering effect, coupled with sudden changes such as jump, steep slopes and valleys further complicate the application of parametric methods.

Wavelet analysis, a signal processing technique commonly used in engineering has the potential to be used as analytical tool for financial time series. Wavelet provides an efficient way

to represent time series with complex dynamics by decomposing the time series into time-frequency space and at the same time preserve both temporal and spectral information. This property enables analysts to identify the dominant modes (spectral information) of a time series and observe how those modes vary over time (temporal information). These information is important as most financial time series such as stock market and foreign exchange consist of heterogeneous investors operating at different timescales.

Many studies have proven effectiveness of wavelet in extracting frequency information from data and efficiency of wavelet in describing both transient and long-term signals. Furthermore, wavelet transform is computationally efficient, only a small number of wavelet coefficients are needed to describe complicated signals.

Although the applications of wavelet in financial time series are well researched, there is a need to verify the consistency of wavelet performance in time series analysis. This paper aims to test the performance of wavelet de-noising between time series with diverse characteristics as well as within the same time series that has volatile characteristics. ARIMA model is used as benchmark in this paper.

2. Literature review

The idea of using advanced signal processing techniques to address the issues in financial analysis always catches attention of researchers. According to Iyer, Kamdar, and Soparkar (2015), concepts of adaptive filters and signal modelling have been employed in order to predict future trends and prices. Financial time series and economic data contain information which is sometimes considered to be unwanted. Donoho and Johnstone (1995) as well as Karim, Karim, Ismail, Hasan, and Sulaiman (2011) state that de-noising the data is necessary in order to smooth the data for a better prediction accuracy. Azimifar, Araabi, and Moradi (2015) have used digital low pass filter to successfully remove the transient fluctuations of the technical indicators of different stocks in Tehran Stock Exchange.

Wavelet transform is well-suited for financial time series analysis. It works with data regardless whether the data is stationary or non-stationary, it provides information from both time-domain and frequency-domain and it possesses the ability to locate precisely the discontinuities in the data. With these characteristics, wavelet is useful in performing static and dynamic correlation analysis and establishing causality relationship between financial or economic variables (Nguyen and He 2015). Besides, wavelet is widely used in the forecasting of indices (Jothimani, Shankar, and Yadav 2015), exchange rates and stock markets (Saravanan and Mala 2018). Rostan and Rostan (2017) have proven the pertinence of spectrum analysis to the forecasting of economic or financial time series. Caus, Badulescu, and Gherman (2017) used the wavelet transform as the filtering technique to find the relation between nominal salary and inflation. Their results shown that the de-noise data able to perform well in the period with low volatility and smaller fluctuations compared with the unfiltered data.

Many choices of mother wavelet functions can be applied while performing wavelet analysis. Haar wavelet is deemed to suit the purpose of the studies in this paper. This choice is justifiable by the facts that the graphs of financial time series tend to be jagged, non-cyclical in predictable manner and cannot be well represented by simple combination of sin and cos terms

3. Methodology

3.1. The data under studied

In this paper, two time series were examined, namely Kijang Emas Daily Prices (KEmas) and BitCoin-USD Daily Exchange Rate (BCoin). The KEmas data was dated from 18/07/2001

to 21/12/2018, while the BCoin data was dated from 16/07/2010 to 20/12/2018. The KEmas data was obtained from Bank Negara (2019) while the BCoin data was obtained from Yahoo Finance (2019). Referring to figure 1 and figure 2, it is obvious that KEmas and BCoin data exhibit very diverse time series characteristics. In addition, the statistical characteristics at different time frames of each time series change significantly.
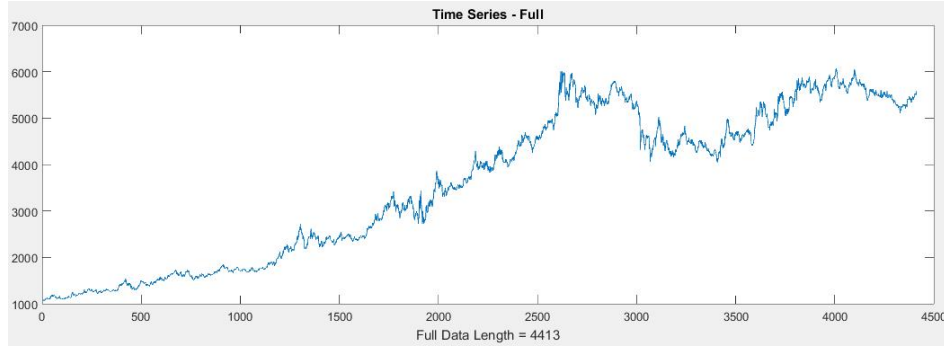

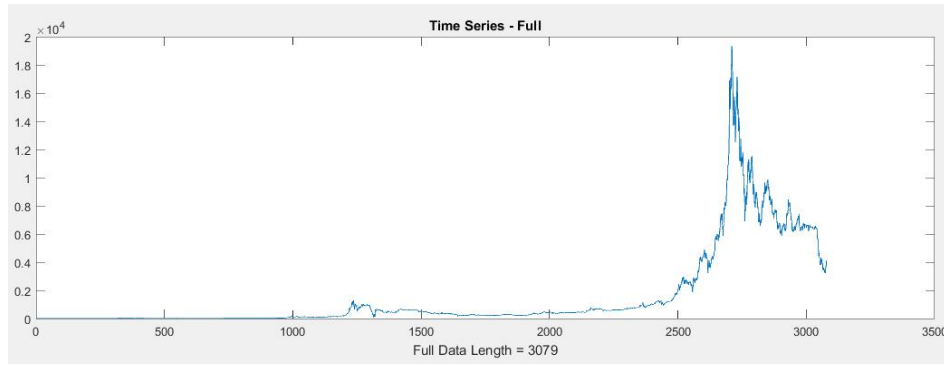Figure 1: Kijang Emas time series from 18/07/2001 to 21/12/2018.

Figure 2: BitCoin-USD time series from 16/07/2010 to 20/12/2018.

The last 20 points of the KEmas and BCoin time series were reserved and used as test data to perform 20-points dynamic forecasting. Training data of four variable lengths (ranging from the latest 25%, 50%, 75% to 100% of remaining data) were used in both time series. Purpose of this arrangement is to test the effect of training data length to forecasting accuracy as well as to compare sensitivity of training data length to forecasting accuracy of imitate-wavelet technique and ARIMA model. Figure 3 illustrates the scenarios of the forecasting performed in KEmas time series where 25%, 50%, 75% and all (100%) of remaining data were used as training data.

3.2. The forecasting process

The flow of the simulation performed in this paper is illustrated in figure 4. First of all, the last 20 points of time series were taken out and used as test data. Part or all of the remaining data (the most recent 25%, 50%, 75% and 100%) were then used as training data in four different forecasting scenarios for KEmas and BCoin time series respectively. The training data were used to perform wavelet decomposition for extraction of spectral information and to find the best fit ARIMA model. Forecasting on the last 20 points were performed dynamically with application of either imitate-wavelet technique or ARIMA modeling on training data of various lengths. For imitate-wavelet technique forecasting, the point being forecasted would first imitate the value of previous point, wavelet decomposition of all levels were then performed to
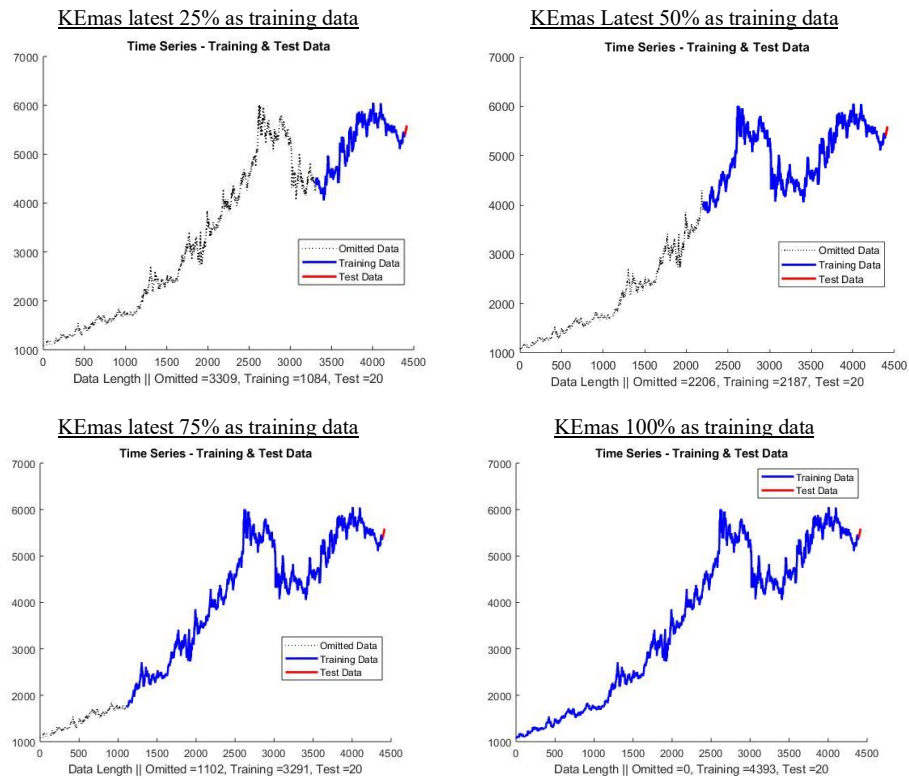
Figure 3: KEmas time series training data of variable lengths.

de-noise the forecasted point. The best forecasted result (with minimum signal to noise ratio) was kept as final result. The present work adopted Discrete Wavelet Transform (for noise removal) using Haar wavelet as the mother wavelet. For ARIMA modeling technique, the best fit model was selected by referring to the Akaike Information Criteria (AIC). The forecasted data were then compared with test data to calculate the mean absolute percentage error (MAPE). Comparison was then made between forecasting performance of imitate-wavelet technique and ARIMA model.

4. Simulation outcomes

Figures in the appendix show the dynamic forecasting outcomes of both KEmas and BCoin time series. Titles of the figures state the conditions, methods and the parameters used in forecasting simulations. The solid line indicates actual values while the dash line and dotted line indicate the forecasted values of ARIMA and imitate-wavelet respectively. MAPE of both techniques used to perform the forecasting are stated below the graphs. The calculated MAPE for best fit ARIMA and imitate-wavelet technique are juxtaposed in table 1.

Referring to table 1, it can be observed that the forecasting results obtained through imitate-wavelet technique yield comparable MAPE with forecasting results of ARIMA modeling. Especially in BCoin time series, imitate-wavelet technique performs better than ARIMA model in cases where 25% and 100% of remaining data were use as training data. The next point to be noted is that there is no observable correlation between training data length and forecasting accuracy in KEmas time series with either imitate-wavelet or ARIMA forecasting techniques. This may be due to the fact that even the shortest training data length (25% with approximate 1,100 points) is suffice to perform dynamic forecasting of 20 points. In the case of BCoin time series, lengthier training data did improve the ARIMA forecasting performance gradually; however, it has no effect on the imitate-wavelet technique. Lastly, the simulation results shown that imitate-wavelet technique performs well with mostly level
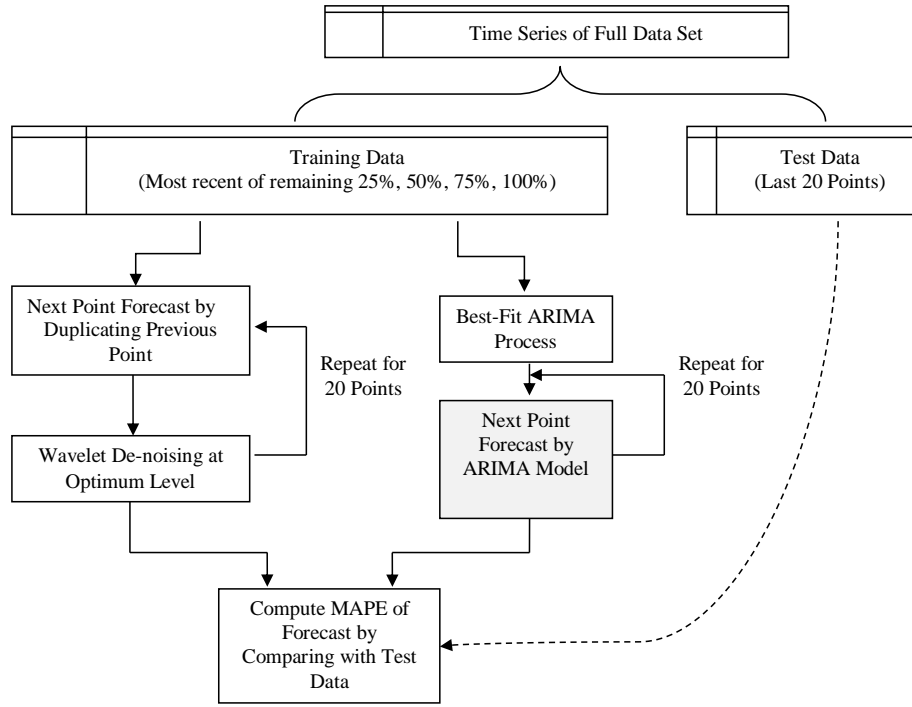
Figure 4: Simulation process flow.

Table 1: MAPE for 20-points dynamic forecasting

Time Series & Training Data	MAPE	
	ARIMA	**Imitate-Wavelet**
Kijang Emas (25% Training Data)	0.3522 ARIMA(0,1,0)	0.3558 Haar Level 1
Kijang Emas (50% Training Data)	0.3533 ARIMA(0,1,0)	0.3558 Haar Level 1
Kijang Emas (75% Training Data)	0.3518 ARIMA(0,1,0)	0.3558 Haar Level 1
Kijang Emas ($\approx$100% Training Data)	0.3522 ARIMA(0,1,0)	0.3558 Haar Level 1
BitCoin-USD (25% Training Data)	3.6251 ARIMA(1,1,1)	3.4003 Haar Level 2
BitCoin-USD (50% Training Data)	3.6239 ARIMA(1,1,1)	3.6891 Haar Level 1
BitCoin-USD (75% Training Data)	3.6235 ARIMA(1,1,1)	3.6891 Haar Level 1
BitCoin-USD ($\approx$100% Training Data)	3.6232 ARIMA(1,1,1)	3.4003 Haar Level 2

1 decomposition.

5. Conclusion

The main objective of this paper is to ascertain the suitability of wavelet technique in analysing time series with diverse characteristics. From the simulated dynamic forecasting outcome, the practice of the simple imitation then wavelet de-noising method achieved comparable results using best fit ARIMA model as benchmark. Most remarkable is that this was achieved with rather consistent wavelet analyzing parameters, regardless of the statistical characteristics of the time series being analyzed. Outcomes from the presented works did reinforce the earlier research findings that wavelet is an effective de-noising tools in financial time series analysis. Conclusion can thus be drawn that wavelet de-noising is applicable in the analysis of both volatile time series such as BCoin as well as relatively stable time series such as KEmas.

References

Azimifar M, Araabi BN, Moradi H (2015). "Improving the Performance of Intelligent Stock Trading Systems by Using a High Level Representation for the Inputs." *Signal Processing and Intelligent Systems Conference*, pp. 15–19. URL https://ieeexplore.ieee.org/document/7422304/references.

Bank Negara (2019). URL http://www.bnm.gov.my/index.php?ch=statistic&pg=statistic_kijangemas.

Caus VA, Badulescu D, Gherman MC (2017). "Using Wavelets in Economics - An Application on the Analysis of Wage-price Relation." *Oradea Journal of Business and Economics*, **2**(1), 32–42.

Donoho D, Johnstone I (1995). "Adapting to Unknown Smoothness via Wavelet Shrinkage." *Journal of the American Statistical Association*, **90**(432), 1200–1224.

Iyer S, Kamdar NR, Soparkar B (2015). "Stock Market Prediction Using Digital Signal Processing." *International Journal of Computer Applications*, (0975-8887).

Jothimani D, Shankar R, Yadav SS (2015). "Discrete Wavelet Transform-Based Prediction of Stock Index: A Study on National Stock Exchange Fifty Index." *Journal of Financial Management and Analysis*, **28**(2), 35–49. URL https://www.questia.com/library/p439848/journal-of-financial-management-analysis/i4245812/vol-28-no-2-july-december.

Karim SAA, Karim BA, Ismail MT, Hasan MK, Sulaiman J (2011). "Application of Wavelet Method in Stock Exchange Problem." *Journal of Applied Sciences*, **11**, 1331–1335. URL http://eprints.ums.edu.my/5177/.

Nguyen T, He TX (2015). "Wavelet Analysis and Applications in Economics and Finance." *Research and Reviews: Journal of Statistics and Mathematical Sciences*, **1**(1).

Rostan P, Rostan A (2017). "Forecasting Financial Time Series with Spectrum Analysis." *Proceedings of The IIER International Conference*.

Saravanan S, Mala S (2018). "Stock Market Prediction System: A Wavelet Based Approach." *Applied Mathematics and Information Sciences - An International Journal*, **12**(3), 579–585.
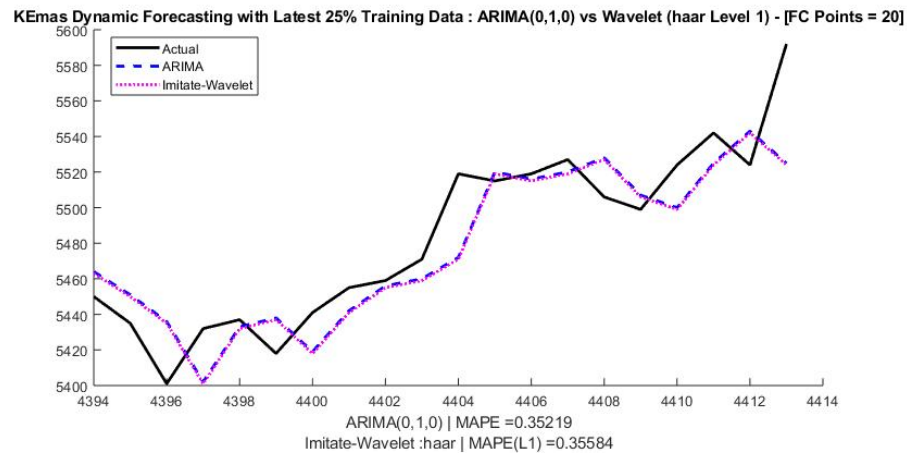
Yahoo Finance (2019). URL https://finance.yahoo.com/quote/BTC-USD/history?p=BTC-USD.

Appendix

KEmas Forecasting Outcomes

KEmas Dynamic Forecasting with Latest 25% Training Data : ARIMA(0,1,0) vs Wavelet (haar Level 1) - [FC Points = 20]

ARIMA(0,1,0) | MAPE =0.35219
Imitate-Wavelet :haar | MAPE(L1) =0.35584

Figure 5: 20-points dynamic forecasting on Kijang Emas 25% training data.
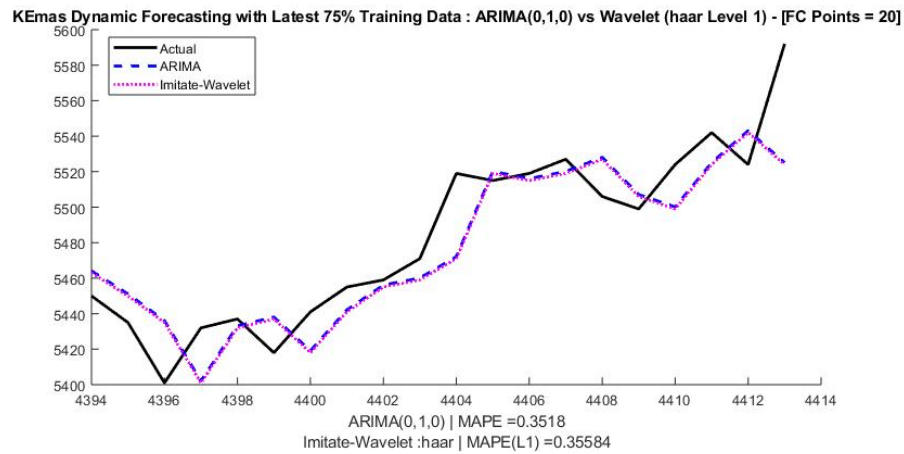
KEmas Dynamic Forecasting with Latest 50% Training Data : ARIMA(0,1,0) vs Wavelet (haar Level 1) - [FC Points = 20]

ARIMA(0,1,0) | MAPE =0.35334
Imitate-Wavelet :haar | MAPE(L1) =0.35584
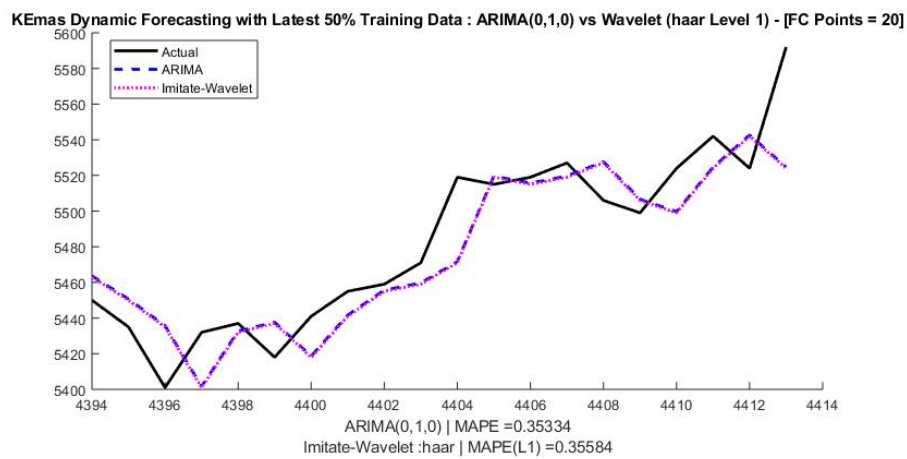
Figure 6: 20-points dynamic forecasting on kijang emas 50% training data.

Figure 7: 20-points dynamic forecasting on Kijang Emas 75% training data.
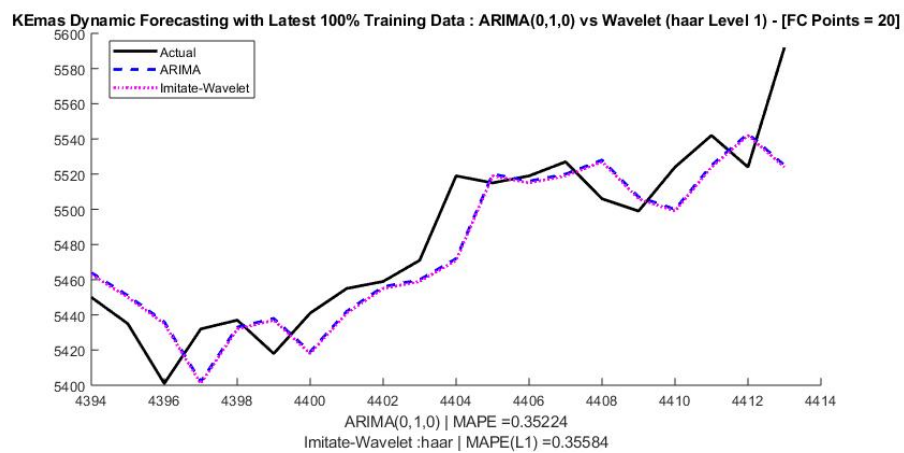
Figure 8: 20-points dynamic forecasting on Kijang Emas 100% training data.

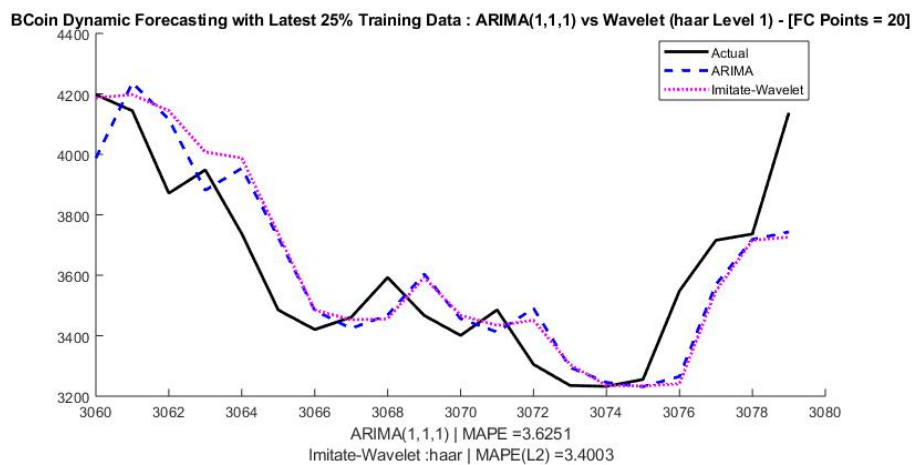## BitCoin Forecasting Outcomes

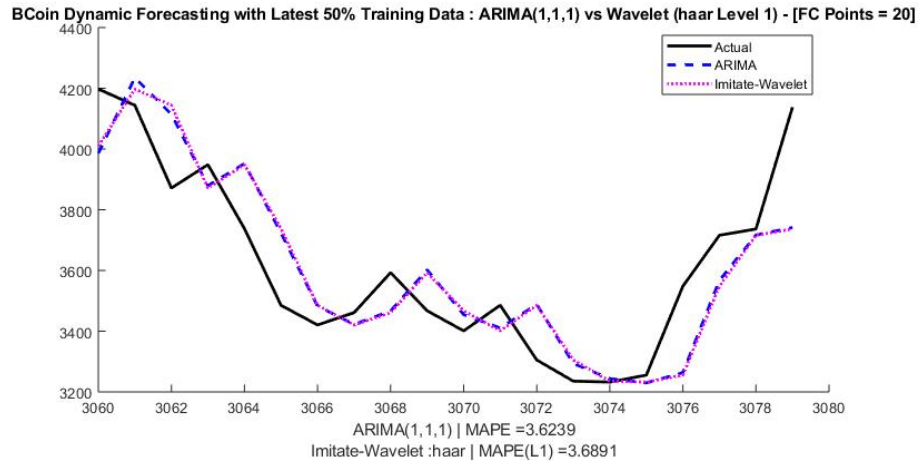Figure 9: 20-points dynamic forecasting on BitCoin-USD 25% training data.

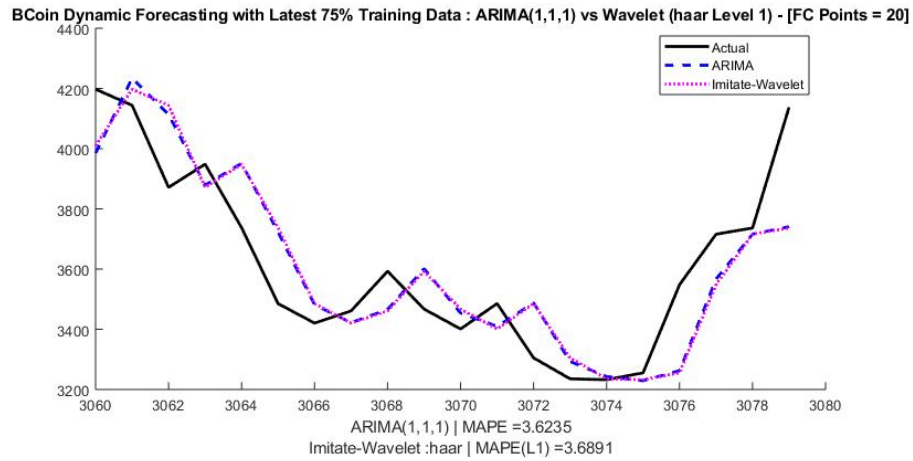Figure 10: 20-points dynamic forecasting on BitCoin-USD 50% training data.

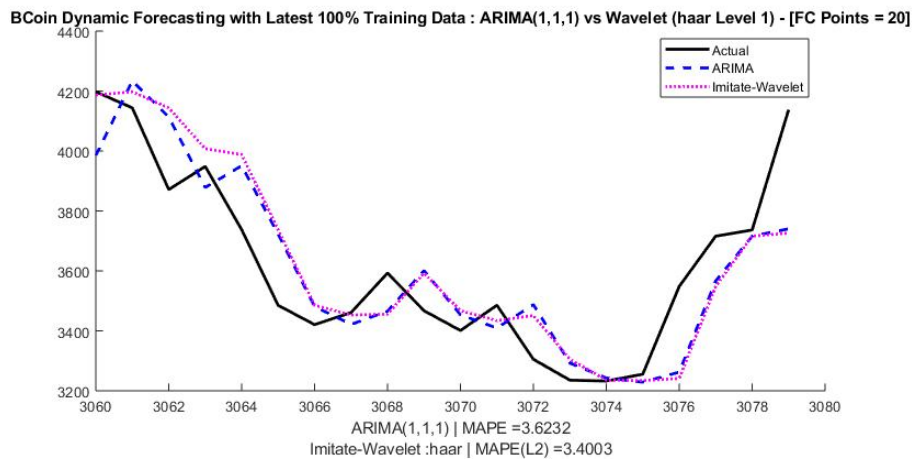Figure 11: 20-points dynamic forecasting on BitCoin-USD 75% training data.

Figure 12: 20-points dynamic forecasting on BitCoin-USD 100% training data.

Affiliation:

Heng Yew Lee
Faculty of Information and Communication Technology
Universiti Tunku Abdul Rahman
Jalan Universiti, Bandar Barat, 31900 Kampar, Perak, Malaysia
E-mail: leehy@utar.edu.my
URL: http://www.utar.edu.my/staffDirDetail.jsp?searchId=10103