Evaluating Weight Estimation Methods for Hybrid Prediction Using Linear Combination Scheme

Mergani A. Khairalla, Xu-Ning, Nashat T. AL-Jallad

***Abstract*—Time series forecasting is recognized as challenging research issue from all temporal data aspects, traditional machine learning algorithms are unable to satisfy the demands of practical prediction applications of time series data. To deal with noise and concept of non-linear pattern in series data, we propose hybrid prediction model based on linear combination method to perform the ensemble of backpropagation neural networks (BPNN), support vector regression (SVR), and integrated auto-regressive moving average (ARIMA) individual predictors in this paper. Equal, inverse MSE, variance-covariance (VACO), and rank methods are used to estimate weight for participating predictors. Experiments on actual datasets Euro/Sudanese pound exchange rate demonstrate the performance of the hybrid algorithm in terms of prediction accuracy, weight estimation, and combination efficiency. The prediction accuracy is further improved by combining the base predictors based on their linear weights. The results indicate that the proposed hybrid models based on VACO yield better accuracy results compared to the best individual predictors and other hybrid models. Though, the linear method can be utilized as optimal combining methods to achieve more significant forecasting accuracy.**

***Index Terms*—Financial Time Series, Linear Combination Scheme, VACO, Rank Weight, Inverse MSE, Exchange Rate.**

I. INTRODUCTION

Currently, financial time series (FTS) data characterize with nonlinear and uncertain behavior which modify crosswise the time. Therefore, the importance of solving highly nonlinear and time variation complications have been overgrowing. Above problems besides other defects of traditional models affected increasing consideration in machine learning methods [1]. Currency Exchange Rate (CER) is exposed at the highest of the challenging research in FTS field [2]. Even so, traditional statistical models for instance, ARIMA model [3] unable to capture the complexity and non-stationary pattern of non-linear time series [4]. Numerous researchers have presented innovative nonlinear techniques based on machine learning algorithms such as Artificial Neural Network (ANN) models [5]. Support Vector Machines (SVM) methods [6] and data mining models such as K-Nearest Neighbor (KNN) algorithm [7] to predict FTS.

Previous studies mentioned that combining results from heterogeneous models can increase forecast accuracy, this finding motivated by Zhang study [8]. Moreover, no particular method has been established to outperform others in all situations. Combination of forecasts has become a significant issue afterward Bates and Granger study [9] that proposed various combination methods in numerous financial aspects. Thus, a robust method in forecasting studies has been applied to combine the forecasts outcome from single models using several combination methods. This lets the concluding forecast outcome to plagiarize strength from the individual forecasting methods, a characteristic that cannot be reached by a single method, as mentioned in [10] and [11]. The strength of individual forecasting outcomes from the ARIMA, SVM and ANN models is motivating us in this study to explore the effectiveness of these models in the hybrid forecast. Instead of that, nonlinearity is avoidance characteristic of financial time series, and hence, approximating it through a linear model is often insufficient to be detected.

The ANNs algorithm is a class of promising alternatives to learners in this condition. ANNs increased varied attention in the field of FTS forecasting [12] due to their nonlinear, nonparametric, self-adaptive and noise-tolerant properties as aforementioned in [13]. ARIMA investigated for time series data, whereas, an ANN and SVR model offers superior results in a surplus. SVM provided an alternative technique to deal with the lacks of ANNs such as complexity, overfitting to training data and extended training times [14].

Moreover, several researchers have proposed financial pricing tools using SVMs[15]. However, ARIMA model has been efficiently used in forecasting and time series analysis combined with SVM[16], a linear approach reached by scientists Jenkins and Box, however, less proficient in fitting with complicated and nonlinear time series[17].

The ANN recognized for its propensity to identify the non-linear characteristics present within the time series data. Similarly, the SVR models were commonly used within the research field of FTS forecasting [5], [18]. Additionally, It has been implemented in multi-layer neural networks model to predict exchange rate which was reached practical significance. However, the disadvantages of using this model in the case of dense time series, these series wherever linear model and nonlinear model at the constant time[5], 18]. Based on that, we can conclude not acceptable to use non-linear models to predict the complicated time series

Foundation item: This work was supported by the Chinese Scholarship Council, China, Wuhan, under Grants CSC N0. 2015-736012.
Mergani A. Khairalla is with Wuhan University of Technology, Wuhan, P. R. China 430070, Doctoral student. (corresponding author to provide phone: 18871490760, e-mail: mirgani2008@gmail.com.com).
Xu-Ning is with Wuhan University of Technology, Wuhan, P. R. China 430070, Professor. (e-mail: cshao@dlut.edu.cn).
Nashat T. AL-Jallad is with Wuhan University of Technology, Wuhan, P. R. China 430070, Doctoral student.

because of these models might not consider the linear abilities that existing in time series[19].

Several combination methods are described in [20], such as least squares estimators for the weight, relative performance weight, minimization of the loss function, non-parametric combination, and pooling several best predictors. The time-varying method is also discussed where the combined weight may change over time.

Recently, Poncela *et al.* [21]combine several dimensional reduction methods, such as principal component analysis, factor analysis, partial least squares and sliced inverse regression, the forecasting results show that partial least squares, principal component regression and factor analysis have similar performances, and better than the usual benchmark models.

Meanwhile, Siwek *et al.* [22] combine prediction results from neural networks using dimensional reduction techniques, namely principal component analysis, and blind source separation. The best results have been obtained with the application of the blind source separation method by decomposing the data into streams of statistically independent components and reconstructing the noise-omitted time series.

Unlike Siwek *et al.* [22] which use the NN for the individual predictor, in this study, NN is utilized not only for one of the individual predictors but also for forecast combination. Furthermore, while Poncela et al. [21] use dimensional reduction method in the individual predictors, this study employs dimensional reduction as well as feature selection in the forecast combination. Thus, the contribution of this study is the use of NN as a nonlinear method for combining several machine learning predictors as well as statistical predictors.

In this paper, we consider a case study of Euro/Sudanese pound exchange rate in Sudan. Recursive out-of-sample from the 3rd of July of 2016 to the 1st of December 2016 forecasts of the exchange rate is generated using the ARIMA, SVR and BPNN models. All the three models have the same data set to forecast 1-step exchange rate. The impression is to acquire an overview of state of the art in current forecasting literature and then to compare with the FTS problem. We present an empirical analysis of the combination approach in the context of FTS.

Finally, the organized report for this paper is summarized as follows: Section. 2 presented materials and methods which is consist of a brief description of linear combination methods, weights estimators and accuracy measures, in Section. 3 describes the ARIMA, SVR and BPNN models; Section. 4 consist of the data set for this study; the experimental finding results and discussion; Section. 5 represented the conclusion of the research and imagined future works.

II. Materials and Methods

A. Individual Models

In this study, a hybrid model has proposed based on the combination of the autoregressive integrated moving average (ARIMA) model, support vector regression (SVR) models. Since there are many types of ANN, this study focused on the backpropagation neural networks (BPNN).

B. Combination Method

The linear combination method is used for hybrid forecast which calculated through a linear function of the single forecasts from the contributing models. Let's, $Y=[y_1,y_2,\ldots,y_N]^T$ denotes the actual time series, under forecasted using n models, the equation $\hat{Y}^{(i)}=[\hat{y}_1^{(i)},\hat{y}_2^{(i)},\ldots,\hat{y}_N^{(i)}]^{-T}$ denoted the forecast acquired from the i^{th} model. Then, a linear combination of these n forecasted series of the original time series yields by $\hat{Y}^{(c)}=[\hat{y}_1^{(c)},\hat{y}_2^{(c)},\ldots,\hat{y}_N^{(c)}]^{-T}$, which is given by $\hat{Y}^{(c)}=f(\hat{y}_k^{(1)},\hat{y}_k^{(2)},\ldots,\hat{y}_k^{(n)})$"k=1,2,...,N .

Where denoted the linear function of the individual forecasts $\hat{y}_k^{(i)}(i=1,2,\ldots n;k=1,2,\ldots,N)$. Therefore, the final equation for linear combination becomes as:

$$\hat{y}^{(c)}=w_1\hat{y}_k^{(1)},w_2\hat{y}_k^{(2)},\ldots,w_n\hat{y}_k^{(n)}=\sum_{i=1}^{n}w_i\hat{y}_k^{(i)} \quad (1)$$

$$\forall k=1,2,\ldots,N$$

Where w_i denoted the weight assigned to the i^{th} forecasting method.

To guarantee unbiasedness, it is often supposed that the weights add up to unity. The framework of the proposed model is shown in (2) and Fig. 1 as follows:

$$\hat{y}=\sum_{i=1}^{n}w_i\hat{y}_t^{(i)}$$

(2)

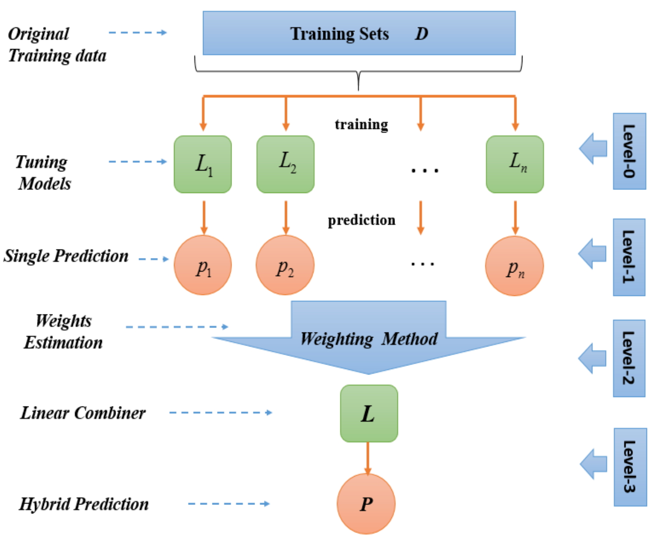

Fig. 1. Proposed Hybrid Model

C. Weights Estimation Techniques

This study investigated four widely used weight estimation methods to assign each model contribution in the hybrid forecast. These methods are briefly defined as follows:

1. Equal-Weights

This method estimates the hybrid forecasts without attention to the historical performance of the individual forecasts; combined weight is allocated equally to each participating model. Where $\hat{y}_t^c$ denoted the combined forecast at time t, where $\hat{y}_t^{(i)}$ denoted the forecast from i^{th} individual forecasting model, and (3) indicated the weight

$$w_i = \frac{1}{m}\sum_{i=1}^{n}\hat{y}_t^{(i)} \quad \textbf{(3)}$$

of the individual forecast for the model i, and m denoted the whole number of single models. Note that other forms of weights are possible as will be seen for the next two methods, but the weights have to satisfy the condition, $\sum_i^m = 1$, $w_i = 1$.

2. Variance-Covariance (VACO) weight

The VACO method determines the weight for each model by considering historical performance of the individual forecasts. This method estimates the weight for a particular model according to (4) as follows:

$$w_i = \frac{\left[\sum_{j=1}^{T}\left(y_j - \hat{y}_j^i\right)^2\right]^{-1}}{\sum_{j=1}^{m}\left[\sum_{j=1}^{T}\left(y_j - \hat{y}_j^i\right)^2\right]^{-1}} \quad \textbf{(4)}$$

Where y_j denoted the j^{th} actual value, where $\hat{y}_j^i$ denoted the j^{th} forecasting value from i^{th} individual forecasting model, and T denoted the total number of out-of-sample points.

D. Inverse MSE weight

This method is assigned a weight of each model by inverse the forecast MSE see (5) of the corresponding model. The Correlations across forecast errors Ignored and set weights relative to the inverse of the models MSE-values, as follows:

$$w_i = \frac{MSE_i^{-1}}{\sum_{i=1}^{m} MSE_i^{-1}} \quad \textbf{(5)}$$

Where MSE denoted mean squared error of model i at the time t .

E. Rank Weights

This method considering the historical performance of individual models, however, forecast models inversely to their rank, the best model gets the first rank, the second best model a rank of 2nd, etc., final weight calculated according to (6) as follows:

$$w_i = \frac{Rank_i^{-1}}{\sum_{i=1}^{m} Rank_i^{-1}} \quad \textbf{(6)}$$

F. Performance Measures

In order to assess and compare the different estimation models, three performance evaluation metrics were considered. To evaluate models performance in this study, both MAPE and RMSE measures considered as error estimators[21] to compare the different experimental results for above models. The following terminology explained that: $y_1 \ldots y_n$ represents a time series, then describes the i^{th} predicted value, where, for $i \le n$, the i^{th} error e_i which is calculated as follows:

$$e_i = y_i - \hat{y} \quad \textbf{(7)}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \quad \textbf{(8)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \quad \textbf{(9)}$$

$$MAPE = \frac{1}{n}\sum_{i=i}^{n}\left|\frac{e_i}{y_i}\right| \times 100\% \quad \textbf{(10)}$$

III. Combination Methods

The main idea of assembling a predictive model by combining diverse models can be described schematically in Fig .1, and the basic framework of the developed combined forecasting model is outlined as follows.

- Initial training data D has m observation and n features so it $(m*n)$.
- Level #0: performing the tuning process to find the optimal model parameters (i.e., SVR model $F(C,k,e)$, ARIMA, and BPNN). Give ranks to each model based on error measures (7-10).

- Level #1: n models (individual) that are trained on D using one method of training (i.e., cross-validation).
- Each model provides a prediction result $p_i i = 1, 2, \ldots, n$ which is then cast into a second level data; the predictions become features for the second level training data.
- Level #2: Here we used four methods for weight estimation techniques in (3-6) to assign each model contribution in the hybrid forecast.
- Level #3: Here we used linear combination method as in (1) to build the hybrid model.
- The hybrid model in (2) can be trained on this data to produce the outcomes which used for final predictions.

A. Case study

The proposed model is verified and validated with real-world time series data sets, the daily exchange rate of the Euro/Sudanese pound (SDG) dataset have investigated. The data has a duration from the 3rd of July of 2016 to the 1st of December 2016. Also, consists of four variables, thus three dependent variables besides one independent variable. Since buying, selling an average exchange rate (Mid-Rate) are financial time series, which is a function over time, and date is the independent variable in this case study.

IV. Results and Discussion

This study applied ARIMA, SVR and BPNN models as benchmark models in the FTS historical data. These models are used separately and integrated to demonstrate their predictability of both linear and nonlinear features in the series. Also, this study proposed four hybrid models based on rank, inverse MSE, VACO and equal weight estimation methods, which they intended to predict SDG next day closing price. To establish the validity of the evaluated practice further procedure prepared by comparing the obtained single models results with obtained results from the hybrid models.

1. Individual Forecasting Results

The individual forecasts obtained from the three forecasting models are generated at 1–ahead step horizons. To ensure consistency of proposed model with previous studies accurate comparison done based on relative error measures, the mean absolute percentage error (MAPE) and root mean squared percentage error (RMSE) see (8), (9), (10). Generally, the SVR is considered to be superior to all of the other models in predicting SDG-EURO exchange. Moreover, the correlation coefficient between SVR and actual data was the highest values. Time-series models (ARIMA) is ranked second, and the third is BPNN model. Also, it can be observed from figure 2,3,4 and Table I all individual models have generated a good forecasting results, the forecast values are so close to the actual values compared to the single predicting models.

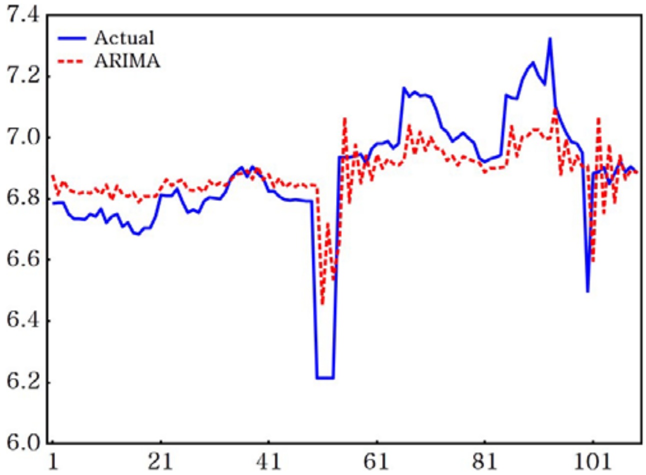

Fig. 2. illustrated comparison between the actual data and predicted data from ARIMA model.

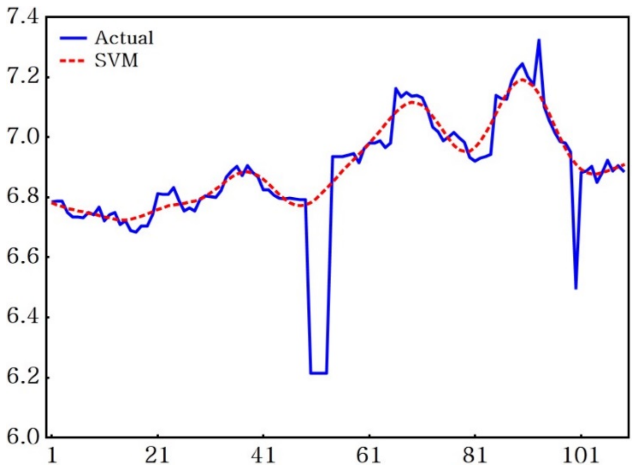

Fig. 3. illustrated comparison between the actual data and predicted data from SVR model.

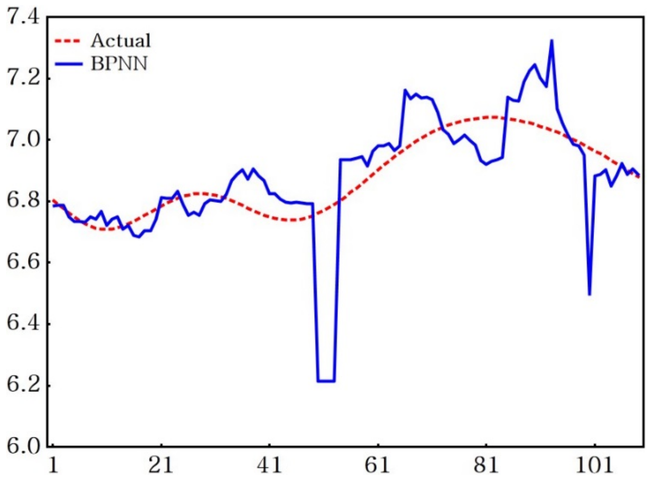

Fig. 4. illustrated comparison between the actual data and predicted data from BPNN model.

TABLE I SUMMARY GOODNESS OF FIT INDIVIDUAL MODELS AS OBSERVED VARIABLE COMPARED WITH ACTUAL EXCHANGE RATE FROM 3RD OF JULY OF 2016 TO THE 1ST OF DECEMBER 2016.

Model	MSE	RMSE	MAPE
SVR	**0.0167**	**0.1254**	**0.0078**
ARIMA	0.0199	0.1412	0.0147
BPNN	0.0204	0.1428	0.0130

Bold numbers represent the best results in individual forecasts.

A. *Weight estimation*

This step conducted after fitting individual models to estimate contribution weight for each model in the hybrid model. As presented in Table II we can observe that individual forecasts effected in VACO, inverse MSE and rank weights the best model gets the higher weight, except equal weight, assign values for respectively model as below:

TABLE II SUMMARY OF ESTIMATED WEIGHTS VALUES FOR ALL MODELS

Benchmark Models	Weight Methods			
	Equal	VACO	Inverse MSE	Rank
ARIMA	0.33	0.31	0.31	0.27
SVR	0.33	0.39	0.39	0.55
BPNN	0.33	0.30	0.30	0.18

B. *Combined Forecasts Result*

This step invistigates after weight assigned respectively to all individual model, the linear combination method in (1) used to combine (ARIMA +SVR+ BPNN) in a hybrid model.

The results of combining forecasts from ARIMA, SVR and BPNN models by using different weight methods are explained in Table IV and Figure 5,6,7,8 respectively. Similarly, Table III reports performance evaluation metrics values for combining methods. The hybrid model based variance-covariance (VACO) combining method performed better than both the individual models and the hybrid models, and also generating large reductions in RMSE of around 12.45% relative to overall other forecasting weights methods.

However, VACO hybrid model can reduce MAPE within 0.95% the obtained forecasting quality and results shown in Fig. 6 and Table III. The result indicates that the VACO hybrid model fitting on the (SDG-EURO) data perform well when measured by different evaluation metrics. Smaller error values mean higher forecasting accuracy. It can be proved that compared to the single forecasting model. The rank combining method performs poorly with less accuracy compared to other combining methods overall forecasting methods with 17.88% for MAPE, the obtained forecasting quality and results shown in Fig. 8 and Table III. Moreover, the predicted values obtained from rank weight method far away from actual values see Table IV, and this point is evident in model plot Fig.8.

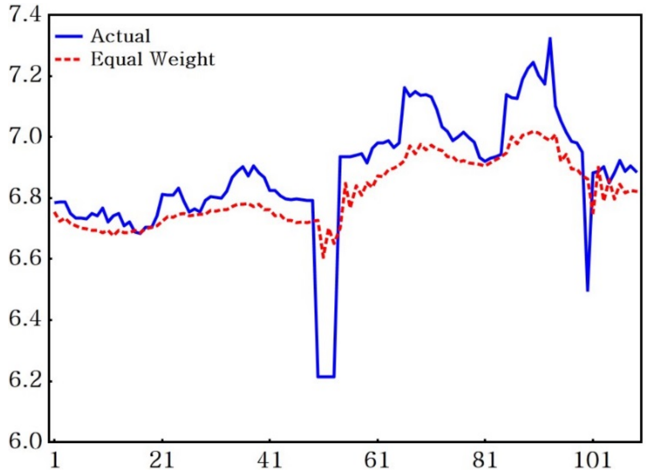

Fig. 5. illustrated comparison between the actual data and predicted data from Equal models.

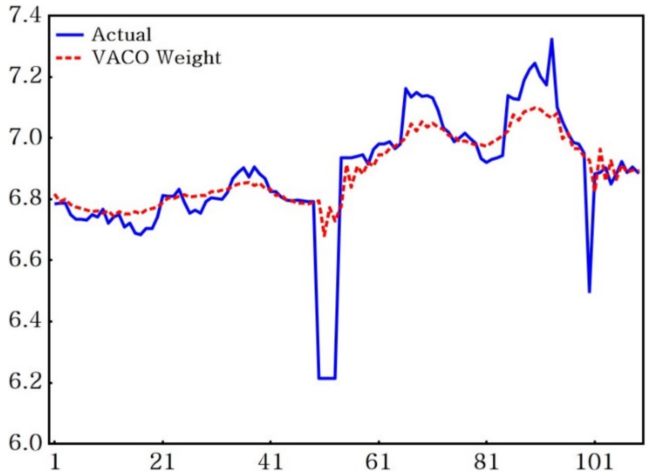

Fig. 6. illustrated comparison between the actual data and predicted data from VACO model.

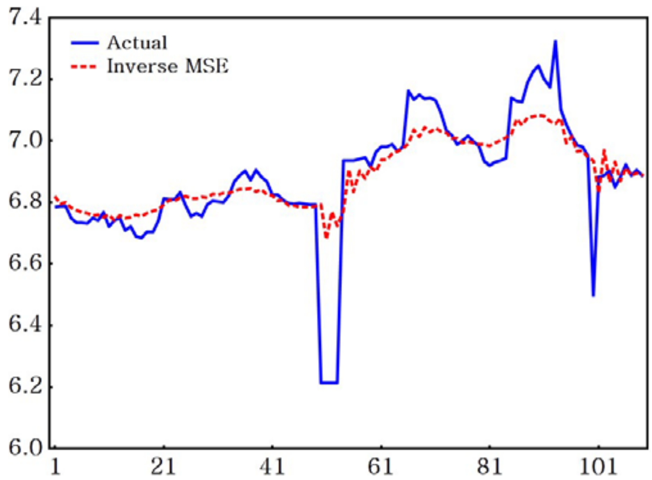

Fig. 7. illustrated comparison between the actual data and predicted data from Inverse MSE model.

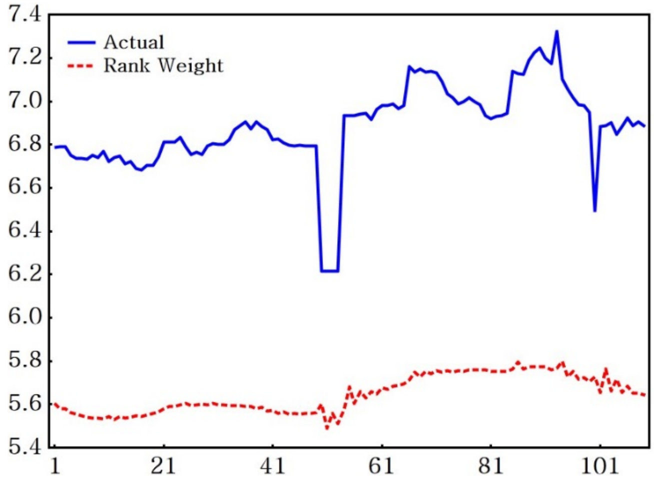

Fig. 8. illustrated comparison between the actual data and predicted data from Rank model.

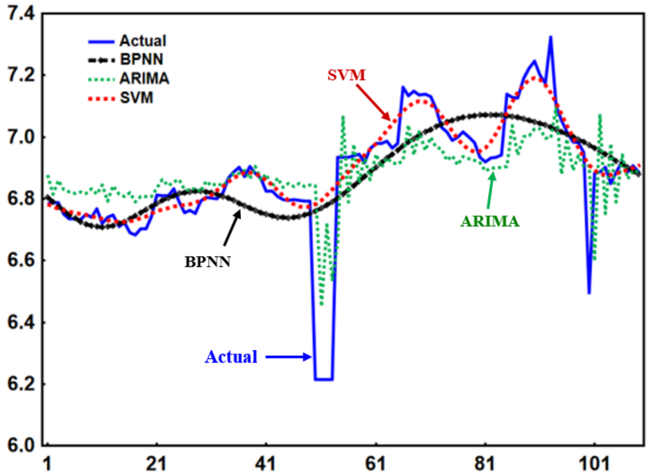

Fig. 9. illustrated comparison between the actual data and predicted data from for single (ARIMA, SVR, and BPNN) models.

TABLE III SUMMARY GOODNESS OF FIT HYBRID MODELS AS OBSERVED VARIABLE COMPARED WITH ACTUAL EXCHANGE RATE FROM 3RD OF JULY OF 2016 TO THE 1ST OF DECEMBER 2016.

Model	MSE	RMSE	MAPE
Equal	0.0195	0.1397	0.0150
VACO	**0.0155**	**0.1245**	**0.0095**
Inverse MSE	0.0159	0.1262	0.0099
Rank	1.5407	1.2412	0.1788

Bold numbers represent best results in combination forecasts

C. Evaluation of Forecast Accuracy

1. Comparing individual model forecasts and combinations

For further analysis, the best performances of the individual models and combinations are evaluated distinguishing among the three prediction error (MSE, RMSE, MAPE) according to the approach followed by [21], all comparisons are performed from two different perspectives. Firstly, we compare the predictive performance of the best individual model (SVR) with that of the best combination (VACO). The results are graphically summarized for the whole period horizon, from Table IV, VACO performs well when forecasting the SDG-EURO data, and the MSE decreased from 1.67% of SVR to 1.55% VACO. Besides, MAPE increases from 0.78% of SVR to 0.95% of VACO. This issue caused due to estimation error and the effects variations of the data generating process have on this error, also the effect of other participation models. In this case, the best performance is obtained with the forecasts combination of three models VACO and Inverse MSE, which outperforms the best individual model SVR.

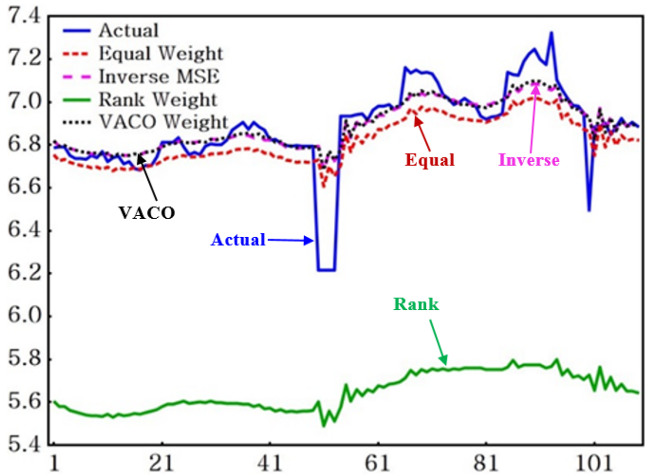

Fig. 10. illustrated comparison between the actual data and predicted data from for single (Equal, VACO, Inverse MSE, and Rank) models.

TABLE IV EFFECT OF CHANGING POINTS OF ACTUAL VALUES OBSERVATIONS FROM JULY 2013 TO DECEMBER 2016.

Actual values		Benchmark Models			Hybrid Models			
Poin t No.	Rate	SVR	ARIMA	BPNN	Equal	VACO	Inverse MSE	Rank
16	6.69	6.73	6.82	6.74	6.69	6.76	6.76	5.55
21	6.81	6.76	6.84	6.79	6.73	6.79	6.79	5.58
35	6.89	6.87	6.88	6.79	6.78	6.85	6.84	5.59
49	6.79	6.78	6.85	6.75	6.72	6.79	6.79	5.56
53	6.21	6.83	6.54	6.79	6.65	6.73	6.72	5.51
66	7.16	7.08	6.92	6.97	6.92	7.00	6.99	5.71
81	6.92	6.97	6.89	7.07	6.91	6.97	6.98	5.75
89	7.22	7.19	7.01	7.06	7.01	7.09	7.07	5.77
93	7.32	7.14	7.17	7.03	6.99	7.07	7.06	5.77
100	6.50	6.91	6.91	6.97	6.86	6.93	6.93	5.73
104	6.85	6.88	6.98	6.93	6.86	6.93	6.93	5.72
Rative Error (%)		3.4	4.3	4.4	4.2	**3.3**	3.4	482.0
Correlation Coefficient (%)		78.5	73.9	68.88	81.88	**81.92**	81.16	75.0

Bold numbers represent best results in combination forecasts

TABLE V COMPARISON OF MAPE OF CANDIDATE FORECASTS AND THEIR COMBINATIONS ERROR FOR EXCHANGE RATE FROM 3RD OF JULY OF 2016 TO THE 1ST OF DECEMBER 2016.

Item		MAPE%				
Model	Weight method	Equal	VACO	Inverse MSE	Rank	Single Error
ARIMA		0. 49	0. 30	0. 31	4. 83	2nd
SVR		0. 49	0. 36	0. 39	9. 98	3rd
BPNN		0. 49	0. 29	0. 30	3. 32	1st
Combined error		1. 47	0. 95	1	18.13	-

2. Hybrid Results Compared to Previous Literature

In this section, the performance of hybrid model based VACO compared to various models mention in literature, such as [8], [22], [20], [23], [24] which are explained in Table VI. Inside the compared error values for all models, the proposed model acquires the lowest MAPE, which is 0.95%. Therefore, we can summarize that the proposed hybrid model outperforms compared to investigative models within the literature. The superior performance of the hybrid model (ARIMA+SVR+BPNN) result will influence each trend and regularity within the original time series, which significantly proved to enhance the financial series prediction with the high-accuracy rate. Besides, was against to conventional ARIMA, SVR and BPNN have a robust ability of generalization, robustness, fault tolerance and convergence ability.

TABLE VI COMPARISON OF MAPES FORECASTING FOR SDG-EURO WITH MODELS IN THE LITERATURE

Model	Dataset	Weight Method	MAPE (%)	Ref.
Hybrid model based on (ARIMA/ANN)	Exchange rate (British pound/US dollar)	N/A	4.99	[8]
Hybrid model based on FLANN-KR	Exchange rate (US/ British Pound)	Kernel	1.15	[22]
Hybrid model based on NLICA-BPN model	Shanghai B-Share stock index	NLICA	0.95	[20]
Hybrid model based on SVR–MFA	Electrical load	MFA	1.69	[23]
Hybrid model based on ARFIMA-ANN	Nordpool electricity market	N/A	6.47	[24]
Proposed Hybrid based VACO model	(SDG-EURO)Exchange rate	VACO	**0.95**	-

Bold numbers represent best results in combination forecasts

D. *Findings from this study*

In this section, we provide some observations based on the experimental analysis of this study to highlight a significant points as follows:

1. The combination method outperforms individual forecasts. Though, improvements in accuracy are adequate even if the individual forecasts are optimized.
2. Combining forecasts from diverse models efficiently reduce the prediction errors and hence provide considerably increased accuracy.
3. The combination leads to significant enhancements in MAPE and MSE. In such cases, the combined forecast can also be used as a benchmark to improve the accuracy of individual predictions.
4. Proved that weight estimate method that related to (mean errors) selects as the best combiner from suggested combination methods so that it best hybrid in a linear architecture.

V. CONCLUSION

This paper evaluated and compared different combiners to optimize the hybrid model for time series prediction. We have conducted an empirical study to assess and compare the performance of all models using exchange rate between Euro/SDG datasets. The main findings of this study are as follows: The results confirm that individual models are not reliable as their performance is inconsistent and unstable across forecasting data. Although none of the hybrid models was consistently the best, many of them were frequently among the best models for each dataset. The hybrid with the linear combiner VACO was the best model when considering the average rank of each model across all performance measures. In individual forecasting models, the empirical results proved that, compared to benchmark models, the VACO model offer accurate predictions that dominate the forecast from the all benchmark model with reductions in MAPE of around 0.78% in all cases and overall forecasting horizons. The study also used some of the recently studied linear combination methods to estimate weights from single models. The empirical results of combining forecasts show that MAPE value for VACO method is relatively smaller than the MAPE of other combining methods which are in turn better than results of the benchmark models. The VACO model as well outperforms the best individual forecasting models (SVR) for all variables and at all forecasting horizons with sizable reductions in RE of around 3.4-3.3% of the RE of the best original forecasts.

Based on the empirical results, we do not recommend using the rank weight method as it was performed worse than the other models across the three performance measures. Moreover, we do not recommend using the equal combiner because this method doesn't consider individual performance.

This study has contributed interesting empirical based insights into the application of different hybrid models in estimating short-term exchange rate. Future works include assessment of varying machine learning models with the non-linear combiner. In addition, interesting future research is to investigate ensemble learning with the best set of base models for heterogeneous ensembles. Another direction of future work is to apply hybrid learning models in other time series prediction problems. This work includes classification and regression problems of energy consumption and changeability prediction.

REFERENCES

[1] Beneki, C. and M. Yarmohammadi, "Forecasting exchange rates: An optimal approach. Journal of Systems Science and Complexity", 2014. 27(1): p. 21-28.

[2] Nair, B. B., Sai, S. G., Naveen, A. N., Lakshmi, A., Venkatesh, G. S., & Mohandas, V. P., "A GA-artificial neural network hybrid system for financial time series forecasting", in Information Technology and Mobile Communication. 2011, Springer. p. 499-506.

[3] Karia, A.A., I. Bujang, and I. Ahmad, "Fractionally integrated ARMA for crude palm oil prices prediction: case of potentially overdifference.", Journal of Applied Statistics, 2013. 40(12): p. 2735-2748.

[4] Dhamija, A. and V. Bhalla, Financial time series forecasting: comparison of neural networks and ARCH models. International Research Journal of Finance and Economics, 2010. 49: p. 185-202.

[5] Khashei, M. and M. Bijari, A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Applied Soft Computing, 2011. 11(2): p. 2664-2675.

[6] Pai, P.-F., et al., Time series forecasting by a seasonal support vector regression model. Expert Systems with Applications, 2010. 37(6): p. 4261-4265.

[7] Ahmed, N. K., Atiya, A. F., Gayar, N. E., & El-Shishiny, H.., "An empirical comparison of machine learning models for time series forecasting. Econometric Reviews", 2010. 29(5-6): p. 594-621.

[8] Zhang, G.P., "Time series forecasting using a hybrid ARIMA and neural network model.", Nerocomputing, 2003. 50: p. 159-175.

[9] Bates, J.M. and C.W. Granger, "The combination of forecasts." Or, 1969: p. 451-468.

[10] Zeng, X., L. Shu, and J. Jiang, "Fuzzy Time Series Forecasting based on Grey Model and Markov Chain.", International Journal of Applied Mathematics, 2016. 46(4).

[11] Stock, J.H. and M.W. Watson, "Combination forecasts of output growth in a seven-country data set.", Journal of Forecasting, 2004. 23(6): p. 405-430.

[12] Hua, X., D. Zhang, and S.C. Leung. "Exchange rate prediction through ANN Based on Kernel Regression.", in Business Intelligence and Financial Engineering (BIFE), 2010 Third International Conference on. 2010. IEEE

[13] Khashei, M. and M. Bijari, "An artificial neural network (p, d, q) model for timeseries forecasting.", Expert Systems with applications, 2010. 37(1): p. 479-489.

[14] Yongpan Ren, Jingli Mao, Yong Liu, and Yingzhe Li, "A Novel DBN Model for Time Series Forecasting," IAENG International Journal of Computer Science, vol. 44, no.1, pp79-86, 2017.

[15] Vapnik, V.N., "An overview of statistical learning theory.", IEEE transactions on neural networks, 1999. 10(5): p. 988-999.

[16] Nie, H., et al., "Hybrid of ARIMA and SVMs for short-term load forecasting.", Energy Procedia, 2012. 16: p. 1455-1460.

[17] Lijie Sun, Cheng Shao, and Li Zhang, "Hybrid Dynamic Continuous Strip Thickness Prediction of Hot Rolling," Engineering Letters, vol. 25, no.3, pp268-276, 2017.

[18] Lu, J., J. Huang, and F. Lu, "Time Series Prediction Based on Adaptive Weight Online Sequential Extreme Learning Machine.", Applied Sciences, 2017. 7(3): p. 217.

[19] Yu, L., S. Wang, and K.K. Lai, "A neural-network-based nonlinear metamodeling approach to financial time series forecasting.", Applied Soft Computing, 2009. 9(2): p. 563-574.

[20] Dai, W., J.-Y. Wu, and C.-J. Lu, "Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes.", Expert Systems with Applications, 2012. 39(4): p. 4444-4452.

[21] Bergmeir, C. and J.M. Benítez, "On the use of cross-validation for time series predictor evaluation.", Information Sciences, 2012. 191: p. 192-213.

[22] Hua, X., D. Zhang, and S.C.H. Leung, "Exchange Rate Prediction through ANN Based on Kernel Regression.", 2010: p. 39-43.

[23] Kavousi-Fard, A., H. Samet, and F. Marzbani, "A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate Short Term Load Forecasting.", Expert Systems with Applications, 2014. 41(13): p. 6047-6056.

[24] Chaâbane, N., "A hybrid ARFIMA and neural network model for electricity price prediction.", International Journal of Electrical Power & Energy Systems, 2014. 55: p. 187-194.