

SIGNAL DECOMPOSITION AND STACKING-ENSEMBLE LEARNING APPROACHES APPLIED TO TIME SERIES FORECASTING

Ramon Gomes da Silva

Industrial and Systems Engineering Graduate Program (PPGEPS), Pontifical Catholic University
of Parana (PUCPR)

Level: Doctorate

Defense date: June 2023

ramongs1406@gmail.com

Leandro dos Santos Coelho

Electrical Engineering Graduate Program (PPGEE), Federal University of Parana (UFPR)

Advisor

leandro.coelho@ufpr.br

Viviana Cocco Mariani

Mechanical Engineering Graduate Program (PGMec), Federal University of Parana (UFPR)

Co-Advisor

viviana.mariani@ufpr.br

1. Introduction

Time series forecasting is an essential approach for businesses and researchers to make informed decisions by predicting future trends and patterns in a given time series data. Moreover, it plays a critical role in diverse fields, such as business planning and strategy, resource optimization, risk management, financial planning, marketing and sales, among others. Time series forecasting plays a significant role in business planning, as it helps predict future trends and patterns in sales and other key business metrics. This information is critical for businesses to make strategic decisions, such as when to launch new products or services, allocate resources, and manage their supply chain [Ma and Fildes, 2021].

Nevertheless, forecasting time series accurately can be challenging. One reason is the complexity of the time series data, making it difficult to interpret. The data may be noisy, with many different patterns and trends that can make identifying the underlying patterns and relationships challenging. Also, real-world time series generally present high fluctuations and nonlinear and non-stationary behavior [Stepchenko et al., 2017; Tealab et al., 2017]. These characteristics mostly lead to imprecise predictions by the statistical models, which could not follow the data extremes and learn from the changes in the data pattern. These imprecise predictions result in low accuracy and, consequently, high forecasting errors.

To overcome this drawback, machine learning (ML) models are proposed for time series forecasting once they can capture the complex and nonlinear relationships in time series data. These models can identify patterns and relationships that may not be immediately apparent by traditional statistical methods. Further, ML models can learn and adapt to new data over time, improving their accuracy and effectiveness.

Besides, the hybridization of forecasting methods has been widely studied to enhance the accuracy of time series forecasting. This hybrid approach usually involves combinations such as signal decomposition methods [Song et al., 2022], optimization algorithms [Jiang et al., 2023], and ensemble learning approaches [Ribeiro et al., 2023]. In general, hybrid approaches perform

better than single forecasting models, primarily due to the diverse characteristics of each employed approach which help each other in identifying the data pattern and behavior, even in nonlinear and high fluctuation data.

Hence, the main objective of this thesis is to evaluate the effectiveness of employing signal decomposition methods coupled with a stacking-ensemble learning (STACK) approach for forecasting time series in real-world applications.

In essence, this thesis has four significant contributions, which were presented by three applications. The first is evaluating the performance of employing signal decomposition methods in the forecasting learning process. The second contribution explores combining the previous techniques with the stacking-ensemble learning approach to enhance forecasting performance. Next, the third one evolves the first contribution by performing a multi-stage signal decomposition strategy and comparing them with single decomposed models. The last contribution lies in combining the STACK approach to improve the forecasting performance of the multi-stage signal decomposition strategy.

Regarding the three applications, the first application of this thesis regards this contribution. The results of this proposal were published in Chaos, Solitons & Fractals Journal [da Silva et al., 2020]. In that study, the forecasting framework applied variational mode decomposition (VMD) to decompose the coronavirus disease (COVID-19) cumulative cases time series in the Brazilian and American context. A second application was conducted, and its results were published in Energy Journal [da Silva et al., 2021]. In that case, the wind energy generation of a turbine in a wind farm in the Northeast Region of Brazil was considered the forecasting target. The third application was published in the International Journal of Electrical Power & Energy Systems [da Silva et al., 2022], which proposed a hybrid forecasting framework that employed multi-stage decomposition strategy and STACK approach to forecast wind speed. Fig. 1 summarizes the structure of the applications presented in the thesis.

	Application 1	Application 2	Application 3
Case Title	Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables	A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting	Multi-step short-term wind speed forecasting based on multi-stage decomposition coupled with stacking-ensemble learning approach
Time Series	COVID-19 cumulative cases (Epidemiological dataset)	Wind energy generation (Renewable energy dataset)	Wind speed (Renewable energy dataset)
Objective	Proposing a forecasting framework based on signal decomposition methods	Combining signal decomposition methods with stacking-ensemble learning	Proposing a multi-stage signal decomposition strategy combined with stacking-ensemble learning

Figure 1: Applications structure

2. Dataset description

For the first application, the collected dataset refers to the COVID-19 cumulative cases that occurred in five states of Brazil and United States of America (USA) until April 20th, 2020. For the Brazilian context, the dataset was collected from an API [Justen, 2020] that retrieves the daily information about COVID-19 cases from all 27 Brazilian State Health Offices, assembles and makes them publicly available. And for USA context, the dataset was collected from “COVID-19 Data Repository” on Github provided by Johns Hopkins University [Center for Systems Science and Engineering (CSSE), 2020].

Regarding the second application, three collected datasets refer to the wind turbine power generation sampled on 10 minutes basis. All variables were related to a wind turbine in a wind farm in Parazinho city, Brazil. The measurement period comprises August, September, and October 2017 for each dataset, respectively.

Last, for the third application, the dataset consists of three months of wind speed measurements. The measurements are made every 10 min at a height 95 meters from ground level, corresponding to a wind generator with a blade diameter of 100 meters. The measurement apparatus is installed to monitor the wind resource for a wind farm, with approximately 150 MW capacity, located in the State of Rio Grande do Norte, in the Northeastern region of Brazil, more specifically at Parazinho city. The measurements are made from March until May 2020.

3. Proposed forecasting framework

Since the thesis shows an evolution of the hybridization of the models, the application 3 present the ultimate forecasting framework, therefore this framework and its results will be presented as following.

The proposed model is made out of two decomposition stages. First, the decomposition using VMD, and, subsequently, the decomposition using Singular Spectrum Analysis (SSA). VMD extracts the trend components from each original dataset output. The remaining signal is decomposed into four other parts using SSA, so each time series is decomposed into five components by VMD and SSA. The Ridge Regression (RIDGE), k -Nearest Neighbor (KNN), Support Vector Regression (SVR) with linear kernel, and Partial Least Squares (PLS) are evaluated as the layer-0 base learners in the training process. In layer-1, the Cubist Regression (CUBIST) model is trained. The predictions of each component of layer-0 are added (for every individual model) to form each model's forecast. The predictions of the base learners are used as input for the layer-1 meta-learner. The predictions of the meta-learner form this study's proposed model named VMD-SSA-STACK. The models' performances are evaluated by using performance metrics. Furthermore, the Diebold-Mariano (DM) hypothesis test is employed to evaluate the statistical significance of the difference in errors of the models. The framework for the proposed model is illustrated in Fig. 2.

4. Main results

This section presents and compares the results of the three conducted experiments. First, the proposed model is compared with other models with dual decomposition but without the use of STACK. Second, the performance of the proposed model is compared with the performance of models with a single-stage decomposition using VMD and SSA. The last experiment, comparisons of the proposed model with non-decomposed models are presented. Through these experiments, the forecasting accuracy of the proposed model can be effectively assessed and compared.

4.1. Comparison with dual decomposed models

In this first experiment, the proposed model (VMD-SSA-STACK) is compared to the other four models that use dual decomposition, namely VMD-SSA-KNN, VMD-SSA-PLS, VMD-SSA-RIDGE, and VMD-SSA-SVR. The STACK method is not used in these models, just the dual

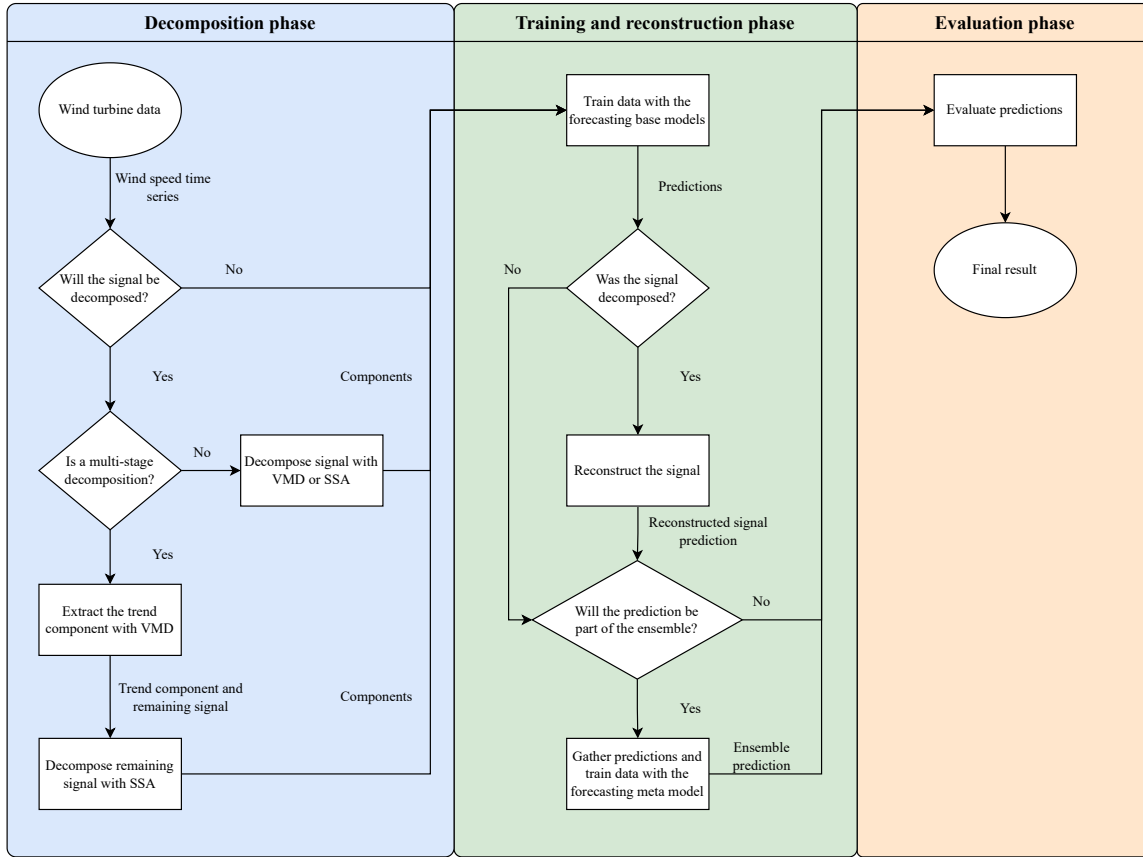


Figure 2: Proposed forecasting framework

decomposition. The results of these models were the ones used as inputs for the proposed VMD–SSA–STACK model. Comparisons are made for all models, in every time horizon, and for all datasets. Furthermore, the model proposed by Moreno [Moreno et al., 2020] (VMD–SSA–LSTM) was also compared.

The proposed forecasting model outperformed the evaluated VMD–SSA models, even though the dual decomposition approach presented low errors (mean absolute percentage error (MAPE) < 8%) and was effectively performed by Moreno in [Moreno et al., 2020] and [Moreno et al., 2021]. This result is due to the divide-and-conquer approach that the STACK method provides. The main gain of the STACK approach is its ability to take advantage of all different learning strategies of every forecasting base model and use it in conjunction to forecast as accurately as possible. Thus, the higher the diversity of the base models, the stronger the STACK becomes, i.e., the variety of the ensemble is more important than the number of models composing the ensemble, as explained by Ribeiro [Ribeiro and Coelho, 2020]. Hence, even though the VMD–SSA–KNN model is not very accurate, its learning characteristics enhance the learning process of the ensemble as a whole.

4.2. Comparison with single decomposed models

The second experiment evaluates the single decomposed models, performed by VMD and SSA approaches. The VMD-based models are named VMD–STACK, VMD–KNN, VMD–PLS, VMD–RIDGE, and VMD–SVR, and the SSA-based models are named SSA–STACK, SSA–KNN, SSA–PLS, SSA–RIDGE, and SSA–SVR.

Despite the consistently accurate performance of VMD and SSA-based models [Moreno and Coelho, 2018; Zhang et al., 2021] and the characteristics of the STACK to potentialize the forecasting, they were not able to overcome the hybridization of the three approaches presented in the VMD–SSA–STACK model. This result is because the VMD technique extracted the trend component from the original time series, and the SSA approach denoised the remaining signal. That enables the training models to deal with the high and low frequencies of the time series. Also, same as the previous experiment, the STACK scheme enhanced the whole forecasting process.

4.3. Comparison with the non-decomposed models

The final experiment compared the proposed approach to models that do not use decomposition. These models were STACK, KNN, PLS, RIDGE, SVR, ANN, ESN, and LSTM.

Although the non-decomposed models usually could not deal with nonlinearity and non-stationary behavior of the wind speed data, these models presented satisfactory results since the MAPE criteria of them were not higher than 13%. Even with promising results related to the non-decomposed models, they could not overcome the VMD–SSA–STACK model. On the other hand, it is essential to emphasize that these models compose the proposed forecasting framework. When these methods are combined, it is possible to build an accurate and efficient forecasting model from the different characteristics and strategies of the approaches.

4.4. Hypothesis tests and graphical analysis

Regarding the DM statistics and p -values of the analysis presented, it is possible to notice that (i) for March 2020 forecasting 10 minutes ahead, the VMD–SSA–STACK is statistically equal to the VMD–SSA–PLS and VMD–SSA–RIDGE models; (ii) for March 2020 forecasting 60 minutes ahead, the VMD–SSA–STACK is statistically equal to VMD–SSA–RIDGE; and (iii) for all other datasets and forecasting horizons, the VMD–SSA–STACK is statistically different from the other models. It is important to emphasize that even for the models where the difference was not statistically significant with a significance of 5%, the proposed model performed better in all performance metrics.

Fig. 3 presented the last 24 hours of observation (last 144 observations) in blue versus the predictions in red for 10, 30, and 60-minutes ahead, for March, April, and May 2020.

Regarding the predicted versus observed values of the wind speed (m/s), the proposed forecasting framework learned the data behavior, which in turn made it able to predict accurate values compatible with the ones from the observed time series. Even in the extremes of the data variability in longer forecasting horizons (i.e., 60 minutes ahead), the proposed model could follow the data behavior and pattern.

5. Publications of this thesis

During the doctorate project, the author made several publications related to the research proposal. The following are presented the main papers which were enumerated according to their publication type, such as peer-reviewed journals publications, papers published in conference proceedings, both national and international events, book chapters, and the awards received.

Peer-Reviewed Journals

1. da Silva, R. G., Ribeiro, M. H. D. M., Mariani, V. C., & Coelho, L. S. (2020). Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. *Chaos, Solitons & Fractals*, 139(110027).
2. Moreno, S. R., da Silva, R. G., Mariani, V. C., & Coelho, L. S. (2020). Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network. *Energy Conversion and Management*, 213(112869).

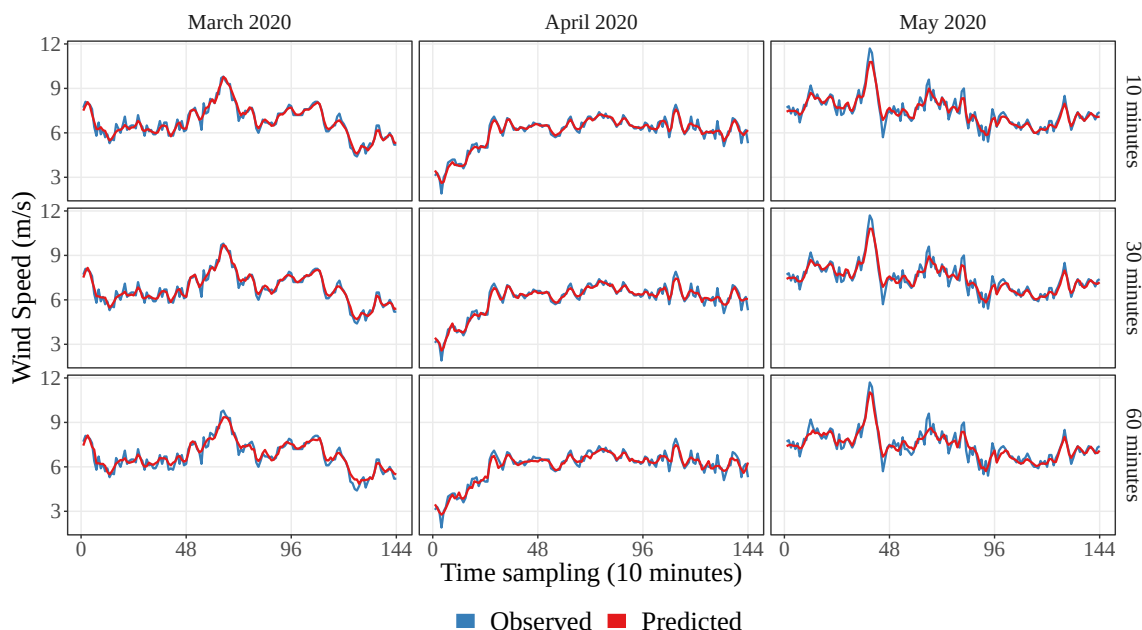


Figure 3: Observed versus predicted time series for the last 24 hours of all datasets in all forecasting horizons

3. Ribeiro, M. H. D. M., da Silva, R. G., Mariani, V. C., & Coelho, L. S. (2020). Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. *Chaos, Solitons & Fractals*, 135(109853).
4. da Silva, R. G., Ribeiro, M. H. D. M., Moreno, S. R., Mariani, V. C., & Coelho, L. S. (2021). A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. *Energy*, 216(119174).
5. da Silva, R. G., Moreno, S. R., Ribeiro, M. H. D. M., Larcher, J. H. K., Mariani, V. C., & Coelho, L. dos S. (2022). Multi-step short-term wind speed forecasting based on multi-stage decomposition coupled with stacking-ensemble learning approach. *International Journal of Electrical Power & Energy Systems*, 143, 108504.
6. Ribeiro, M. H. D. M., da Silva, R. G., Moreno, S. R., Mariani, V. C., & Coelho, L. S. (2022). Efficient bootstrap stacking ensemble learning model applied to wind power generation forecasting. *International Journal of Electrical Power & Energy Systems*, 136(107712).
7. Ribeiro, M. H. D. M., da Silva, R. G., Ribeiro, G. T., Mariani, V. C., & Coelho, L. S. (2023). Cooperative ensemble learning model improves electric short-term load forecasting. *Chaos, Solitons & Fractals*, 166, 112982.

Conference Proceedings

1. Moreno, S. R., da Silva, R. G., Ribeiro, M. H. dal M., Fraccanabbia, N., Mariani, V. C., & Coelho, L. S. (2019). Very short-term wind energy forecasting based on stacking ensemble. *Anais Do 14. Congresso Brasileiro de Inteligência Computacional*, 1-7.

2. da Silva, R. G., Ribeiro, M. H. D. M., Fraccanabbia, N., Mariani, V. C., & Coelho, L. S. (2020). Multi-step ahead Bitcoin price forecasting based on VMD and ensemble learning methods. *International Joint Conference on Neural Networks (IJCNN)*, 1-8.
3. da Silva, R. G., Ribeiro, M. H. D. M., Moreno, S. R., Mariani, V. C., & Coelho, L. S. (2020). Wind energy multi-step ahead forecasting based on variational mode decomposition. *Proceedings of the 18th Brazilian Congress of Thermal Sciences and Engineering*, 1-8.
4. da Silva, R. G., Ribeiro, M. H. D. M., Larcher, J. H. K., Mariani, V. C., & Coelho, L. S. (2020). Forecasting the cumulative cases of COVID-19 in four large Brazilian cities using machine learning approaches. *INnovation for Systems Information and Decision Meeting (INSID)*, 60-67.
5. da Silva, R. G., Ribeiro, M. H. D. M., Larcher, J. H. K., Mariani, V. C., & Coelho, L. S. (2021). Artificial intelligence and signal decomposition approach applied to retail sales forecasting. In C. J. A. B. Filho, H. V. Siqueira, D. D. Ferreira, D. W. Bertol, & R. C. L. de Oliveira (Eds.), *Anais do 15 Congresso Brasileiro de Inteligência Computacional (pp. 1-6)*. SBIC.
6. da Silva, R. G., Moreno, S. R., Ribeiro, M. H. D. M., Mariani, V. C., & Coelho, L. S. (2021). Multi-step wind speed forecasting based on multi-stage decomposition approach. *Proceedings of the 26th International Congress of Mechanical Engineering*.
7. da Silva, R. G., Tafoya-Martinez, V., Silva, F. D., Cardoso, M. A., Severo, E. B., Cardoso, C. Q., Ribeiro, M. H. D. M., Mariani, V. C., & Coelho, L. S. (2023). Beta-hCG Test Demand Forecasting Using Stacking Ensemble-Learning and Machine Learning Approaches. In F. Deschamps, E. Pinheiro De Lima, S. E. Gouvêa Da Costa, & M. G. Trentin (Eds.), *Proceedings of the 11th International Conference on Production Research - Americas* (pp. 274-280). Springer Nature Switzerland.

Book Chapter

1. Ribeiro, M. H. D. M., da Silva, R. G., Larcher, J. H. K., Mariani, V. C., & Coelho, L. S. (2021). Ensemble learning models coupled with urban mobility information applied to predict COVID-19 incidence cases. In A. T. Azar & A. E. Hassanien (Eds.), *Modeling, Control and Drug Development for COVID-19 Outbreak Prevention* (pp. 821-858). Springer International Publishing.

Awards

1. Best Paper Award from the International Conference on Production Research Americas - ICPR 2022 on the paper entitled "Beta-hCG test demand forecasting using stacking ensemble-learning and machine learning approaches".

References

Center for Systems Science and Engineering (CSSE) (2020). Novel coronavirus (COVID-19) cases, provided by JHU CSSE. (accessed in 28 April, 2020).

da Silva, R. G., Moreno, S. R., Ribeiro, M. H. D. M., Larcher, J. H. K., Mariani, V. C., and Coelho, L. d. S. (2022). Multi-step short-term wind speed forecasting based on multi-stage decomposition coupled with stacking-ensemble learning approach. *International Journal of Electrical Power & Energy Systems*, 143:108504. ISSN 01420615.

- da Silva, R. G., Ribeiro, M. H. D. M., Mariani, V. C., and Coelho, L. S. (2020). Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. *Chaos, Solitons & Fractals*, 139(110027).
- da Silva, R. G., Ribeiro, M. H. D. M., Moreno, S. R., Mariani, V. C., and Coelho, L. S. (2021). A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. *Energy*, 216(119174).
- Jiang, P., Nie, Y., Wang, J., and Huang, X. (2023). Multivariable short-term electricity price forecasting using artificial intelligence and multi-input multi-output scheme. *Energy Economics*, 117: 106471. ISSN 01409883.
- Justen, Á. (2020). COVID-19: Coronavirus newsletters and cases by municipality per day. (accessed in 28 April, 2020).
- Ma, S. and Fildes, R. (2021). Retail sales forecasting with meta-learning. *European Journal of Operational Research*, 288(1):111–128. ISSN 03772217.
- Moreno, S. R. and Coelho, L. S. (2018). Wind speed forecasting approach based on singular spectrum analysis and adaptive neuro fuzzy inference system. *Renewable Energy*, 126:736–754.
- Moreno, S. R., da Silva, R. G., Mariani, V. C., and Coelho, L. S. (2020). Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network. *Energy Conversion and Management*, 213(112869).
- Moreno, S. R., Mariani, V. C., and Coelho, L. S. (2021). Hybrid multi-stage decomposition with parametric model applied to wind speed forecasting in Brazilian Northeast. *Renew. Energy*, 164: 1508–1526.
- Ribeiro, M. H. D. M. and Coelho, L. S. (2020). Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Applied Soft Computing*, 86: 105837.
- Ribeiro, M. H. D. M., da Silva, R. G., Ribeiro, G. T., Mariani, V. C., and Coelho, L. S. (2023). Co-operative ensemble learning model improves electric short-term load forecasting. *Chaos, Solitons & Fractals*, 166:112982. ISSN 09600779.
- Song, C., Chen, X., Xia, W., Ding, X., and Xu, C. (2022). Application of a novel signal decomposition prediction model in minute sea level prediction. *Ocean Engineering*, 260:111961. ISSN 00298018.
- Stepchenko, A., Chizhov, J., Aleksejeva, L., and Tolujew, J. (2017). Nonlinear, non-stationary and seasonal time series forecasting using different methods coupled with data preprocessing. *Procedia Computer Science*, 104:578–585. ISSN 18770509.
- Tealab, A., Hefny, H., and Badr, A. (2017). Forecasting of nonlinear time series using ANN. *Future Computing and Informatics Journal*, 2(1):39–47. ISSN 23147288.
- Zhang, S., Chen, Y., Xiao, J., Zhang, W., and Feng, R. (2021). Hybrid wind speed forecasting model based on multivariate data secondary decomposition approach and deep learning algorithm with attention mechanism. *Renew. Energy*.