

Evaluation Study of Linear Combination Technique for SVM related Time Series Forecasting

Xian Cheng
School of Economics and
Management,
Southwest Jiaotong University
beyoungbelong@gmail.com

Ji Wu
Business School,
Sun Yat-sen University
wuji3@mail.sysu.edu.cn

Jin Xu
School of Economics and
Management,
Southwest Jiaotong University
Jinxu@swjtu.edu.cn

Abstract

Time series forecasting and SVM are widely used in many domains, for example, smart city and digital services. Focusing on SVM related time series forecasting model, in this paper we empirical investigate the performance of eight linear combination techniques by using M3 competition dataset which includes 3003 time series. The results reveals that the “forecast combination puzzle” is not exist for combining SVM related forecasting model as the simple average is almost the worst combination technique.

1. Introduction

Smart city and digital services call for advanced and efficient decision support techniques. Time series forecasting, which is concerned about the prediction of future values based on a series of historical observed data, is a crucial component in decision making process and widely used in smart city and digital services, for example, traffic flow prediction in intelligent transportation system for smart city[1]. Forecasting models devoting to reasonable accuracy are an important but quite difficult work which has received a considerable attention during the past several decades. There are two main branches for these forecasting models, one is the traditional statistics model, such as exponential smoothing[2], autoregressive integrated moving average (ARIMA)[3], and autoregressive conditional heteroskedasticity (ARCH)[4], the other is the emerging artificial intelligence based model, such as neural network[5], support vector machine (SVM)[6], and K-nearest neighbor model[7]. In these time series forecasting models, researchers often concentrate on identifying the best individual model. However, combination forecasting which integrates several individual models

has been widely proved to a highly successful forecasting strategy in many fields[8, 9], as combination forecasting significantly improves the forecasting accuracy as well as often produces better results than the best individual forecasting model[10].

Indeed, the individual forecasting models are often problem-specific and none can be viewed as the robust and absolutely optimal model for all situations. So the best strategy is combining or aggregating multiple forecasting models instead of for choosing a single forecasting model. Combination forecasting by taken into account alternative time series models started in 1960's with the original work of Bates and Granger[11] and since then it has been extensively studied in the domain of forecasting. The combining technique can be classified into linear combination[12] and nonlinear combination[13]. To linearly combine constituent forecasting model, by assigning a suitable weight to each model, is the most intuitive and popular strategy for forecasting[14]. There are a large number of literatures aiming at deriving a weight for constituent forecasting, e.g., the simple average, the trimmed mean[15], the Winsorized mean[15], the median[12], the Bates-Granger method[11]etc. However, a vast body of empirical study and extensive simulations found repeatedly that the simple average combination forecast is a difficult benchmark to beat, and commonly outperforms many sophisticated combining techniques[16], this is known as the “forecast combination puzzle”.

In this paper, we empirical investigate this puzzle by comprehensive study the linear combination techniques in the field of artificial intelligence based time series forecasting model, particularly, the SVM related models. In detail, we build eight constituent forecasting models based on different parameterizations of the SVM model, and then test eight different linear combination techniques on the renowned M3 competition dataset[17] to check whether the simple average forecast combination is

always best. The M3 competition dataset contains 3003 time series data (Micro, Industry, Macro, finance, etc.) and different time intervals for successive observations, including 645 yearly series, 756 quarterly series, 1428 monthly series and 174 other series. As the diversification of the data type, the M3 dataset has become an important dataset for comparing alternative forecasting model. According to the comprehensive study, we find that the “forecast combination puzzle” is not exist in the situation for combining SVM related forecasting model as the simple average is almost the worst combination technique. We also find that the best combination technique is different for different type time series.

The rest of the paper is organized as follow. Section 2 presents a brief description of the related theory including SVM in forecasting and combination forecasting technique. Section 3 and 4 contain the setup and results of the empirical investigation. Section 5 provides a brief discussion, followed by the conclusions and implications.

2. Related Theory

2.1. SVM in Forecasting

As in this heading, they should be Times 11-point boldface, initially capitalized, flush left, with one blank line before, and one after.

SVM technique, as a well-known statistical learning algorithm, was originally developed for classification problems in pattern recognition domain. By the introduction of Vapnik’s in sensitive loss function, this technique can be used for non-linear regression estimation[18]. As an outstanding method using a high-dimensional feature space, and penalizing the ensuing complexity using a penalty term, SVM takes into account the global and unique solutions and do not suffer from multiple local minima. So it reveals a remarkable ability of balancing model accuracy and model complexity in the field of regression [6, 19].

SVM related regression seeks to estimate functions: $f(\mathbf{X}) = (\mathbf{w}, \mathbf{X}) + \mathbf{b}$, where $\mathbf{w}, \mathbf{X} \in \mathbf{R}^n$, $\mathbf{b} \in \mathbf{R}$, based on data $(\mathbf{X}_1, \mathbf{y}_1), \dots, (\mathbf{X}_n, \mathbf{y}_n) \in \mathbf{R}^n \times \mathbf{R}$, by minimizing the regularized risk functional $\frac{\|\mathbf{w}\|^2}{2} + \mathbf{C} \times \mathbf{R}_{train}^\epsilon$, where \mathbf{C} is a constant determining the trade-off between minimizing the training error $\mathbf{R}_{train}^\epsilon = \frac{1}{l} \sum_{i=1}^l |\mathbf{y}_i - f(\mathbf{X}_i)|_\epsilon$ and the model complexity term $\|\mathbf{w}\|^2$. In here the so-called ϵ -insensitive loss function $|\mathbf{y}_i - f(\mathbf{X}_i)|_\epsilon = \max\{0, |\mathbf{y}_i - f(\mathbf{X}_i) - \epsilon|\}$. it does not penalize errors below some $\epsilon > 0$ which are chosen a priori. There are two types of SVM regression: ϵ -SVM regression[20] and ν -SVM regression [21]. Compared

with ν -SVM regression, the ϵ -SVM regression model is extremely sensitive to the choosing of the ϵ parameter and requires the desired accuracy of the approximation to be specified beforehand. What’s more, to conduct non-linear regression using SVM, it is necessary to adopt a kernel function that satisfies Mercer’s conditions. There are several kernel functions that satisfy Mercer’s conditions, such as Gaussian, polynomial, and hyperbolic tangent [22].

SVM regression models are widely used in time series forecasting domain. Such as, in the paper of [23], the authors examined the feasibility of SVM in financial time series forecasting by comparing with neural network model. For the paper of [24], a SVM regression model was used to predict the stock price index. Furthermore, Sapankevych and Sankar presented a comprehensive review on SVM related time series forecasting model [22]. Besides, by implementing a hybrid chaos-based SVM model, [25] predicted the EUR/USD, GBP/USD, NZD/USD, AUD/USD, JPY/USD and RUB/USD exchange rates, and compared to chaos-based NNs model and several traditional non-linear forecast models. Recently, in the paper of [6], a hybrid rolling genetic algorithm based SVM model was applied to forecast the EUR/USD, EUR/GBP and EUR/JPY exchange rates.

2.2. Combination Forecasting

It is well known that no individual forecasting model work best for all dataset, the diversification of data, including data type, data size, normality, linearity, and correlation may affect the performance of forecasting model considerably[26]. So, instead of finding the best accurate individual forecasting model, building multiple forecasting models and combining them has been found to be effective way to improve the overall forecasting accuracy. As a widely used strategy, the linear combination has been well applied in time series forecasting domain. In this paper, we mainly focus on eight linear combination techniques which are presented as follow.

At first, we have to define some variables being considered in this paper. In-sample or training time series composed of t successive observations: $\mathbf{y}_{train} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t)'$, Out-of-sample or test time series with maximum horizon H : $\mathbf{y}_{test} = (\mathbf{y}_{t+1}, \mathbf{y}_{t+2}, \dots, \mathbf{y}_{t+H})'$. N individual models produce a forecast vector at time $t + h$ ($h \leq H$), $\hat{\mathbf{y}}_{t+h|t} = (\hat{\mathbf{y}}_{t+h|t,1}, \hat{\mathbf{y}}_{t+h|t,2}, \dots, \hat{\mathbf{y}}_{t+h|t,N})'$ based on in-sample time series \mathbf{y}_{train} . A linear combination of N individual forecast models is obtained based the following equation.

$$\hat{\mathbf{y}}_{t+h|t}^C = \sum_{i=1}^N \mathbf{w}_i \hat{\mathbf{y}}_{t+h|t,i} \quad (1)$$

where $\sum_{i=1}^N \mathbf{w}_i = \mathbf{1}$, and $\mathbf{w}_i \geq 0$.

2.2.1. Simple Average. As the simplest form of linear combination, simple average is a robust combination technique that is considered to be “hard to beat” [12, 27]. However, simple averaging may not be a suitable combination method when some of the predictors are biased [28]. For simple average combination, the weight for all the constituent forecasting models is always the same, regardless of the horizon H:

$$\hat{y}_{t+h|t}^C = \frac{1}{N} \sum_{i=1}^N \hat{y}_{t+h|t,i} \quad (2)$$

2.2.2. Median Method, Trimmed Mean and Winsorized Mean. The median combination method has been proposed by authors such as [12, 29], it is an appealing simple, rank-based combination method which gives a weight of 1 to the median forecasting and a weight of 0 to all other constituent forecasting models. The trimmed mean combination technique is an interpolation between the simple average and the median, and is less sensitive to outliers than simple average. By using a trim factor λ (i.e., the top/bottom $\lambda\%$ are trimmed), the trimmed mean has the ability of excluding the worst performing of the constituent forecast models. Like the trimmed mean, the winsorized mean which integrate another trim technique, is a robust statistic that is less sensitive to outliers than the simple average. The trimmed mean and winsorized mean are also widely used in forecasting literatures, such as [15], [12] and [27].

The median method:

$$\hat{y}_{t+h|t}^C = \text{median}(\hat{y}_{t+h|t}) \quad (3)$$

The trimmed mean:

$$\hat{y}_{t+h|t}^C = \frac{1}{N(1-2\lambda)} \sum_{i=\lambda N+1}^{(1-\lambda)N} \hat{y}_{t+h|t,i} \quad (4)$$

The winsorized mean: $\hat{y}_{t+h|t}^C = \frac{1}{N} [k\hat{y}_{t+h|t,k} + \sum_{i=k+1}^{N-K} \hat{y}_{t+h|t,i} + k\hat{y}_{t+h|t,N-k}]$ (5)

2.2.3. Bates & Granger Technique. Based on the portfolio diversification theory, Bates and Granger [11] has introduced a combination technique by using the diagonal elements of the estimated mean squared forecast error matrix to obtain the combining weights. Due to difficulties in precisely estimating covariance matrix, the Bates & Granger Technique ignores correlation between forecasting models. This technique is present as follow:

$$w_i^{BG} = \frac{MSE_i^{-2}}{\sum_j^N MSE_j^{-2}} \quad (6)$$

$$\hat{y}_{t+h|t}^C = \sum_{i=1}^N w_i^{BG} \hat{y}_{t+h|t,i} \quad (7)$$

Where the MSE_i is the estimated mean squared forecasting error for the forecast model i .

2.2.4. Inverse Rank Combination Technique.

The inverse rank combination technique, which is proposed by [30], computes the combination weights which are inversely proportional to the rank of the constituent forecasting models.

the weight is obtained by:

$$w_i^{Inv} = \frac{Rank_i^{-1}}{\sum_j^N Rank_j^{-1}} \quad (8)$$

and then the combination forecast is obtained by:

$$\hat{y}_{t+h|t}^C = \sum_{i=1}^N w_i^{Inv} \hat{y}_{t+h|t,i} \quad (9)$$

2.2.5. Eigenvector Combination Technique.

The standard eigenvector combination technique which is proposed by [31], retrieves the combining weights based on the sample estimated mean squared forecast error matrix as follows: Suppose Σ is the squared forecast error matrix, and the N positive eigenvalues are then arranged in ascending order ($\Phi_1, \Phi_2, \dots, \Phi_N$), and w^j is defined as the eigenvector corresponding to Φ_j , and the combining weight vector is chosen corresponding to the minimum of $(\frac{\Phi_1}{d_1^2}, \frac{\Phi_2}{d_2^2}, \dots, \frac{\Phi_N}{d_N^2})$, where the $d_j = e^1 w^1$.

An variant of the standard eigenvector combination is the bias-corrected eigenvector approach which is proposed also by [31] based on the idea that if one or more forecast model produce biased predictions, the accuracy of the standard eigenvector combination can be improved by excluding the bias. Comparing with standard eigenvector combination, the bias-corrected eigenvector approach applied the centered MSPE matrix after extracting the bias.

3. Parameter Setting

For SVM regression model, the key parameters, which control the complexity of the model, are ϵ , ν , C and σ_{kernel} , as there are several kernel functions (Linear kernel, Polynomial kernel, Radial Basis kernel, and so on). many literatures focus on the analysis of the parameters setting, such as [32] and [18]. Considering that the M3 competition dataset is large (3003 time series), it is very hard and impossible to setting the suitable parameter for all series. So we set several discrete values for these key parameters: $C = \{1, 5, 10, 20, 100\}$, $\epsilon = \nu = \{0.01, 0.1, 0.4, 0.9\}$, $\sigma_{\text{kernel}} = \{0.1, 1, 10\}$, and we also take into account three types of kernel functions: Linear kernel, Polynomial kernel, Radial Basis kernel.

For one time series data, we partition it into two parts: training set and testing set, and use the training data to train a series of SVM regression models obtained by adjusting the key parameters. We select the best 8 SVM regression models based on the in-

sample mean squared forecasting error (MSE) to form the 8 constituent forecasting models. At last, the 8 constituent forecast models are linear combined based on different combination techniques as presented in subsection 2.2, the performance of combination

techniques is evaluated by the testing data. Figure 1 shows the research framework of the linear combination.

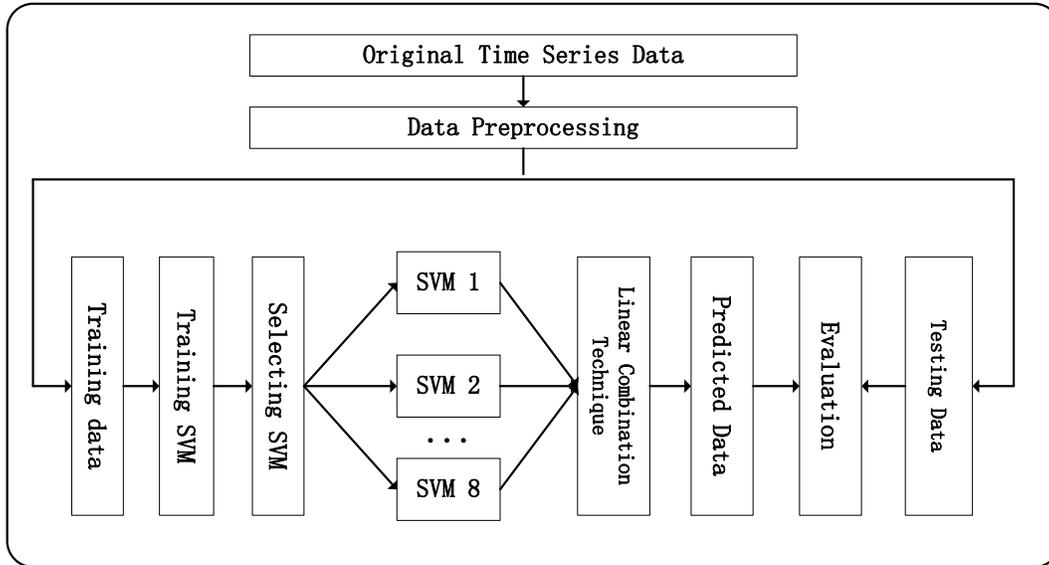


Figure 1. The Research Framework of Linear Combination

4. Empirical Investigation

4.1. Design

This section presents the empirical study, performed to investigate the “forecast combination puzzle” by comparing the performance of the eight linear combination techniques. In order to make a comprehensive investigation, we adopt the M3 competition dataset which contains 3003 time series. This dataset has become an important benchmark and widely used stage for testing forecasting model. Table 1 presents the M3-competition. For the 3003 time series, the data preprocessing is conducted based on log transformation, deseasonalization and scaling. The testing data for each time series is the last 6 / 8 / 18 / 8 observations corresponding to yearly series, quarterly series, monthly series, and other series, respectively.

The rest of each series is consider as the training data. A number of SVM model is trained by the training data and the best 8 SVM model is selected as the constituent forecasting models. And then the constituent forecasting models are combined by combination techniques. The testing data is then used to evaluate the performance of the eight different combination techniques which are abbreviated as SA(simple average), MED(median method), TM(trimmed mean), WM(winsorized mean), BG(Bates & Granger technique), INV(inverse rank combination technique), SEC(standard eigenvector combination technique), BEC(bias-corrected eigenvector combination Technique). In order to evaluate the performance, here we use three accuracy measures— Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE), which are widely used in forecast domain.

Table 1. The 3003 Time Series of the M3-Competition (source: <https://forecasters.org>)

Types of Time Series Data							
Interval	Micro	Industry	Macro	Finance	Demog	Other	Total
Yearly	146	102	83	58	245	11	645
Quarterly	204	83	336	76	57	0	756
Monthly	474	334	312	145	111	52	1428

Other	4	0	0	29	0	141	174
Total	828	519	731	308	413	204	3003

4.2. Results

Table 2 presents the performance of the linear combination techniques for yearly series. For each row, the best and second-best combination techniques are presented in boldface and italics respectively. Here the $RMSE_{rank}$ is the average ranking of all the yearly data, the ranking order based on RMSE for the eight combination technique is quite different for different series, so we sum the ranking order for all yearly series,

and then calculate the average score. The smaller of $RMSE_{rank}$, the better for the corresponding combination technique. Similarly, The MAE_{rank} and the $MAPE_{rank}$ are obtained based on the same procedure. Table 2 reveals that the BEC (bias-corrected eigenvector combination Technique) is the best combination technique for yearly series data, and the SEC (standard eigenvector combination technique) is the second-best technique. However, the SA (simple average combination) is almost the worst combination technique.

Table 2. The Performance of Linear Combination Technique for Yearly Series (#obs=645)

Measures	BG	SEC	BEC	INV	MED	SA	TM	WM
$RMSE_{rank}$	3.09	2.39	1.67	3.35	6.33	6.29	6.17	6.20
$RMSE_{average}$	219.08	<i>131.77</i>	116.38	250.17	323.74	325.84	323.25	323.33
MAE_{rank}	3.03	2.53	1.92	3.34	5.99	6.52	6.07	6.11
$MAE_{average}$	154.43	<i>101.69</i>	90.51	174.51	217.99	226.91	221.50	221.64
$MAPE_{rank}$	3.00	2.57	1.96	3.34	6.01	6.41	6.10	6.11
$MAPE_{average}$	4.88	<i>3.01</i>	2.73	5.35	6.84	6.87	6.83	6.83

Table 3 shows the performance of the linear combination technique for quarterly series data. From this table, we conclude several interesting results: on one hand, the INV (inverse rank combination technique) and BG (Bates & Granger technique) are the best and second best combination technique, respectively, when considering any of the three rank measures— $RMSE_{rank}$, MAE_{rank} and $MAPE_{rank}$. On the other

hand, by considering any of the three average measures— $RMSE_{average}$, $MAE_{average}$ and $MAPE_{average}$, the best and second best combination technique are the BEC (bias-corrected eigenvector combination Technique) and SEC (standard eigenvector combination technique) combination technique, respectively.

Table 3. The Performance of Linear Combination Technique for Quarterly Series (#obs=756)

	BG	SEC	BEC	INV	MED	SA	TM	WM
$RMSE_{rank}$	2.75	3.84	3.23	2.54	6.10	5.99	5.73	5.81
$RMSE_{average}$	165.07	<i>147.59</i>	144.52	173.86	211.00	211.23	210.24	210.41
MAE_{rank}	<i>2.61</i>	4.03	3.74	2.50	5.79	6.02	5.65	5.67
$MAE_{average}$	111.39	<i>110.28</i>	109.08	117.57	141.54	145.53	142.71	142.84
$MAPE_{rank}$	2.63	3.95	3.65	2.56	5.79	6.03	5.69	5.71
$MAPE_{average}$	3.25	<i>2.73</i>	2.69	3.37	3.95	4.05	3.98	3.98

Similar insights are provided by table 4, where the result of linear combination techniques for monthly series data is presented. Taken into account any of the three rank measures, the best technique is still

INV (inverse rank combination technique), and closely followed by BG (Bates & Granger technique). However, there is a slight difference when consider the three average measures. The best one and second best one

are the BEC(bias-corrected eigenvector combination Technique) and SEC(standard eigenvector combination technique) for the $RMSE_{average}$ and $MAE_{average}$, but the order for the best and second best techniques is

inverted by considering the $MAPE_{average}$. The major reason of the similar insights for table 3 and 4 is that the quarterly and monthly series always contains the seasonal and trend patterns.

Table 4. The Performance of Linear Combination Techniques for Monthly Series (#obs=1428)

	BG	SEC	BEC	INV	MED	SA	TM	WM
$RMSE_{rank}$	2.86	3.88	3.24	2.44	6.47	5.52	5.86	5.73
$RMSE_{average}$	299.34	252.42	247.62	328.97	410.28	399.66	402.83	402.41
MAE_{rank}	2.72	4.06	3.94	2.33	5.95	5.83	5.65	5.52
$MAE_{average}$	196.79	179.26	178.48	213.35	263.17	266.30	262.14	261.90
$MAPE_{rank}$	2.72	4.10	3.94	2.35	5.97	5.81	5.63	5.50
$MAPE_{average}$	5.39	5.09	5.20	5.82	7.24	7.31	7.20	7.19

Table 5 reports the performance of linear combination techniques for other series data. First, we find that the INV(inverse rank combination technique) and BG(Bates & Granger technique) are the best and second best technique when considering the rank based measures. Second, when taking into account either $RMSE_{average}$ or $MAE_{average}$, we find that the

conclusion is inverted. Furthermore, when considering the , we conclude that the best one is BG(Bates & Granger technique) and the second best one is the SEC (standard eigenvector combination technique) and BEC(bias-corrected eigenvector combination Technique).

Table 5. The Performance of Linear Combination Techniques for other Series (#obs=174)

	BG	SEC	BEC	INV	MED	SA	TM	WM
$RMSE_{rank}$	2.17	6.07	5.32	1.49	6.32	4.49	5.19	4.93
$RMSE_{average}$	98.24	105.12	104.53	102.12	121.53	117.99	119.82	119.77
MAE_{rank}	2.14	5.65	5.53	1.83	6.26	4.17	5.34	5.07
$MAE_{average}$	66.80	73.36	73.27	69.75	83.07	80.99	82.31	82.31
$MAPE_{rank}$	2.16	5.59	5.48	1.89	6.29	4.09	5.41	5.09
$MAPE_{average}$	1.24	1.35	1.35	1.52	1.92	1.90	1.93	1.93

5. Conclusions

In this paper, we conduct a comprehensive investigation on the performance of eight linear combination techniques by using M3 competition dataset, particularly, we concentrate on SVM related forecasting model. Several conclusions are obtained from our study. First, we find that the “forecast combination puzzle” is not exist in the SVM related forecast model as the simple average is almost the worst combination technique for all the situation in our empirical investigation. This result coincides with several related literatures, such as [3], [5] and [31]. Besides, we also find that the best combination technique is different for either different type series or based on different accuracy measures. For example,

when accuracy measure $MAPE_{average}$ is considered, the best combination techniques for quarterly and monthly series are BEC and SEC, respectively. Lastly, we conclude that the recommended combination techniques for yearly series are BEC and SEC, but for quarterly and monthly series are BG, SEC, BEC and INV. These results are very useful in the domain of time series forecasting which is a crucial component in decision making process for a variety of fields. However, there are also some shortcomings. We only consider the SVM related forecasting model, but how about statistics forecasting model, or other artificial intelligence based model, such as neural network [5] and K-nearest neighbor model [7]. We also only consider eight linear combination techniques, how about the performance of other line combination techniques, or the non-linear combination techniques.

All the mentioned questions are very interesting and should be consider in future study.

6. Acknowledgements

This study is supported by grants from National Natural Science Foundation of China (NO. 71701172, 71472158, 71490725). We would like to thank the anonymous HICSS-49 reviewers whose constructive comments let us to improve this paper.

7. References

- [1]Y. S. Lv, et al., "Traffic Flow Prediction With Big Data: A Deep Learning Approach," *Ieee Transactions on Intelligent Transportation Systems*, vol. 16, pp. 865-873, Apr 2015.
- [2]F. Petropoulos, et al., "Exploring the sources of uncertainty: why does bagging for time series forecasting work?," *European Journal of Operational Research*, 2018/02/02/ 2018.
- [3]V. G. Azevedo and L. M. S. Campos, "Combination of forecasts for the price of crude oil on the spot market," *International Journal of Production Research*, vol. 54, pp. 5219-5235, 2016.
- [4]A. Aue, et al., "Functional generalized autoregressive conditional heteroskedasticity," *Journal of Time Series Analysis*, vol. 38, pp. 3-21, 2017.
- [5]R. D. V. dos Santos and M. Vellasco, "Neural Expert Weighting: A NEW framework for dynamic forecast combination," *Expert Systems with Applications*, vol. 42, pp. 8625-8636, Dec 2015.
- [6]G. Sermpinis, et al., "Modeling, forecasting and trading the EUR exchange rates with hybrid rolling genetic algorithms-Support vector regression forecast combinations," *European Journal of Operational Research*, vol. 247, pp. 831-846, Dec 2015.
- [7]J. Andrada-Félix, et al., "Combining nearest neighbor predictions and model-based predictions of realized variance: Does it pay?," *International Journal of Forecasting*, vol. 32, pp. 695-715, 2016/07/01/ 2016.
- [8]M. Podsiadlo and H. Rybinski, "Financial time series forecasting using rough sets with time-weighted rule voting," *Expert Systems with Applications*, vol. 66, pp. 219-233, 2016/12/30/ 2016.
- [9]R. Adhikari, "A neural network based linear ensemble framework for time series forecasting," *Neurocomputing*, vol. 157, pp. 231-242, 2015.
- [10]S. Cang and H. N. Yu, "A combination selection algorithm on forecasting," *European Journal of Operational Research*, vol. 234, pp. 127-139, Apr 2014.
- [11]J. M. Bates and C. W. J. Granger, "The Combination of Forecasts," *OR*, vol. 20, pp. 451-468, 1969.
- [12]J. H. Stock and M. W. Watson, "Combination forecasts of output growth in a seven-country data set," *Journal of Forecasting*, vol. 23, pp. 405-430, 2004.
- [13]R. Adhikari, "A mutual association based nonlinear ensemble mechanism for time series forecasting," *Applied Intelligence*, pp. 1-18, 2015/01/30 2015.
- [14]R. Adhikari and R. K. Agrawal, "Performance evaluation of weights selection schemes for linear combination of multiple forecasts," *Artificial Intelligence Review*, vol. 42, pp. 529-548, 2014/12/01 2014.
- [15]V. R. R. Jose and R. L. Winkler, "Simple robust averages of forecasts: Some empirical results," *International Journal of Forecasting*, vol. 24, pp. 163-169, 2008.
- [16]J. G. De Gooijer and R. J. Hyndman, "25 years of time series forecasting," *International Journal of Forecasting*, vol. 22, pp. 443-473, 2006.
- [17]S. Makridakis and M. Hibon, "The M3-Competition: results, conclusions and implications," *International Journal of Forecasting*, vol. 16, pp. 451-476, 2000.
- [18]V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 17, pp. 113-126, 2004/01/01/ 2004.
- [19]J. A. K. Suykens, et al., "Weighted least squares support vector machines: robustness and sparse approximation," *Neurocomputing*, vol. 48, pp. 85-105, 2002/10/01/ 2002.
- [20]V. Vapnik, *The nature of statistical learning theory*: Springer science & business media, 2013.
- [21]B. Schölkopf, et al., "New support vector algorithms," *Neural computation*, vol. 12, pp. 1207-1245, 2000.
- [22]N. I. Sapankevych and R. Sankar, "Time series prediction using support vector machines: a survey," *IEEE Computational Intelligence Magazine*, vol. 4, 2009.
- [23]F. E. H. Tay and L. Cao, "Application of support vector machines in financial time series forecasting," *Omega*, vol. 29, pp. 309-317, 2001/08/01/ 2001.
- [24]K.-j. Kim, "Financial time series forecasting using support vector machines," *Neurocomputing*, vol. 55, pp. 307-319, 2003/09/01/ 2003.
- [25]S.-C. Huang, et al., "Chaos-based support vector regressions for exchange rate forecasting," *Expert Systems with Applications*, vol. 37, pp. 8590-8598, 2010/12/01/ 2010.
- [26]M. Y. Kiang, "A comparative assessment of classification methods," *Decision Support Systems*, vol. 35, pp. 441-454, 2003.
- [27]A. Timmermann, "Chapter 4 Forecast Combinations," in *Handbook of Economic Forecasting*. vol. Volume 1, C. W. J. G. G. Elliott and A. Timmermann, Eds., ed: Elsevier, 2006, pp. 135-196.
- [28]F. C. Palm and A. Zellner, "To combine or not to combine? Issues of combining forecasts," *Journal of Forecasting*, vol. 11, pp. 687-701, 1992.
- [29]R. T. Clemen, "Combining forecasts: A review and annotated bibliography," *International Journal of Forecasting*, vol. 5, pp. 559-583, 1989.
- [30]M. Aiolfi and A. Timmermann, "Persistence in forecasting performance and conditional combination strategies," *Journal of Econometrics*, vol. 135, pp. 31-53, 2006.
- [31]C. Hsiao and S. K. Wan, "Is there an optimal forecast combination?," *Journal of Econometrics*, vol. 178, pp. 294-309, 2014.
- [32]A. Chalimourda, et al., "Experimentally optimal v in support vector regression for different noise models and parameter settings," *Neural Networks*, vol. 17, pp. 127-141, 2004/01/01/ 2004.