CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 10 (2) (2025), Pages 1266-1276
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Empirical Evaluation of Wavelet Filter and Wavelet Decomposition Level on Time Series Forecasting

Mira Andriyani* and Dewi Retno Sari S.

Mathematics Department, Universitas Sebelas Maret, Surakarta, Indonesia

Abstract

Time series forecasting is essential for anticipating future outcomes and supporting decision-making, yet achieving high predictive accuracy remains challenging. Wavelet-based approaches, particularly the Maximal Overlap Discrete Wavelet Transform (MODWT), offer potential improvements, although limited studies have systematically compared wavelet filter types and decomposition levels. This study evaluates several wavelet filters and decomposition levels combined with ARIMA models across six datasets exhibiting varying temporal characteristics. Forecasting accuracy was measured using the Mean Absolute Error (MAE) and Symmetric Mean Absolute Percentage Error (SMAPE). For the datasets analyzed, the Haar filter yielded the lowest MAE and SMAPE values, a result supported by the Kruskal–Wallis test and Dunn's test, which indicated significant differences in accuracy across filters. In contrast, differences in decomposition levels were not statistically significant, suggesting that decomposition level played a limited role in forecasting performance within this dataset context. These findings provide empirical, dataset-specific evidence regarding filter selection in MODWT–ARIMA modeling and highlight the comparatively minor influence of decomposition level on forecasting accuracy.

Keywords: ARIMA; level; MODWT; wavelet filter

1 Introduction

Time series data is a collection of observational data organized chronologically [1]. The time intervals between observations are fixed, such as seconds, minutes, hours, days, weeks, months, and so on. One application for time series data is to forecast future events using time series forecasting methods [1]. Future occurrences can be predicted via time series forecasting, allowing actions to be planned ahead of time to produce favorable results [2]. To be truly useful, time series forecasting predictions must be highly accurate, with minimal errors [3]. Root Mean Squared Error (RMSE), mean absolute error (MAE), and symmetric mean absolute percentage error (SMAPE) as criteria for choosing the optimal forecasting model [3], [4], [5], [6], [7], [8], [9]. Lower RMSE, MAE, or SMAPE values suggest better forecasting outcomes. Finding the forecasting approach that yields the fewest possible errors is still an open problem.

Several forecasting models have been developed and used to analyze data with varying properties. Stationary data can be forecasted using the Autoregressive (AR), Moving Average (MA), and Autoregressive Moving Average (ARMA) models [10]. The ARIMA (Autoregressive

*Corresponding author. E-mail: miraandriyani10@staff.uns.ac.id

Submitted: September 9, 2025 Reviewed: September 28, 2025 Accepted: November 21, 2025
DOI: https://doi.org/10.18860/cauchy.v10i2.36440

Integrated Moving Average) model can be utilized with non-stationary data since it combines Autoregressive (AR), Moving Average (MA), and Integrated (I). The integrated (I) value reflects the number of differencing steps used to render the data stationary [5]. If the data is non-stationary and incorporates seasonal patterns, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model can be applied. Other approaches, such as machine learning, like artificial neural networks (ANN) and convolutional neural network (CNN), can be employed to deal with nonlinear data [11]. Hybrid methods, which integrate numerous techniques or procedures, are also being developed.

Wavelet decomposition is one such approach that enhances forecasting accuracy by transforming non-stationary series into several stationary components [10]. The Maximal Overlap Discrete Wavelet Transform (MODWT) is preferred over the standard Discrete Wavelet Transform (DWT) because it can handle arbitrary data lengths [12], [13]. Previous studies have demonstrated the usefulness of MODWT in improving forecasts. For example, Ndlovu and Chikobvu [12] applied MODWT with Haar and Daubechies wavelets to predict cryptocurrency risks, obtaining improved accuracy. Other works also confirmed that wavelet-based preprocessing can significantly enhance time series forecasting results [3], [10], [13], [6], [14], [15], [16], [17], [18], [7], [8]. Various wavelet filters have been used, including Haar, Daubechies, Least Asymmetric , Best Localized, Coiflet, and symlet4 [12], [13], [14], [19]. However, most studies adopted these filters without explaining the rationale behind their selection, even though the choice of filter substantially affects the resulting forecast.

Similarly, the decomposition level which determines how many scales the signal is broken into varies widely across studies. For instance, Al-Wadi et al. [20] used the first level, Farajpanah et al. [21] the third, and Panja et al. [22] and Hermansah et al. [23] the sixth level, yet none provided justification for their choices. This raises a critical question: Does a higher decomposition level necessarily yield more accurate forecasts, or is the opposite true?

This study addresses this research gap by systematically comparing multiple wavelet filters and decomposition levels within the MODWT-ARIMA framework. Six datasets, both stationary and non-stationary, were analyzed using seven wavelet filters (Haar, Daubechies 4, Daubechies 6, Least Asymmetric 8, Least Asymmetric 16, Best Localized 14, and Best Localized 20) at all possible decomposition levels. The results provide empirical evidence and guidelines for selecting appropriate wavelet filters and decomposition levels to achieve optimal forecasting accuracy.

2 Methods

This section details the methodological framework used to evaluate the effect of wavelet filters and decomposition levels on forecasting accuracy. First, we describe the research design, including the characteristics of the six time series, the rolling-origin evaluation scheme, and the computation of MAE and SMAPE as accuracy measures. We then outline the MODWT–ARIMA modeling procedure, specifying how the MODWT decomposition and subsequent ARIMA modeling are combined to generate one-step-ahead forecasts for each filter–level combination.

2.1 Research Framework

This is a quantitative study that compares the empirical performance of wavelet filters and wavelet decomposition levels according to RMSE, MAE, and SMAPE values. In this study, six datasets were used, namely simulated observations, monthly rainfall data for Gunung Kidul Regency from January 2018 to December 2024, traffic accident deaths in the United States from 1973 to 1978 [24], the number of goods by rail between 1987 and 2022 [25], tobacco production statistics in the United States from 1871 to 1984 [1], and international airline passenger data (Airline Data) from January 1949 to December 1960 [1]. The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test was conducted to determine the stationarity of the six data sets [5]. All datasets are

initially split into two parts using an 80% training and 20% testing ratio under a rolling-origin scheme. The number of origins increases by one in each subsequent iteration, and the maximum number of origins is equal to the total number of observations minus one. At each origin, the training data are decomposed using the MODWT, producing several sub-series. Each sub-series is then modeled using the auto.arima function, and the resulting ARIMA model is used to generate a one-step-ahead forecast (h = 1).

The wavelet filters employed in the MODWT decomposition include Haar, Daubechies 4 (d4), Daubechies 6 (d6), Least Assymmetric 8 (la8), Least Assymmetric 16 (la16), Best Localized 14 (bl14), and Best Localized 20 (bl20) filters. The decomposition levels considered are all feasible levels, which depend on the length of the data. Within the waveslim package in R, the maximum decomposition level is specified as $\lfloor \log_2(N) \rfloor$ with N is the length of the data.

Regarding ARIMA modeling, auto.arima in R returns the model with the lowest AIC as the best candidate. The auto.arima function does not only generate non-seasonal ARIMA models, but it is also capable of automatically fitting SARIMA models when the argument seasonal = TRUE is specified. However, auto.arima is unable to detect the seasonal period automatically, and by default, the seasonal period is set to 1. Therefore, when the time series plot clearly indicates the presence of seasonality, the seasonal period must be explicitly defined within the auto.arima function prior to model estimation.

For evaluation, RMSE, MAE, and SMAPE are computed from the one-step-ahead forecasts, where RMSE, MAE, and SMAPE are defined as in Eq. 1, Eq. 2, and Eq. 3, where N is the number of data points, X_t is the actual data, and $\hat{X}_t$ is the prediction result.

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(X_t - \hat{X}_t)^2} \tag{1}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|X_t - \hat{X}_t| \tag{2}$$

$$SMAPE = \frac{100\%}{N}\sum_{t=1}^{N}\frac{|X_t - \hat{X}_t|}{(|X_t| + |\hat{X}_t|)/2} \tag{3}$$

In the case of h=1, the RMSE value becomes identical to the MAE value, as both reduce to the absolute error. At each origin, one value of RMSE, MAE, and SMAPE is obtained. The performance metrics used for evaluation are the average RMSE, average MAE, and average SMAPE across all origins (iterations). Given that h=1, the average RMSE is equivalent to the MAE, or the average MAE. Therefore, only MAE and SMAPE are used as the evaluation measures for all possible combinations of wavelet filters and decomposition levels in each dataset. These values are then compared, and the combination yielding the smallest value is considered to produce the best forecasting model.

In addition to selecting the minimum error values, the conclusions are further validated using the Kruskal–Wallis test followed by Dunn's post-hoc test. Before conducting the Kruskal–Wallis test, the MAE and SMAPE values were first normalized to ensure that the magnitudes of the values were comparable and not excessively different from one another. The Kruskal–Wallis test, as applied in Akila's study [26], does not require the MAE or SMAPE distributions to be normally distributed. The test indicates that at least one pair of filters or decomposition levels differs significantly when the p-value is less than 0.05, leading to the rejection of the null hypothesis [27]. If the Kruskal–Wallis test reveals significant differences, Dunn's test is subsequently performed, following the procedure described by Dunn (1964) and referenced in [28]. In this study, the Holm correction is employed for the multiple-comparison adjustment.

2.2 MODWT-ARIMA

MODWT-ARIMA modelling is the ARIMA model of MODWT decomposition data. The MODWT-ARIMA forecasting results are the total of the ARIMA forecasting results, both detailed and smooth. The original data is the sum of the MODWT outcomes, as calculated in Eq. 4.

$$Z_t = \sum_{j=1}^{J} D_{j,t} + S_{J,t}, t = 1, 2, \ldots, N \tag{4}$$

If $D_{j,t}$ is modelled using ARIMA(p, d, q), the model in Eq. 5 is achieved [10]. Meanwhile, Eq. 6 is used to obtain the model for $S_{J,t}$. The MODWT-ARIMA forecast result is calculated by adding the ARIMA results for detail and smoothing.

$$\phi_{j,p}(B)(1-B)^d D_{j,t} = \theta_0 + \theta_{j,q}(B)a_t \tag{5}$$

$$\phi_{J,p}(B)(1-B)^d S_{J,t} = \theta_0 + \theta_{J,q}(B)a_t \tag{6}$$

3 Results and Discussion

This section presents the empirical findings obtained from applying the MODWT–ARIMA framework across six datasets using various combinations of wavelet filters and decomposition levels. The analysis is organized sequentially, beginning with an overview of the datasets and their stationarity properties, followed by the decomposition outcomes, the resulting ARIMA models, and the forecasting accuracy summarized through MAE and SMAPE. The discussion then evaluates the comparative performance of filters and decomposition levels and concludes with statistical validation using the Kruskal–Wallis and Dunn tests. Together, these results provide a comprehensive assessment of how filter choice and decomposition depth influence forecasting accuracy.

The datasets utilized in this study are presented in Fig. 1. Each dataset displays distinct temporal characteristics, including stationary behavior, non-stationary trends, and seasonal patterns, which collectively provide a diverse empirical basis for evaluating the performance of different wavelet filters and decomposition levels.

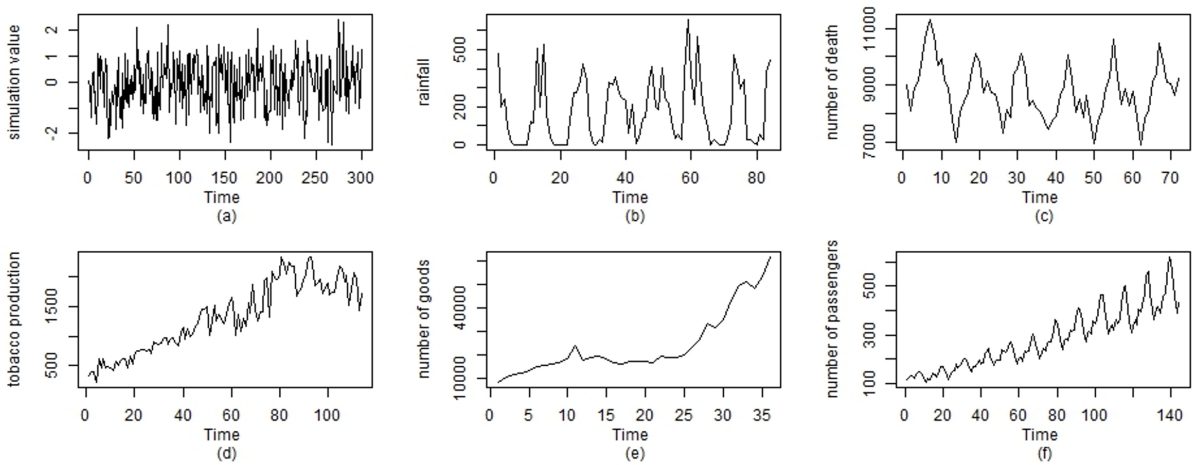

Figure 1: Graph of Six Data Sets Used in the Study

In Fig. 1 dataset (a), is a simulated dataset, dataset (b) is rainfall dataset, dataset (c) contains number of death, datasets (d) is a tobacco production dataset; dataset (e) is number of goods dataset; and dataset (f) is an airline dataset. Based on the observed plots, datasets (a), (b), and (c) appear to be stationary, whereas datasets (d), (e), and (f) exhibit an upward trend, indicating

non-stationarity. In addition, dataset (f) displays a clear seasonal pattern with a period of 12. Dataset (c) also shows a nearly repeating pattern that suggests seasonality, although its period is not clearly identifiable. To formally assess the stationarity of all six datasets, the KPSS test was conducted, and the results are summarized in the Table 1.

Table 1: The Results of KPSS Test

Data	p-value	Decision
Simulation	0.1	Stationary
Rainfall	0.1	Stationary
Death	0.08853	Stationary
Tobacco Production	0.01	Non-Stationary
Number of Goods by Rail	0.01592	Non-Stationary
Airlines	0.01	Non-Stationary

From Table 1, three datasets, (a), (b), and (c) are stationary and the other three are non-stationary. Under the rolling-origin scheme, the training dataset is decomposed using MODWT at each iteration, producing several sub-datasets. For example, in the airlines dataset, applying MODWT with the Haar filter at level 3 yields four sub-datasets—D1, D2, D3, and S3—as illustrated in Fig. 2. MODWT decomposition is applied to all datasets using seven wavelet filters, namely Haar, d4, d6, la8, la16, bl14, and bl20, across levels 1 up to the maximum feasible level.

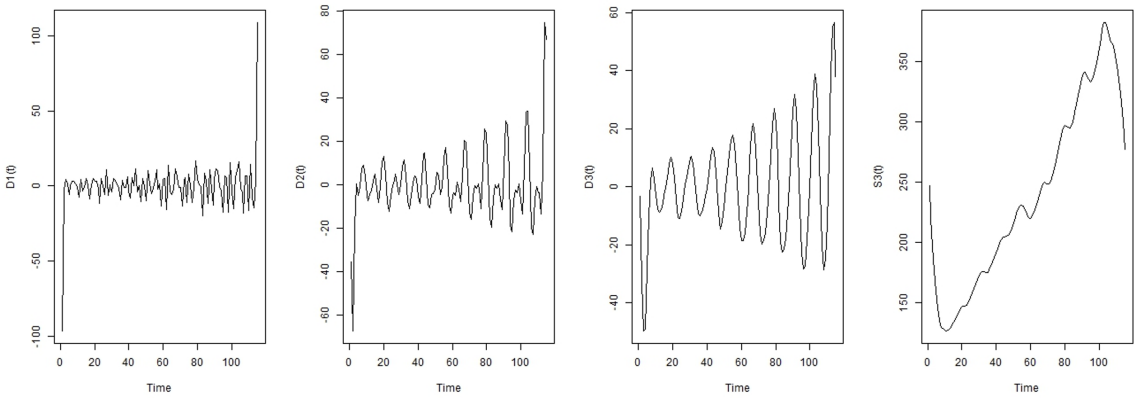

Figure 2: The Decomposition Results of Airlines Dataset by Haar at the Third Level

The y-axis labels in Fig. 2, such as D1(t), represent the values of D1 at time t. For the simulated dataset with 300 observations, the initial training size of 240 (0.8×300) allows a maximum level of 7. The maximum decomposition levels for the other datasets—rainfall, death, tobacco, number of goods, and airlines—are 6, 5, 6, 4, and 6, respectively.

The total number of decompositions for each dataset equals ($7 \times$ maximum level). For each combination of dataset, filter, and level, each sub-dataset generated through MODWT is modeled using auto.arima, producing as many ARIMA models as there are sub-datasets. As an example, Table 2 presents the four ARIMA models generated from the sub-datasets resulting from the decomposition of the airlines dataset using the Haar filter at level 3 with origin 115 (initial-origin).

Table 2: ARIMA models of the sub-datasets from the airlines dataset using the Haar filter at level 3

Subdataset	ARIMA model
D1	ARIMA(0,0,2)(1,0,0)[12]
D2	ARIMA(3,0,0)(0,1,0)[12]
D3	ARIMA(3,0,0)(0,1,1)[12]
S3	ARIMA(2,1,0)(1,0,0)[12]

Based on Table 2, the auto.arima function used in this study can automatically construct SARIMA models for datasets exhibiting seasonal trends. It should be noted, however, that the seasonal period must be specified in advance. Because each combination produces multiple models, it is not practical to report all of them, and representative examples are therefore provided.

It is important to note that the ARIMA models used in this study were not subjected to detailed diagnostic checking, such as residual autocorrelation or normality tests. This decision was taken to maintain the primary focus of the analysis, which was to compare forecasting accuracy across different wavelet filters and decomposition levels rather than to optimize or evaluate the ARIMA models themselves. Although diagnostic tests can further validate model adequacy, the consistent modeling framework applied across all filters ensures that the comparative insights drawn in this study remain reliable. Nonetheless, incorporating full diagnostic procedures represents a valuable direction for future research, particularly for studies aiming to assess both filter performance and model specification in greater depth.

In this framework, the expected value of the original series is reconstructed by summing the expected values of the ARIMA-fitted sub-datasets. The same set of ARIMA models is then used to generate one-step-ahead forecasts, and the final predicted value is obtained by aggregating the forecasts from all sub-models.

After obtaining the one-step-ahead forecasts at each origin, the MAE and SMAPE values were calculated. The MAE and SMAPE values reported in this section represent the averages across all forecasts at each origin. Therefore, each combination of dataset, filter, and level is associated with a single MAE and SMAPE value. For instance, in the simulated dataset with a total of 300 observations and an initial training size of 240, there are 60 forecasts, resulting in 60 MAE values. The reported MAE is the average of these 60 values. Table 3 presents a snapshot of the table showing the average MAE and SMAPE values for each combination of dataset, filter, and level.

Table 3: Illustration of the mean MAE and SMAPE obtained for each dataset, filter, and decomposition level combination

Name of Dataset	Filter	Level	MAE	SMAPE
Simulation	Haar	1	0.804035	177.333947
Simulation	Haar	2	0.820109	155.498955
Simulation	Haar	3	0.828560	161.147645
⋮	⋮	⋮	⋮	⋮
Airlines	bl20	5	255.858626	82.832744
Airlines	bl20	6	257.321961	83.479226

Based on the results in Table 3, it is possible to identify the filter–level combinations that yield the lowest MAE or SMAPE for each dataset. Table 4 and Table 5 provide a consolidated summary of the minimum MAE and SMAPE values obtained across all datasets.

Table 4: Summary of Minimum MAE

Name of Dataset	Minimu MAE	Filter	Level
Simulation	0.804035	Haar	1
Rainfall	231.712661	Haar	1
Death	679.467340	Haar	5
Tobacco	1288.178639	Haar	1
Number of Goods	15463.125093	Haar	1
Airlines	236.958006	Haar	6

Table 5: Summary of Minimum SMAPE

Name of Dataset	Minimu MAE	Filter	Level
Simulation	155.498956	Haar	2
Rainfall	117.054792	Haar	1
Death	7.830627	Haar	5
Tobacco	102.809825	Haar	1
Number of Goods	39.971473	Haar	1
Airlines	73.277248	Haar	6

Based on Table 4 and Table 5, the Haar filter consistently yields the smallest MAE and SMAPE values across all evaluated datasets, regardless of whether the series exhibits stationary

or non-stationary behavior. There is no difference in the selected filter when the selection is based on MAE or SMAPE. However, a discrepancy arises in the choice of decomposition level for the simulated dataset: MAE identifies level 1 as the most accurate, whereas SMAPE indicates that level 2 provides superior performance. The most frequently selected decomposition level across the six datasets was Level 1 (low level). Higher levels were selected only for the death and airlines datasets, both at Level 5. As shown in Figure 1, these two datasets exhibit clear seasonal patterns. If the significance testing had indicated that decomposition level exerts a statistically significant influence on forecasting accuracy, one possible interpretation would be that datasets with pronounced seasonality may require higher decomposition levels to achieve improved accuracy.

Subsequently, validation is performed using the Kruskal–Wallis (KW) test and Dunn's post hoc test to determine whether the observed differences in MAE and SMAPE are statistically significant. Both the KW and Dunn tests are conducted using the R software. Prior to conducting the Kruskal–Wallis significance tests, the datasets containing the MAE and SMAPE values were normalized. This normalization step was necessary to prevent large differences in scale across datasets from influencing the statistical comparison. The KW tests include: (i) a KW test on MAE values across filters, (ii) a KW test on MAE values across decomposition levels, and (iii) a KW test on SMAPE values across filters. The p-values obtained from the Kruskal–Wallis tests for MAE and SMAPE are 2.2e-16 and 8.884e-11, respectively. These results lead to the rejection of the null hypothesis, indicating that there are significant differences in both MAE and SMAPE across the filters. To identify which specific pairs of filters differ significantly, Dunn's post hoc test was performed, and the results are presented in Table 6. In Table 6, a comparison is considered

Table 6: Result of Dunn Test for Filter Based on SMAPE

Comparison	P.adj	Conclusion
bl14 - bl20	0.952246172	not significant
bl14 - d4	1	not significant
bl20 - d4	1	not significant
bl14 - d6	0.856675677	not significant
bl20 - d6	0.819622238	not significant
d4 - d6	0.02784277	significant
bl14 - haar	2.81E-05	significant
bl20 - haar	3.57E-05	significant
d4 - haar	0.008836356	significant
d6 - haar	1.08E-09	significant
bl14 - la16	1	not significant
bl20 - la16	1	not significant
d4 - la16	0.52668853	not significant
d6 - la16	1	not significant
haar - la16	7.11E-07	significant
bl14 - la8	1	not significant
bl20 - la8	1	not significant
d4 - la8	0.118405471	not significant
d6 - la8	1	not significant
haar - la8	2.30E-08	significant
la16 - la8	1	not significant

statistically significant when the adjusted p-value (P.adj) is less than 0.05. A significant result indicates that the pair of filters being compared exhibits a statistically meaningful difference in their SMAPE values. The table also shows that the Haar filter displays significant differences when compared with each of the other six filters included in the analysis. The Dunn post hoc test for the MAE values across filters produced results consistent with those obtained for SMAPE, indicating that the Haar filter exhibits significantly different MAE values compared with the other filters.

Meanwhile, the Kruskal–Wallis test conducted to examine differences in MAE across decomposition levels yielded a p-value of 0.9983 (> 0.05). This result indicates that the MAE values do not differ significantly across levels, suggesting that the decomposition level does not exert a significant effect on forecasting accuracy. However, these findings negate the earlier assumption and demonstrate that, even for datasets exhibiting seasonal patterns, the choice of decomposition level does not have a significant impact on forecasting accuracy.

Overall, the findings from both the accuracy evaluation and the statistical tests consistently highlight the superior performance of the Haar filter across various data characteristics (whether stationary or non-stationary, and whether exhibiting trend or seasonality) and decomposition levels. The convergence of results from MAE, SMAPE, and subsequent post hoc analyses provides strong empirical support for the robustness of this filter in improving forecasting accuracy.

4 Conclusion

Based on the analyses conducted on the six datasets in this study, the Haar filter consistently yielded the lowest forecasting errors and demonstrated statistically significant differences compared with the other filters. These findings should be understood as dataset-specific insights rather than universal claims, as the datasets used represent a limited set of characteristics, including stationary, non-stationary, trending, and seasonal series. Although different decomposition levels emerged across datasets, the significance tests indicated that decomposition level did not exert a meaningful influence on forecasting accuracy within this empirical context. Overall, the results offer a practical empirical guideline suggesting that the Haar filter may be a strong candidate for forecasting tasks involving datasets with similar properties.

This study has several limitations. In particular, the ARIMA models employed were not subjected to diagnostic checks, such as residual autocorrelation or normality testing, which could influence the reliability of the model estimates. Future research may extend this work by incorporating comprehensive diagnostic testing to ensure model adequacy, examining a broader variety of time-series structures, exploring additional wavelet families, and assessing the interaction between filter choice, decomposition level, and forecasting model performance. Such extensions would help strengthen and potentially generalize the empirical observations reported here.

CRediT Authorship Contribution Statement

Mira Andriyani: Conceptualization, Methodology, Data Curation, Formal Analysis, Software, Validation, Visualization, Writing–Original Draft, Writing–Review & Editing.
Dewi Retno Sari S.: Supervision.

Declaration of Generative AI and AI-assisted technologies

During the preparation of this text, the authors used ChatGPT-5 to help create some LaTeX scripts in R and sentences in more academic English. The author used DeepL to translate and QuillBot to paraphrase. Meanwhile, the author relied on R 4.4.2 and RStudio for image visualization and calculations.

Declaration of Competing Interest

The authors declare no competing interests, financial or personal, that could have influenced the work presented in this study.

Funding and Acknowledgments

This research was supported by Universitas Sebelas Maret through Agreement on the Assignment of Research Implementation of Non-State Budget Funds (APBN) 2025 with contract number 371/UN27.22/PT.01.03/2025.

Data and Code Availability

The code analyzed during the current study are publicly available in the GitHub[1]. All datasets utilized in this study can be found in the sources cited in Section 2.

References

[1] W. W. Wei, *Time Series Analysis Univariate and Multivariate Methods*. Pearson Education, Inc., 2006, vol. 2. Available online.

[2] S. Adamala, "Time series analysis: A hydrological prospective," *American Journal of Scientific Research and Essays*, vol. 1, no. 1, pp. 31–40, 2016. Available online.

[3] E. Aladag, "Forecasting of particulate matter with a hybrid arima model based on wavelet transformation and seasonal adjustment," *Urban Climate*, vol. 39, no. June, pp. 100 930–100 945, 2021. DOI: 10.1016/j.uclim.2021.100930.

[4] J. Bruzda, "The haar wavelet transfer function model and its applications," *DYNAMIC ECONOMETRIC MODELS*, vol. 11, pp. 141–153, 2011. DOI: 10.12775/DEM.2011.010.

[5] H. P. M., M. Z. Rehman, A. A. Dar, and T. W. A., "Forecasting co2 emissions in india: A time series analysis using arima," *Processes*, vol. 12, no. 12, pp. 2699–2714, 2024. DOI: 10.3390/pr12122699.

[6] S. Al Wadi, A. Hamarsheh, and H. Alwadi, "Maximum overlapping discrete wavelet transform in forecasting banking sector," *Applied Mathematical Sciences*, vol. 7, no. 80, pp. 3995–4002, 2013. DOI: 10.12988/ams.2013.36305.

[7] N. A. Yaacob, J. J. Jaber, D. Pathmanathan, S. Alwadi, and I. Mohamed, "Hybrid of the lee-carter model with maximum overlap discrete wavelet transform filters in forecasting mortality rates," *Mathematics*, vol. 9, no. 18, pp. 2295–2305, 2021. DOI: 10.3390/math9182295.

[8] M. U. YOUSUF, I. AL-BAHADLY, and E. AVCI, "Short-term wind speed forecasting based on hybrid modwt-arima-markov model," *IEEE Access*, vol. 9, no. May, pp. 4803–4820, 2021. DOI: 10.1109/ACCESS.2021.3084536.

[9] A. Jierula, S. Wang, T.-M. OH, and P. Wang, "Study on accuracy metrics for evaluating the predictions of damage locations in deep piles using artificial neural networks with acoustic emission data," *Applied Science*, vol. 11, p. 2314, 2021. DOI: 10.3390/app11052314.

[10] L. Zhu, Y. Wang, and Q. Fan, "Modwt-arma model for time series prediction," *Applied Mathematical Modelling*, vol. 38, no. 5–6, pp. 1859–1865, 2014. DOI: 10.1016/j.apm.2013.10.002.

[11] K. Szostek, D. Mazur, G. Drałus, and J. Kusznier, "Analysis of the effectiveness of arima, sarima, and svr models in time series forecasting: A case study of wind farm energy production," *Energies*, vol. 17, no. September 2024, pp. 4803–4820, 2024. DOI: 10.3390/en17194803.

[1] https://github.com/mir2876/modwt-arima-rolling-origin

[12] T. Ndlovu and D. Chikobvu, "A wavelet-decomposed wd-arma-garch-evt model approach to comparing the riskiness of the bitcoin and south african rand exchange rates," *data*, vol. 8, no. 7, pp. 122–145, 2023. DOI: 10.3390/data8070122.

[13] A. A. A. Dghais and M. T. Ismail, "A study of stationarity in time series by using wavelet transform," in *Proceedings of the 21st National Symposium on Mathematical Sciences (SKSM21)*, vol. 1605, AIP Publishing, 2014, pp. 798–804. DOI: 10.1063/1.4887692.

[14] T. S. ALSHAMMARI, M. T. ISMAIL, S. AL-WADI, M. H. SALEH, and J. J. JABER, "Modeling and forecasting saudi stock market volatility using wavelet methods," *Journal of Asian Finance, Economics and Business*, vol. 7, no. 11, pp. 83–93, 2020. DOI: 10.13106/jafeb.2020.vol7.no11.08.

[15] L.-W. Lin and X.-H. Zhou, "Multiscale forecasting approach of property insurance income via wavelet method," *Mathematical Problems in Engineering*, vol. 2022, no. 1, p. 9 554 695, 2022. DOI: 10.1155/2022/9554695.

[16] P. Mittal, "Wavelet transformation and predictability of gold price index series with arma model," *International Journal of Experimental Research and Review (IJERR)*, vol. 30, no. April, pp. 127–133, 2023. DOI: 10.52756/ijerr.2023.v30.014.

[17] P. Mittal, "Forecasting of crude oil prices using wavelet decomposition based denoising with arma model," *Asia Pacific Financial Markets*, vol. 31, pp. 355–365, 2024. DOI: 10.1007/s10690-023-09418-7.

[18] J. Quilty and J. Adamowski, "A maximal overlap discrete wavelet packet transform integrated approach for rainfall forecasting – a case study in the awash river basin (ethiopia)," *Environmental Modelling and Software*, vol. 144, no. July, pp. 105 119–105 133, 2021. DOI: 10.1016/j.envsoft.2021.105119.

[19] G. Chiranjivi and R. Sensarma, "International review of financial analysis the effects of economic and financial shocks on private investment: A wavelet study of return and volatility spillovers," *International Review of Financial Analysis*, vol. 90, no. January, p. 102 936, 2023. DOI: 10.1016/j.irfa.2023.102936.

[20] S. Al Wadi, O. Al Singlawi, J. J. Jaber, M. H. Saleh, and A. A. Shehadeh, "Enhancing predictive accuracy through the analysis of banking time series: A case study from the amman stock exchange," *Journal of Risk and Financial Management*, vol. 17, no. 3, p. 98, 2024. DOI: 10.3390/jrfm17030098.

[21] H. Farajpanah, A. Adib, M. Lotfirad, H. Esmaeili-Gisavandani, M. M. Riyahi, and A. Zaerpour, "A novel application of waveform matching algorithm for improving monthly runoff forecasting using wavelet–ml models," *Journal of Hydroinformatics*, vol. 26, no. 7, pp. 1771–1789, 2024. DOI: 10.2166/hydro.2024.128.

[22] M. Panja, T. Chakraborty, U. Kumar, and N. Liu, "Epicasting: An ensemble wavelet neural network for forecasting epidemics," *Neural Networks*, vol. 165, no. June, pp. 185–212, 2023. DOI: 10.1016/j.neunet.2023.05.049.

[23] Hermansah, D. Rosadi, H. Utami, Abdurakhman, and G. Darmawan, "Hybrid modwt-ffnn model for time series data forecasting," in *AIP Conference Proceedings*, vol. 2192, 2019. DOI: 10.1063/1.5139175.

[24] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-29854-2.

[25] Badan Pusat Statistik, *Jumlah penumpang dan barang melalui transportasi kereta api indonesia tahun 1987-2022*, https://www.bps.go.id/id/statistics-table/1/MTQxNCMx/jumlah-penumpang-dan-barang-melalui-transportasi-kereta-api-indonesia-tahun-1987-2022.html, Date Accessed: July 14th, 2025.

[26] A. D. Kayit and M. T. Ismail, "Advancing stock price prediction through the development of hybrid ensembles: A comprehensive comparative analysis of machine learning approaches," *Journal of Big Data*, vol. 12, p. 232, 2025. DOI: 10.1186/s40537-025-01185-8.

[27] D. Ramadhani, A. M. Soleh, and Erfiani, "Characteristics of machine learning-based univariate time series imputation method," *JUITA: Jurnal Informatika*, vol. 12, pp. 279–288, 2024. DOI: 10.30595/juita.v12i2.23453.

[28] A. Dinno, "Cnonparametric pairwise multiple comparisons in independent groups using dunn's test," *The Stata Journal*, vol. 15, pp. 292–300, 2015.