ELECTRIMACS 2026 – Palermo, Italy, 18-21 May 2026

Simple vs Complex: Evaluating Deep Learning Architectures for Decomposition-Based Electrical Load Time Series Forecasting

Paolo Fazzini · **Giuseppe La Tona** · **Marco Montuori** · **Matteo Diez** · **Maria Carmela Di Piazza**

Abstract Decomposition methods, such as Variational Mode Decomposition (VMD), convert complex, nonstationary time series into band-limited Intrinsic Mode Functions (IMFs), enabling the isolation of interpretable temporal components and the analysis of trends, seasonality, and oscillatory behavior. These techniques have gained increasing attention in the scientific community and are commonly employed as preprocessing steps in hybrid forecasting frameworks. To evaluate the performance of different categories of forecasting algorithms on band-limited components, this study investigates nine neural architectures, including recurrent, feedforward, basis expansion, and transformer models, applying them to both synthetic IMFs and VMD-decomposed real-world electrical load data. Results show that simple or lightweight models, such as Multi-Layer Perceptrons (MLPs) and Temporal Convolutional Networks (TCNs), often match or outperform more complex transformers, indicating that architectural innovations designed for raw time series do not necessarily improve IMF forecasting.

P. Fazzini · G. La Tona · M.C. Di Piazza
Institute of Marine Engineering (INM)
National Research Council (CNR)
90146 Palermo, Italy
e-mail:
paolo.fazzini@cnr.it,
giuseppe.latona@cnr.it,
mariacarmela.dipiazza@cnr.it

M. Diez
Institute of Marine Engineering (INM)
National Research Council (CNR)
00128 Rome, Italy
e-mail: matteo.diez@cnr.it

M. Montuori
Institute of Complex Systems (ISC)
National Research Council (CNR)
00185 Rome, Italy
e-mail: marco.montuori@cnr.it

1 Introduction

Decomposition-based forecasting leverages algorithms such as Empirical Mode Decomposition (EMD) [7], Variational Mode Decomposition (VMD) [3], and Synchrosqueezing Wavelet Transform (SWT) [2] to transform complex time series into more tractable components. Among these, VMD and SWT stand out as methods with rigorous mathematical convergence guarantees, producing band-limited Intrinsic Mode Functions (IMFs) characterized by slowly varying amplitude and frequency relative to their phase.

The field of neural forecasting has evolved from Multi-Layer Perceptrons (MLPs) and Long Short-Term Memory (LSTM) networks to sophisticated transformer-based models such as Informer [14], FEDformer [15], and PatchTST [8]. Interpretable architectures like N-BEATS [9] and N-HiTS [11] have also demonstrated competitive performance. While these newer architectures excel on complex datasets, their effectiveness on IMF signals remains largely unexplored. This raises a critical question: do the architectural innovations of modern deep learning maintain their advantages when applied to band-limited, quasi-periodic signals? The fundamental preprocessing strategy of decomposition trades complex, nonstationary trends for well-behaved components. If simpler architectures prove sufficient for IMF forecasting, this would suggest that decomposition contributes primarily by interacting favorably with the ANN training phase, thereby improving overall predictive performance.

Despite extensive literature on decomposition-enhanced forecasting across domains, including electric power [6,1,4], renewable energy [13,12], and network throughput [10], a systematic evaluation comparing how different architectural paradigms perform specifically on IMFs is absent.

To fill this gap, in this paper we propose a methodological contribution: by evaluating and comparing nine diverse forecasting algorithms on synthetic IMFs and real-world decomposed signals, we show that architectural complexity does not consistently translate to improved performance in multi-step-ahead forecasting on band-limited components.

2 Methodology

2.1 Evaluated Algorithms

The NeuralForecast framework [9] was employed to assess nine forecasting algorithms spanning four architectural categories (basis expansion models, transformer-based models, feedforward models, and recurrent models), as summarized in Table 1.

Basis expansion models, such as N-BEATS and N-HiTS, decompose the forecast into interpretable components using learnable basis functions across hierarchical temporal scales.

Transformer-based models, including Informer, FEDformer, and PatchTST, leverage attention mechanisms to model long-range dependencies, while incorporating architectural optimizations to improve computational efficiency.

Feedforward models, exemplified by MLP and Temporal Convolutional Networks (TCN), operate over temporal windows using either fully connected layers or dilated causal convolutions.

Finally, recurrent models, including LSTM and DeepAR, process sequences sequentially through gated memory cells, enabling the effective capture of temporal dependencies.

In terms of forecasting strategy, LSTM and DeepAR adopt a recursive approach, where each predicted value is iteratively fed back as input for subsequent steps. All other models utilize direct multi-output forecasting, producing all future values simultaneously in a single forward pass.

Table 1 Overview of evaluated forecasting algorithms

Category	Algorithm	Key Feature
Basis expansion	N-BEATS	Backcast-forecast blocks with learnable basis functions
Basis expansion	N-HiTS	Multi-rate sampling and hierarchical interpolation
Transformer-based	PatchTST	Patch-based transformer with channel independence
Transformer-based	Informer	ProbSparse attention, $\mathcal{O}(L\log L)$ complexity
Transformer-based	FEDformer	Frequency-enhanced attention
Feed-forward	TCN	Dilated causal convolutions
Feed-forward	MLP	Fully connected layers on flattened windows
Recurrent	LSTM	Gated memory cells (forget, input, output gates)
Recurrent	DeepAR	LSTM-based probabilistic autoregression

2.2 Experimental Configuration and Evaluation Metrics

The "Auto" variants of the NeuralForecast framework were used, allowing a fair comparison across methods by allocating equal computational budgets (12,000 gradient steps in total) through controlled hyperparameter optimization. All models were configured with: forecast horizon $h = 48$ time steps, input window size $2h$, robust scaling for normalization, early stopping with patience of 2 steps, and initial learning rate of 10^{-3}. Models were trained using either Normal or Student's t-distribution losses, with MAE used for Informer and FEDformer where distribution-based losses performed poorly. This work evaluates only point forecasts. For models that produce full predictive distributions, we converted those distributions into point forecasts by taking their expected value.

We employed two normalized error metrics for fair comparison across different signal scales, i.e., the Normalized Root Mean Squared Error (NRMSE) and the Normalized Mean Absolute Error (NMAE), defined as follows:

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\sigma_y} \tag{1}$$

$$\mathrm{NMAE} = \frac{\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|}{\sigma_y} \tag{2}$$

where y_i are actual values, $\hat{y}_i$ are predictions, σ_y is the standard deviation of actual values, and n is the number of actual points.

3 Datasets

The selected neural models were evaluated using two distinct datasets.

The first dataset was synthetically generated by constructing signals that are mathematically consistent with the definition of IMFs [2]; these signals are referred to, in the following, as synthetic IMFs.

The second dataset was derived from real-world time series data consisting of measured shipboard load power consumptions, which were decomposed using the VMD technique [4]. In both cases, the same division of data into training and test sets was applied to assess the performance of the forecasting models. Further details of the datasets used are provided in the following subsections.

3.1 Synthetic IMFs

Four artificial IMFs were generated according to the mathematical definition in [2]. A function $f(t) = A(t)e^{i\phi(t)}$ is an Intrinsic Mode Function with accuracy $\varepsilon > 0$ if:

1. $A \in C^1(\mathbb{R}) \cap L_\infty(\mathbb{R})$, $\phi \in C^2(\mathbb{R})$

2. $\inf_{t\in\mathbb{R}} \phi'(t) > 0$, $\sup_{t\in\mathbb{R}} \phi'(t) < \infty$
3. $|A'(t)|, |\phi''(t)| \leq \varepsilon|\phi'(t)|, \forall t \in \mathbb{R}$
4. $\sup_{t\in\mathbb{R}} |\phi''(t)| < \infty$

The key constraint (3) requires that amplitude and instantaneous frequency vary slowly relative to the phase.
Our synthetic IMFs take the real-valued form:

$$\mathrm{IMF}(t) = A(t)\cdot\cos(\phi(t)) \tag{3}$$

where $A(t) = 2+\cos^2(\gamma t/2) + A\cos(17\gamma t/2)$ and $\phi(t) = \Gamma(\sin(t)+bt)/b$. The parameters were varied to produce four distinct configurations (Table 2) all satisfying the IMF constraints. Fig. 1 shows that the rate of change of amplitude and frequency is substantially lower than that of phase across all configurations, confirming that the four synthetic signals employed reflect the typical characteristics of IMFs. The first 10,000 samples of each IMF were used as training set and the next 144 samples were used as test set.

Table 2 Parameters for synthetic IMF configurations

Config	γ	b	Γ	A
#1	0.01	4.7	1.1	1.0
#2	0.015	4.3	0.7	1.0
#3	0.005	5.5	0.9	1.5
#4	0.017	5.1	0.8	1.3

3.2 VMD-Decomposed Real-World Data

We also evaluated the performance of the selected forecasting models on real-world electrical load power consumption data measured on a passenger vessel; these data were decomposed by VMD, as described in [4].

The electrical load power dataset comprises four power consumption variables (P_1–P_4), each with 10,144 observations sampled at ten-minute intervals, spanning approximately 71 days. The first 10,000 samples were used as training set and the remaining day was used as test set.

The data were decomposed into $K = 12$ IMFs, a configuration shown to yield accurate reconstruction in previous studies [4,5]. Among these components, we focused on the lowest-frequency mode because it captures the dominant trend of the electrical load and has the greatest impact on forecasting accuracy. This mode also shows strong spectral localization and closely reflects the band-limited, quasi-harmonic behavior of ideal IMFs.

4 Results

The forecast results for the nine evaluated neural models are summarized in Tables 3 and 4, which correspond to the aggregated metrics for synthetic IMFs and for the VMD-decomposed modes of the measured power consumption data, respectively. Specifically, each neural model is tested on every signal from the two datasets and, based on the results of each experiment, the mean and standard deviation of the metrics are computed.

Across both datasets, AutoMLP consistently achieves the best performance, exhibiting the lowest NRMSE and NMAE values along with moderate variability. AutoNHITS and AutoPatchTST also demonstrate strong performance, ranking immediately after AutoMLP, whereas AutoDeepAR exhibits the poorest results, with substantially higher errors and variability in both settings.

In the synthetic IMF experiments, AutoMLP reaches an average NRMSE of 0.073, which is 6.4% lower than AutoNHITS (0.078) and 29.8% lower than AutoPatchTST (0.104). AutoInformer and AutoNBEATS achieve intermediate errors (0.139–0.143), while AutoDeepAR exhibits the highest NRMSE (0.996), an order of magnitude above the top-performing models.

For the VMD-decomposed real-world data, AutoMLP maintains the lowest average NRMSE (0.419), 19.6% lower than AutoNHITS (0.521) and 28.3% lower than AutoTCN (0.585). AutoLSTM (0.730) outperforms AutoInformer (0.811) and AutoFEDformer (0.928) in terms of both mean error and stability, demonstrating more consistent performance.

Overall, fully connected and temporal convolution-based architectures, despite being simpler and computationally lighter, achieve comparable or more consistent forecasting performance, while attention-based models tend to exhibit higher variability and moderate underperformance, particularly on real-world VMD-decomposed load power consumption data.

Fig. 2 reports the average NRMSE and the corresponding standard deviation obtained for each neural model across all eight experiments, including both synthetic IMFs and real-world data decomposed via VMD. This aggregated analysis clearly demonstrates the superior performance of the AutoMLP model, whereas transformer-based architectures yield higher average errors and exhibit greater performance variability.

Table 3 Summary of forecasting performance metrics on synthetic IMF experiments

Model	NRMSE	NMAE
AutoMLP	0.073 ± 0.041	0.061 ± 0.034
AutoNHITS	0.078 ± 0.035	0.065 ± 0.030
AutoPatchTST	0.104 ± 0.027	0.084 ± 0.017
AutoInformer	0.139 ± 0.045	0.110 ± 0.038
AutoNBEATS	0.143 ± 0.062	0.119 ± 0.056
AutoTCN	0.209 ± 0.169	0.175 ± 0.150
AutoLSTM	0.224 ± 0.075	0.196 ± 0.074
AutoFEDformer	0.265 ± 0.127	0.208 ± 0.092
AutoDeepAR	0.996 ± 0.329	0.777 ± 0.298

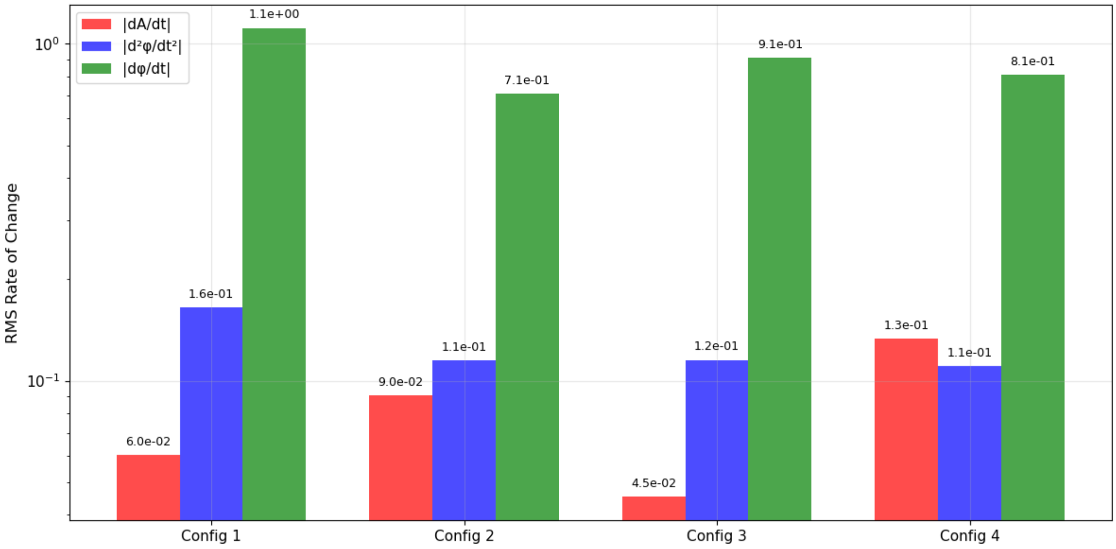

Fig. 1 Validation of synthetic IMFs: RMSE of amplitude (A) and frequency derivatives relative to phase (φ) derivatives, confirming compliance with the IMF definition.

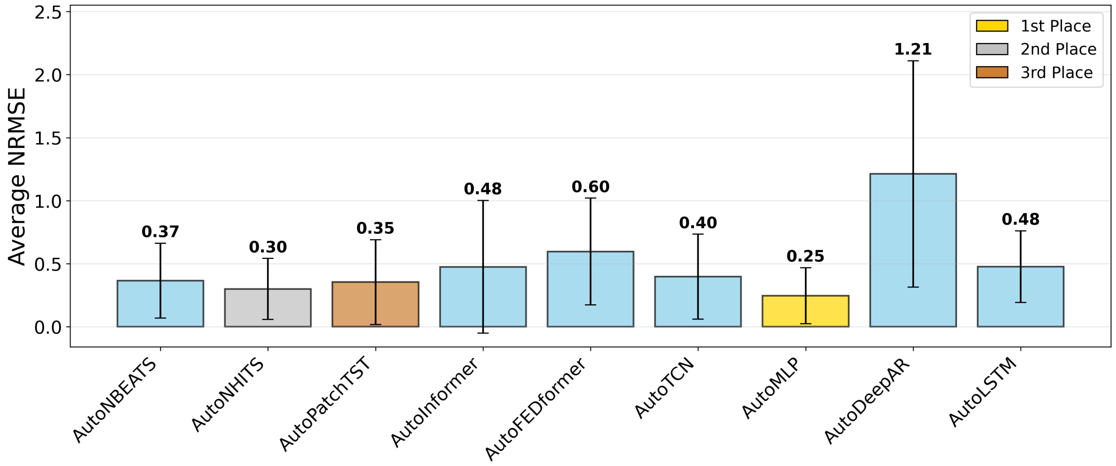

Fig. 2 Average NRMSE across all experiments comparing algorithm performance on synthetic IMFs and VMD-decomposed data.

Table 4 Summary of forecasting performance metrics on VMD-decomposed real-world data

Model	NRMSE	NMAE
AutoMLP	0.419 ± 0.195	0.337 ± 0.155
AutoNHITS	0.521 ± 0.133	0.425 ± 0.100
AutoTCN	0.585 ± 0.357	0.495 ± 0.339
AutoNBEATS	0.588 ± 0.272	0.496 ± 0.221
AutoPatchTST	0.604 ± 0.317	0.473 ± 0.250
AutoLSTM	0.730 ± 0.165	0.633 ± 0.146
AutoInformer	0.811 ± 0.573	0.712 ± 0.562
AutoFEDformer	0.928 ± 0.351	0.773 ± 0.374
AutoDeepAR	1.428 ± 1.189	1.102 ± 1.021

The superior performance of simpler architectures on band-limited IMF signals can be attributed to the alignment between their constrained capacity and the regular, quasi-harmonic structure of decomposed signals. Lacking complex mechanisms for multi-scale temporal patterns in raw time series, these models naturally excel at spectrally concentrated behavior.

PatchTST offers a nuanced case: although it is a full transformer, it avoids explicit seasonal-trend decomposition (unlike FEDformer) and uses standard self-attention rather than specialized mechanisms like Informer's ProbSparse attention. Whether its success is due to architectural simplicity, reduced sequence length through patching, or both is an interesting open question for future investigation.

DeepAR performs poorly because its probabilistic framework, designed for complex, noisy series, is ill-suited to quasi-deterministic IMFs. Its additional parameters for distribution estimation introduce unnecessary complexity, and its recursive prediction strategy compounds errors over the forecast horizon, highlighting a mismatch between model

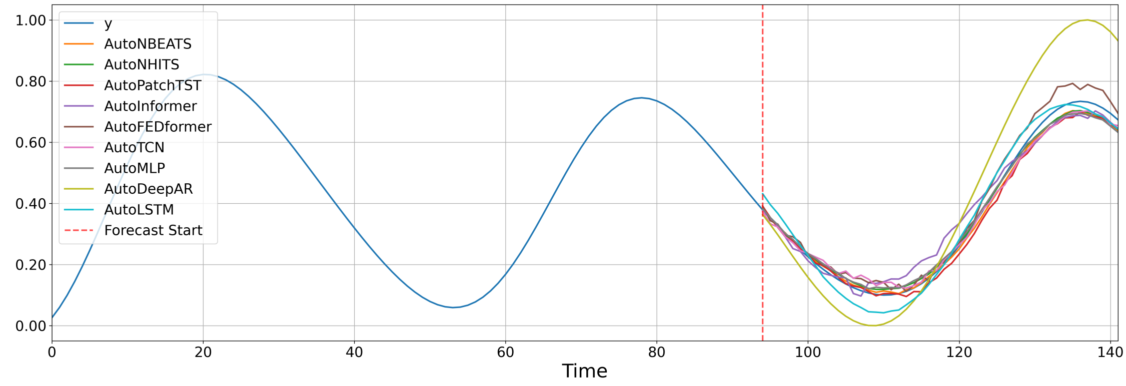

Fig. 3 Forecast results for synthetic IMF #1.

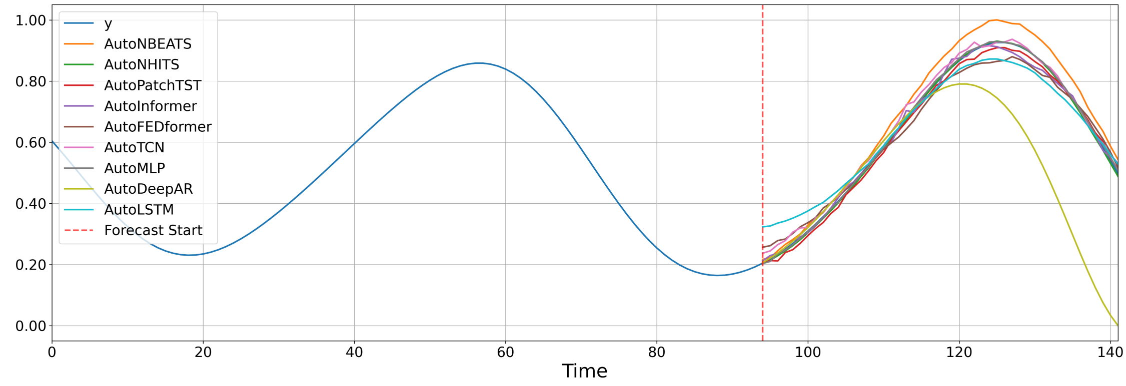

Fig. 4 Forecast results for synthetic IMF #3.

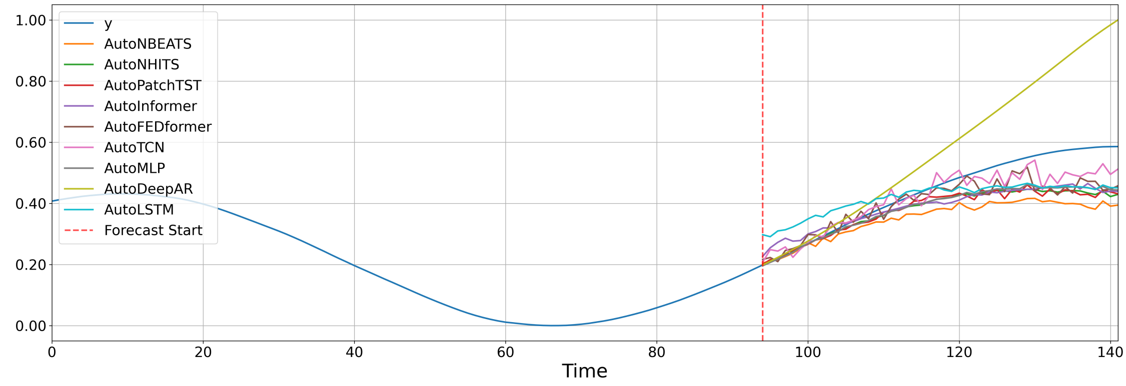

Fig. 5 Forecast results for VMD-decomposed signal P_1.

assumptions and signal characteristics. All these considered, features crafted for raw, non-stationary time series, such as specialized attention mechanisms and multi-scale decomposition, do not offer benefits for band-limited signals. This suggests that the computational cost of complex transformers is not justified for forecasting IMFs.

To illustrate the qualitative behavior of the forecasts produced by the different neural models, Figures 3, 4, 5, and 6 present representative examples of the forecast results for synthetic IMF#1 and IMF#3 and for the VMD decomposed signals P_1 and P_3, respectively. In all plots, both types of data are normalized to enable a more direct comparison and improve readability.

5 Conclusions and Future Directions

The comparative study presented in this paper suggests that, for band-limited signals such as IMFs obtained through decomposition, simpler neural architectures tend to achieve performance comparable to, or slightly better than more complex models. In particular, a basic MLP delivered the best

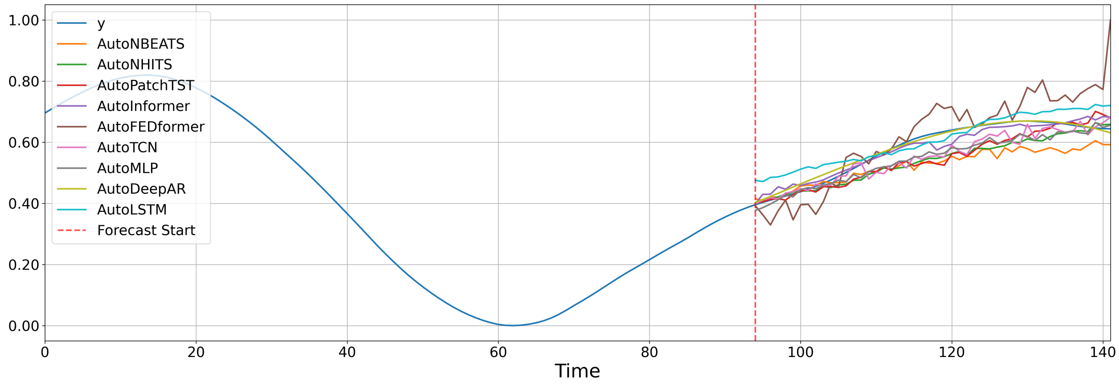

Fig. 6 Forecast results for VMD-decomposed signal P_3.

overall results in our experiments, despite its relatively dated and straightforward design. For practitioners employing preprocessing based on decomposition within hybrid forecasting frameworks, these findings indicate that simpler neural models may provide a more favorable cost–benefit tradeoff than state-of-the-art transformer-based architectures. Therefore, the selection of forecasting models should take into account not only benchmark performance on raw time series but also the spectral and mathematical properties of the preprocessed signals.

Future work will examine the performance of neural models across the full range of IMF components, the influence of decomposition parameters (e.g., number of modes K and regularization) on model ranking, and the potential of hybrid approaches that combine simple per-IMF predictors with learned aggregation to outperform end-to-end complex models. Additionally, multivariate VMD decompositions, which introduce cross-dependencies among variables, will be explored to assess whether attention-based models can more effectively capture inter-series relationships.

Acknowledgments This research was funded by the Italian Ministry of University and Research (MUR): Sustainable Mobility Center (CNMS) Spoke 3, project id CN00000023, CUP: B43C22000440001.

References

1. Chao, H.C., Lin, F.I., Pan, J.S., Chien, W.T., Lai, C.S.: Power load forecasting based on vmd and attention-lstm. In: Proceedings Of The 3rd International Conference On Data Science And Information Technology, pp. 19–25 (2020). DOI 10.1145/3414274.3414277. URL https://dl.acm.org/doi/10.1145/3414274.3414277
2. Daubechies, I., Lu, J., Wu, H.T.: Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. Appl. Comput. Harmon. Anal. **30**, 243–261 (2011)
3. Dragomiretskiy, K., Zosso, D.: Variational mode decomposition. IEEE Trans. Signal Process. **62**, 531–544 (2014)
4. Fazzini, P., La Tona, G., Diez, M., Di Piazza, M.C.: Sailing towards efficiency: A variational mode decomposition based approach to forecasting shipboard electrical power consumption. In: 2024 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC), pp. 1–6 (2024). DOI 10.1109/ESARS-ITEC60450.2024.10819873
5. Fazzini, P., Tona, G.L., Diez, M., Piazza, M.C.D.: Enhanced forecasting of shipboard electrical power demand using multivariate input and variational mode decomposition with mode selection. Scientific Reports **15**(1), 23941 (2025). DOI 10.1038/s41598-025-06153-z. URL https://doi.org/10.1038/s41598-025-06153-z
6. Haykal, V., Cardot, H., Ragot, N.: A combination of variational mode decomposition with neural networks on household electricity consumption forecast. In: Proceedings Of ITISE 2017 (2017)
7. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. A **454**, 903–995 (1998). DOI 10.1098/rspa.1998.0193
8. Nie, Y., et al.: A time series is worth 64 words: Long-term forecasting with transformers. In: ICLR (2023)
9. Oreshkin, B.N., Carpov, D., Chapados, N., Bengio, Y.: N-beats: Neural basis expansion analysis for interpretable time series forecasting (2020). URL https://arxiv.org/abs/1905.10437
10. Shi, J., Zhou, J., Feng, J., Chen, H.: Network traffic prediction model based on improved vmd and pso-elm. International Journal Of Communication Systems **36**(5) (2023). DOI 10.1002/dac.5448
11. Stefenon, S.F., Seman, L.O., Yamaguchi, C.K., Coelho, L.D.S., Mariani, V.C., Matos-Carvalho, J.P., Leithardt, V.R.Q.: Neural hierarchical interpolation time series (nhits) for reservoir level multi-horizon forecasting in hydroelectric power plants. IEEE Access **13**, 54853–54865 (2025). DOI 10.1109/ACCESS.2025.3554446
12. Xia, X., Wang, X.: A novel hybrid model for short-term wind speed forecasting based on twice decomposition, psr, and imvo-elm. Complexity **2022**, 1–21 (2022). DOI 10.1155/2022/4014048
13. Zhang, D., Peng, X., Pan, K.: A novel online method for wind speed forecasting based on hybrid decomposition and osorelm. In: 2018 International Conference On Power System Technology (POWERCON), pp. 1797–1804 (2018). DOI 10.1109/POWERCON.2018.8601706. URL https://ieeexplore.ieee.org/document/8601706/
14. Zhou, H., et al.: Informer: Beyond efficient transformer for long sequence time-series forecasting. In: Proceedings of AAAI (2021)
15. Zhou, T., et al.: Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting. In: ICML (2022)