ESANN 2022 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 5-7 October 2022, i6doc.com publ., ISBN 978287587084-1. Available from http://www.i6doc.com/en/.

Wind power forecasting based on bagging extreme learning machine ensemble model

Matheus Henrique Dal Molin Ribeiro[1], Sinvaldo Rodrigues Moreno[2], Ramon Gomes da Silva[3], José Henrique Kleinubing Larcher[2], Cristiane Canton[4], Viviana Cocco Mariani[2,5] and Leandro dos Santos Coelho[3,5] *

1 - Federal University of Technology - Parana (UTFPR). Industrial and Systems Engineering Graduate Program (PPGEPS) - Pato Branco, PR, Brazil.

2 - Pontifical Catholic University of Parana (PUCPR). Mechanical Engineering Graduate Program (PPGEM) - Curitiba, PR, Brazil.

3 - Pontifical Catholic University of Parana (PUCPR). Industrial and Systems Engineering Graduate Program (PPGEPS). - Curitiba, PR, Brazil.

4 - University Center of Pato Branco (UNIDEP) Pato Branco, PR, Brazil.

5 - Federal University of Parana (UFPR). Department of Electrical Engineering. - Curitiba, PR, Brazil.

Abstract. The wind energy forecast is an useful tool for wind farm production planning, and operation, facilitating decision making in terms of maintenance, electricity market clearing, and load sharing. This study proposes a cooperative ensemble learning model, using time series preprocessing, multi-objective optimization, and artificial intelligence to forecast wind energy generation in two wind farms in Brazil. Multi-objective optimization is employed to combine variational mode decomposition-based components of a model with bootstrap aggregation (bagging) and extreme learning machine models. Forecasting accuracy is evaluated through the root mean squared error, mean absolute error, mean absolute percentage error, and Diebold-Mariano hypothesis test. The empirical results suggest that proposed ensemble learning model achieved better forecasting performance than bootstrap stacking, machine learning, artificial neural networks, and statistical models, with values of approximately 12.76%, 25.25%, 31.91%, and 34.76%, respectively, in terms of root mean squared errors reduction for out-of-sample forecasting.

1 Introduction

Renewable energies are featured as non-polluting and clean energy resources. Despite the power demand reduction observed over the last two years, due to strategies to mitigate the new coronavirus pandemic, the generation from renewable energy recorded its largest increase [1]. As a consequence of that, those cities with wind and solar farms had a positive impact on their Human Development

*The authors thank the National Council of Scientific and Technologic Development of Brazil – CNPq (Grants: 307958/2019-1-PQ, 307966/2019-4-PQ, 404659/2016-0-Univ, 405101/2016-3-Univ, 408164/2021-2-Univ), PRONEX '*Fundação Araucária*' 042/2018, and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* - Brasil (CAPES) - Finance Code 001 for support.

ESANN 2022 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 5-7 October 2022, i6doc.com publ., ISBN 978287587084-1. Available from http://www.i6doc.com/en/.

Index and Gross Domestic Product [2]. Wind power generation is influenced by climatic and demographic [3] changes, as well as by the wind farms layout [4]. To ensure reliable forecasting results, the combination of individual models provides improvements in forecasting performance by averaging model errors, thus reducing the uncertainty [5]. Cooperative ensemble learning methods can be applied to the forecasting of time series. It leverages the use of artificial intelligence (AI) models, decomposition, optimization, and feature selection. They work in the divide-to-conquer scheme, where a set of base models are adopted to solve the same task. [6].

The proposed ensemble learning model is a cooperative ensemble forecasting model, which integrates different learning strategies to achieve wind power generation forecasting. It proposes a framework by employing variational mode decomposition (VMD) and bagging ensemble learning with moving block bootstrapping (MBB) for resampling the VMD residual [7]. The extreme learning machine (ELM) handles the VMD components and bootstrap samples. This choice is based on recent studies where the effectiveness in solving different problems is highlighted [8, 9]. The median operator is adopted as proposed in Meira et al [10]. The multi-objective optimization was explored to tune the weights of the proposed ensemble learning model. The results of the proposed model are compared with those presented in Ribeiro et al. [11] for 10, 30, 60, and 120-minutes-ahead predictions for wind power generation considering real-time series from two wind farms located in Bahia State, Brazil. The forecasting accuracy of the proposed approach and for the compared models were evaluated applying the mean absolute error (MAE), RMSE, mean absolute percentage error (MAPE), and Diebold-Mariano (DM) hypothesis test.

The main contributions can be summarized as follows: The first is related to evaluating the use of a cooperative ensemble learning model integrating time series decomposition, bagging, and artificial neural networks (ANNs) for wind power generation forecasting. Second, comparisons with several types of models such as statistical, and ANNs are performed. Last, this study evaluates the proposed framework forecasting in a multi-step ahead forecasting strategy, considering very-short (10 and 30-minutes-ahead) and short-term (60 and 120-minutes ahead) horizons, applied to wind time series in the renewable energy context.

2 Proposed cooperative ensemble learning model for forecasting

This section presents all steps adopted to develop the proposed cooperative ensemble learning model.

Step 1: Firstly, the VMD pre-processing is used to obtain four components and one residual. Each one of these components represent different features of the data.

Step 2: Next, considering that the VMD residuals have different sources of uncertainty, this component is re-sampled through bootstrap strategy proposed

ESANN 2022 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 5-7 October 2022, i6doc.com publ., ISBN 978287587084-1. Available from http://www.i6doc.com/en/.

by Bergmeir et al. [7]. In this study 30 samples were generated taking into account the results achieved by Ribeiro et al. [11]. In that opportunity, the authors observed no significant forecasting accuracy improvement by increasing the samples, and the set of samples assessed were 30, 50, or 100 samples, all of them generated through moving block bootstrap (MBB). This process is conducted using the `bld.mbb.bootstrap` function from the `forecast` package [12] in R software.

Step 3 The s samples obtained in the previous step and for the VMD components are used to train the ELM model. Moreover, time series cross-validation is employed in training stage using the training set, helping to prevent overfitting. The hyperparameters for each forecasting model were obtained during the training process through an automatic grid-search using the `caret` package [13] from R software. The recursive strategy is adopted to conduct multi-step-ahead forecasting using each model, the past lagged values are used as input to forecast one-step-ahead. In a trial-and-error strategy, the five past values are used as inputs. For the desired forecasting window (H = 1, 3, 6, and 12), the forecasts are computed according to Eq. (1), as follows:

$$\hat{y}_{t+h} = \begin{cases} f\left[y_t, y_{t-1}, \ldots, y_{t-n_y+1}\right] & \text{if } h = 1 \\ f\left[\hat{y}_{t+h-1}, \ldots, \hat{y}_{t+1}, y_t, \ldots, y_{t+h-n_y}\right] & \text{if } h \in [2, \ldots, n_y] \\ f\left[\hat{y}_{t+h-1}, \ldots, \hat{y}_{t+h-n_y}\right] & \text{if } h \in [n_y+1, \ldots, H], \end{cases} \tag{1}$$

where $\hat{y}_{t+h}$ is the forecast value at time t and the forecast horizon up to h, y_{t+h-n_y} and $\hat{y}_{t+h-n_y}$ are the previously observed and forecast power generation lags for $n_y = 1, 2, 3, 4, 5, 7, 9, 11, 15, 24$.

Step 4: After the training for s samples obtained through an MBB using ELM for the VMD residuals, the final forecasts for this component is defined by the median operator. Finally, the forecasting for the proposed ensemble is computed by a weighted sum of the predictions for the modes and residual. [14]. When all weights are defined as equal to one, it lead to the direct integration. Non-Dominated Sorting Genetic Algorithm – version II (NSGA-II) [15] and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) [16] are adopted to apply the ensemble weight tuning by minimizing the MAE and forecasting errors variance simultaneously.

Figure 1 shows the flowchart for the proposed cooperative ensemble learning model. To access the models' performances, MAE, RMSE, and MAPE criteria are used. The DM test [17] is also applied to compare the forecasting errors.

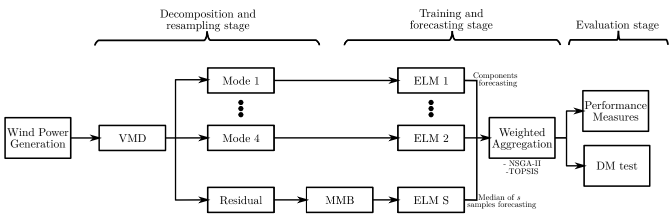


Fig. 1: Flowchart of proposed cooperative ensemble learning model.

ESANN 2022 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 5-7 October 2022, i6doc.com publ., ISBN 978287587084-1. Available from http://www.i6doc.com/en/.

3 Results

The proposed model was applied to forecast wind power production for two different wind farms, considering 10 minutes up to 120 minutes ahead. The ground-truth for comparisons with predicted values, both are presented in Figures 2-a and 2-b. The proposed model is compared with stacking ensemble learning and competitive bagging stacking ensemble learning model, support vector regression (SVR), k-nearest neighbors (KNN), random forest (RF), and eXtreme gradient boosting (XGBoost). Next, comparisons are performed with statistical models named autoregressive integrated moving average (ARIMA), a model specified by seasonal components based on a Fourier and Box-Cox transforms, autoregressive moving average errors, trend, and seasonal component (TBATS), theta model, and the persistence named Naïve. Finally, some ANNs gated recurrent unit (GRU), recurrent neural network (RNN), LSTM, CNNs, ELM, and artificial neuro-fuzzy inference system (ANFIS) are also used for comparison.

Comparing the 10 minutes ahead forecasting results with the ground-truth for both wind farms, can be inferred that the values related to the peaks and valleys were not well assimilated by the model, although the tendency and seasonality were predicted with excellent accuracy as presented in Table 1. Analysing the results for 30 and 60 minutes-ahead, it can be noticed that the tendency and seasonality remain aligned (predicted and real values), and the peak and valleys values became wiggly with amplitude higher than the observed values. This behavior is related to the recursive forecasting strategy adopted, leading to propagating the errors from the past predicted values to the next. On the other hand, results for 120-minutes-ahead, show an increase in amplitude and reduction of the oscillatory pattern. Even though, the seasonality and tendency could be predicted with excellent accuracy.

Through the DM test, it can be stated that in all comparisons, the proposed model reached statistically lower errors than the other models. The AI models based approaches reached high differences between the errors to a significance of 1%, and the cooperative model bagging stacking based approach reached similar errors regarding the proposed framework. These findings are valid for all wind farms and forecasting horizons.

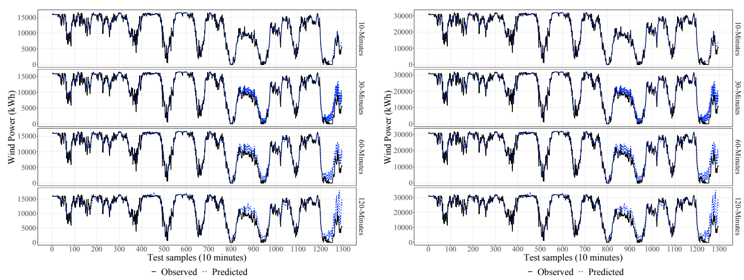
Table 1: Out of sample results of proposed model versus other models.

WDF	Forecasting Horizon	Criteria	Proposed	Artificial Intelligence [11]						Statistical Models [11]				Artificial Neural Networks [11]					
				MBB. Stacking	Stacking	SVR	KNN	RF	XGBoost	ARIMA	TBATS	THETA	NAIVE	GRU	LSTM	RNN	CNNs	ELM	ANFIS
1	10-Minutes	RMSE	**912.02**	1098.64	1200.15	1191.70	1262.64	1239.33	1320.83	1198.56	1192.26	1196.07	1241.66	1437.94	1160.86	1485.66	1178.78	1206.84	1264.44
		MAE	**663.08**	826.52	861.61	842.86	903.00	903.92	954.12	837.07	839.83	838.16	853.51	1090.65	844.50	1077.81	836.14	861.36	893.51
		MAPE	**13.22%**	13.89%	14.75%	14.97%	16.02%	15.67%	18.34%	34.78%	41.42%	35.29%	31.42%	22.53%	14.03%	21.81%	18.92%	14.8%	42.54%
	30-Minutes	RMSE	**1205.31**	1254.00	1452.63	1441.12	1517.22	1499.55	1584.20	1662.62	1642.53	1655.33	1872.08	1795.24	1546.00	1989.43	1581.33	1459.24	1521.89
		MAE	**816.97**	938.33	1053.04	1030.43	1076.50	1086.04	1138.98	1188.51	1187.00	1178.43	1323.84	1385.49	1133.43	1464.27	1162.84	1059.56	1075.25
		MAPE	**13.99%**	15.10%	16.41%	16.39%	18.16%	17.63%	20.16%	50.20%	69.53%	51.32%	58.61%	24.76%	18.23%	25.51%	19.15%	16.65%	20.56%
	60-Minutes	RMSE	**1373.39**	1559.15	1806.98	1787.23	1882.32	1860.82	1927.94	2222.61	2186.13	2225.37	2407.73	2151.59	2077.69	2458.85	2029.35	1831.20	1877.14
		MAE	**950.88**	1156.46	1325.05	1293.97	1345.91	1371.64	1399.59	1601.64	1599.04	1597.64	1730.28	1664.95	1491.97	1817.70	1505.87	1351.44	1358.54
		MAPE	**14.99%**	17.13%	19.04%	18.91%	22.27%	21.82%	23.60%	71.82%	117.14%	74.52%	81.31%	27.97%	27.63%	28.85%	22.94%	20.14%	22.76%
	120-Minutes	RMSE	**1756.43**	2154.69	2422.32	2430.56	2499.99	2520.79	2490.30	3277.92	3146.21	3288.62	3260.79	2881.30	2984.38	3392.19	2733.52	2462.14	2521.37
		MAE	**1143.90**	1554.80	1745.51	1733.35	1744.84	1801.24	1731.29	2336.35	2278.82	2339.45	2310.57	2253.81	2173.40	2453.57	2044.01	1784.64	1781.71
		MAPE	**16.29%**	21.00%	22.04%	22.20%	27.60%	26.90%	34.43%	166.88%	252.41%	169.16%	157.42%	31.98%	34.19%	33.29%	25.39%	24.74%	25.39%
2	10-Minutes	RMSE	**1763.24**	2122.72	2320.20	2303.97	2441.07	2392.08	2571.47	2316.55	2305.50	2312.44	2400.53	2817.19	2493.86	2874.02	2279.21	2333.23	2445.54
		MAE	**1281.95**	1596.52	1665.00	1629.55	1745.76	1743.97	1837.56	1617.87	1624.28	1620.46	1650.12	2094.93	1864.39	2084.89	1615.88	1665.31	1728.68
		MAPE	**13.32%**	13.87%	14.76%	14.97%	16.02%	15.67%	18.31%	34.89%	41.43%	35.29%	31.42%	22.04%	19.97%	21.81%	18.29%	14.76%	42.57%
	30-Minutes	RMSE	**2330.27**	2422.31	2809.51	2786.18	2933.27	2881.23	3075.11	3216.19	3179.78	3200.28	3619.36	3617.00	3341.58	3848.29	3058.10	2821.20	2942.33
		MAE	**1579.47**	1812.34	2035.63	1992.21	2081.24	2081.23	2189.43	2299.06	2298.70	2278.26	2559.43	2729.07	2523.47	2832.04	2245.01	2048.48	2078.81
		MAPE	**14.05%**	15.09%	16.41%	16.39%	18.16%	17.54%	19.98%	50.21%	73.20%	51.32%	58.61%	25.25%	23.56%	25.51%	19.09%	16.65%	16.05%
	60-Minutes	RMSE	**2655.23**	3013.41	3492.37	3455.34	3639.20	3575.85	3757.41	4297.59	4227.69	4302.29	4654.95	4348.95	4262.30	4756.14	3928.28	3540.32	3629.15
		MAE	**1838.36**	2228.96	2560.36	2501.74	2602.14	2632.79	2721.93	3097.76	3096.52	3088.72	3345.21	3291.36	3241.74	3515.35	2901.65	2612.79	2626.52
		MAPE	**15.09%**	17.09%	19.03%	18.91%	22.27%	21.65%	24.21%	72.23%	118.23%	74.52%	81.31%	27.98%	27.26%	28.85%	22.97%	20.14%	22.75%
	120-Minutes	RMSE	**3395.76**	4152.88	4685.49	4699.10	4833.33	4762.98	5060.18	6338.75	6086.57	6357.92	6304.20	5767.75	6204.91	6560.75	5292.01	4760.14	4874.65
		MAE	**2211.54**	3005.33	3375.37	3351.22	3373.45	3438.96	3565.92	4518.28	4417.86	4522.89	4467.10	4466.13	4687.94	4744.70	3949.40	3450.30	3444.64
		MAPE	**16.59%**	20.84%	22.05%	22.20%	27.60%	26.68%	36.07%	167.47%	252.95%	169.15%	157.42%	32.42%	33.68%	33.29%	25.34%	24.74%	27.53%

ESANN 2022 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 5-7 October 2022, i6doc.com publ., ISBN 978287587084-1.
Available from http://www.i6doc.com/en/.

Considering all forecasting horizons and all performance metrics presented in Table 1, when comparing the proposed model in terms of MAPE with AI, statistical, and ANNs, errors reduction spans 16.63%–37.50%, 62.64%–90.96%, and 31.12%–43.85%, respectively. In all scenarios, the second-best model (smaller MAPE compared to the proposed model) was the MBB.stacking model, with an average MAPE of 16.75%. In contrast, the worse model (higher average MAPE) was the TBATS.

Figures 2-a and 2-b illustrate that the proposed ensemble learning model that integrates VMD, bagging, ELM, and MOO in the same framework learns the data behavior, being able to obtain forecast prices similar to observed values.



a) **b)**

Fig. 2: Observed versus predicted values for a) WDF1 and b) WDF2 in test set.

4 Conclusion

The proposed ensemble learning model was able to reduce the forecasting errors in a average of 26.71% of the comparisons. Indeed, the results demonstrate that proposed ensemble improves the forecasting accuracy by 12.76% in terms of the RMSE reduction in relation the bootstrap stacking ensemble learning model. These results showed that the use of time series decomposition, and in this study, the VMD allow to enhance the forecasting accuracy by dealing with several features of the data through of VMD components. In addition, compared with single stacking and ELM model, the reduction in the forecasting errors spans 23.13%-23.87% for compared forecasting horizons and wind farms. Compared with statistical models and ANNs, the proposed forecasting model improves the forecasting accuracy in terms of the RMSE by 34.76% and 31.91%, respectively. The results of this study offer insight into the field of wind energy forecasting could be used in the preventive maintenance of wind turbines, electricity market clearing, and reload sharing. For futer works it is intended to integrate the use of optimization techniques such as Cheetah Based Optimization Algorithm [18] to improve the hyperparameter tuning of the proposed model.

References

[1] BP. Statistical review of world energy, 2022. (accessed 19 January 2022).

ESANN 2022 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 5-7 October 2022, i6doc.com publ., ISBN 978287587084-1. Available from http://www.i6doc.com/en/.

[2] Brazilian Wind Energy Association (ABEEólica). Impactos Socieconômicos e ambientais da geração de energia eólica no Brasil.L, July 2020. (accessed 19 January 2022, in Portuguese).

[3] Ramon Gomes da Silva, Matheus Henrique Dal Molin Ribeiro, Sinvaldo Rodrigues Moreno, Viviana Cocco Mariani, and Leandro Santos Coelho. A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. *Energy*, 216(119174), 2021.

[4] Souma Chowdhury, Jie Zhang, Achille Messac, and Luciano Castillo. Unrestricted wind farm layout optimization (UWFLO): Investigating key factors influencing the maximum power generation. *Renewable Energy*, 38(1):16–30, 2012.

[5] J. M. Bates and C. W. J. Granger. The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468, 1969.

[6] João Mendes-Moreira, Carlos Soares, Alípio Mário Jorge, and Jorge Freire De Sousa. Ensemble approaches for regression: A survey. *ACM Computing Survey*, 45(1):10:1–10:40, December 2012.

[7] Christoph Bergmeir, Rob J. Hyndman, and José M. Benítez. Bagging exponential smoothing methods using STL decomposition and Box–Cox transformation. *International Journal of Forecasting*, 32(2):303–312, 2016.

[8] S.M. Sulaiman, P. Aruna Jeyanthy, D. Devaraj, and K.V. Shihabudheen. A novel hybrid short-term electricity forecasting technique for residential loads using empirical mode decomposition and extreme learning machines. *Computers & Electrical Engineering*, 98(107663), 2022.

[9] Jujie Wang, Quan Cui, and Maolin He. Hybrid intelligent framework for carbon price prediction using improved variational mode decomposition and optimal extreme learning machine. *Chaos, Solitons & Fractals*, 156(111783), 2022.

[10] Erick Meira, Fernando Luiz Cyrino Oliveira, and Lilian de Menezes. Forecasting natural gas consumption using bagging and modified regularization techniques. *Energy Economics*, 106(105760), 2022.

[11] Matheus Henrique Dal Molin Ribeiro, Ramon Gomes da Silva, Sinvaldo Rodrigues Moreno, Viviana Cocco Mariani, and Leandro Santos Coelho. Efficient bootstrap stacking ensemble learning model applied to wind power generation forecasting. *International Journal of Electrical Power & Energy Systems*, 136(107712), 2022.

[12] Rob Hyndman, George Athanasopoulos, Christoph Bergmeir, Gabriel Caceres, Leanne Chhay, Mitchell O'Hara-Wild, Fotios Petropoulos, Slava Razbash, Earo Wang, and Farah Yasmeen. *forecast: Forecasting functions for time series and linear models*, 2021. R package version 8.15.

[13] Max Kuhn. Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5):1–26, 2008.

[14] Jingjing Song, Jianzhou Wang, and Haiyan Lu. A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting. *Applied Energy*, 215:643–658, 2018.

[15] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multio-bjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[16] Ching-Lai Hwang, Young-Jou Lai, and Ting-Yun Liu. A new approach for multiple objective decision making. *Computaters & Operations Research*, 20(8):889–899, 1993.

[17] Francis X Diebold and Roberto S Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263, 1995.

[18] Carlos Eduardo Klein, Viviana Cocco Mariani, and Leandro dos Santos Coelho. Cheetah based optimization algorithm: A novel swarm intelligence paradigm. In *ESANN*, pages 685–690. Bruges, Belgium, 2018.