

Comparative Analysis of the Outcomes of Differing Time Series Forecasting Strategies

Shamsul Masum, Ying Liu and John Chiverton

School of Engineering

University of Portsmouth

Anglesea Building, Anglesea Road,

Portsmouth (UK) PO1 3DJ

Email: shamsul.masum@port.ac.uk, ying.liu@port.ac.uk, john.chiverton@port.ac.uk

Abstract—Time series forecasting of data from various domains has become an increasingly interesting research subject in recent times. Prediction of future sample values is the main goal of time series forecasting. There are mainly two classes of time series forecasting, namely, single step and multi-step forecasting. There are various machine learning approaches that can be used along with various forecasting strategies which exist, some of which can be utilized in and by time series forecasting. However, inappropriate use of strategies can limit the applicability of the techniques to real world problems. This paper provides a review of time series forecasting, a classification of time series forecasting and also approaches and strategies of time series forecasting. It also demonstrates how inappropriate use of point to point rolling forecast strategy for forecasting could lead to unrealistic outcomes and how a multiple step forecasting strategy called Direct H step strategy could help to overcome this issue. Comparative analysis of two strategies using Auto-Regressive Integrated Moving Average (ARIMA) approach are demonstrated.

Index Terms—Time series forecasting; Single step forecast strategy; Multi step forecast strategy; ARIMA.

I. INTRODUCTION

Time series forecasting is considered to be an important area of machine learning with an ultimate goal of predicting the future. It is also called “forecasting by exploring the pattern from past data” [1]. Forecasting of future samples plays a vital role in guiding the decision making of selected areas. Data that is required for time series forecasting are classified as two types: one is time series data and another one is data with time points [2]. Time series data can be described as

$$X = (x_t; t = 1, \dots, N) \quad (1)$$

where X = time series, t = time, N = observations during that time, x_t = valued measure at time instant t .

Shumway and Stoffer [3] defined time series as “a collection of random variables indexed according to the order they are obtained in time”. Time series forecasting has been extensively used in the medical sector for forecasting different events. Events can include dialysis in critically ill patients, mortality risk predictions in paediatric critical care, hypotensive episode predictor for intensive care and prediction of morbidity of tuberculosis (MTB); where it can be seen that

time series forecasting have been used [6], [7], [8], [9]. Time series forecasting has also been given priority in the finance sector where datasets of different finance applications have been investigated for forecasting index prices, stock closing price, production revenues and sales volume [12], [13], [31], [32]. To forecast the load of electricity in advance, time series forecasting has also been used [14], [15], [30]. Time series forecasting has also found use in weather prediction for forecasting maximum temperature, rainfall, evaporation and wind speed [10], [11]. Machine learning approaches such as Auto-Regressive Integrated Moving Average (ARIMA) and Recursive Neural Network (RNN) along with different forecasting strategies are sometimes used for training from past data to predict the future [4], [5].

Short term forecasting has been implemented in many applications with different strategies. Deng and Jirutitijaroen have used multiplicative decomposition and seasonal ARIMA strategy for short term electricity load forecasting [14]. Next day electricity price has been forecast using RNN by Mandal et al. [15]. Cyprich et al. have used ARIMA strategy to forecast short term passenger demand [19]. Short term glucose prediction has been forecast using a neural network, where CGM reading and meal information is used as features by Zecchin et al. [29]. Ren et al. have forecast short term electricity load demand using a Random Vector Functional Link (RVFL) network where RVFL network was found to outperform several other competing methods [30]. Gupta et al. have forecast next day stock index price using feed-forward neural networks [31]. Next 5 days price percentage oscillator (PPO) has been forecast using a support vector regression optimised by genetic algorithms (SVR-GA) and multilayer perceptron also optimised with GA (MLP-GA) by Zhu et al. [32].

Long-term forecasting has been implemented in different areas using several forecasting strategies. Cococcioni et al. has forecast energy production in solar PV systems for 24 hours using a feed forward neural network [20]. Multi-step ahead forecasting has been computed to forecast the wind power using three different classes of artificial neural network known as neural network with time delay (LNNTD), feed forward neural network (FFNN) and Elman recurrent neural network (ERNN) by Saroha and Aggawal [21]. Guanqun et al.

used Backpropagation neural network to forecast stock price considering time series data [22]. A long-term prediction of time series datasets has been conducted by Park [33] using a BiLinear Recurrent Neural Network (BLRNN) to make predictions more robust compared to traditional MultiLayer Perceptron Type Neural Network (MLPNN). Xiong et al. [34] explored multi step predictions of crude oil prices using a feed-forward neural network where the performance of several strategies was also compared in order to find the best. Chang et al. [35] presented a multi-step-ahead flood forecasting using Multi-layer feed-forward networks. Multi-input Multi-output (MIMO), Multi-input single-output (MISO) and serial-propagated structure of multi step forecasting strategies were compared to find the best strategies.

This paper provides a review of ARIMA techniques along with a forecasting strategy called point to point rolling forecast strategy and Direct H step forecast strategy [18], [37]. Both of these have been used for time series forecasting in different applications. A comparative analysis of two strategies (point to point rolling forecast strategy and Direct H step forecast strategy) is also presented in this paper. Preliminary results of our work suggest that the point to point rolling forecast could generate a somewhat misleading outcome for long-term forecasting. Instead the multi-step forecast strategy can replace the point to point rolling forecast strategy. Point to point rolling forecast provides a forecast where ARIMA model is recreated after receiving a new observation the underlying noise free test data is stored. This dependence on knowledge of the true data values assists the forecasting however it is not always applicable to many real world scenarios. For example, knowledge of the true data may not be possible, even for past values. Whereas, multi step forecast strategy called direct H step strategy creates a forecast where a single forecast is computed for H step forecast and the model does not depend on the underlying test data.

The paper is organised as follows. Section II provides a review of two classes of time series forecasting along with the review of a point to point single step and direct multi-step forecasting strategy. Section III provides a review of the ARIMA approach from [24] that is used for this paper. Section IV describe the comparative analysis with results. Finally, the conclusion is provided in section V.

II. CLASSIFICATION AND FORECASTING STRATEGY OF TIME SERIES

Forecasting class and forecasting strategy of time series forecast are correlated. Their correlation depends on the duration of the forecast. Single point strategy is used for single step forecasting where only one step ahead is forecast. There are different multi-step forecast strategies available but in this paper, direct H step strategy for forecasting more than one step ahead is employed [16].

A. Single Step Forecasting with single step Strategy

Selection of forecasting class and its strategy depends on the forecasting requirement of the application field and data

frequency of collected data. Single step forecasting is applicable where short term forecasting is required. For example, several minutes, hours or days could be considered as short term duration. For such a scenario, computing one step ahead forecast is useful. One step forecast ($t + 1$) is achieved by passing the current and past observations ($t, t - 1, \dots, t - n$) to a chosen model, as shown in equation 2:

$$F(t + 1) = M(o(t), o(t - 1), o(t - 2), \dots, o(t - n)) \quad (2)$$

where F is the forecast point, M is the model and o is the observation.

B. Multi Step Forecasting with Direct H Step Strategy

Multi-step forecasting is useful where the field of application requires long-term duration forecasting. This class of forecasting is very challenging and there is a lack of studies available that consist of machine learning algorithms and methodologies for this class. It has also been found that lack of collaborations between these different fields is creating a barrier to further developments [17]. For such a scenario, multiple-step ahead forecast computation is required and to do so Ben Taieb et al. [18] described five multi-step strategies, among them Direct Multi-Step Strategy was considered, also considered here in the work here. The direct multi-step strategy can be expressed as

$$y_{t+h} = f_h(y_t, \dots, y_{t-n+1}) \quad (3)$$

where h is the number of steps to forecast into the future, n is the autoregressive order of the model, f_h is any arbitrary learner.

III. ARIMA APPROACH OF TIME SERIES FORECASTING

There are many different approaches available for time series forecasting which have been described in [23]. ARIMA approach is considered to be the most common approach for time series forecasting. The ARIMA approach was introduced by Box and Jenkins. It has been found to be efficient for short term forecasting [24]. Acronym of ARIMA is meaningful as it represents the key characteristics of the model which are: AR(Autoregression), relating relationship between current observation and past observations; I(Integrated): differencing of actual observations in order to make time series stationary; MA(Moving Average): Lags of the forecast errors of the moving average model.

These component are included in an ARIMA model as a parameter. The standard notation for the ARIMA model is usually given as ARIMA(p, d, q); where p is the number of lag observations, d is the degree of differencing and q is the size of the moving average window. In the ARIMA model, the forecast value is a linear compound of past values and past errors, expressed as follows, where: y_t actual value, ε_t random error at t , φ_i and θ_j are known as coefficients, p and q are autoregressive and moving average.

$$y_t = \theta_0 + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (4)$$

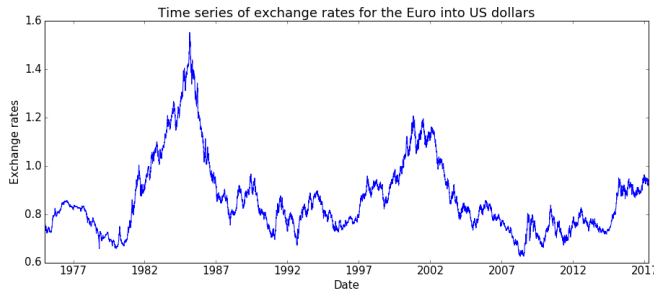


Fig. 1. Daily exchange rate of Euro into US dollars (USD): 1975-01-02 - 2017-04-16

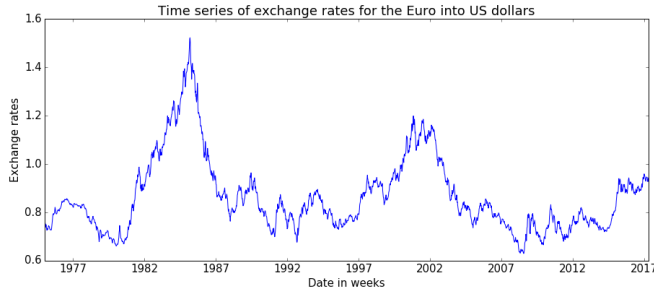


Fig. 2. Weekly exchange rate of Euro into US dollars (USD): 1975-01-02 - 2017-04-16

IV. EXPERIMENTS AND RESULTS: COMPARATIVE ANALYSIS

Some comparative analyses are now performed between the single point rolling forecast and the Direct H step forecast on selected time series data.

A. Dataset

The data set used is from Bank of England's official statistics on spot exchange rates for the Euro into US dollars (USD) [25]. The dataset consists data of exchange rates for the Euro into US dollars (USD) from 1975-01-02 to 2017-04-16 which comprise of 10693 data points. Times series of the data set is visualised shown on Fig. 1.

B. ARIMA model formulation

The sampling frequency of the time series is too granular so down sampling of time series data from days to weeks has been conducted using resampling technique [36]. Down sampled data of original data are shown in Fig. 2.

Applying Dickey-Fuller test [26] it is found that time series data is non-stationary where as it is important to have the time series data to be stationary for applying ARIMA model in order to perform forecasting [27]. Nonlinear log transformation was first applied followed by first order differencing of the log-transformed time series data to make non-stationary time series stationary following the procedure as outlined in [27]. Table I shows that before doing the first order differencing ADF Statistic found to be more than the 5% critical value and the p-value was larger than 0.05 so the null hypothesis of the Dickey-Fuller test cannot be rejected. Whereas after first order

TABLE I
DICKEY-FULLER TEST RESULT COMPARISON.

Original dataset	Differenced dataset
ADF Statistic: -2.025002	ADF Statistic: -36.364359
p-value: 0.275719	p-value: 0.000000
Critical Values:	Critical Values:
10%:-2.567	10%:-2.567
1%:-3.433	1%:-3.433
5%:-2.863	5%:-2.863

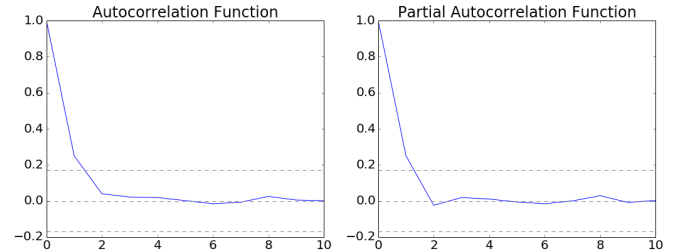


Fig. 3. Autocorrelation Function and Partial Autocorrelation Function plots.

difference ADF Statistic is found less than the 5% critical value and the p-value is lower than 0.05 which suggest that the d parameter of ARIMA model should at least be a value of 1.

Next, for the other two parameters p and q of ARIMA model that had to be found and for this, the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were computed and plotted following the procedure as outlined in [28]. These can be seen in Fig. 3. Autocorrelation describes how the observations of a time series are related to each other which is calculated using observations with previous observation. Autocorrelation Function (ACF) shows the autocorrelation of a time series by a given lag. Whereas a Partial Autocorrelation Function (PACF) computes the relationship between an observation with observations at previous time steps with the relationships of the removed intervening observations.

From the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots, the ' p ' and ' q ' values can be identified as follows:

- p : lag value where PACF cuts off first, so $p=2$.
- q : lag value where ACF chart crosses the upper confidence interval first, so $q=1$.

The optimal values of ARIMA model parameters are found as $p=2$, $d=1$ and $q=1$. So ARIMA (2, 1, 1) model is used for the dataset.

C. Comparative analysis on single step rolling forecast vs direct H step forecast

To perform out of sample forecasting using single step rolling forecast strategy and Direct H step forecast strategy, the dataset has been split into train and test dataset. Where the training dataset is used to train the ARIMA model and test dataset is used for testing the model performance. For data analysis, the last 15 weeks data of the dataset have been used for testing and the rest of the data have been used for

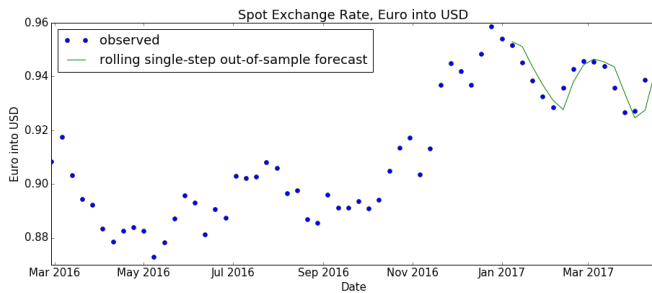


Fig. 4. Forecast response using single step rolling forecast

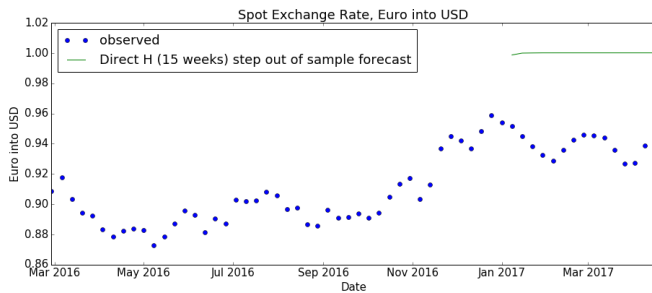


Fig. 5. Forecast response using Direct H step forecast

training. Direct H step forecast and one step forecast rolling strategy have been used to forecast the last 15 weeks exchange rates for the Euro into US dollars (USD). To measure the performance of the ARIMA model, Root Mean Square Error (RMSE) performance measure technique was used. RMSE can be expressed and calculated using the equation given below:

$$RMSE = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m}} \quad (5)$$

where, y_i is actual value, \hat{y}_i is forecast value and m is the number of target output data.

Forecast response from ARIMA (2,1,1) model on the dataset for single step rolling forecast and direct H step forecast is provided in Fig.4 and Fig.5 respectively.

ARIMA (2,1,1) model performance result on the dataset for both single step rolling forecast and direct H step forecast is given in Table 2.

D. Discussion

RMSE performance indicator from Table II shows that single step rolling forecast produces a better outcome with RMSE of 0.005761 compare to direct H step forecast strategy where the RMSE is found 0.063770. Fig.4 shows that running

TABLE II

ARIMA (2,1,1) MODEL PERFORMANCE ON FORECAST STRATEGIES.

Single step rolling forecast	Direct H step forecast
Model: ARIMA (2,1,1)	Model: ARIMA (2,1,1)
RMSE: 0.005761	RMSE: 0.063770

the dataset on ARIMA (2,1,1) model using rolling single step forecast strategy produces a forecast that matches well with test data. This is a deceptive outcome because the rolling forecast recreates the ARIMA model for every time a single point is forecast. It also keeps a track record of all test observations in the list which is seeded with training data to which new observations are appended for each iteration. Such methodology helps to create a forecast line that can be observed in Fig.4. It is made up of single forecast points which have the support of a whole prior true history window behind every single forecast points. Due to this reason, the model does not require to learn from time series rather, each next point would not be far from the last point. Even if the model gets the forecast point wrong, actual forecast value will be seeded to the training data for next forecast which helps to ignore an incorrect forecast at worst.

The observed forecast line can be the case of a scenario of a random walk where forecasting the next point on a Gaussian distribution which allows a random forecast to not drift too widely from the test data. To avoid such scenarios which can be considered to be somewhat deceptive, a forecast of a sequence of points rather a single point needs to be computed. To do such direct H step strategy was used where a sequence of points for last 15 weeks of the dataset was computed. Fig.5 shows the forecast response of direct H step strategy where it is visible that forecast of exchange rates values drifts from the actual forecast thus reduce the model performance with RMSE of 0.063770.

V. CONCLUSIONS

In this paper, a comparative analysis on strategies of time series forecasting has been conducted to explore the outcome of single point rolling forecast for time series forecasting in comparison to direct H step forecasting strategy. Single point rolling forecast strategy and direct H step strategy have been chosen to forecast 15 weeks ahead using the past data of exchange rates. According to our comparative analysis, rolling single point forecast strategy provides a somewhat deceptive outcome for the case that we have considered and direct H step forecast strategy result does help to verify this finding. These strategies have been implemented and tested using the Python ecosystem [36].

REFERENCES

- [1] S. M. Molaei and M. R. Keyvanpour, "An analytical review for event prediction system on time series," 2nd International Conference on Pattern Recognition and Image Analysis (IPRIA), pp. 16, 2015.
- [2] B. Minaei-Bidgoli and S. B. Lajevardi, "Correlation Mining between Time Series Stream and Event Stream," Fourth International Conference on Networked Computing and Advanced Information Management, Gyeongju, pp. 333-338, 2008.
- [3] I. N. Soyiri and D. D. Reidpath, "An overview of health forecasting," Environmental Health and Preventive Medicine, 18(1), pp. 19, 2013.
- [4] M. B. Li, J. W. Zhang and S. Q. Zheng, "Time Series Forecasting for Density of Wood Growth Ring using ARIMA and Neural Networks," International Conference on Machine Learning and Cybernetics, Hong Kong, pp. 2816-2820, 2007.

- [5] S. V. Rodygina and A. V. Rodygin, "Comparative analysis of electrical load forecasting models of industrial plants," *Dynamics of Systems, Mechanisms and Machines (Dynamics)*, Omsk, pp. 1-4, 2016.
- [6] F. Ongenaes, S. Van Looy, D. Verstraeten, T. Verplancke, D. Benoit, F. De Turck, T. Dhaene, B. Schrauwen, and J. Decruyenaere, "Engineering Applications of Artificial Intelligence Time series classification for the prediction of dialysis in critically ill patients using echo state networks," *Eng. Appl. Artif. Intell.*, vol. 26, no. 3, pp. 984996, 2013.
- [7] M. Aczon, D. Ledbetter, L. Van Ho, A. Gunny, A. Flynn, J. Williams, R. Wetzel, T. L. P, L. K. Whittier, V. Pediatric, and I. Care, "Dynamic Mortality Risk Predictions in Pediatric Critical Care Using Recurrent Neural Networks," arXiv preprint arXiv:1701.06675, pp. 118,2017.
- [8] J. Lee and R. G. Mark, "A Hypotensive Episode Predictor for Intensive Care based on Heart Rate and Blood Pressure Time Series," *InComputing in Cardiology*, pp. 8184, 2010.
- [9] C. Eswaran and R. Logeswaran, "An Adaptive Hybrid Algorithm for Time Series Prediction in Healthcare," *In Computational Intelligence, Modelling and Simulation (CIMSIM)*, Second International Conference on, pp. 21-26. IEEE, 2010.
- [10] O. Folorunsho, "Application of Data Mining Techniques in Weather Prediction and Climate Change Studies" *International Journal of Information Engineering and Electronic Business, MECS*, volJ, pp.51-59,2012.
- [11] N. Radzuan, Z. Othman and A. Bakar, "Uncertain Time Series in Weather Prediction," *In Proceedings of Elsevier The 4th International Conference on Electrical Engineering and Informatics, Procedia Technology 11*, pp.557-564, 2013.
- [12] M. Khashei, M. Bijari and S. R. Hejazi, "Combining Seasonal ARIMA Models with Computational Intelligence Techniques for Time Series Forecasting" Springer- Verlag, Berlin, Heidelberg, pp.1091-1105, 2012.
- [13] C. Li and J. W. Hu, "A New Arima-based Neuro-fuzzy Approach and Swarm Intelligence for Time series Forecasting," *Journal of Engineering Applications of Artificial Intelligence 25*, Elsevier, pp. 295-308, 2012.
- [14] J. D. J. Deng and P. Jirutitjaroen, "Short-term load forecasting using time series analysis: A case study for Singapore," *Cybern. Intell. Syst. (CIS)*, 2010 IEEE Conf., no. 1, pp. 16, 2010.
- [15] P. Mandal, T. Senjyu, N. Urasaki, A. Yona, T. Funabashi and A. K. Srivastava, "Price Forecasting for Day-Ahead Electricity Market Using Recursive Neural Network," *IEEE Power Engineering Society General Meeting, Tampa, FL*, pp. 1-8, 2007.
- [16] A. Sorjamaa and A. Lendasse, "Time series prediction using DirRec strategy," *In ESANN*, vol. 6, pp. 143-148, 2006.
- [17] Ch. Haibin and T. Pang-Ning, "Semi-supervised Learning with Data Calibration for Long-Term Time Series Forecasting," *ACM, PPI-9*, 2008.
- [18] S. Ben Taieb, G. Bontempi, A. F. Atiya, and A. Sorjamaa, "A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 70677083, 2012.
- [19] O. Cyprich, V. Konecny and K. Kilianova, "Short-term passenger demand forecasting using univariate time series theory," *Promet- Traffic & Transportation*, Vol. 25, ppJ3-41, 2013.
- [20] M. Cococcioni, E. D'Andrea and B. Lazzarini, "24-hour-ahead forecasting of energy production in solar PV systems," *11th International Conference on Intelligent Systems Design and Applications, Cordoba*, pp. 1276-1281, 2011.
- [21] S. Saroha and S. K. Aggawal, "Multi-step-ahead forecasting of wind-power by different class of neural networks," *In proceeding of IEEE Conference on Engineering and Computational Sciences, Chandigarh*, pp. 1-6,2014.
- [22] D. Guanqun, K. Fataliyev and W. Lipo , "One step and Multi-step-ahead stock prediction using back propagation neural networks," *In proceeding of IEEE International Conference on Information, Communications and Signal Processing , Tainan*, pp. 1-5, 2013.
- [23] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting," *Applied Soft Computing*, Elsevier, pp.2664-2675, 2011.
- [24] A. Meyler, G. Kenny, and T. Quinn, "Forecasting Irish inflation using ARIMA models," *Cent. Bank Financ. Serv. Auth. Irel. Tech. Pap. Ser.*, vol. 3, no. July, pp. 148, 1998.
- [25] [Quandl.com,XUDLERD-Spot-Exchange-Rate-Euro-Into-USD](https://www.quandl.com/data/BOE/XUDLERD-Spot-Exchange-Rate-Euro-Into-USD), [Online]. Available: <https://www.quandl.com/data/BOE/XUDLERD-Spot-Exchange-Rate-Euro-Into-USD>
- [26] D. Dickey and W. A. Fuller, "Distribution of the estimates for autoregressive time series with a unit root," *J. th Am. Stat. Assoc.*, vol. 74, no. 366, pp. 427431, 1979.
- [27] E. P. B. George, M. J. Gwilym, C. R. Gregory, M. L. Greta, *Time Series Analysis: Forecasting and Control*, New Jersey: Wiley Publisher, 2015.
- [28] Chen, Peng, Hongyong Yuan, and Xueming Shu. "Forecasting crime using the arima model," *Fuzzy Systems and Knowledge Discovery, Fifth International Conference on*. Vol. 5. IEEE, 2008.
- [29] Zecchin, C., et al."A new neural network approach for short-term glucose prediction using continuous glucose monitoring time-series and meal information," *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011.
- [30] Ren, Ye, et al. "Random vector functional link network for short-term electricity load demand forecasting," *Information Sciences* , 367, pp.1078-1093, 2016.
- [31] Gupta, Shekhar, and L. P. Wang. "Stock forecasting with feedforward neural networks and gradual data sub-sampling," *Australian Journal of Intelligent Information Processing Systems 11*, no. 4, 2010.
- [32] Zhu, Ming, and Lipo Wang. "Intelligent trading using support vector regression and multilayer perceptrons optimized with genetic algorithms," *In Neural Networks (IJCNN), The 2010 International Joint Conference on*, pp. 1-5. IEEE, 2010.
- [33] Park, Dong-Chul. "A time series data prediction scheme using bilinear recurrent neural network," *In Information Science and Applications (ICISA), 2010 International Conference on*, pp. 1-7. IEEE, 2010.
- [34] Xiong, T., Bao, Y. and Hu, Z. "Beyond one-step-ahead forecasting: evaluation of alternative multi-step-ahead forecasting models for crude oil prices," *Energy Economics*, 40, pp.405-415, 2013.
- [35] Chang, F.J., Chiang, Y.M. and Chang, L.C. "Multi-step-ahead flood forecasting by neural networks," *Hydrological Sciences Journal*, 52(1), pp.114-130, 2007.
- [36] W. McKinney, *Python for Data Analysis*, Sebastopol, CA: O'Reilly, 2013.
- [37] J. Brownlee, *Introduction to Time Series Forecasting with Python*, 2017.