

# Artificial Intelligence and Signal Decomposition Approach Applied to Retail Sales Forecasting

Ramon Gomes da Silva\*, Matheus Henrique Dal Molin Ribeiro<sup>†</sup>, José Henrique Kleinübing Larcher<sup>‡</sup>,  
Viviana Cocco Mariani<sup>‡§</sup>, and Leandro dos Santos Coelho<sup>\*§</sup>

\*Industrial and Systems Engineering Graduate Program (PPGEPS), Pontifical Catholic University of Parana (PUCPR),  
Curitiba, Parana, Brazil. 80215-901

<sup>†</sup>Department of Mathematics, Federal University of Technology - Parana (UTFPR),  
Pato Branco, Parana, Brazil. 85503-390

<sup>‡</sup>Mechanical Engineering Graduate Program (PPGEM), Pontifical Catholic University of Parana (PUCPR),  
Curitiba, Parana, Brazil. 80215-901

<sup>§</sup>Department of Electrical Engineering, Federal University of Parana (UFPR),  
Curitiba, Parana, Brazil. 81530-000

Emails: gomes.ramon@pucpr.edu.br, mribeiro@utfpr.edu.br, jose.kleinubing@pucpr.edu.br,  
viviana.mariani@pucpr.br, leandro.coelho@pucpr.br

**Abstract**—Sales forecasting is essential for decision-making and are crucial in many areas of a firm, such as planning and scheduling, resource management, marketing, logistics, and supply chain. Due to the fluctuations in retail sales, prediction with high accuracy is a challenging task. In this context, this study proposes a framework that combines ensemble empirical mode decomposition (EEMD) based on artificial intelligence models to forecast the retail sales of a Rossmann Store, using a multi-step-ahead forecasting strategy, in the task of time series forecasting with one, seven, and fourteen-days-ahead. The forecasting models of the retail sales time series are Bayesian Regularization of Artificial Neural Networks, Cubist Regression, and Support Vector Regression. The performance of the proposed forecasting models were evaluated by using two performance metrics: mean absolute percentage error and root mean squared percentage error. The EEMD models outperform the single models in all forecasting horizons, with a performance improvement that ranges 1.30% - 76.25%. Indeed, EEMD models are efficient and accurate models for retail sales forecasting.

**Index Terms**—retail sales, time series, artificial intelligence, ensemble empirical mode decomposition, artificial neural networks, cubist regression, support vector regression

## I. INTRODUCTION

Sales forecasting is essential for decision-making, once these forecasts are important inputs to managerial decisions, such as pricing, store space allocation, ordering, inventory management for an item, and others [1]. Further, sales predictions are crucial for firms in many areas, such as planning and scheduling, resource management, marketing, logistics, and supply chain [2].

Despite its relevance, sales forecasting is a complex subject because the sales' success of a product is highly dependent on the personal taste of consumers, which varies greatly [3]. Moreover, the behavior of retail sales fluctuates widely,

making it difficult to make predictions as accurate as possible. So much so that companies create competitions to solve their prediction problems [4], such as Rossmann [5], Walmart [6], [7], Corporación Favorita [8], and others.

Due to these factors, forecasting sales is a challenging task. Also, the importance of sales forecasting for a firm brings attention not only from practitioners but academics as well. Artificial intelligence approaches can be handy tools to solve this problem. Loureiro et al. [3] proposed a deep learning framework that combines a deep neural network with decision trees, random forest (RF), support vector regression (SVR), and artificial neural network (ANN) to forecast the retail sales of a company that operates in the fashion market. Wang and Yun [9] explored the demand planning and sales forecasting for computer component manufacturers by performing autoregressive integrated moving average (ARIMA), vector autoregression, ensemble learning algorithms, such as RF and gradient boosting, and Lotka-Volterra model. Sohrabpour and colleagues [2] performed Genetic Programming, which is an evolutionary computation approach, to forecast the sales of an export company. Also, Güven and Şimşir [10] proposed a forecasting framework that used ANN and SVR models to predict the sales of retail apparel industry. Many others have proposed forecasting frameworks to solve sales problem using diverse approaches, such as [11]–[13]

Hence, the objective of this study is to propose a forecasting framework combining Ensemble Empirical Mode Decomposition (EEMD) with artificial intelligence models. The model will be applied to train the intrinsic mode functions (IMFs) generated by the decomposition step aiming to forecast the retail sales of a Rossmann store in a multi-step-ahead forecasting strategy (one, seven, and fourteen days ahead). The retail sales historical series will be decomposed into five different IMFs and one residual by EEMD. The training process will be performed by Bayesian Regularization of Artificial Neural Networks (BRNN), Cubist Regression (CUBIST), and SVR

The authors would like to thank the National Council of Scientific and Technologic Development of Brazil – CNPq (Grants number: 307958/2019-1-PQ, 307966/2019-4-PQ, 404659/2016-0-Univ), PRONEX 'Fundação Araucária' 042/2018, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 for financial support of this work.

with Radial Basis Function kernel, respectively, applied into each component. The components predictions are summed giving three different predictions, one for each model, named as EEMD–BRNN, EEMD–CUBIST, and EEMD–SVR, respectively.

The main contributions of this study can be summarized as follows: (i) the first contribution is related to evaluate the use of signal decomposition method for retail sales forecasting; (ii) as the second contribution, we can highlight the use of exogenous input signal (number of customers) coupled with past (delayed) retail sales observations (input lags), to provide additional information to the forecaster model; and (iii) this study evaluates the proposed framework forecasting in a multi-step-ahead forecasting strategy (one, seven, and fourteen days ahead).

The remainder of this paper is structured as follows. Section II illustrates the dataset adopted in this study. Section III defines the models used in this study. Section IV details the procedures of the proposed model framework. Section V presents the results obtained and discussions. Finally, Section VI concludes with the final considerations and some proposals of future works.

## II. DATASET DESCRIPTION

The dataset refers to the Rossmann Store Sales dataset available on Kaggle [5]. For this study, the Rossmann Store #415 was randomly chosen, and the time horizon was from January 1<sup>st</sup> to July 31<sup>st</sup> 2015. Besides, the days which the store was closed, whether for holidays or Sundays, were suppressed from the analysis due to no sales. Also, the only numerical input (number of customers) was taken into account. The dataset was divided into training and test in the proportion of 70% and 30%, respectively.

Fig. 1 illustrates the retail sales and the number of customers according to the day of the week. It is possible to observe that sales are high on Mondays, descend through the week, and tend to ascend after Wednesdays.

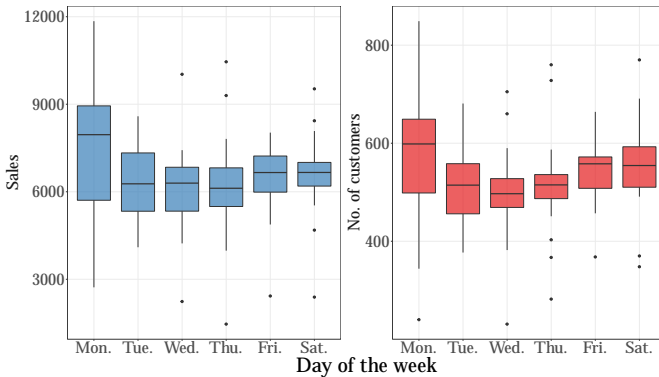


Fig. 1. Sales and number of the customers according to the days of the week

Moreover, the Tab. I presents the summary of the statistical indicators of the output (retail sales) and the input (number of customers) of the Rossmann Store #415.

TABLE I  
SUMMARY OF THE STATISTICAL INDICATORS OF THE OUTPUT AND INPUT OF THE ROSSMANN STORE #415

Variable	Samples	Minimum	Mean	Maximum	Standard deviation
Retail Sales	Whole	1464	6,520.22	11,847	1,494.13
	Training	3981	6,675.03	11,847	1,364.18
	Test	1464	6,140.76	10,228	1,729.61
Number of Customers	Whole	231	533.54	849	89.99
	Training	344	543.36	849	86.21
	Test	231	509.47	694	95.27

Further, Fig. 2 presents the historical daily retail sales (output) and the number of customers (input) of the Rossmann Store #415.

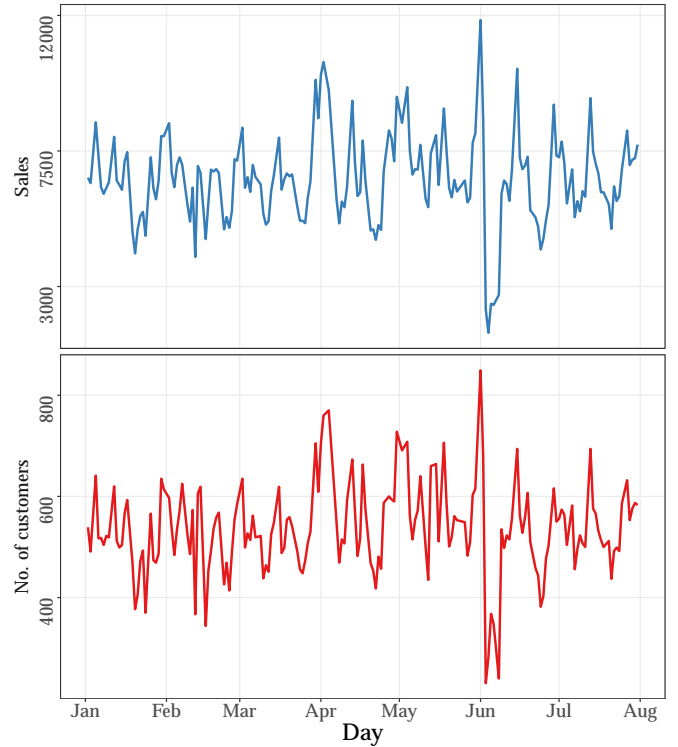


Fig. 2. Sales and customers numbers - historical series of the Rossmann Store #415

## III. METHODS

This section presents the main aspects of the methods proposed in this study. The EEMD method is presented followed by the description of the evaluated forecasting models.

### A. Ensemble empirical mode decomposition

The Empirical Mode Decomposition (EMD) [14] decomposes the original signal into IMFs and one residual component. The main drawback of this decomposition is the mode mixed problem (MMP) which is characterized by the fact that different scales (not homogeneous data scaling) could appear in one IMF. To overcome this disadvantage, EEMD was proposed by Wu and Huang [15], which allows extracting several

features of the data, such as trend, high, and low frequencies [16]. Although EEMD can effectively handle MMP issues, the residue noise in signal reconstruction has been raised, and the noise is independent and identically distributed [17]. In this approach, EMD is performed  $k$  times, and different white noise (a random signal that follows a normal distribution with zero mean and constant variance) is added to the data in each trial. Two disadvantages can be stated for EEMD, as follows: (i) extra noise exists in the reconstructed signal and (ii) it needs more computational resources than EMD [18]. For more details, [15] and [16] can be consulted.

### B. Bayesian regularized neural networks

BRNN is a kind of feedforward artificial neural network, a two-layer neural network, composed of one input and one hidden layer, which uses the Bayesian methods, such as empirical Bayes, for parameter estimation, to avoid overfitting [19]. In the BRNN formulation, the variances are regularization parameters, in which the trade-off between goodness-of-fit and smoothing can be controlled. Also, in this approach, the method of Nguyen and Widrow [20] is used to assign initial weights of neural network and the Gauss-Newton training algorithm to perform the optimization.

### C. Cubist regression

CUBIST is a rule-based algorithm used to build forecasting models (in the time series field) based on the analysis of input data [21]. It estimates the target values by establishing regression models with one or more rules (committee/ensemble of rules) based on the input set. These rules are employed based on a combination of conditions with a linear function (in general linear regression). When the rule satisfies all conditions defined in the learning process, this approach can execute multiple rules once and find different linear functions suitable to forecast retail sales. However, if the standard deviation reduction value is smaller or equal to the expected error for sub-tree, some leaves are pruned to avoid overfitting [22].

### D. Support vector regression

SVR is a type support vector machine that consists of determining support vectors close to a hyperplane, which maximizes the margin between two-point classes obtained from the difference between the target value and a threshold. To deal with non-linear problems SVR takes into account kernel functions, which calculates the similarity between two observations through the inner product. In this study, the radial basis function kernel is adopted. The main advantages of the use of SVR lie in its capacity to capture the predictor non-linearity and then use it to improve the forecasting cases [23], [24].

## IV. PROPOSED MODEL FRAMEWORK

This section describes the main steps in the data analysis adopted by BRNN, CUBIST, SVR, and EEMD based forecasting models.

**Step 1:** First, the dataset output variable is decomposed into five IMFs and Residual by performing EEMD, as presented in Fig. 3. The lag equals 5 was chosen by grid-search, applied on IMFs creating five inputs from the lags, and applied on the exogenous inputs too. Further, the new data is split into training and test sets. The test set consists of the last 51 observations, corresponding to 30% of the dataset, and the training set defined by the remaining samples. In the training state, a 5-fold cross-validation was adopted, such as developed by [25] and [26].

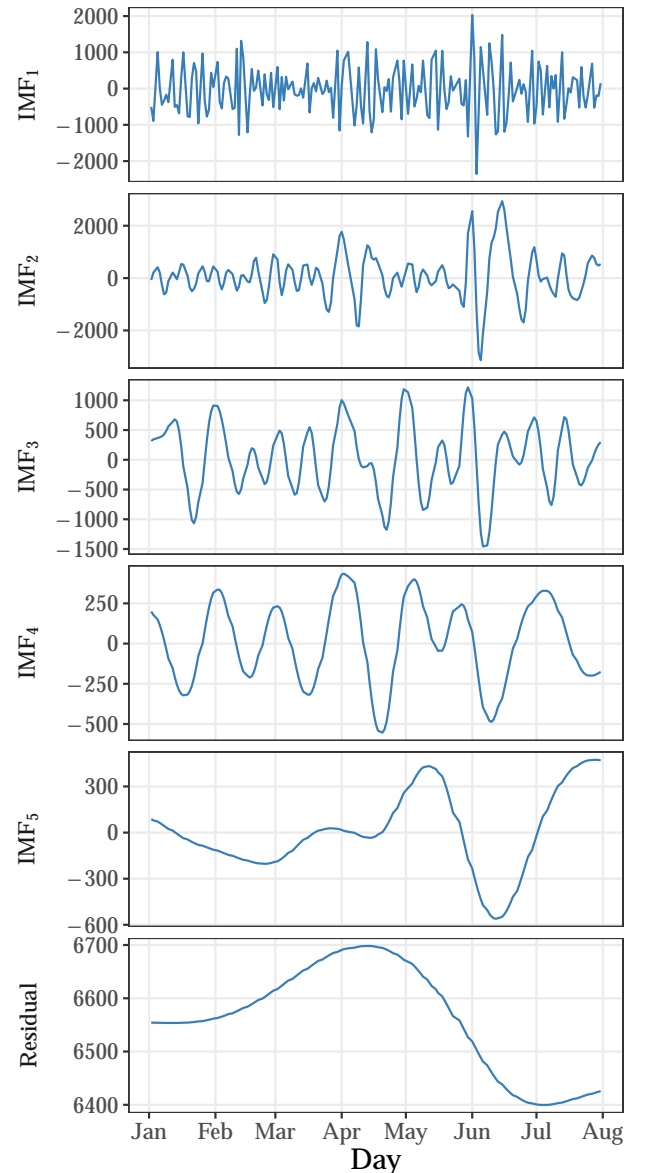


Fig. 3. Dataset decomposed into IMFs and Residual by EEMD

**Step 2:** Each IMF is trained with each model described in Section III using a 5-fold cross-validation approach. Next, the IMF predictions were reconstructed by a simple summation-grouping model, in other words, the IMF is trained by the same model and is summed. Then, three prediction outputs

were generated named EEMD–BRNN, EEMD–CUBIST, and EEMD–SVR.

**Step 3:** A recursive strategy is employed to develop multi-step-ahead retail sales forecasting [27], [28]. Regarding this, one model is fitted for one-step-ahead forecasting, then the recursive strategy uses this forecasting result as an input for the same model to forecast the next step, continuing until the desirable forecasting horizon. In this study, the aim is to obtain the cases up to  $\hat{H}$  next steps, especially up to 1 (1 day ahead), 7 (7 days ahead), and 14-steps-ahead (14 days ahead), respectively. The following structures are considered,

$$\hat{y}_{(t+h)} = \begin{cases} \hat{f} \{y_{(t+h-1)}, \dots, y_{(t+h-5)}, x_{(t+h-1)}\} & \text{if } h = 1, \\ \hat{f} \{\hat{y}_{(t+h-1)}, \dots, \hat{y}_{(t+h-5)}, x_{(t+h-7)}\} & \text{if } h = 7, \\ \hat{f} \{\hat{y}_{(t+h-1)}, \dots, \hat{y}_{(t+h-5)}, x_{(t+h-14)}\} & \text{if } h = 14, \end{cases} \quad (1)$$

where  $\hat{f}$  is a function that maps the retail sales,  $\hat{y}_{(t+h)}$  is the forecast of retail sales in horizon  $h = 1, 7,$  and  $14,$   $y_{(t+h-1)}$  to  $y_{(t+h-5)}$  are the previous observed,  $\hat{y}_{(t+h-1)}$  to  $\hat{y}_{(t+h-5)}$  are the predicted retail sales,  $x_{(t+h-n_x)}$  is the number of customers at the maximum lag of inputs ( $n_x = 1$  if  $h = 1,$   $n_x = 7$  if  $h = 7,$  and  $n_x = 14$  if  $h = 14$ ). The analyses are developed using `{caret}` package [29] in R software [30]. Moreover, Tab. II presents the hyperparameters of the forecasting models for each component and non-decomposed models. A grid-search defined the best tunes for the models.

TABLE II  
CONTROL HYPERPARAMETERS FOR THE FORECASTING MODELS

Component	BRNN		CUBIST		SVR	
	No. of neurons	No. of committees	No. of instances	Cost	Sigma	Kernel
IMF <sub>1</sub>	3	1	9	0.1533	1	
IMF <sub>2</sub>	1	1	0	0.3457	1	
IMF <sub>3</sub>	1	1	0	0.1875	1	
IMF <sub>4</sub>	3	10	0	0.2387	1	Radial
IMF <sub>5</sub>	2	20	9	0.7939	1	
Residual	2	1	5	0.6423	1	
Non-decomposed	1	1	0	0.2191	0.5	

**Step 4:** To evaluate the effectiveness of adopted models, from obtained forecasts out-of-sample (test set), improvement percentage index (IP) (2), mean absolute percentage error (MAPE) (3), and root mean squared percentage error (RMSPE) (4), respectively, were performed.

$$IP = \frac{M_c - M_b}{M_c} \times 100, \quad (2)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100, \quad (3)$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{y_i - \hat{y}_i}{y_i} \right)^2}, \quad (4)$$

where  $n$  is the number of observations,  $y_i$  denotes the sales of a single store on a single day, and  $\hat{y}_i$  denotes the corresponding prediction. Also, the  $M_c$  and  $M_b$  represent the performance measure of compared and best models, respectively.

**Step 5:** Last, a hypothesis test named Diebold-Mariano (DM) test [31] was conducted to determine whether time-series forecasts have the same or significantly different accuracy. The DM test will verify whether the forecasting errors are lower concerning each other. A hypothesis test is conducted, where the null hypothesis says that there is no difference between the forecasting errors of the models compared, and the alternative hypothesis says that the forecasting error of the model proposed is lower than compared one. By using the hypothesis defined, the aim is to know if the errors for the proposed model are lower than the compared model. If the null hypothesis is rejected, it is possible to say that there is statistical evidence that

there is a reduction in the errors of the proposed model regarding the compared model at the  $\alpha$  level of significance.

The proposed model framework as described in this section is illustrated in Fig. 4.

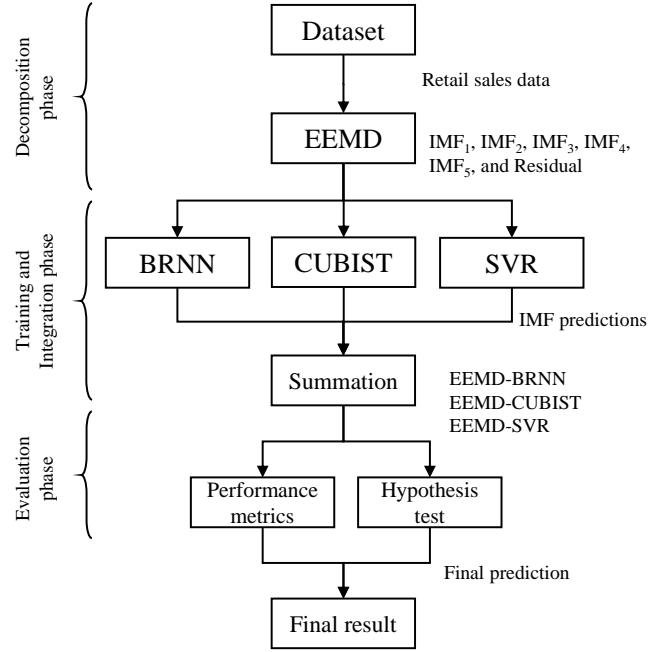


Fig. 4. Framework of the proposed forecasting models

## V. RESULTS

This section describes the main results obtained by the proposed framework forecasting model for retail sales multi-step-ahead forecasting. The performance metrics of the proposed and compared models are presented in Tab. III for all forecasting horizons (one, seven, and fourteen-steps ahead), and all performance criteria. The best performance accuracy is highlighted in bold. The experiment compared the decomposed and non-decomposed models' performance. In all cases, EEMD models outperformed the single models for both performance criteria.

TABLE III  
PERFORMANCE MEASURES RESULTS OF THE MODELS

Models	1-day-ahead		7-days-ahead		14-days-ahead	
	MAPE	RMSPE	MAPE	RMSPE	MAPE	RMSPE
EEMD-BRNN	<b>8.27%</b>	<b>16.77%</b>	17.77%	31.85%	20.66%	33.77%
EEMD-CUBIST	8.95%	20.14%	<b>17.54%</b>	<b>28.89%</b>	19.35%	<b>30.75%</b>
EEMD-SVR	15.08%	63.27%	18.44%	65.37%	<b>18.85%</b>	65.59%
BRNN	16.16%	54.65%	19.10%	75.94%	21.06%	76.19%
CUBIST	15.90%	47.86%	19.38%	77.67%	19.67%	77.67%
SVR	16.43%	70.63%	19.14%	70.23%	21.64%	70.64%

Regarding the one-day-ahead forecasting, the performance metric improvement of the EEMD–BRNN model ranged between 7.61% - 76.25%, where the SVR model presented the worse performance for both criteria, and EEMD–CUBIST model presented the smaller difference from the EEMD–BRNN results. Regarding the seven-days-ahead, the improvement of the EEMD–CUBIST ranged between 1.30% - 62.80%, where the CUBIST model presented the worse performance for both criteria, and EEMD–BRNN model presented the smaller difference from the EEMD–CUBIST results. Last, regarding the fourteen-days-ahead forecasting and MAPE criterion, EEMD–SVR presented a performance improvement that ranges from

2.59% to 12.92%, where EEMD–CUBIST presented the smaller difference from the best model. Regarding the fourteen-days-ahead forecasting the RMSPE criterion, EEMD–CUBIST’s improvement ranges between 8.92% - 60.41%, and EEMD–BRNN presented the smaller difference from the EEMD–CUBIST results.

These results are due to the EEMD approach, which deals with the non-linearity and non-stationarity behaviors of the retail sales data [16]. Also, the CUBIST model has been effective in dealing with the non-stationarity behavior of unstable signals, due to the ensemble of rules based on a combination of conditions with a linear function that the algorithm performs. BRNN takes advantage of its trade-off between goodness-of-fit and smoothing. Last, SVR deals with the non-linear behavior of the series due to the radial basis function kernel adopted in this study. The combination of the strengths of these approaches made it possible for the EEMD to outperform the single models.

Moreover, the results of the DM hypothesis test are detailed with the purpose of performing statistical comparisons between errors of the proposed and compared models described in previous sections. Tab. IV details the statistic of the DM test, as well as when the comparisons are statistically significant.

TABLE IV  
STATISTICS OF DM TEST FOR STATISTICAL COMPARISON OF PROPOSED APPROACH VERSUS OTHER MODELS

Model	One-day-ahead	Seven-days-ahead	Fourteen-days-ahead
	(A)	(B)	(B)
(A) EEMD–BRNN	-	-0.4019	-1.0146
(B) EEMD–CUBIST	-0.7518	-	-
(C) EEMD–SVR	-3.3240*	-0.8293	-0.2719
(D) BRNN	-3.8176*	-0.8977	-0.5537
(E) CUBIST	-3.2382*	-0.8866	-0.3867
(F) SVR	-3.2614*	-0.7028	-0.5462

Note: \*1% significance level.

By the results depicted in Tab. IV, it can be stated that for short term horizon (one-day-ahead) the model EEMD–BRNN has error statistically lower than the other models, but achieved errors statistically equal to EEMD–CUBIST. For longer-term horizons (seven and fourteen-days-ahead), EEMD–CUBIST model presented lower statistical errors than other models, however, its errors are statistically equal to all other models. This is because wider forecasting time windows tend to accumulate errors when forecasting too many steps.

Therefore, based on the RMSPE performance criterion and DM test analysis, EEMD–BRNN model is the most accurate model for forecasting 1-day-ahead, and EEMD–CUBIST is the most suitable model for forecasting 7 and 14-days-ahead. Fig. 5 presents the observed time-series (blue line) versus the predictions given by the proposed model (red line), for forecasting horizon of 1, 7, and 14-days-ahead, respectively.

Besides, is possible to notice that for all forecasting horizons the EEMD–BRNN and EEMD–CUBIST models could learn accurately the pattern and behavior of the data in both training and test sets, allowing to predict accurate values compatible with the observed time-series.

Also, it is important to highlight that, even though the proposed model presented difficulties to follow the extremes of the data variability, the results indicate that the model is capable of predicting accurate values for retail sales in different forecasting horizons.

Analyzing the errors’ standard deviation of the proposed and compared model, the Fig. 6 were developed. The models are labeled in the Fig. 6 as: (A) EEMD–BRNN, (B) EEMD–CUBIST, (C) EEMD–SVR, (D) BRNN, (E) CUBIST, and (F) SVR, same as presented in Tab. IV.

According to the radar plots of errors’ standard deviation, the EEMD proposed models presented the higher stability and lower errors in all forecasting horizons. This means that the proposed fore-

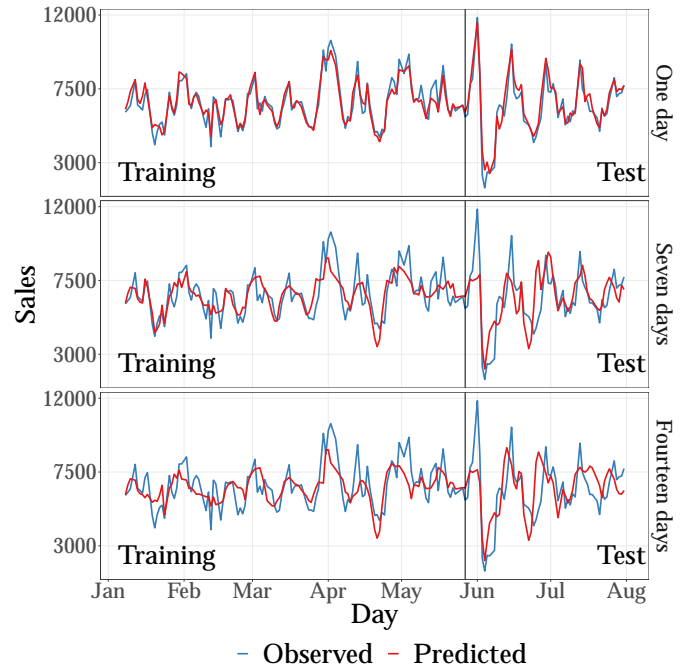


Fig. 5. Observed versus predicted values in all forecasting horizons

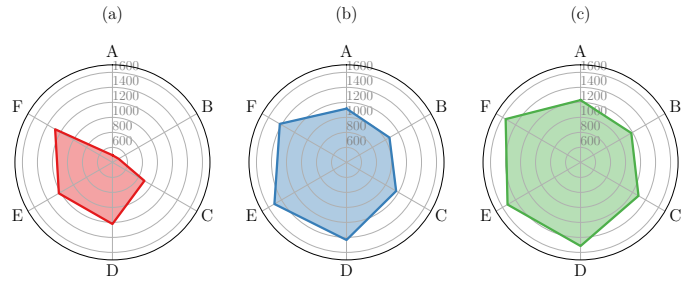


Fig. 6. Errors’ standard deviation for each model for (a) one-day-ahead, (b) seven-days-ahead, and (c) fourteen-days-ahead

casting framework learned better than any other developed forecasting model in this study.

## VI. CONCLUSION

In this study, artificial intelligence approaches named BRNN, CUBIST, and SVR, as well as the EEMD approach, were employed in the task of forecasting one, seven, and fourteen days ahead of the retail sales of the Rossmann Store #415. EEMD approach decomposed the retail sales data into six components (five IMFs and one Residual). Each component was trained and fitted with three well-known artificial intelligence models. The predictions of the components were grouped by trained models and summed to generate three models. The IP, MAPE, and RMSPE criteria were adopted to evaluate the performance of the compared approaches.

The results show that the EEMD models can accurately forecast daily retail sales. The proposed EEMD models outperformed single model approaches in terms of MAPE and RMSPE performance criteria. Even in large forecasting windows, i.e., forecasting 14 days-ahead, the maximum error signal reached by EEMD–SVR was 18.85% in MAPE criterion, showing that the approach is promising for forecasting daily retail sales multi-step ahead.

For future works, it is intended to adopt (i) stacking ensemble learning approach, (ii) different signal decomposition approaches, and

(iii) multi-objective optimization approach to tune hyperparameters of forecasting models.

## REFERENCES

- [1] S. Ma and R. Fildes, "Retail sales forecasting with meta-learning," *European Journal of Operational Research*, vol. 288, no. 1, pp. 111–128, 2021.
- [2] V. Sohrabpour, P. Oghazi, R. Toorajipour, and A. Nazarpour, "Export sales forecasting using artificial intelligence," *Technological Forecasting and Social Change*, no. 120480, 2020.
- [3] A. L. Loureiro, V. L. Miguéis, and L. F. da Silva, "Exploring the use of deep neural networks for sales forecasting in fashion retail," *Decision Support Systems*, vol. 114, pp. 81–93, 2018.
- [4] C. S. Bojer and J. P. Meldgaard, "Kaggle forecasting competitions: An overlooked learning opportunity," *International Journal of Forecasting*, 2020, in Press.
- [5] Kaggle, "Rossmann Store Sales: Forecast sales using store, promotion, and competitor data," 2015. [Online]. Available: <https://www.kaggle.com/c/rossmann-store-sales>
- [6] —, "Walmart Recruiting - Store Sales Forecasting," 2014. [Online]. Available: <https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting>
- [7] —, "Walmart Recruiting II: Sales in Stormy Weather," 2015. [Online]. Available: <https://www.kaggle.com/c/walmart-recruiting-sales-in-stormy-weather>
- [8] —, "Corporación Favorita Grocery Sales Forecasting," 2018. [Online]. Available: <https://www.kaggle.com/c/favorita-grocery-sales-forecasting>
- [9] C.-H. Wang and Y. Yun, "Demand planning and sales forecasting for motherboard manufacturers considering dynamic interactions of computer products," *Computers & Industrial Engineering*, vol. 149, no. 106788, 2020.
- [10] İ. Güven and F. Şimşir, "Demand forecasting with color parameter in retail apparel industry using artificial neural networks (ann) and support vector machines (svm) methods," *Computers & Industrial Engineering*, vol. 147, no. 106678, 2020.
- [11] C.-J. Lu and L.-J. Kao, "A clustering-based sales forecasting scheme by using extreme learning machine and ensembling linkage methods with applications to computer server," *Engineering Applications of Artificial Intelligence*, vol. 55, pp. 231–238, 2016.
- [12] M. Abolghasemi, J. Hurley, A. Eshragh, and B. Fahimnia, "Demand forecasting in the presence of systematic events: Cases in capturing sales promotions," *International Journal of Production Economics*, vol. 230, no. 107892, 2020.
- [13] J. Huber and H. Stuckenschmidt, "Daily retail demand forecasting using machine learning with emphasis on calendric special days," *International Journal of Forecasting*, vol. 36, no. 4, pp. 1420–1438, 2020.
- [14] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [15] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.
- [16] M. H. D. M. Ribeiro, V. C. Mariani, and L. S. Coelho, "Multi-step ahead meningitis case forecasting based on decomposition and multi-objective optimization methods," *Journal of Biomedical Informatics*, vol. 111, no. 103575, 2020.
- [17] M. H. D. M. Ribeiro, R. G. da Silva, C. Canton, N. Fraccanabbia, V. C. Mariani, and L. S. Coelho, "Electricity energy price forecasting based on hybrid multi-stage heterogeneous ensemble: Brazilian commercial and residential cases," in *International Joint Conference on Neural Networks (IJCNN)*, Glasgow, United Kingdom, July 2020, pp. 1–8.
- [18] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, "A review and discussion of decomposition-based hybrid models for wind energy forecasting applications," *Applied Energy*, vol. 235, pp. 939–953, 2019.
- [19] D. J. C. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, no. 3, pp. 415–447, 1992.
- [20] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *IJCNN International Joint Conference on Neural Networks*, vol. 3, San Diego, USA, Jun 1990, pp. 21–26.
- [21] J. R. Quinlan, "Learning with continuous classes," in *5th Australian Joint Conference on Artificial Intelligence*, vol. 92. Hobart, Tasmania: World Scientific, 1992, pp. 343–348.
- [22] M. H. D. M. Ribeiro, R. G. da Silva, V. C. Mariani, and L. S. Coelho, "Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil," *Chaos, Solitons & Fractals*, vol. 135, no. 109853, Jun 2020.
- [23] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems 9*, M. C. Mozer, M. I. Jordan, and T. Petsche, Eds. MIT Press, 1997, pp. 155–161.
- [24] R. G. da Silva, M. H. D. M. Ribeiro, V. C. Mariani, and L. S. Coelho, "Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables," *Chaos, Solitons & Fractals*, vol. 139, no. 110027, oct 2020.
- [25] M. H. D. M. Ribeiro and L. S. Coelho, "Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series," *Applied Soft Computing*, vol. 86, no. 105837, 2020.
- [26] R. G. da Silva, M. H. D. M. Ribeiro, S. R. Moreno, V. C. Mariani, and L. S. Coelho, "A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting," *Energy*, vol. 216, no. 119174, feb 2021.
- [27] S. R. Moreno, R. G. da Silva, V. C. Mariani, and L. S. Coelho, "Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network," *Energy Conversion and Management*, vol. 213, no. 112869, Jun 2020.
- [28] R. G. da Silva, M. H. D. M. Ribeiro, N. Fraccanabbia, V. C. Mariani, and L. S. Coelho, "Multi-step ahead bitcoin price forecasting based on VMD and ensemble learning methods," in *International Joint Conference on Neural Networks (IJCNN)*. Glasgow, United Kingdom: IEEE, jul 2020, pp. 1–8.
- [29] M. Kuhn, *caret: Classification and Regression Training*, 2020, r package version 6.0-86. [Online]. Available: <https://CRAN.R-project.org/package=caret>
- [30] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2020. [Online]. Available: <https://www.R-project.org/>
- [31] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *Journal of Business & Economic Statistics*, vol. 20, no. 1, pp. 134–144, 2002.