27th ABCM International Congress of Mechanical Engineering
December 04-08, 2023, Florianópolis-SC, Brazil

COB-2023-1027 ECHO STATE NETWORKS APPLIED TO HYDROLOGICAL SERIES FORECASTING

José Henrique Kleinübing Larcher
Luiz Eduardo Thomaz
Mechanical Engineering Graduate Program (PPGEM). Pontifical Catholic University of Paraná, Curitiba, Brasil
jose.kleinubing@pucpr.edu.br;luizethomaz@hotmail.com

Leandro dos Santos Coelho
Industrial and Systems Engineering Graduate Program (PPGEPS). Pontifical Catholic University of Paraná, Curitiba, Brasil
Federal University of Paraná, Department of Electrical Engineering, Curitiba, PR, Brazil
leandro.coelho@pucpr.br

Viviana Cocco Mariani
Mechanical Engineering Graduate Program (PPGEM). Pontifical Catholic University of Paraná, Curitiba, Brasil
Federal University of Paraná, Department of Electrical Engineering, Curitiba, PR, Brazil
viviana.mariani@pucpr.br

Abstract. *Time series analysis and forecasting are essential for effective water resource management, especially in countries like Brazil that heavily rely on hydroelectric power. This study investigates the effectiveness of employing various decomposition methods in conjunction with Echo State Networks (ESNs) to forecast natural water flow in two Brazilian hydroelectric reservoirs: Itaipu and Furnas. The research aims to evaluate ESNs in hydrological series forecasting and enhance accuracy by integrating decomposition methods with ESNs. The tested forecasting horizons are 7, 14 and 21 days ahead. The study utilizes ESNs tuned with the Coyote Optimization Algorithm (COA) and integrates them with the decomposition methods Variational Mode Decomposition (VMD), Empirical Wavelet Transform, and Empirical Mode Decomposition. The performance of these models is compared using metrics mean absolute error (MAE), mean absolute percentage error, root mean square error, and root mean square logarithmic error. The results demonstrate that incorporating decomposition methods into ESNs enhances their accuracy in forecasting hydrological series. VMD consistently outperforms other methods across all forecast horizons, reducing MAE by 43.70% to 88.88% compared to other models. The statistical significance of the improvement achieved by employing VMD is confirmed by the Diebold-Mariano test at a 1% level.*

Keywords: *Time series forecast, signal decomposition, echo state networks, hydroelectric reservoirs.*

1. Introduction

Time series forecasting plays a crucial role in managing water resources, including predicting floods, ebbs, and reservoir levels in hydrological series. Brazil, with its hydroelectric potential, relies heavily on accurate hydrological forecasting, which is challenging due to non-stationarity, seasonality, trend, noise, and non-linear characteristics (Coulibaly, 2010). Machine learning (ML) methods offer advantages over traditional approaches, handling non-linearity and learning from data. Among ML techniques, echo state neural networks (ESNs) have been extensively studied (Tealab, 2018). ESNs are artificial neural networks with randomly connected neurons and fixed weights in a reservoir. They have an output layer trained for time series forecasting and offer benefits like efficient training, internal memory, and handling non-linear series. Signal decomposition techniques, such as empirical mode decomposition (EMD), variational mode decomposition (VMD), and empirical wavelet transform (EWT) (Liu *et al.*, 2021), break down series into components for preprocessing and enhancing forecast accuracy.

This study explores the use of ESNs in combination with VMD, EMD, and EWT for forecasting hydrological time series in two Brazilian hydroelectric reservoirs: Furnas (Minas Gerais) and Itaipu (Paraná). Forecast evaluation employs metrics like root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square logarithmic error (RMSLE) for horizons of 7, 14, and 21 days. The Diebold-Mariano test compares the error signal of the best method with other approaches, determining statistical significance.

The remaining sections of this paper are organized in the following manner. Section 2.introduces the data and techniques used in this study, including ESN and decomposition methods. In Section 3.the methodology for obtaining the results is outlined and the findings of the study are presented. Finally, Section 4.offers conclusions and insights gained from this work, as well as directions for future research.

2. Data Set and Forecasting Models

In this section, the data used in this study are presented, as well as the ESN and the decomposition methods that were employed in this work.

2.1 Data set

This study analyzes natural water flow data from two Brazilian reservoirs, Furnas and Itaipu, obtained from the Brazilian National Water Agency through the Reservoir Monitoring System (SAR) (SAR, 2023). Natural water flow refers to the quantity of water that would naturally flow in any watercourse had the flow not been affected by human interference or human intervention. The use of natural water flow data, unadulterated by human activities, ensures that the predictive models are capable of capturing the true hydrological dynamics of the specific system under investigation. Natural water flow is preferred is this study over other flow data such as inflow and outflow because it is not affected by human intervention and provides a more accurate representation of the natural water cycle.

The research focuses on the Furnas hydroelectric facility in Minas Gerais state and the Itaipu hydroelectric plant on the Brazilian-Paraguayan border. The data covers the period from January 1st 1993 to December 31, 2021, with daily sampling. Summary statistics for the analyzed period are presented in Table 1, and Figure 1 depicts the natural water flow in both reservoirs during that period.

Reservoir	**Measurement Date**		**Natural Flow (m^3/s)**			
	Minimum	Maximum	Minimum	Maximum	Mean	Standard Deviation
Furnas	1993-01-01	2021-12-31	70.45	5016.61	761.297976	624.543194
Itaipu	1993-01-01	2021-12-31	3473.38	42321.80	11393.285259	5090,057085

Table 1: Summary of the data for the hydrological series

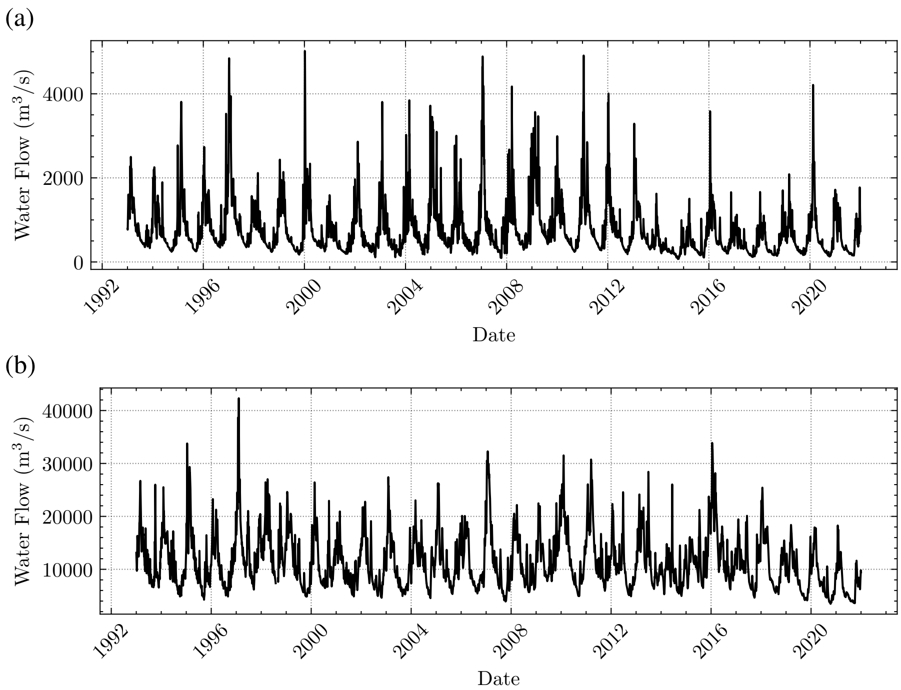

Figure 1: Natural water flow for a) Furnas and b) Itaipu

2.2 Echo State Networks

ESNs were initially introduced by Jaeger (Jaeger, 2001) and gained prominence in 2004 (Jaeger and Haas, 2004). They are reservoir computing ANNs designed for sequence data processing. ESNs possess a unique structure with sparse internal connections and random weights. Input layer weights are also randomly assigned. Unlike traditional ANNs, ESNs have a specific training approach where only the output layer is trained, while the internal layers remain fixed. Training is accomplished using mathematical techniques such as Moore-Penrose pseudoinverse or Tikhonov regularized regression (Tikhonov *et al.*, 1995). This structure simplifies the optimization process and expedites training. Refer to

Fig. 2 for an ESN diagram.

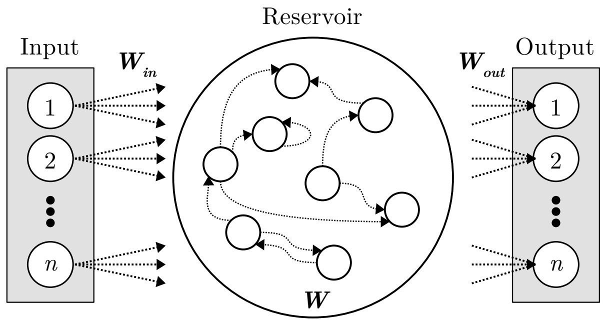

Figure 2: Diagram showing the structure of an ESN.

The reservoir is governed by some key parameters, such as (Lukoševičius, 2012) its size, the sparsity of the connections, and the spectral radius scaling. The effect of a previous state on the next state is determined by the leaking rate (α) of the ESN. The i-th $\mathbf{x}_i$ state of the reservoir is calculated by

$$\mathbf{x}_i = (1-\alpha)\mathbf{x}_{i-1} + \alpha f\left(\mathbf{W}_{in}\mathbf{u}_i + \mathbf{W}_{res}\mathbf{x}_{i-1}\right), \tag{1}$$

where $\mathbf{u}_i$ is the input in its i-th position, $\mathbf{W}_{in}$ refers to the weights of the input layer , $\mathbf{W}_{res}$ refers to the internal layer weights, and f is an activation function. The output of the reservoir is calculated by

$$\hat{\mathbf{y}}_i = \mathbf{W}_{out}\left(\mathbf{u}_i, \mathbf{x}_i\right) \tag{2}$$

where $\mathbf{y}_i$ is the output of the reservoir in its i-th position, $\mathbf{W}_{out}$ represents the weights of the output layer, and $(\mathbf{u}_i, \mathbf{x}_i)$ is the concatenation of the input and reservoir states.

When using Ridge Regression to train the output layer of the ESN, a regularization parameter (λ) is also needed. The weights (W_{out}) of the output layer, are calculated by

$$\mathbf{W}_{out} = \mathbf{Y}\mathbf{X}^T\left(\mathbf{X}\mathbf{X}^T + \lambda\mathbf{I}\right)^{-1} \tag{3}$$

where the matrix $\mathbf{Y}$ contains the known values, the matrix $\mathbf{X}$ represents the states, and $\mathbf{I}$ is the identity matrix.

2.3 Empirical Mode Decomposition

EMD is a data decomposition technique that breaks down a signal into its intrinsic mode functions (IMFs). It is a non-linear and non-stationary method based on the Hilbert-Huang Transform (HHT) (Huang *et al.*, 1998), an adaptive, data-driven approach to decomposing signals into components that are meaningful in terms of their physical properties.

A sifting process is applied iteratively to the original signal in order to obtain each IMF. This process starts by choosing the signal to be decomposed. Then, the local maximum and minimum points of the signal are computed. After that, the local maxima and minima are connected with a cubic spline. The mean M_i of the two curves is computed in the i-th iteration. The IMF is the difference between the signal in the first step and the mean curve, so that $\text{IMF}_i = S_i - M_i$, where S_i is the signal to be decomposed in the i-th iteration and M_i is the mean curve of the correspondent iteration. The steps 2 to 5 are repeated until i is equal to k, in which k is the desired number of IMFs. On the first iteration S_i is the original signal. In the next iterations, S_i is the residual component of the previous iteration. A residual component from the complete process can be obtained by adding the IMFs and subtracting from the original signal. Furthermore, the steps can also be repeated until the residual component is a constant signal.

2.4 Empirical Wavelet Transform

The Empirical Wavelet Transform (EWT) is similar to EMD in that it is a technique for breaking down time series into various frequency components. The technique was proposed by Gilles (2013). Like the EMD, the EWT aims to divide the signal into Amplitude-Modulated and Frequency-Modulated (AM-FM) parts, however, instead of using the average between the maximum and minimum curves, it uses a filter bank of wavelets determined based on the original time series.

To perform EWT first the local maxima are determined using the normalized signal in the frequency spectrum. Then, the maxima are arranged in descending order and boundaries are determined based on subsequent maxima. The nth segment boundary is denoted by ω_n, and the nth segment is denoted by $\wedge_n$, so that $\wedge_n = [\omega_{n-1}, \omega_n]$. A transition region between the $\wedge_n$ is defined with width $2\tau_n$, with $\tau_n = \gamma\omega_n$, where γ is an adjustment parameter between 0 and 1, smoothing the transition between the segments. The Meyer wavelet is used to define the wavelets in the segments, and the empirical scaling function is used to define the wavelets in the transition regions as described in Gilles (2013), thus building the filter bank. The signal is then filtered by each of the bank's filters, generating a set of AM-FM components. The sum of the individual obtained components is the reconstructed signal. A residual component can be obtained by subtracting the reconstruction from the original one.

2.5 Variational Mode Decomposition

VMD (Variational Mode Decomposition) is a signal decomposition method introduced by Dragomiretskiy and Zosso (2013). It decomposes the original signal into Intrinsic Mode Functions (IMFs), which are sub-signals or modes. The number of components is specified as a parameter for the decomposition. VMD can serve as a pre-processing technique for feature creation or as a noise reduction method by summing the IMFs. In VMD, the IMFs are treated as AM-FM signals, with modulated amplitude ($C_i(t)$) and modulated frequency ($\frac{\mathrm{d}}{\mathrm{d}t}\varphi_i(t)$) (Daubechies *et al.*, 2011). VMD decomposes the original signal ($h(t)$) into k IMFs, such that the sum of the IMFs reconstructs the original signal.

To determine the bandwidth of an IMF (Dragomiretskiy and Zosso, 2013) first the Hilbert transform is calculated for each IMF μ_i, obtaining the analytic signal with a one-sided frequency spectrum. Then, by combining the original signal with an exponential signal that has a frequency equivalent to the center frequency of the IMF, the frequency spectrum of each IMF is shifted to its baseband. Finally, the bandwidth is estimated using a Gaussian smoothing H^1 of the demodulated signal, for example using the gradient norm L^2.

2.6 Performance Evaluation Metrics

The performance evaluation metrics used in this work are the mean absolute error (MAE), the mean absolute percentage error (MAPE), the root mean square error (RMSE), and the root mean square logarithmic error (RMSLE). The MAE is defined as (Klaar *et al.*, 2023):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} \left|\hat{y}_i - y_i\right|, \tag{4}$$

where n is the number of samples, $\hat{y}_i$ is the predicted value, and y_i is the observed value. The MAPE is defined as (da Silva *et al.*, 2022):

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n} \left|\frac{\hat{y}_i - y_i}{y_i}\right|. \tag{5}$$

The RMSE is defined as (Ribeiro *et al.*, 2023):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(\hat{y}_i - y_i\right)^2}. \tag{6}$$

The RMSLE is defined as:

$$\text{RMSLE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(\log\left(\hat{y}_i + 1\right) - \log\left(y_i + 1\right)\right)^2}, \tag{7}$$

where $\log(x)$ corresponds to the natural logarithm of x.

3. Results

In this section, the proposed methodology is presented and the obtained results are shown and discussed. The proposed methodology is applied to the signals of the dataset described in Section 2.1

The proposed methodology was used to produce and compare the forecasts of ESN alone (ESN-COA), ESN with EMD decomposition (ESN-EMD-COA), ESN with EWT decomposition (ESN-EWT-COA), and ESN with VMD decomposition (ESN-VMD-COA). The summary of the proposed methodological approach is shown in Figure 3. Forecasts for each combination are generated for forecasting with horizons of 7, 14, and 21 steps ahead (also days ahead).

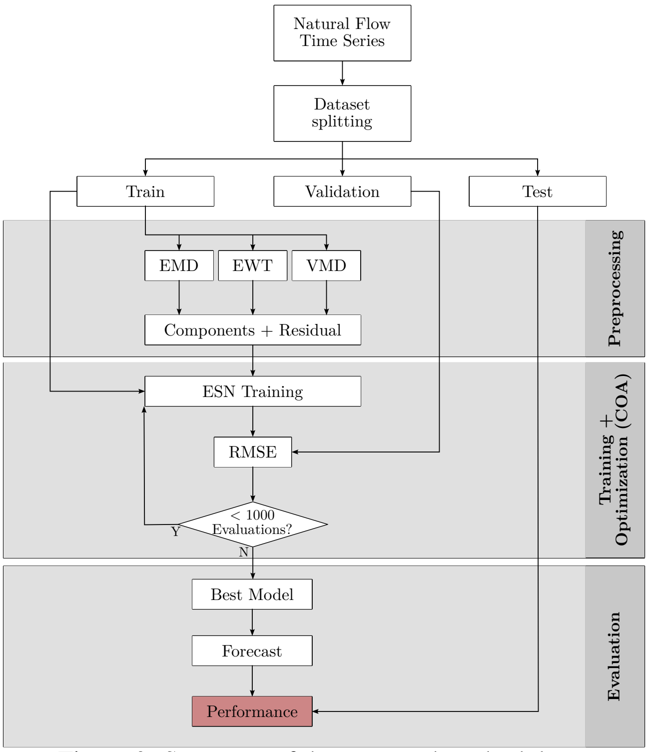

Figure 3: Summary of the proposed methodology.

The hyperparameters of the ESNs were optimized using the Coyote Optimization Algorithm (COA) (Pierezan and Coelho, 2018; Pierezan *et al.*, 2019b; de Souza *et al.*, 2020; Tong *et al.*, 2022; Pierezan *et al.*, 2019a) in all cases, for a total of 1000 iterations, to evaluate the models in the same conditions. The ESN hyperparameters were optimized for each model and for each number of steps ahead. The parameters for the COA optimization were 10 packs with 5 coyotes each. Prior to the start of the experiment, the hold-out method was used separating the data set into 60% for training, 20% for validation, and 20% for testing. The training and validation data sets were employed to train the models and optimize the hyperparameters. The testing data was used for out-of-sample tests, such as evaluating the metrics and performing DM statistical tests.

For the EMD the number of components was not assigned, with the number of components being decided by the EMD method. For the EWT the number of components was 12. For the VMD the number of components was 5. These values were defined as per the number of components that could describe the signal (in the training set) without leaving out much information in the residuals. The input features of the ESN-COA are the last available values of the signal before the step that is meant to forecast, and the output values are the predicted values h steps ahead, making it a Single Input Single Output (SISO) architecture. In the case of ESN-EMD-COA, ESN-EWT-COA, and ESN-VMD-COA the input is the conjunction of the last available values of each component. The output values for the decomposed values are the predicted values h steps ahead for each component, making it a Multi Input Multi Output (MIMO) architecture. To produce the forecasts for the original signal, the decomposed values are summed.

The forecasts are produced in a direct manner, such that the forecast is the direct output of the ESN for the horizon h, without recursion. The implementation was performed in Python 3.7.4 using the library `ReservoirPy` (Trouvain *et al.*, 2020) for the ESNs. To perform the decompositions, the libraries `pyemd`[1] (Pele and Werman, 2008), `ewtpy`[2] and `vmdpy`[3] (Carvalho *et al.*, 2020) were used for the decomposition methods EMD, EWT and VMD, respectively. For the COA optimization the library `COA`[4] (Pierezan and Coelho, 2018) was used.

The data is scaled based on the training dataset using Min-Max scaling. The data is scaled before entering the ESNs, and the predictions are scaled back to the original scale. The scaling is performed for each component of the decomposed signals.

The ESNs were tuned for the following hyperparameters: reservoir size (`units`), leaking rate (`leak_rate`), spectral radius (`spectral_radius`), input scaling (`input_scaling`), density (`density`), input connectivity (`input_connectivity`), and Tikhonov regularization (`regularization`).

[1]Github: `https://github.com/wmayner/pyemd`
[2]Github: `https://github.com/vrcarva/ewtpy`
[3]Github: `https://github.com/vrcarva/vmdpy`
[4]Github: `https://github.com/jkpir/COA`

3.1 Evaluation

Table 2 shows the evaluation metrics RMSE, MAE, MAPE, and RMSLE of the all models for each reservoir and forecast horizon. The results expose that ESN-VMD-COA outperforms the other models in all cases for every metric. ESN-EMD-COA comes in second place in every case, except for forecasts 21 days ahead in the Furnas reservoir, where ESN-EWT-COA outperforms ESN-EMD-COA for every metric. ESN-EWT-COA outperforms ESN-COA in all cases, except for forecasts 7 days ahead in the Itaipu reservoir, where ESN-COA outperforms ESN-EWT-COA for every metric. Comparing with other models, ESN-VMD-COA decreases MAE from 43.70% (compared with ESN-EMD-COA for Furnas 7 days ahead) to 88.88 % (compared with ESN-COA for Itaipu 21 days ahead). Bold values indicate the best performance for each reservoir and forecast horizon in each metric.

Reservoir	Steps ahead	Model	RMSE	MAE	MAPE	RMSLE
Furnas	7	COA	257.57	158.38	0.2856	0.3309
		ESN-COA-EMD	129.98	71.37	0.1172	0.1569
		ESN-COA-EWT	235.04	143.37	0.2586	0.3301
		ESN-COA-VMD	**71.65**	**40.18**	**0.0652**	**0.0848**
	14	COA	333.17	215.16	0.4153	0.4422
		ESN-COA-EMD	179.24	107.93	0.2055	0.2494
		ESN-COA-EWT	231.3	141.72	0.2653	0.3132
		ESN-COA-VMD	**78.32**	**46.69**	**0.0823**	**0.1018**
	21	COA	358.76	246.06	0.4999	1.4406
		ESN-COA-EMD	410.27	366.47	1.0946	0.7669
		ESN-COA-EWT	245.08	157.11	0.3104	0.6928
		ESN-COA-VMD	**93.02**	**56.31**	**0.1014**	**0.1226**
Itaipu	7	COA	1764.4	1349.6	0.1496	0.1734
		ESN-COA-EMD	818.71	592.08	0.0625	0.079
		ESN-COA-EWT	2013.03	1542.22	0.1652	0.2047
		ESN-COA-VMD	**283.66**	**189.61**	**0.0205**	**0.0283**
	14	COA	2647.85	2154.8	0.2554	0.2723
		ESN-COA-EMD	1409.28	1017.71	0.1056	0.1306
		ESN-COA-EWT	1993.59	1520.51	0.1597	0.1962
		ESN-COA-VMD	**373.09**	**255.49**	**0.0275**	**0.0373**
	21	COA	3509.34	2808.12	0.338	0.3465
		ESN-COA-EMD	1791.68	1335.94	0.1436	0.1853
		ESN-COA-EWT	2054.58	1625.89	0.1782	0.2048
		ESN-COA-VMD	**446.32**	**312.38**	**0.0337**	**0.0454**

Table 2: Performance metrics for each model for the reservoirs Furnas and Itaipu.

Figure 4 shows the plot of the results for the best model, ESN-VMD-COA, for the reservoirs Furnas on the left and Itaipu on the right, and every forecast horizon.

As VMD had the best performance in all metrics, it remains to evaluate if the differences with the other forecasts are statistically significant. To accomplish this, a Diebold-Mariano (DM) test (Diebold and Mariano, 2002) was performed. The DM test is a non-parametric statistical analysis that compares two forecasts. The null hypothesis of the test is such that the error signal of the two forecasts is the same. So if the null hypothesis is rejected, the forecasts are statistically different. After comparing ESN-VDM-COA with the other forecasts with the corresponding reservoir and steps ahead, results showed that the ESN-VDM-COA is statistically different from the other forecasts for all reservoirs and steps ahead with a significance level of 1% in all cases.

4. CONCLUDING REMARKS AND DIRECTIONS FOR FUTURE RESEARCH

This study aimed to investigate the impact of employing different decomposition techniques, namely EMD, WDT, and VMD, on the forecasting performance of hydroelectric reservoirs using ESNs. As it was observed, the incorporation of these decomposition methods yields improved forecasting performance for ESNs, suggesting their potential to enhance hydrological forecasting with ESNs. Notably, the combined utilization of VMD, ESN, and COA consistently outperforms alternative approaches across all evaluated metrics and forecasting horizons. This superiority is further supported by the statistically significant forecasting difference observed in the ESN-VMD-COA model compared to other forecasting techniques, with a significance level of 1%. These findings highlight the considerable efficacy of VMD as a robust and powerful tool for hydrological forecasting.

It is noteworthy that the VMD and EWT decompositions, being parametric methods, may be influenced by the choice of parameters, which can potentially impact the final results. To address this, potential solutions include employing grid search or incorporating the parameter selection within the optimization method's search space, serving as avenues for future research. Although ESNs demonstrate promising results independently, further investigations are required to

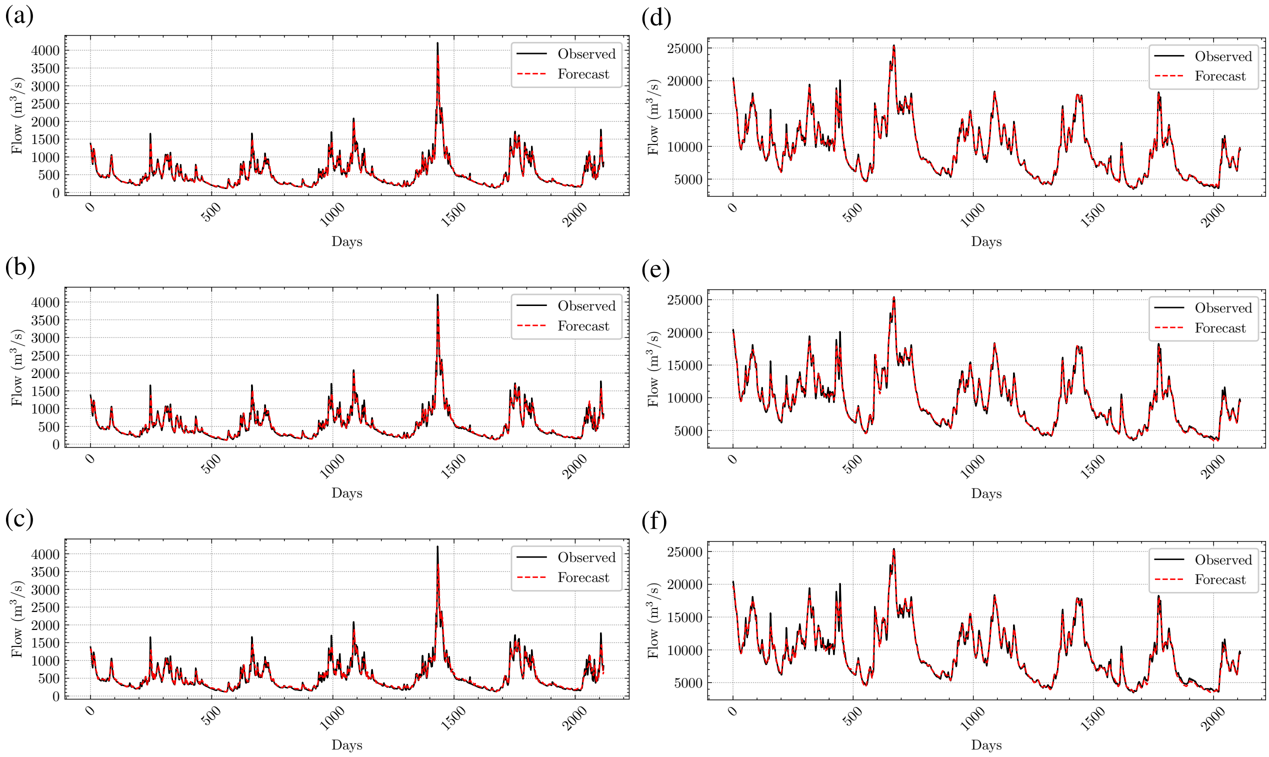

Figure 4: Forecasts using ESN-COA-VMD for the reservoirs Furnas (a) 7, (b) 14 an (c) 21 steps ahead, and Itaipu (d) 7, (e) 14 and (f) 21 steps ahead.

substantiate this claim, involving comparative analyses with alternative forecasting techniques. Moreover, the exploration of other decomposition techniques, such as EMD variants (e.g., extended EMD, complete EMD with adaptive noise) and Singular Spectrum Analysis (SSA), along with the examination of alternative optimization methods like the Falcon Optimization Algorithm (de Vasconcelos Segundo *et al.*, 2019), are also contemplated as potential avenues for future investigations.

5. ACKNOWLEDGMENTS

This research was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (Grants number: 307958/2019-1-PQ, 307966/2019-4-PQ, 404659/2016-0-Univ, 408164/2021-2-Univ), PRONEX 'Fundação Araucária' 042/2018, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES (Finance Code 001).

6. REFERENCES

Carvalho, V.R., Moraes, M.F., Braga, A.P. and Mendes, E.M., 2020. "Evaluating five different adaptive decomposition methods for eeg signal seizure detection and classification". *Biomedical Signal Processing and Control*, Vol. 62, p. 102073. doi:10.1016/j.bspc.2020.102073.

Coulibaly, P., 2010. "Reservoir computing approach to great lakes water level forecasting". *Journal of Hydrology*, Vol. 381, No. 1, pp. 76–88. ISSN 0022-1694. doi:10.1016/j.jhydrol.2009.11.027.

da Silva, R.G., Moreno, S.R., Ribeiro, M.H.D.M., Larcher, J.H.K., Mariani, V.C. and Coelho, L.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*, Vol. 143, p. 108504. ISSN 0142-0615. doi:10.1016/j.ijepes.2022.108504.

Daubechies, I., Lu, J. and Wu, H.T., 2011. "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool". *Applied and Computational Harmonic Analysis*, Vol. 30, No. 2, pp. 243–261. doi:10.1016/j.acha.2010.08.002.

de Souza, R.C.T., de Macedo, C.A., Coelho, L.S., Pierezan, J. and Mariani, V.C., 2020. "Binary coyote optimization algorithm for feature selection". *Pattern Recognition*, Vol. 107, p. 107470. ISSN 0031-3203. doi:10.1016/j.patcog.2020.107470.

de Vasconcelos Segundo, E.H., Mariani, V.C. and Coelho, L.S., 2019. "Design of heat exchangers using falcon optimization algorithm". *Applied Thermal Engineering*, Vol. 156, pp. 119–144. ISSN 1359-4311. doi:10.1016/j.applthermaleng.2019.04.038.

Diebold, F.X. and Mariano, R.S., 2002. “Comparing predictive accuracy”. *Journal of Business & Economic Statistics*, Vol. 20, No. 1, pp. 134–144. doi:10.1198/073500102753410444.

Dragomiretskiy, K. and Zosso, D., 2013. “Variational mode decomposition”. *IEEE Transactions on Signal Processing*, Vol. 62, No. 3, pp. 531–544. doi:10.1109/TSP.2013.2288675.

Gilles, J., 2013. “Empirical wavelet transform”. *IEEE Transactions on Signal Processing*, Vol. 61, No. 16, pp. 3999–4010. doi:10.1109/TSP.2013.2265222.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C. and Liu, H.H., 1998. “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. doi:10.1098/rspa.1998.0193.

Jaeger, H., 2001. “The "echo state" approach to analysing and training recurrent neural networks – with an erratum note”. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, Vol. 148.

Jaeger, H. and Haas, H., 2004. “Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication”. *Science*, Vol. 304, No. 5667, pp. 78–80. ISSN 0036-8075, 1095-9203. doi:10.1126/science.1091277.

Klaar, A.C.R., Stefenon, S.F., Seman, L.O., Mariani, V.C. and Coelho, L.d.S., 2023. “Optimized ewt-seq2seq-lstm with attention mechanism to insulators fault prediction”. *Sensors*, Vol. 23, No. 6. ISSN 1424-8220. doi:10.3390/s23063202. URL `https://www.mdpi.com/1424-8220/23/6/3202`.

Liu, H., Yan, G., Duan, Z. and Chen, C., 2021. “Intelligent modeling strategies for forecasting air quality time series: A review”. *Applied Soft Computing*, Vol. 102, p. 106957. ISSN 1568-4946. doi:10.1016/j.asoc.2020.106957.

Lukoševičius, M., 2012. “A Practical Guide to Applying Echo State Networks”. In *Neural Networks: Tricks of the Trade: Second Edition*, Springer, Lecture Notes in Computer Science, pp. 659–686. ISBN 978-3-642-35289-8. doi:10.1007/978-3-642-35289-8_36.

Pele, O. and Werman, M., 2008. “A linear time histogram metric for improved sift matching”. In *Computer Vision (ECCV)*. Marseille, France, pp. 495–508. doi:10.1007/978-3-540-88690-7_37.

Pierezan, J. and Coelho, L.S., 2018. “Coyote optimization algorithm: A new metaheuristic for global optimization problems”. In *Congress on Evolutionary Computation (CEC)*. Rio de Janeiro, Brazil, pp. 1–8. doi:10.1109/CEC.2018.8477769.

Pierezan, J., Coelho, L.S., Mariani, V.C. and Lebensztajn, L., 2019a. “Multiobjective coyote algorithm applied to electromagnetic optimization”. In *2019 22nd International Conference on the Computation of Electromagnetic Fields (COMPUMAG)*. Paris, France, pp. 1–4. doi:10.1109/COMPUMAG45669.2019.9032768.

Pierezan, J., Maidl, G., Yamao, E.M., Coelho, L.S. and Mariani, V.C., 2019b. “Cultural coyote optimization algorithm applied to a heavy duty gas turbine operation”. *Energy Conversion and Management*, Vol. 199, p. 111932. ISSN 0196-8904. doi:10.1016/j.enconman.2019.111932.

Ribeiro, M.H.D.M., da Silva, R.G., Ribeiro, G.T., Mariani, V.C. and dos Santos Coelho, L., 2023. “Cooperative ensemble learning model improves electric short-term load forecasting”. *Chaos, Solitons & Fractals*, Vol. 166, p. 112982. ISSN 0960-0779. doi:10.1016/j.chaos.2022.112982.

SAR, 2023. “Sistema de acompanhamento de Reservatórios - ana”. `https://www.ana.gov.br/sar0/MedicaoSin`. Accessed: 2023-01-03. In Portuguese.

Tealab, A., 2018. “Time series forecasting using artificial neural networks methodologies: A systematic review”. *Future Computing and Informatics Journal*, Vol. 3, No. 2, pp. 334–340. ISSN 2314-7288. doi:10.1016/j.fcij.2018.10.003.

Tikhonov, A.N., Goncharsky, A.V., Stepanov, V.V., Yagola, A.G., Tikhonov, A.N., Goncharsky, A.V., Stepanov, V.V. and Yagola, A.G., 1995. “Regularization methods”. In *Numerical Methods for the Solution of Ill-Posed Problems*, Springer Netherlands, pp. 7–63. doi:10.1007/978-94-015-8480-7_2.

Tong, H., Zhu, Y., Pierezan, J., Coelho, L.S. and Mariani, V.C., 2022. “Chaotic coyote optimization algorithm”. *Journal of Ambient Intelligence and Humanized Computing*, Vol. 13, pp. 2807–2827. doi:10.1007/s12652-021-03234-5.

Trouvain, N., Pedrelli, L., Dinh, T.T. and Hinaut, X., 2020. “Reservoirpy: An efficient and user-friendly library to design echo state networks”. In *Artificial Neural Networks And Machine Learning (ICANN)*. Bratislava, Slovakia, pp. 494–505. ISBN 978-3-030-61616-8. doi:10.1007/978-3-030-61616-8_40.

7. RESPONSIBILITY NOTICE

The authors are solely responsible for the printed material included in this paper.