DEVELOPING A FRAMEWORK FOR FORECASTING FINANCIAL TIME SERIES USING WAVELET-LSTM HYBRID MODEL

Submitted to

American University of Armenia
Manoogian Simone College of Business and Economics

In partial fulfillment of the requirements for the degree of BA in Business

By: Robert Yenokyan

Supervisor: Dr. Gayane Barseghyan

Yerevan 2018

Abstract

In this paper, 4 models are fitted in EUR/USD, GBP/USD, and USD/JPY rates. First LSTM model is fitted to forecast the price at the next step and is compared with Wavelet-LSTM hybrid model. The results show that the hybrid model performs better than LSTM model without wavelet decomposition. However, mean absolute percentage error (MAPE) shows that the hybrid model performs poor for trading purposes. Then a classification problem is considered. The objective of the problem is to predict the direction of the return on the next timestamp. Again LSTM is compared with Wavelet-LSTM model. The results show that Wavelet-LSTM performs better than LSTM. Considered simple trading algorithm achieves annualized return of 20% with Wavelet-LSTM hybrid model.

Key words: Financial Time Series, Wavelet, LSTM, Deep Learning, Forex, Regression, Classification

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Gayane Barseghyan for her support and patience during my work.

All the remaining errors in this paper are mine.

Table of contents

1. Introduction

Forex market includes traders with diverse risk tolerance and investment horizon. Market participants include central banks, institutional investors, hedge funds, pension funds and even proprietary traders. Because of the different investment horizon ranging from a couple of minutes to years, the relationship between variables in forex market differs based on the time scale. Wavelet analysis makes it possible to decompose time series into different time scales and study each one separately depending on one's needs and objectives.

The models are developed keeping in mind the aspect of practicality. If a model is time consuming to create or is not suitable for trading or investment purposes, then the model is not considered. Also only models that perform well on the short term are considered.

2. Literature Review

Financial time series are a complex set of signals. Autoregressive integrated moving average (ARIMA) and traditional econometrics models assume stationarity and linearity [1]. Thus they are not a good fit for the research. Artificial neural network (ANN) models showed high performance in forecasting financial time series [2-4]. Also, [4] compared ANNs with ARIMA models and showed that ANNs perform better in forecasting financial time series.

Studies show that long-short term memory (LSTM) models perform better than ARIMA models by a significant margin [7]. In time series forecasting LSTM models perform better than ANNs, as LSTM models are designed just for sequence modeling.

[8] analyzed stock market and concluded that wavelet decomposition helps to detect structures in time series that are non-detectable without decomposition. [9] showed that wavelet decomposition improves the performance of ARIMA model in forecasting Amman stock index.

[10] showed that using ARIMA-GARCH model to forecast monthly gold return has better performance if the data is preprocessed with wavelets.

3. Data

The data is downloaded from Dukascopy[1], a Swiss forex bank. The data ranges from January 1, 2004 till December 31, 2017. The data has hourly frequency. The training set ranges from January 1, 2004 till December 31, 2012 and consists of 56,352 observations. The testing set ranges from January 1, 2013 till December 31, 2017 and consists of 33,362 observations. The data is collected for the following forex pairs: EUR/USD, GBP/USD, and USD/JPY. These 3 forex pairs are among the most traded ones and have high liquidity. That is the main reason that only these 3 pairs are collected.

The mid price is chosen to train the models. The mid price is defined as:

$$midprice = \frac{bid + ask}{2}$$

The reason mid price is chosen is that the ask price is the lowest price a seller is willing to accept, and the bid price is the highest price that a buyer is willing to pay. As the model is developed for trading purpose, both buying and selling of the currency pair will occur. Thus it is important to choose the mid price in order to account for this fact.

4. Methodology

As the models in this paper are developed for practical purposes, I will not develop ARIMA model. First, as discussed in the literature review section, ARIMA models perform the

[1] www.dukascopy.com

worst compared to other machine learning models. Second, each sub-series require careful selection of parameters. And when the number of sub-series is large, designing ARIMA models become inefficient and time consuming for practical purposes. Finally, ARIMA models assume time series to be stationary and linear. However, financial time series in form of price are mostly non stationary and not linear.

LSTM model

Each time series is scaled to be between the range (-1,1) with mean 0.

$$scaled\ data_i = \frac{data_i - mean_{data}}{sd_{data}}$$

where $mean_{data}$ is the mean of the time series, sd_{data} is the standard deviation of the time series, and $data_i$ and $scaled\ data_i$ are the i-th observation of the original and the scaled time series respectively.

The input is the last 24 observations from t_k to t_{k+24} of the scaled series.

For regression problem, the target is t_{k+25}-th observation of the original time series. For classification problem, the target is a binary variable: 1 if the sign of $t_{k+25} - t_{k+24}$ is positive and 0 otherwise.

Wavelet-LSTM hybrid model

Each time series is decomposed into 16 sub-series using Haar wavelet. The input is the last 24 observations from t_k to t_{k+24} of the scaled series. Then the inputs (16 sub-series) are scaled to be between the range (-1,1) with mean 0.

Training and testing datasets are processed separately in order not to include look-ahead bias. LSTM model is trained simultaneously taking the 16 sub-series as input.

For regression problem the target is the original price at time t_{k+25}. LSTM model has two layers each with 32 neurons. The number of epochs is adjusted for each time series. The loss function is chosen to be MAE (mean absolute error) and the network is optimized to minimize MAE.

For the classification problem the target is a binary variable, the same as for LSTM model without wavelet decomposition. LSTM model has two layers with 32 neurons each. The number of epochs is adjusted for each time series. The loss function is binary cross entropy.

Wavelet decomposition

[5-6] use Haar and Daubechies LA8 (least asymmetric wavelet filter of length 8) wavelets for decomposing time series. Haar wavelet is good for edge detection; it can capture the oscillation between adjacent observations. LA8 wavelet is hypothesizes to be a good fit for time series. In this research the Haar wavelets are used.

Figure 3 shows the wavelet decomposition of EUR/USD pair from January 1, 2004 to December 31, 2012 (see Appendix).

5. Results

Tables 1-3 present the results. Wavelet-LSTM models perform better on all three datasets compared with only LSTM models both in regression and classification problems. However, in regression problems MAPE (mean absolute percentage error) is high for the purposes of trading.

In classification problems, the Wavelet-LSTM models have accuracy of higher than 0.72. The results can be used to design profitable trading algorithms.

Table 1 Regression and classification results on USD/JPY pair.

	USD/JPY regression			
	Wavelet-LSTM		**LSTM**	
	Train	**Test**	**Train**	**Test**
MSE	5.16	109.9	6.4	115.8
MAE	1.18	8.48	1.95	12.61
MAPE	2.25	7.6	3.11	9.7
	USD/JPY classification			
Accuracy	0.8	0.72	0.74	0.53

Table 2 Regression and classification results on EUR/USD pair.

	EUR/USD regression			
	Wavelet-LSTM		**LSTM**	
	Train	**Test**	**Train**	**Test**
MSE	7.34e-05	0.0021	8.49e-06	0.0048
MAE	0.0074	0.014	0.018	0.094
MAPE	0.56	6.5	1.1	11.3
	EUR/USD classification			
Accuracy	0.81	0.74	0.79	0.6

Table 3 Regression and classification results on GBP/USD pair.

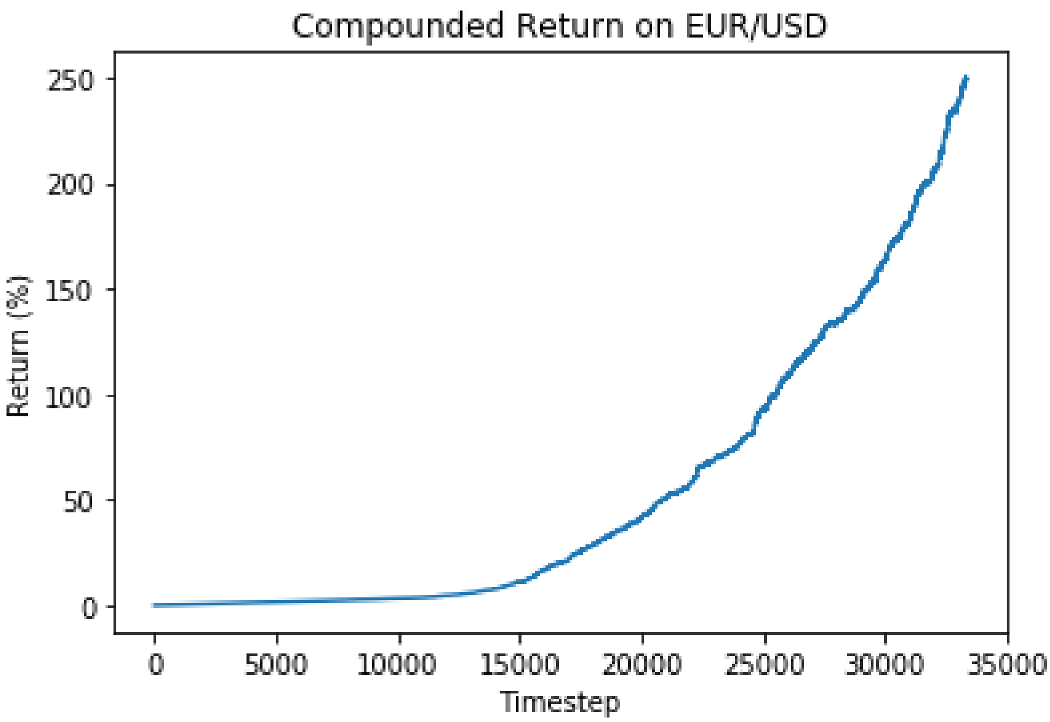
	GBP/USD regression			
	Wavelet-LSTM		**LSTM**	
	Train	**Test**	**Train**	**Test**
MSE	2.67e-05	0.0073	4.84e-05	0.015
MAE	0.0043	0.15	0.01	0.98
MAPE	0.27	8.4	1.08	13
	GBP/USD classification			
Accuracy	0.85	0.79	0.76	0.61

Trading Algorithm

To validate the good performance of classification models, I created a simple trading algorithm for EUR/USD pair. The algorithm takes a long position[2] for one period (1 hour) and closes it at the end of the period if the output of Wavelet-LSTM classification model is 1 (that is the model predicts the price to rise). The algorithm takes a short position[3] for one period and closes it at the end of the period if the output of Wavelet-LSTM classification model is 0 (that is the model predicts that the price will not rise). The algorithm invests the entire portfolio every time it opens a position. Transaction costs are not included in calculating the return figures.

The trading algorithm achieves 249.3% compounded return during the period January 1, 2013 to December 31, 2017. Annualized return is 20%. The equity curve is presented in figure 1.

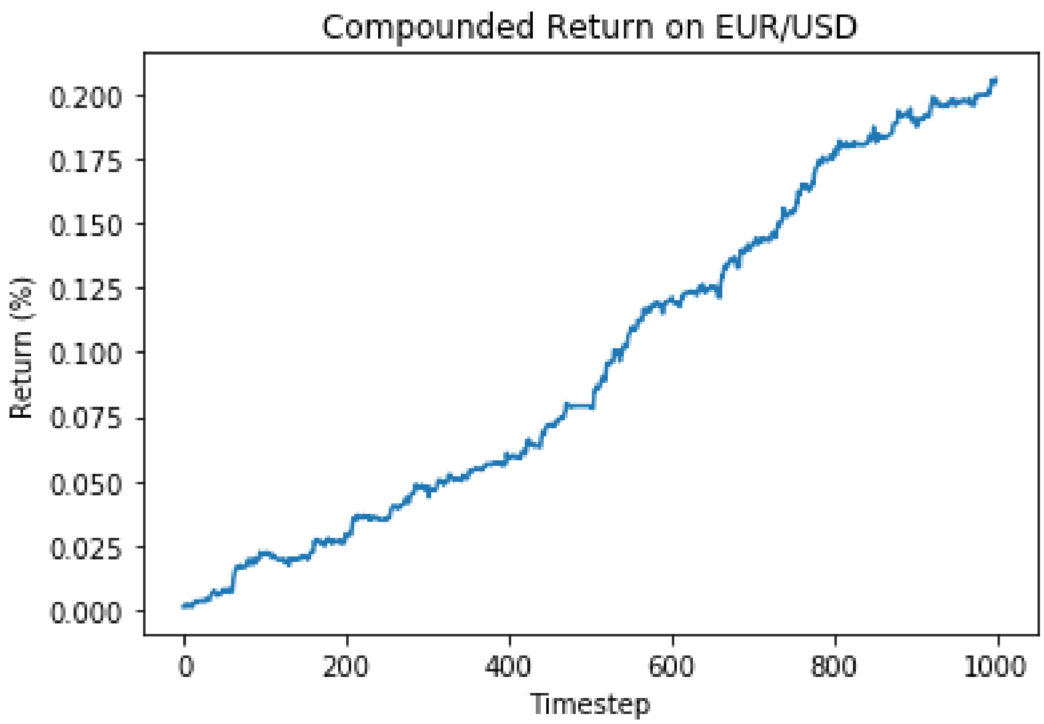
Figure 1 Compounded return on EUR/USD pair from January 1, 2013 to December 31, 2017.



[2] A long position in a security means buying the security
[3] A short position means borrowing the security and selling in the market

It may be difficult to see the variance due to relatively high number of observations. Figure 2 shows the first 1000 observations as in figure 1.

Figure 2 A close up on compounded return on EUR/USD pair (the first 1000 time steps).



6. Discussion

Because financial time series has high complexity, it is difficult to accurately forecast it. The developed regression models showed high MAPE, which means that these models cannot be used for trading applications. Then a classification problem is considered. The results show that it is possible to develop profitable trading algorithm using the proposed Wavelet-LSTM hybrid model.

Further research can be done to improve the results. First, a different wavelet filter can be considered from Coiflet or Daubechies families. Second, bigger LSTM architectures can be considered, which might be able to find more complex relationship in the data. Third, other inputs can be considered such as traded volume. And lastly, a different frequency of price data can be used.

As the models are trained on data up to December 31, 2012 and are tested up to December 31, 2017, the performance of the models can deteriorate over time because of regime shifts or some other fundamental changes (similar to 2008 financial crisis). One improvement is to train and test the model using moving windows. That is to train the model every time new data is available. This way the model will take into account all the recent available data, thus possibly improving the results.

7. References

1. Box, G. E. P., Jenkins, G. M., Reinsel, G.C.: (2008) Time Series Analysis: Forecasting and Control. the 4 th ed. John Wiley & Sons.

2. Yao, J., Li, Y., Tan, C.L.: (2000) Option price forecasting using neural networks. OMEGA: International Journal of Management Science, Vol. 28, No.4, 455-466.

3. Yao, J., Tan, C.L.: (2000) A case study on using neural networks to perform technical forecasting of forex. Neurocomputing, Vol. 34, 79-98.

4. Kamruzzaman, J., Sarker, R.: (2003) Comparing ANN based models with ARIMA for prediction of forex rates. ASOR Bulletin, Vol. 22, 2-11.

5. Minu, K.K., Lineesh, M.C., Jessy John, C.: (2010) Wavelet Neural Networks for Nonlinear Time Series Analysis. Applied Mathematical Sciences, Vol. 4, No. 50, 2485-2595.

6. Fan, S., Ji, T., Gordon, W., Rickard, B.: (2013) Forecasting Baltic Dirty Tanker Index by Applying Wavelet Neural Networks. Journal of Transportation Technologies, Vol. 3, No.1, 68-87.

7. Siami-Namini S., Siami Namin A.: (2018) Forecasting Economics and Financial Time Series: ARIMA vs. LSTM. https://arxiv.org/abs/1803.06386v1

8. Ramsey, J. B., Usikov, D., & Zaslavsky, G. M.: (1995). An analysis of U.S. stock price behavior using wavelets. Fractals, 3, 377-389

9. Ababneh, F., Wadi, S. A., & Ismail, M. T.: (2013). Haar and Daubechies wavelet methods in modeling banking sector. International Mathematical Forum, 8(12), 551-566.

10. Nouri, M., Oryoie, A. R., & Fallahi, S.: (2012). Forecasting gold return using wavelet analysis. World Applied Sciences Journal, 19(2), 276-280.

8. Appendix

Figure 3 EUR/USD (training set) wavelet decomposition using Haar filter. From top to bottom are the series with frequency components in descending order. The first one contains the highest frequency components and the last one is the slowest moving part (trend).