BAREKENG: Journal of Mathematics and Its Applications
December 2026 Volume 20 Issue 4 Page 2783-2798
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss4pp2783-2798

A HYBRID ARIMA-DISCRETE WAVELET TRANSFORM MODEL FOR ENHANCED COASTAL SEA LEVEL FORECASTING

Hakami Ali Mohammed[1], **Muhammad Fadhil Marsani**[2*], **Mohd Shareduwan Mohd Kasihmuddin**[3], **Farid Zamani Che Rose**[4], **Basri Badyalina**[5]

[1,2,3] School of Mathematical Sciences, Universiti Sains Malaysia
Penang, 11800, Malaysia

[4]Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia
[4]Institut Penyelidikan Matematik (INSPEM), Universiti Putra Malaysia
UPM Serdang, Selangor Darul Ehsan, 43400, Malaysia

[5]Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara
Cawangan Johor Kampus Segamat, Johor, 85000, Malaysia

*Corresponding author e-mail: * **fadhilmarsani@usm.my***

Article Info

Article History:

Received: 5th September 2025
Revised: 2nd April 2026
Accepted: 16th May 2026
Published: 1st July 2026

Keywords:

coastal resilience;
Hybrid ARIMA-DWT Model;
non-linear trends
sea level forecasting;
Time Series Decomposition.

ABSTRACT

Accurate sea level forecasting is essential for coastal resilience, as rising sea levels threaten coastal residents, infrastructure, and ecosystems, heightening the hazards of floods and erosion; yet, conventional models struggle with non-linear trends. This study proposes a hybrid ARIMA-DWT (Autoregressive Integrated Moving Average with Discrete Wavelet Transform) model to enhance prediction accuracy by decomposing sea level data into linear (ARIMA) and non-linear (DWT) components. The hybrid model, utilizing monthly sea level data from Penang, Malaysia (1984–2018; N=408), diminished errors by 20% compared to the standalone ARIMA, attaining out-of-sample RMSE=81.76, MAE=68.40, and MAPE=2.45%. The ARIMA-DWT framework effectively captures long-term climate trends and short-term variations, providing a computationally efficient resource for coastal planners. This method demonstrates enhanced efficacy in situations characterized by significant seasonality, offering practical insights for climate adaptation. This research enhances the existing body of work on hybrid forecasting techniques by providing a scalable framework for areas vulnerable to sea level rise.

How to cite this article:

H. A. Mohammed, M. F. Marsani, M. S. M. Kasihmuddin, F. Z. C. Rose and B. Badyalina., "A HYBRID ARIMA-DISCRETE WAVELET TRANSFORM MODEL FOR ENHANCED COASTAL SEA LEVEL FORECASTING", *BAREKENG: J. Math. & App.,* vol. 20, no.4, pp. 2783-2798, Dec, 2026.

1. INTRODUCTION

The sea level, a crucial indicator of the coastal environment's condition, exhibits spatiotemporal dynamics. These attributes are influenced by a range of general or localized hydrodynamic factors. Although projected sea-level rise poses a substantial threat to coastal communities, current knowledge does not effectively address coastal adaptation, which requires customized services. Services should describe uncertainties, high-end estimates, accurate storm and flood modeling, shoreline change projections, and current practice-based adaptation solutions. Sea data are used for environmental studies and navigation. Observations alone provide a historical record, but they cannot forecast future events; therefore, sea level observations require predictive models. Predictive technologies help policymakers and researchers anticipate and be prepared for extreme events that may impact coastal areas. Global sea levels have risen at unprecedented rates, resulting in severe physical and socioeconomic impacts on coastal populations [1], [2]. Coastal areas such as Penang are susceptible to compounded flooding hazards, with a 50 cm rise potentially submerging 12% of essential infrastructure [3]. Sea-level rise causes water and soil degradation, loss of life and property, infrastructure damage, and depletion of agricultural resources [4]. Therefore, evaluating sea level rise and accurately forecasting future sea level variations is essential for the sustainable growth of coastal communities.

Researchers are using hybrid models based on the Autoregressive Integrated Moving Average (ARIMA) framework to address challenges in forecasting sea levels. For a long time, ARIMA has been a basic tool for various time-series analyses and modeling tasks. Recent developments, however, have focused on overcoming its limitations by combining it with signal processing methods. The ARIMA model, on the other hand, requires data to be stationary, and its predictions are usually more accurate for short- to medium-term forecasts [5]. The model does a good job of capturing seasonality and short-term changes in the data, but its accuracy tends to decline for long-term forecasting [6]. The complex nature of sea level forecasting indicates that a standalone ARIMA approach is rarely adequate for comprehensive coastal resilience planning [7]. Environmental time series, such as coastal sea levels, exhibit complex trends and seasonal patterns that contradict stationary behavior, making standalone ARIMA models inappropriate for capturing real-world non-linear dynamics [8].

Hybrid models are developed to address the limitations of a single algorithm in sea level and hydrological forecasting. Recent hybrid models include ARIMA and Long Short-Term Memory (LSTM) [10], and ARIMA and genetic algorithm BP (GABP) [11]. These models use ARIMA to identify linear, stationary trends in the data, then pass the residuals to a machine learning or deep learning model to capture non-linear patterns.

The Discrete Wavelet Transform (DWT) has become very popular for improving prediction performance [12]. This method works in many different areas of hydrology. For example, [13] used a combination of DWT and artificial intelligence (AI) to predict changes in groundwater levels. In [14] the DWT was integrated as a data preprocessing approach with AI forecasting models to anticipate ocean wave behavior. However, these models struggle to manage the non-stationary, chaotic nature of externally driven sea levels. These hybrid models have improved accuracy in individual case studies, but there is a lack of a comprehensive study analyzing their robustness across climate and regional variability, making data non-linear and adaptable.

The ARIMA-DWT hybrid model is better than the mentioned current models in some ways. For instance, ARIMA-SVR only models the final error in the data, whereas DWT mathematically decomposes the raw data into high- and low-frequency components [15]. This allows ARIMA to accurately predict the sea level independently. The ARIMA-LSTM model requires extensive high-frequency datasets for effective weight training; however, a monthly sea level dataset often contains only a limited number of data points, leading to overfitting [16]. The ARIMA-GABP requires substantial computing power, time, and trial-and-error with hyperparameters [11]. However, ARIMA-DWT is theoretically limited, which makes it very stable and accurate for short, monthly time-series datasets.

The authors in [17] used machine learning and deep learning approaches to estimate sea level variance along the West Peninsular Malaysia coastline. The ARIMA, Support Vector Regression (SVR), and LSTM neural network models were trained using four scenarios with various variables, including sea surface temperature, salinity, density, and sea level data. The optimal model for predicting urban land surface temperature in the Jakarta area was identified by [18] as a crucial step in comprehending and addressing the effects of climate change. The authors in [19] claimed that integrating wavelet decomposition with SVM significantly enhances the precision of sea-level forecasting using essential variables.

A novel forecasting method using DWT with ARIMA to separate time series datasets into linear and nonlinear components was proposed by [20]. This model uses DWT to partition the time series in-sample training dataset into detailed and approximate (A) portions. The suggested method tactically uses DWT, ARIMA, and Artificial Neural Network (ANN) strengths to increase forecasting accuracy. However, ANN components, especially deep networks, can overfit to tiny or noisy datasets, and without regularization, the model may perform poorly on unseen data. [21] claimed that LSTM had challenges in forecasting sea level influenced by various interacting causes. DWT exhibits edge effects that lead to non-adaptive decomposition, complicating the interpretation and comparison of raw sea level time series and increasing computational complexity. In [22], the authors analyze climatic changes and compare machine learning methods for climate prediction. They showed that support vector regression (SVR) and random forest (RF) outperformed other machine learning models, especially for temperature-related variables. Climate change changes the nature of floods, making predictions more difficult. KNN and SVM models were extensively employed for weekly water level forecasting in [23], while ensemble empirical mode decomposition (EEMD) was used to decompose the dataset into intrinsic mode functions (IMFs) to reduce complexity and remove periodicity. The EEMD-LSSVM-PSO model significantly improves forecasting accuracy compared to standard single models. A novel ICML model was introduced in [24] by integrating an SVM and an ANFIS model to enhance forecasting accuracy through weight optimization, resulting in improved predictive performance. In [25], the authors introduce an ELSSVM model that incorporates an additional bias-error control term. This enhancement aims to better manage noise and variations in daily water-level data.

The authors in [26] integrated ANN, ARIMA, and WT to accurately predict seasonal rainfall. [27] presented a W-EEMD-ARIMA hybrid model for drought forecasting. They incorporated Ensemble Empirical Mode Decomposition (EEMD) and wavelet transformation. [28] proposed a hybrid STL-ARIMA-LSTM model for forecasting heatwaves with seasonal adjustment. [29] demonstrated the efficacy of traditional and deep learning time series methods, including SARIMA, LSTM neural networks, and gated recurrent units (GRU), in estimating present and global sea level rise. They concluded that LSTM possesses a greater capacity to reliably forecast future sea levels. [30] proposed a hybrid ARIMA-ANN model to enhance predictive accuracy. The initial dataset is processed using ARIMA, and its residuals are utilized for the subsequent analysis with ANN.

Although numerous hybrid models have been developed to improve performance, few have been applied to sea level prediction. To overcome those limitations, numerous hybrid models have been developed; however, integrating all these methods increases complexity, requires extreme-event forecasting, and increases computational time [31]. The study proposes a hybrid ARIMA-DWT model that offers reduced complexity and enhanced computational speed, thereby improving forecasting performance. ARIMA-based models for linear trends and wavelet-based techniques for noise reduction and feature extraction have been studied separately. Few works have successfully merged ARIMA with DWT to leverage both approaches' strengths: ARIMA for linear forecasting and DWT for nonlinearities and noise. Our proposed hybrid model using these methods may increase forecasting accuracy, especially for complicated, non-stationary datasets.

This research presents a hybrid model that integrates ARIMA and DWT to address challenges such as non-stationarity, complex seasonality, and external influences in sea level prediction. The proposed hybrid model provides an optimal algorithm for addressing challenges in sea level forecasting. The paper is structured as follows: Section 2 presents the methodology and data used in the study; Section 3 outlines the statistical analysis; Section 4 presents the results and discussion; and Section 5 concludes the work.

2. RESEARCH METHODS

This study used a hybrid methodology, specifically the ARIMA-DWT, to estimate sea level in Penang. We have gathered daily sea-level data from Penang tidal stations of the National Oceanography Center spanning [32] years from 1984 to 2018, with a total population size of 12,784. The outliers have been removed, and normalization has been implemented exclusively using Z-score standardization on the DWT components. Although ARIMA models effectively capture linear trends [33], they fail to account for non-linear accelerations in sea level data characterized by high-frequency noise, as demonstrated in Fig. 2. The existing DWT model treats all components as uniformly addressed, neglecting that D1-D2 (sub-daily) often includes instrument noise, while D3-D5 (seasonal) has consistent tidal harmonics. So, DWT was utilized instead of Empirical Mode Decomposition (EMD) and Variational Mode Decomposition (VMD since Eq. (1) makes it easier to understand the scale-aligned components, Eq. (2) is deterministic and can be reproduced, and Eq. (3) is faster for operational forecasting.

For the signal decomposition, the Daubechies-4 (*db4*) is used for optimal smoothness and localization with 5 levels $(D1 - D5)$. The db4 wavelets, with four vanishing moments, effectively mitigate short-term noise while preserving climatic trends. For the time series at the initial scale factor be $s(t)$, DWT calculates the scale factor with a low-pass filter as follows;

$$s(t) = F(t), \tag{1}$$

$$s_{i+1}(t) = \sum_{n=-\infty}^{\infty} h(n) s_i\left(t + 2^i n\right), \tag{2}$$

where, $F(t)$ is the original time series and t is the time step of the series.

The detail factor can be calculated from the scale factor as follows:

$$d_{i+1}(t) = s_i(t) - s_{i+1}(t), \tag{3}$$

where $s_i(t)$ and $s_{i+1}(t)$ are the scale factors at decomposition levels i and $i + 1$ respectively. These signals are outputs of the low-frequency trends of the series.

The wavelet decomposition of the original series at resolution q is represented as $\{d_1, d_2, \dots, d_q, c_q\}$, where d_i (i = 1, 2, ⋯, q) denotes a detailed signal at level i, and c_q signifies an approximation signal, given as follows:

$$F(t) = s_q(t) + \sum_{i=1}^{q} d_i(t), \tag{4}$$

where $s_q(t)$ is the final residual approximation signal at the maximum specified decomposition level q.

We use $h(n) = \left(\frac{1}{2}, \frac{1}{2}\right)$ as the low-pass filter, so $s_i(t)$ and $d_i(t)$ can be written as

$$s_{i+1}(t) = \frac{1}{2}\left(s_i\left(t - 2^i\right) + s_i(t)\right, \tag{5}$$

$$d_{i+1}(t) = s_i(t) - s_{i+1}(t). \tag{6}$$

2.1 Stationarity Testing

It is almost impossible for a time series to be stable for the ARMA model. Differencing, thus, makes series non-stationary. Stationarity is a crucial prerequisite for ARIMA models. If the data is non-stationary, differentiation of the certain order is required to achieve stationarity [34], [35]. Stationarity can be assessed by the Augmented Dickey-Fuller (ADF) test. ADF test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are used to determine stationarity. The ADF test for the time series X_t can be expressed as follows;

$$\Delta X_t = \sigma X_{t-1} + \sum_{i=1}^{n} \beta_i X_{t-i} + \epsilon_t, \tag{7}$$

$$\Delta X_t = \alpha + \beta t + \sigma X_{t-1} + \sum_{i=1}^{n} \beta_i X_{t-1} + \epsilon_t. \tag{8}$$

Here, ϵ_t has been assumed to have a mean of 0 and a standard deviation equivalent to the random error value of σ, which is considered constant for every t. The original and alternative hypotheses for the KPSS test are presented as follows. The null hypothesis H_0: $\sigma = 0$ indicates that the original series possesses a unit root and is non-stationary. Alternative hypothesis If H_1: $\sigma < 0$, the original series lacks a unit root and is stationary.

2.2 The ARIMA Modeling

Subsequent to the confirmation of stationarity, the sub-signals are modeled. The equation for the ARIMA model is as follows:

$$X_t = \varphi_0 + \sum_{i=1}^{p} \varphi_1 \ X_{t-i} + \varepsilon_0 - \left(\sum_{j=1}^{q} \theta_j \ \varepsilon_{t-i} \right), \tag{9}$$

where ε_t is stationary white noise with zero mean and a constant variance of σ^2, φ_i is the autoregressive (AR) coefficient, and θ_j represents the moving average (MA) coefficient.

Least Squares estimation is employed to ascertain model parameters. The significance of parameters is assessed using the t-test, and the model is evaluated using the minimum Akaike Information Criterion (AIC) to determine the optimal fit. The diagnostic evaluation of the model generated during the estimating phase is essential. The Ljung-Box Test is frequently used to assess whether the residuals satisfy the assumptions of the error model, including the assumption of white noise. White noise possesses a mean of 0 and a variance of 1, conforming to conventional model evaluation criteria [34] and [36].

2.3 *Evaluation Criteria*

To assess the efficacy of the suggested hybrid model, three standard statistical indicators, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), were used. Each metric provides a different view of how well the model is working: RMSE significantly penalizes larger forecasting errors (crucial for predicting extreme sea level events or storm surges), MAE quantifies the average magnitude of errors irrespective of their direction, and MAPE conveys accuracy as a relative percentage for clear interpretation by coastal planners. The formulas for these statistical indicators are provided as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}, \tag{10}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right|, \tag{11}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}|, \tag{12}$$

where $\widehat{y_i}$ denotes the model's anticipated sea level, y_i represents the actual sea level readings, and n indicates the sample size to be forecasted.

Fig. 1 below shows the methodology for forecasting sea level, including preprocessing, discrete wavelet decomposition, stationarity testing, ARIMA modeling, diagnostic checking, signal reconstruction, and final model evaluation.

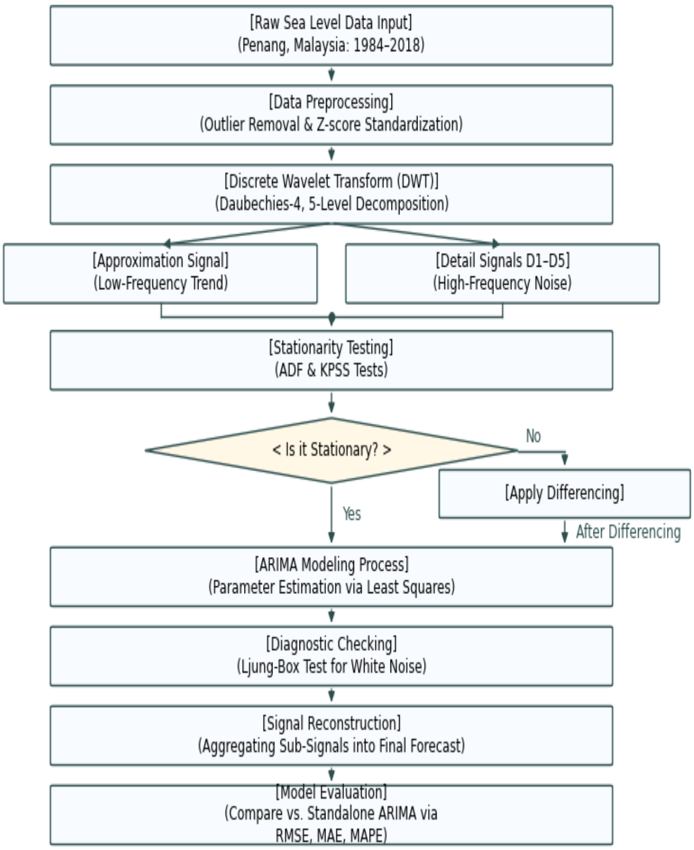


Figure 1. Flowchart of the DWT-ARIMA Hybrid Forecasting Methodology for Sea Level Data**.**

3. RESULTS AND DISCUSSION

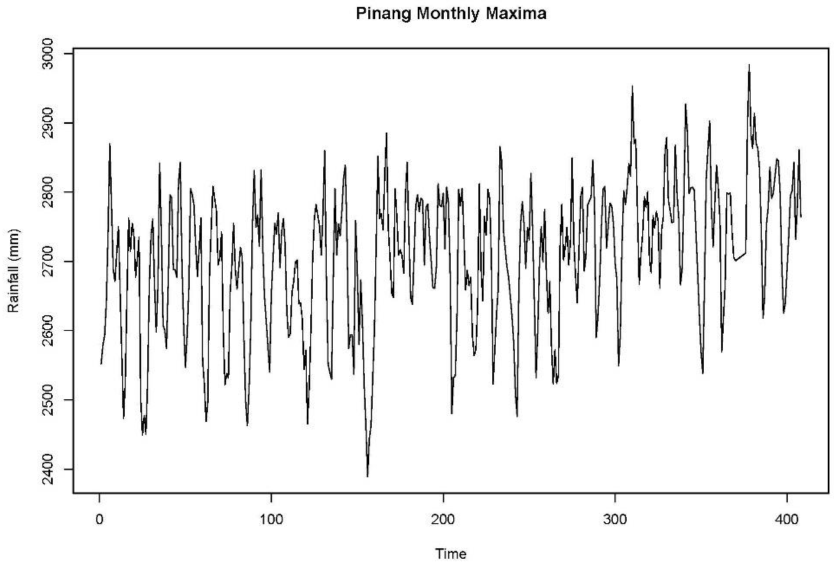
In this study, we utilized time series data, namely sea level data from Penang, Malaysia, obtained from January 1984 to 2018. We gathered sea level data for Penang and categorized it by month, quarter, and year. Our model is the first application of Penang's mixed sea level data, which includes complex and monsoon influences. The data are adjusted and differentiated between stationary and non-stationary using the DWT model. Penang serves as an optimal site for sea level research owing to its extensive data availability, exposure to various marine and climatic influences, significant susceptibility, and pressing policy considerations. Investigating sea level variations in this region can enhance global climate science and facilitate local adaptation efforts. Our study demonstrates the efficacy of the hybrid model for predicting sea level changes over short-, medium-, and long-term time scales.

Table 1 below presents descriptive statistics for sea level; it summarizes sea level data (408 values) for monthly, quarterly, and annual data. Where mean denotes the average value, *Std* is the standard deviation (σ), *Skew* denotes skewness, *Kurt* is the kurtosis, *Min* is the minimum value, *Q1* and *Q3* are the first quartile and third quartile respectively.

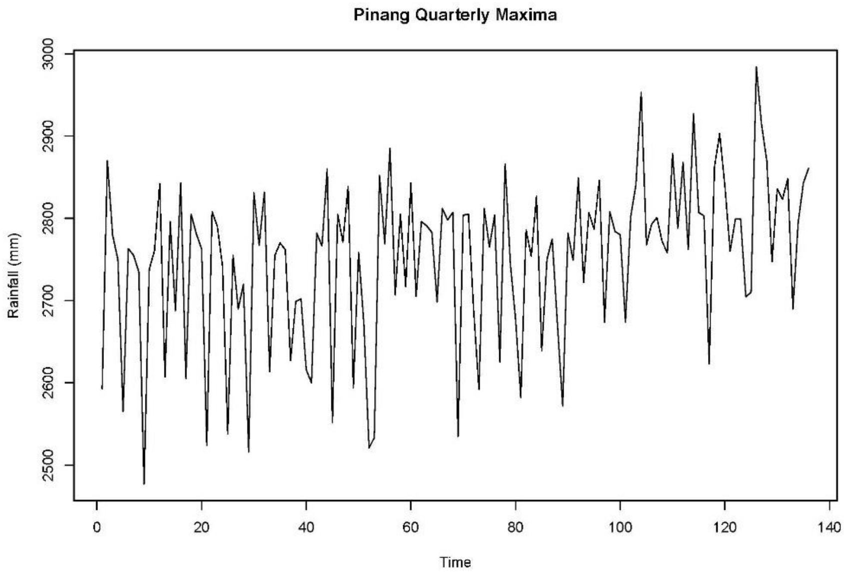
Table 1. Descriptive Statistics of the Sea Level

	Num	**Mean**	**Std**	**Skew**	**Kurt**	**Min**	**Q1**	**Mid**	**Q3**	**Max**
Monthly (1984-2018)	408	2707.54	102.11	-0.49	-0.02	2389	2653.75	2720	2780	2984
Quarterly (1984-2018)	136	2755.75	93.93	-0.72	0.54	2477	2720	2769.5	2807	2984
Annual (1984-2018)	34	2833.15	57.45	0.39	0.41	2702	2802	2835.5	2860.75	2984

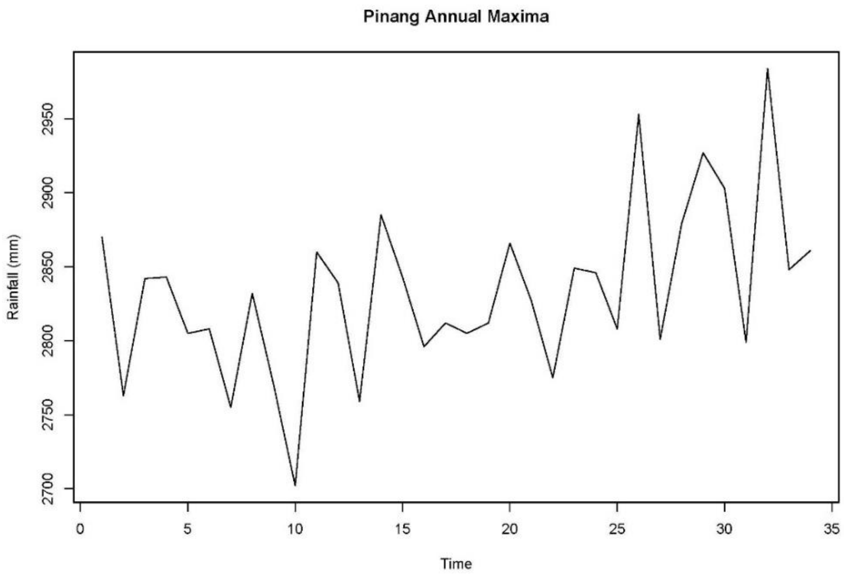
As granularity decreases from monthly to annual, the standard deviation decreases by 43% (102.11 to 57.45 mm), and the distribution becomes more symmetric (skewness: -0.49 to +0.39). Monthly data exhibits negative skewness (-0.49) and a platykurtic distribution (kurtosis = -0.02). The table indicates that the high-frequency components reflect monthly fluctuation, whereas the approximation corresponds with annual trend stability.



(a)



(b)



(c)

Figure 2. Time Series Plots of Original Sea Level Data for Penang, Malaysia (1984–2018), Aggregated (a) Monthly, (b) Quarterly, and (c) Annually.

Fig. 2 illustrates that the monthly series (top) displays significant high-frequency variations and increased variability, especially evident between 1995 and 2000. This five-year window is emphasized because short-term sea level variability typically manifests over interannual to multi-year cycles (e.g., Niño-Southern Oscillation (ENSO) events or regional monsoon effects), making a 5-year span sufficient to capture pronounced fluctuations without being masked by longer-term rising trends. Such variability during this period suggests possible external drivers such as climate oscillations or storm surges.

It is necessary to check that the data is stationary before moving on to signal decomposition and ARIMA model estimation, as ARIMA frameworks require a constant mean and variance over time. To ascertain this, both the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests were employed. It is critical to note that these tests operate under inverse null hypotheses: the ADF test assumes a unit root (non-stationary) as its null hypothesis, whereas the KPSS test assumes the series is inherently stationary.

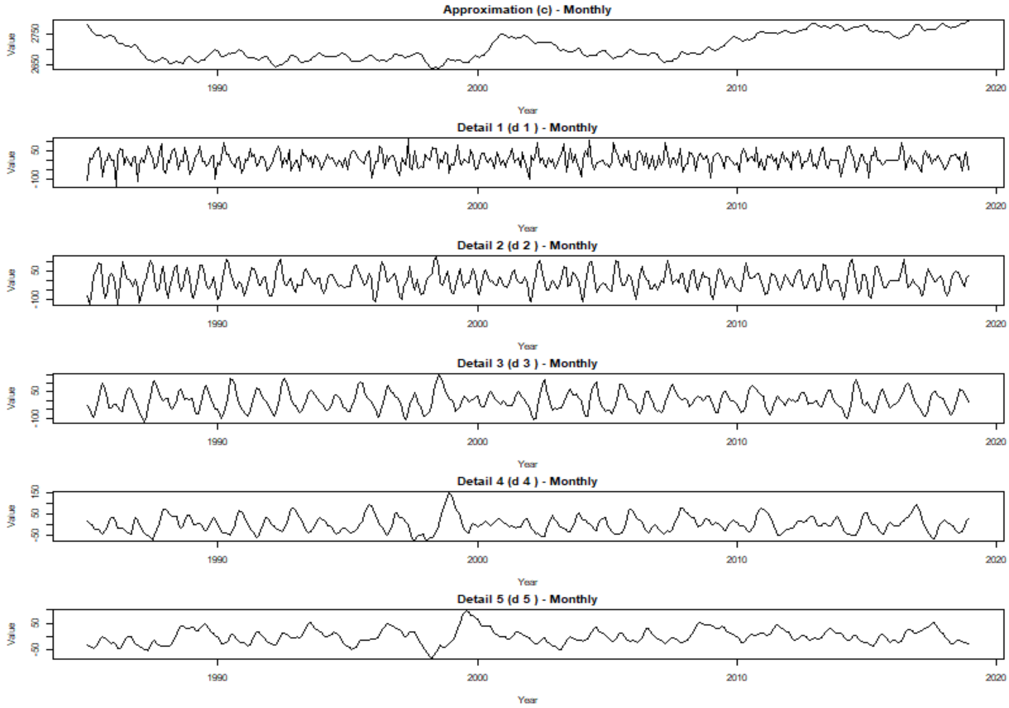
Table 2 shows that the ADF test produced inconsistencies for the original sequences because of deterministic trends, whereas the KPSS test consistently rejected the hypothesis of stationarity at all temporal granularities (KPSS $p < 0.01$). The contradictory outcomes from the ADF and KPSS tests clearly show that the original sea level series is trend-non-stationary. The direct modeling of non-stationary data results in erroneous forecasts, hence mathematically substantiating the need for first-order differencing and Discrete Wavelet Transform (DWT) decomposition. Following these changes, both tests verified that the resultant sub-signals attained stringent stationarity (KPSS $p > 0.05$), thereby affirming their suitability as ideal inputs for ARIMA parameter estimation.

Table 2. Statistics of Stationarity Test

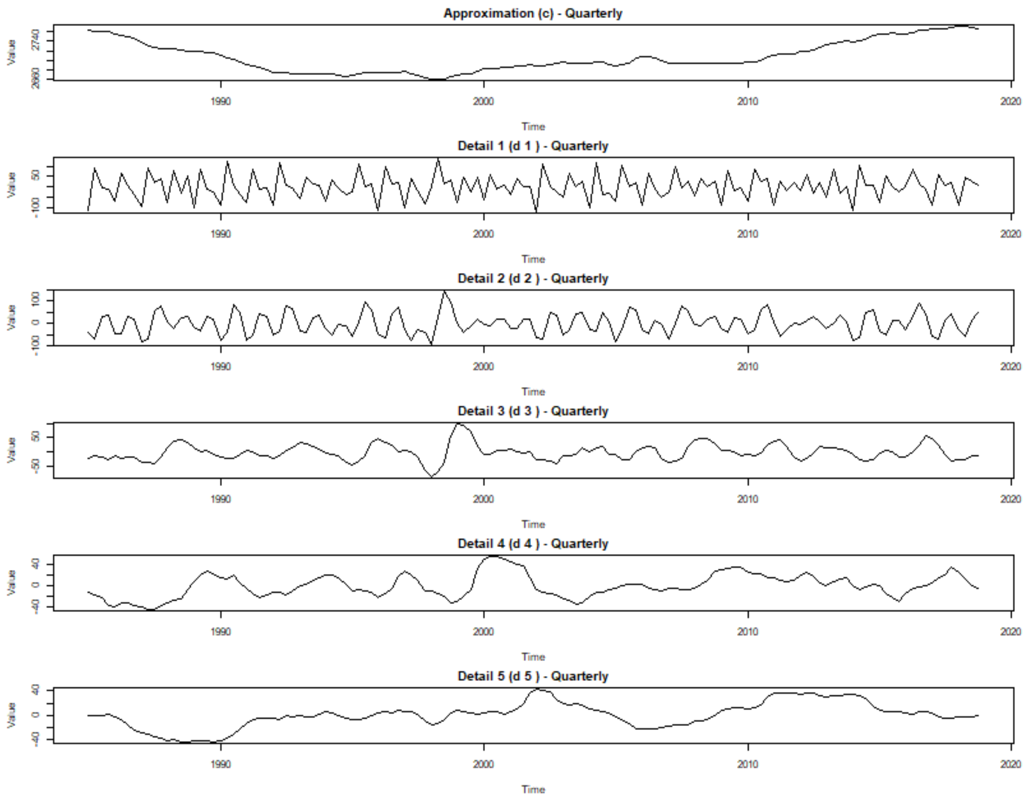
Test	Statistic	P_Value	Conclusion	Type
ADF (Original)	-3.84652	0.002466	Stationary	Monthly
KPSS (Original)	1.78742	0.01	Non-Stationary	Monthly
ADF (Differenced)	-7.43394	0.0	Stationary	Monthly
KPSS (Differenced)	0.025013	0.1	Stationary	Monthly
ADF (Original)	-1.12397	0.705419	Non-Stationary	Quarterly
KPSS (Original)	1.72446	0.01	Non-Stationary	Quarterly
ADF (Differenced)	- 4.88586	0.000037	Stationary	Quarterly
KPSS (Differenced)	0.06582	0.1	Stationary	Quarterly
ADF (Original)	-3.0907	0.027249	Non-Stationary	Annual
KPSS (Original)	0.735181	0.010347	Non-Stationary	Annual
ADF (Differenced)	-7.35405	0.0	Non-Stationary	Annual
KPSS (Differenced)	0.229138	0.1	Stationary	Annual

To address the non-stationarity and isolate the underlying physical drivers, the Discrete Wavelet Transform (DWT) breaks the time series frequency into components (approximate level C and detailed levels D1–D5). Due to its multiple levels of detail, DWT places a heavy load on computers, especially at high resolutions. ARIMA models each decomposed element (low-frequency trend and high-frequency noise) separately, requiring several model estimations. This boosts memory usage, parameter optimization, and time complexity.

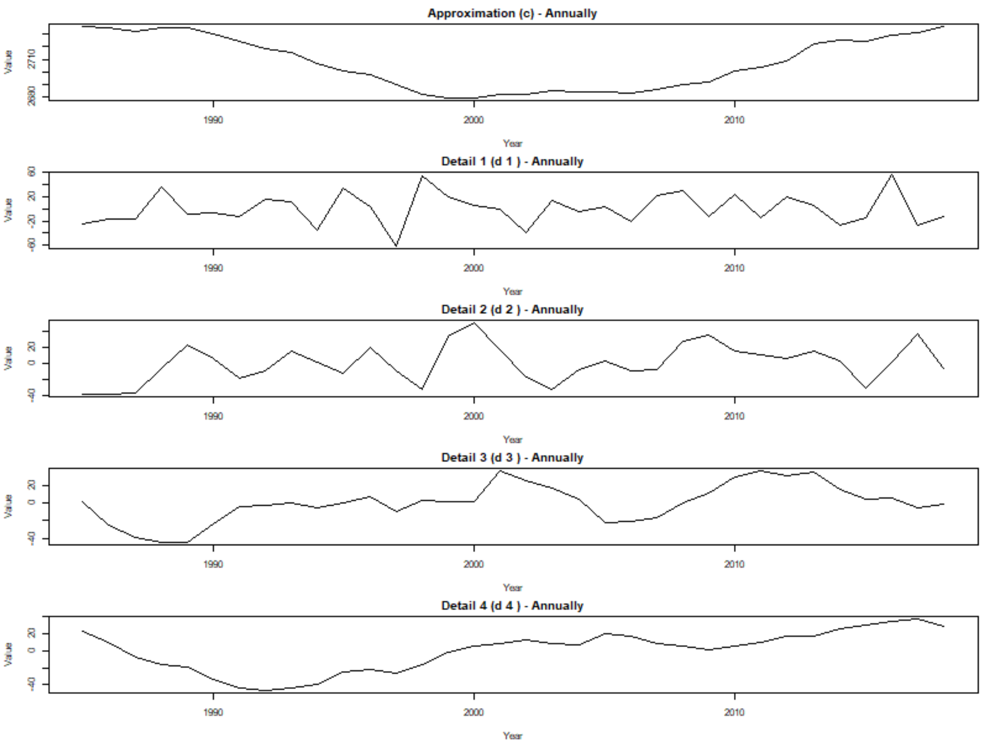
Fig. 3 (a), (b), and (c) illustrate the decomposition of the monthly, quarterly, and annual time series, respectively. Signal analysis theory posits that the approximate signal represents the low-frequency component of the original series, typically denoting its long-term trend, whereas the fine signal embodies the high-frequency component, usually signifying short-term fluctuations.


Approximation (c) - Monthly
Detail 1 (d 1) - Monthly
Detail 2 (d 2) - Monthly
Detail 3 (d 3) - Monthly
Detail 4 (d 4) - Monthly
Detail 5 (d 5) - Monthly
Value
Year
1990
2000
2010
2020

(a)


Approximation (c) - Quarterly
Detail 1 (d 1) - Quarterly
Detail 2 (d 2) - Quarterly
Detail 3 (d 3) - Quarterly
Detail 4 (d 4) - Quarterly
Detail 5 (d 5) - Quarterly
Value
Time
1990
2000
2010
2020

(b)



(c)

Figure 3. Decomposed sequences by the DWT method for (a) monthly, (b) quarterly, and (c) annual data.

The DWT decomposes the time series into frequency components (approximate and detailed levels such as D1–D5), increasing preparation time and computational complexity. Due to its multiple levels of detail, DWT places a heavy load on computers, especially at high resolutions. ARIMA models each decomposed element (low-frequency trend and high-frequency noise) separately, requiring several model estimations. This boosts memory usage, parameter optimization, and time complexity.

When the dataset is small or noisy, hybrid models such as DWT-ARIMA-LSTM may overfit, especially at high decomposition levels [37]. Overfitting is especially problematic in AI or deep learning-modeled components; however, this work uses ARIMA. Deconstructed components may show random changes instead of the signal. Modeling these as distinct ARIMA systems risks uncovering non-generalizable patterns. A study on hybrids found that without regularization, these models overfit the training data and make poor out-of-sample predictions.

Table 3 presents simulated wavelet analysis statistics for sea level data. The table indicates that level C exhibits a high correlation of 0.995 and a variance proportion of 0.98, which shows the long-term trend. Conversely, the high-frequency detail levels (D1 to D5) show minimal variance proportions between 0.005 and 0.025 and low correlations ranging from 0.05 to 0.18. According to standard statistical thresholds [38], correlations in this range are classified as negligible to weak, confirming that these detail signals are largely stripped of the primary secular trend and primarily contain high-frequency noise. Furthermore, it is notable that all detail levels exhibit positive skewness (e.g., Level D skewness = 0.448) alongside high kurtosis (> 3). In the physical context of coastal hydrodynamics, this positive skewness indicates that the short-term anomalies are characterized by extreme positive spikes rather than symmetric fluctuations. This makes sense statistically since Penang has transitory, high-impact coastal phenomena like storm surges or monsoonal wind sets that raise water levels for a short time before going away.

C alone accounts for 98% of the total variance and exhibits a 0.995 correlation with the original signal, thereby elucidating the predominant trend in sea level data. The application of ARIMA to model C effectively addresses both trend and seasonality, rendering it the principal factor in enhancing accuracy. D1–D5 encapsulate residual structures, collectively representing approximately 2% of the total variance, yet they contribute to minor short-term fluctuations. Their elevated kurtosis and low correlation suggest they are noisy and volatile, albeit not entirely random; they exhibit micro-patterns that, when modeled, help minimize prediction error, particularly for outliers and abrupt short-term shifts. The combination of C and D1–D5 facilitates multi-scale learning. Modeling C with an ARIMA captures the overarching structure, while separately modeling D1–D5 accommodates localized behaviors without introducing noise into the trend model.

Table 3. Signal Sequence Decomposed by WT.

Level	Pearson	Proportion Variance	Skewness	Kurtosis
C	0.995	0.980	0.300	0.500
D	0.200	0.020	0.448	3.525
D1	0.050	0.005	0.800	3.200
D2	0.070	0.008	0.900	4.100
D3	0.100	0.010	1.200	4.800
D4	0.150	0.015	1.600	6.000
D5	0.180	0.025	1.800	7.000

To construct the hybrid forecasting framework, an individual ARIMA model was fitted to the Approximation (C) and Detail (D1–D5) sub-signals. The optimal autoregressive (p), differencing (d), and moving average (q) orders for each component were determined by minimizing the Akaike Information Criterion (AIC). Following model identification, the coefficients for each parameter were calculated using Least Squares estimation. The fitted ARIMA models, alongside their respective parameter estimates and statistical significance (p-values), are detailed in Table 4. All selected parameters demonstrated statistical significance ($p < 0.05$), ensuring the mathematical robustness of the reconstructed forecast.

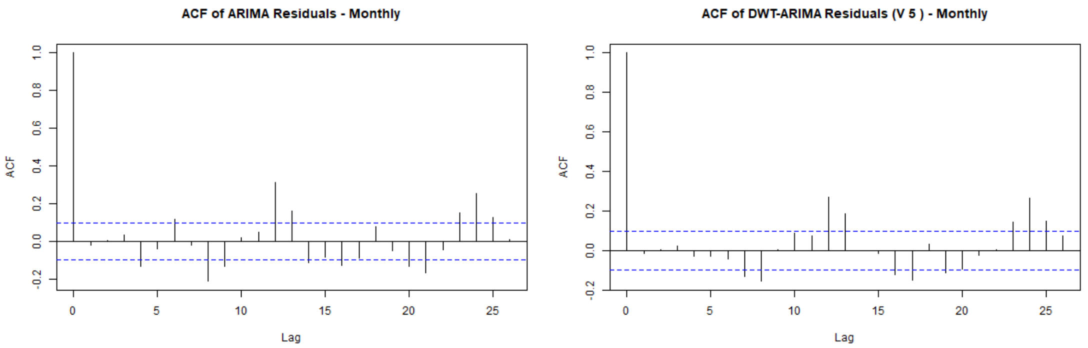
Table 4. Optimal ARIMA (p, d, q) Orders and Parameter Estimates

Decomposed Signal	ARIMA Order (p, d, q)	Parameter	Coefficient Estimate	Standard Error	p-value
Approximation (C)	(0, 1, 0)	Variance	24036.8559	7029.4042	< 0.001
Detail (D5)	(0, 0,0)	Constant	-34.1255	35.7173	0.339
		Variance	19874.6039	6408.7688	0.002
Detail (D4)	(0, 0, 0)	Constant	-15.3414	20.7771	0.460
		Variance	12230.1029	2795.8918	< 0.001
Detail (D3)	(2, 0, 2)	Constant	9.0344	14.0345	0.520
		AR (1)	-0.9946	0.0188	< 0.001
		AR (2)	-0.9932	0.0082	< 0.001
		MA (1)	0.8274	0.0789	< 0.001
		MA (2)	0.9225	0.1003	< 0.001
		Variance	9176.7599	2073.5719	< 0.001
Detail (D2)	(3, 1, 3)	AR (1)	-1.4305	0.1505	< 0.001
		AR (2)	-1.3974	0.1489	< 0.001
		AR (3)	-0.4225	0.1420	0.003
		MA (1)	0.1556	0.1386	0.262
		MA (2)	-0.2830	0.1049	0.007
		MA (3)	-0.6634	0.0956	< 0.001
		Variance	2596.4145	358.7711	< 0.001
Detail (D1)	(2, 0, 2)	Constant	0.6816	2.3341	0.770
		AR (1)	-0.9401	0.0553	< 0.001
		AR (2)	-0.9130	0.0519	< 0.001
		MA (1)	0.7437	0.0885	< 0.001
		MA (2)	0.7656	0.0826	< 0.001
		Variance	1376.4415	143.7523	< 0.001

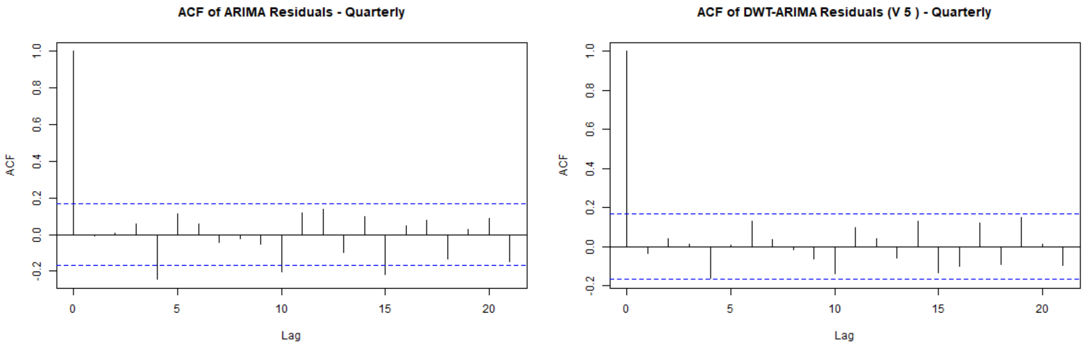
The parameter estimation, as shown in Table 4, states that the decomposed sea level signals have different underlying characteristics. The optimal ARIMA (p, q, d) orders varied significantly across the sub-signals, confirming that the DWT successfully isolated different physical processes. For instance, the Approximation (C) signal, which shows the overall climate trend, was best approximated as an ARIMA (0, 1, 0) process. This shows that there is a random walk structure, a simple way to model the long-term rise in sea levels that doesn't remain constant over time.

the other hand, higher-order models were needed to accurately track short-term coastal noise and tide changes using high-frequency signal detail. For instance, D1 and D3 were best fitted with ARIMA (2, 0, 2) models, while D2 required an ARIMA (3, 1, 3) structure to account for its unique oscillatory behavior. Furthermore, almost all of the autoregressive (AR) and moving average (MA) coefficients yielded highly significant p-values (< 0.001). This confirms that the selected parameters are statistically significant and that the hybrid framework relies on robust mathematical relationships rather than false noise.

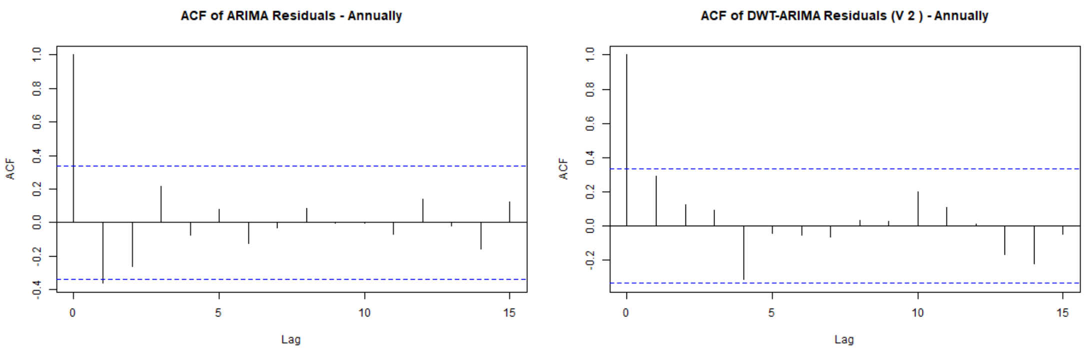
To accurately characterize the models' anticipated performance, Fig. 4 illustrates the Autocorrelation Function (ACF) plots of the residuals for the monthly, quarterly, and annual sea level forecast series. ACF plots are used to assess whether the residuals from forecasting models exhibit characteristics of white noise, a fundamental requirement for the validity of ARIMA-based models. The X-axis denotes time lags, while the Y-axis shows the correlation strength between residuals at various lags; the blue-shaded area indicates the 95% confidence interval. ACF values outside this range indicate statistically significant autocorrelation.



(a)



(b)



(c)

Figure 4. Residual Series Autocorrelation Test: (a) Monthly Series, (b) Quarterly Series, and (c) Annual Series.

The forecasting performance of ARIMA and ARIMA-DWT for sea level data is compared in Table 5. ARIMA-DWT outperforms in-sample (MAE=43.22 vs. 54.02) and out-of-sample (MAE=68.40 vs. 85.50). The enhancement is approximately 20% across all metrics (RMSE, MAE, and MAPE). This means the 20% improvement in accuracy is achieved by reducing RMSE relative to the standard ARIMA, thereby eliminating noise (D1–D2).

Table 5. Predictive Accuracy Statistics.

	MAE in-sample	MAPE in-sample	RMSE in-sample	MAE out-sample	MAPE out-sample	RMSE out-sample
ARIMA	54.02	2.01	68.13	85.50	3.06	102.21
ARIMA-DWT	43.22	1.61	54.51	68.40	2.45	81.76

Some hybrid models, such as the wavelet-ARIMA-LSTM model, may be insufficient for sea level data because they do not exhibit the high-frequency noise and intricate nonlinearity that wavelets and LSTMs are intended to address [39]. Less complex models can account for the majority of variability in sea level data while minimizing the risk of overfitting. The supplementary components fail to deliver significant enhancements and may impair performance by adding superfluous complexity. DWT separates long-term trends from short-term fluctuations (D1-D5), enabling ARIMA to predict each component independently. This is especially beneficial for non-stationary environmental data such as sea levels.

The 20% decrease in error metrics indicates that the hybrid method captures both linear trends (ARIMA) and residual non-linearities (DWT) more effectively than ARIMA alone. Unlike other data that exhibit significant volatility and nonlinearity, sea level data reveals distinct trends and seasonality, making ARIMA-DWT a more appropriate model than complex alternatives such as LSTM. Comparing the predictive performance of the original ARIMA model with that of the proposed ARIMA-DWT hybrid model reveals that incorporating data decomposition techniques significantly enhances the model's predictive capabilities, as measured by MAE, MAPE, and RMSE. The hybrid prediction model using the DWT approach marginally outperformed those using the ARIMA method across most samples, suggesting that the DWT hybrid method more effectively captures the data's varying frequency properties. In conjunction with Fig. 3 along with the pertinent information in Table 5, this paper concludes that while ARIMA exhibits more pronounced edge effects, the DWT method offers superior denoising capabilities, and the approximate signal obtained through the DWT method more effectively represents the long-term trend of the sea level data series. The ARIMA-DWT model decreases forecasting errors by around 20% (MAE, RMSE). This indicates a need for more accurate forecasting of short-term water-level surges, particularly during monsoon seasons or catastrophic weather events. Additionally, reducing false alarms and missed notifications improves emergency readiness and fosters public trust.

Finally, it is important to explicitly acknowledge the limitations of this study to provide a clear context for these findings. The hybrid methodology encompasses several phases: wavelet decomposition, individual component modeling via ARIMA, and signal reconstruction, all of which elevate computational expenses. This may impede the model's scalability for real-time or high-frequency data applications unless optimal computing procedures are implemented. The model was created and validated utilizing sea level data from Penang, Malaysia a comparatively stable coastal region with extensive long-term data availability. The efficacy in more volatile or data-sparse places (e.g., areas affected by extreme weather events or tectonic movements) remains unexamined. Application to these places may necessitate the inclusion of supplementary meteorological or geophysical data. The efficacy of both the DWT and ARIMA components is contingent upon parameter selections, including the degree of wavelet decomposition and the selection of ARIMA (p, q, d) orders. Insufficient tuning may lead to overfitting or suboptimal generalization. The absence of automated or adaptive parameter tuning in this study constrains reproducibility and wider applicability. Forecasts are presented as point estimates devoid of prediction intervals, so constraining the evaluation of forecast confidence and hindering risk-informed decision-making. This may be resolved in future research by integrating probabilistic forecasting methods.

4. CONCLUSION

The study presented a hybrid ARIMA-DWT model for predicting sea level changes in Penang, Malaysia, by combining discrete wavelet transform (DWT) with the ARIMA model to account for both long-term trends and short-term variability. Compared with the single ARIMA model, the hybrid approach has consistently shown improved predictive performance across all temporal resolutions. The hybrid model decreased out-of-sample MAE from 85.5 to 68.4 and RMSE from 102.21 to 81.76 for monthly data, resulting in an improvement of almost 20%. In the quarterly forecasts, the Mean Absolute Error (MAE) decreased from 65.54 to 50.54, and the Root Mean Square Error (RMSE) diminished from 76.32 to 59.47. In annual projections, MAE fell from 32.16 to 13.54, and RMSE decreased from 41.35 to 16.80, demonstrating the model's resilience across multiple temporal scales. The findings validate that disaggregating sea level data into various frequency components enables ARIMA to model each layer effectively, leading to greater accuracy and reduced error. The low-frequency (C) component, accounting for around 98% of the variance, emerged as the primary contributor to prediction reliability, but high-frequency components (D1–D5) improved responsiveness to short-term volatility.

Author Contributions

Hakami Ali Mohammed: Conceptualization, Methodology, Writing Original Draft, Software, Validation. Muhammad Fadhil Marsani: Data Curation, Resources, Draft Preparation, Supervision. Mohd Shareduwan Mohd Kasihmuddin: Formal Analysis, Validation. Farid Zamani Che Rose: Validation, Writing, Review and Editing. Basri Badyalina: Software, Visualization. All authors discussed the results and contributed to the final manuscript.

Funding Statement

This work was supported by Universiti Sains Malaysia through the School of Mathematical Sciences

Acknowledgment

The authors would like to express their sincere gratitude to the School of Mathematical Sciences, Universiti Sains Malaysia (USM), for the support and resources provided in facilitating the dissemination of research findings by its lecturers and students.

Declarations

The authors declare no competing interests.

Declaration of Generative AI and AI-assisted technologies

Generative AI tools (e.g., ChatGPT) were used solely for language refinement (grammar, spelling, and clarity). The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. The authors reviewed and approved all final text.

REFERENCES

[1] W. Shi, C. Lu, and A. D. Werner, “ASSESSMENT OF THE IMPACT OF SEA-LEVEL RISE ON SEAWATER INTRUSION IN SLOPING CONFINED COASTAL AQUIFERS,” *J. Hydrol.*, vol. 586, p. 124872, Jul. 2020, doi: https://doi.org/10.1016/j.jhydrol.2020.124872.

[2] M. F. Marsani *et al.*, “THE APPLICATION OF SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA) MODEL IN FORECASTING MALAYSIA MEAN SEA LEVEL,” *AIP Conf. Proc.*, vol. 3123, no. 1, p. 070005, Aug. 2024, doi: https://doi.org/10.1063/5.0223836.

[3] H. A. Rahman, “WATER ISSUES IN MALAYSIA,” *Int. J. Acad. Res. Bus. Soc. Sci.*, vol. 11, no. 8, pp. 860–875, Aug. 2021, doi: https://doi.org/10.6007/IJARBSS/v11-i8/10783.

[4] N. Zheng *et al.*, “HOURLY SEA LEVEL HEIGHT FORECAST BASED ON GNSS-IR BY USING ARIMA MODEL,” *Int. J. Remote Sens.*, vol. 43, no. 9, pp. 3387–3411, May 2022, doi: https://doi.org/10.1080/01431161.2022.2091965.

[5] A.J.B. Alhallaf *et al*., "PREDICTING SEA LEVEL VARIATIONS FOR EARLY WARNING USING SARIMA MODEL AND DEEP LEARNING TECHNIQUES IN THE NORTHWEST ARABIAN GULF." *Discover Geosci.,* vol. 3, no. 1, p.158, Oct. 2025, doi: https://doi.org/10.1007/s44288-025-00270-y.

[6] J.J. Selvaraj *et al*., "TIME-SERIES MODELING OF FISHERY LANDINGS IN THE COLOMBIAN PACIFIC OCEAN USING AN ARIMA MODEL." *Regional Studies in Marine Science*., vol. 39, p.101477, Sept. 2020, doi: https://doi.org/10.1016/j.rsma.2020.101477.

[7] N. Alenezi, Nasser, A. Abdalrahman, and A. Mohamad, "PREDICTION OF SEA LEVEL IN THE ARABIAN GULF USING ARTIFICIAL NEURAL NETWORKS." *Journal of Marine Science and Engineering*., vol. 11, no. 11, p.2052, Sept. 2023, doi: https://doi.org/10.3390/jmse11112052.

[8] J. Kaur, K.S. Parmar, and S, Singh, “AUTOREGRESSIVE MODELS IN ENVIRONMENTAL FORECASTING TIME SERIES: A THEORETICAL AND APPLICATION REVIEW,” *Environmental Science and Pollution Research*, vol. 30, no.8, pp.19617-19641, Jan. 2023.doi: https://doi.org/10.1007/s11356-023-25148-9

[9] S. Nie *et al.* "INTEGRATION OF ARIMA AND LSTM MODELS FOR REMAINING USEFUL LIFE PREDICTION OF A WATER HYDRAULIC HIGH-SPEED ON/OFF VALVE." *Applied Sciences*, vol. 12, no. 16, p.8071, Aug. 2022, doi: https://doi.org/10.3390/app12168071.

[10] L. Rubio, and K. Alba, “FORECASTING SELECTED COLOMBIAN SHARES USING A HYBRID ARIMA-SVR MODEL.” *Mathematics*, vol. 10, no. 13, p.2181, Jun. 2022, doi: https://doi.org/10.3390/math10132181.

[11] X. Chen, *et al.* "A HYBRID ARIMA-GABP MODEL FOR PREDICTING SEA SURFACE TEMPERATURE." *Electronics* vol. 11, no. 15, p.2359, Jul. 2022, doi: https://doi.org/10.3390/electronics11152359.

[12] C. Song *et al.*, “APPLICATION OF A NOVEL SIGNAL DECOMPOSITION PREDICTION MODEL IN MINUTE SEA LEVEL PREDICTION,” *Ocean Eng.*, vol. 260, p. 111961, 2022, doi: https://doi.org/10.1016/j.oceaneng.2022.111961.

[13] S. Momeneh and V. Nourani, “FORECASTING OF GROUNDWATER LEVEL FLUCTUATIONS USING A HYBRID OF MULTI-DISCRETE WAVELET TRANSFORMS WITH ARTIFICIAL INTELLIGENCE MODELS,” *Hydrol. Res.*, vol. 53, no. 6, pp. 914–944, Jun. 2022, doi: https://doi.org/10.2166/nh.2022.035.

[14] A. Altunkaynak, A. Çelik, and M. B. Mandev, “DYNAMIC ADAPTIVE WAVELET-BASED FUZZY FRAMEWORK FOR EXTENDED SIGNIFICANT WAVE HEIGHT FORECASTING,” *Ocean Eng.*, vol. 295, p. 116814, Mar. 2024, doi: https://doi.org/10.1016/j.oceaneng.2024.116814.

[15] Y. Ni, L. Shuang, and G. Ping, "DISCRETE WAVELET INTEGRATED CONVOLUTIONAL RESIDUAL NETWORK FOR BEARING FAULT DIAGNOSIS UNDER NOISE AND VARIABLE OPERATING CONDITIONS." *Scientific Reports*, vol. 15, no. 1, p. 16185, May 2025, doi: https://doi.org/10.1038/s41598-025-99346-5.

[16] Z. Li, H. Jing, and S. Yuping, "ON THE FORECASTING OF HIGH-FREQUENCY FINANCIAL TIME SERIES BASED ON ARIMA MODEL IMPROVED BY DEEP LEARNING." *Journal of Forecasting*, vol. 39, no. 7, pp. 1081-1097, Feb. 2020, doi: https://doi.org/10.1002/for.2677.

[17] A.-L. Balogun and N. Adebisi, “SEA LEVEL PREDICTION USING ARIMA, SVR AND LSTM NEURAL NETWORK: ASSESSING THE IMPACT OF ENSEMBLE OCEAN–ATMOSPHERIC PROCESSES ON MODELS’ ACCURACY,” *Geomatics, Nat. Hazards Risk*, vol. 12, no. 1, pp. 653–674, Mar. 2021, doi: https://doi.org/10.1080/19475705.2021.1887372.

[18] B. A. P. Suhendi *et al.*, “COMPARISON OF ARIMA AND MACHINE LEARNING METHODS FOR PREDICTING URBAN LAND SURFACE TEMPERATURE IN JAKARTA,” *J. Comput. Sci. Informatics Eng.*, vol. 8, no. 2, Dec. 2024, doi: https://doi.org/10.29303/jcosine.v8i2.579.

[19] S. Sithara, *et al*., "SEA LEVEL PREDICTION USING CLIMATIC VARIABLES: A COMPARATIVE STUDY OF SVM AND HYBRID WAVELET SVM APPROACHES." *Acta Geophysica*., vol. 68, no.6, pp. 1779-1790, Sept. 2020, doi: https://doi.org/10.1007/s11600-020-00484-3.

[20] W. Zhang, Z. Lin, and X. Liu, “SHORT-TERM OFFSHORE WIND POWER FORECASTING—A HYBRID MODEL BASED ON DISCRETE WAVELET TRANSFORM (DWT), SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA), AND DEEP-LEARNING-BASED LONG SHORT-TERM MEMORY (LSTM),” *Renew. Energy*, vol. 185, pp. 611–628, Mar. 2022, doi: https://doi.org/10.1016/j.renene.2021.12.100.

[21] L. B. Sina, C. A. Secco, M. Blazevic, and K. Nazemi, “HYBRID FORECASTING METHODS—A SYSTEMATIC REVIEW,” *Electronics*, vol. 12, no. 9, p. 2019, Apr. 2023, doi: https://doi.org/10.3390/electronics12092019.

[22] F. Z. Che Rose, N. A. K. Rosili, and M. F. Marsani, “COMPARISON OF MACHINE LEARNING MODEL PERFORMANCE FOR PREDICTING THE CLIMATE VARIABLES IN JOHOR BAHRU, MALAYSIA,” *Sci. Rep.*, vol. 15, no. 1, p. 23465, Jul. 2025, doi: https://doi.org/10.1038/s41598-025-08033-y.

[23] V. Someetheram *et al.*, “HYBRID DOUBLE ENSEMBLE EMPIRICAL MODE DECOMPOSITION AND K-NEAREST NEIGHBORS MODEL WITH IMPROVED PARTICLE SWARM OPTIMIZATION FOR WATER LEVEL FORECASTING,” *Alex. Eng. J.*, vol. 115, pp. 423–433, Mar. 2025, doi: https://doi.org/10.1016/j.aej.2024.12.035.

[24] K. Abdualkarim, M. S. M. Kasihmuddin, and M. F. Marsani, “ENHANCING FLOOD FORECASTING ACCURACY THROUGH IMPROVED SVM AND ANFIS TECHNIQUES,” *Math. Model. Comput.*, vol. 12, no. 2, pp. 447–460, 2025, doi: https://doi.org/10.23939/mmc2025.02.447.

[25] V. Someetheram *et al.*, “DOUBLE DECOMPOSITION WITH ENHANCED LEAST-SQUARES SUPPORT VECTOR MACHINE TO PREDICT WATER LEVEL,” *J. Water Clim. Change*, vol. 15, no. 6, pp. 2582–2594, Jun. 2024, doi: https://doi.org/10.2166/wcc.2024.558.

[26] R. Farooq and M. A. Imteaz, “SEASONAL RAINFALL FORECASTING USING WAVELET-BASED HYBRID MODELS,” in *Proc. 2023 IEEE Eng. Informatics Conf.*, IEEE, 2023, pp. 1–7, doi: https://doi.org/10.1109/IEEECONF58110.2023.10520402.

[27] R. Rezaiy and A. Shabri, “AN INNOVATIVE HYBRID W-EEMD-ARIMA MODEL FOR DROUGHT FORECASTING USING THE STANDARDIZED PRECIPITATION INDEX,” *Nat. Hazards*, vol. 120, no. 14, pp. 13513–13542, Jul. 2024, doi: https://doi.org/10.1007/s11069-024-06758-z.

[28] M. M. U. Qureshi *et al.*, “DEVELOPING A SEASONAL-ADJUSTED MACHINE-LEARNING-BASED HYBRID TIME-SERIES MODEL TO FORECAST HEATWAVE WARNING,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–17, Mar. 2025, doi: https://doi.org/10.1038/s41598-025-93227-7

[29] S. Tumse and U. Alcansoy, “STATISTICAL AND DEEP LEARNING APPROACHES IN ESTIMATING PRESENT AND FUTURE GLOBAL MEAN SEA LEVEL RISE,” *Nat. Hazards*, vol. 121, no. 9, pp. 10377–10404, Mar. 2025, doi: https://doi.org/10.1007/s11069-025-07203-5.

[30] A. Atesongun and M. Gulsen, “A HYBRID FORECASTING STRUCTURE BASED ON ARIMA AND ARTIFICIAL NEURAL NETWORK MODELS,” *Appl. Sci.,* vol. 14, no. 16, p. 7122, Aug. 2024., doi: https://doi.org/10.3390/app14167122

[31] S. Makridakis, S. Evangelos, and V. Assimakopoulos. "STATISTICAL AND MACHINE LEARNING FORECASTING METHODS: CONCERNS AND WAYS FORWARD." *PloS one*, vol.13, no.3, p. e0194889, Mar. 2018, doi: https://doi.org/10.1371/journal.pone.0194889

[32] “NATIONAL OCEANOGRAPHY CENTER.” National Oceanography Center. https://www.noc.ac.uk/ (accessed Jul. 26, 2025)

[33] J. Kaur *et al*. "AUTOREGRESSIVE MODELS IN ENVIRONMENTAL FORECASTING TIME SERIES: A THEORETICAL AND APPLICATION REVIEW." *Environ. Sci. Pollut. Res.*, vol. 30, no. 8, Jan. 2023, pp. 19617-19641, doi: https://doi.org/10.1007/s11356-023-25148-9.

[34] M. F. Marsani and A. Shabri, “RANDOM WALK BEHAVIOUR OF MALAYSIA SHARE RETURN IN DIFFERENT ECONOMIC CIRCUMSTANCE FORMULA,” *Mat. Malaysian J. Ind. Appl. Math.*, vol. 35, no. 3, , Dec. 2019, doi: https://doi.org/10.11113/matematika.v35.n3.1105

[35] M. F. Marsani, A. Shabri, and B. Badyalina, “EFFICIENT MARKET HYPOTHESIS FOR MALAYSIAN EXTREME STOCK RETURN: PEAKS OVER A THRESHOLD METHOD,” *Matematika*, vol. 38, no. 2, pp. 141–155, Aug.2022, doi: https://doi.org/10.11113/matematika.v38.n2.1396

[36] M. F. Marsani and A. Shabri, “NON-STATIONARY IN EXTREME SHARE RETURN: WORLD INDICES APPLICATION,” *ASM Sci. J.*, vol. 26, pp. 1–9, 2020, doi: https://doi.org/10.32802/asmscj.2020.sm26(1.16)

[37] Y. Chen*, et al*. "DISCRETE WAVELET TRANSFORM APPLICATION FOR BIKE SHARING SYSTEM CHECK-IN/OUT DEMAND PREDICTION." *Transportation Letters*, vol. 16, no. 6, pp. 554-565, Jul. 2024, doi: https://doi.org/10.1080/19427867.2023.2219045.

[38] P. Schober, *et al*. "CORRELATION COEFFICIENTS: APPROPRIATE USE AND INTERPRETATION." *Anesthesia & analgesia* vol. 126, no. 5, pp. 1763-1768, May 2018, doi: https://doi.org/10.1213/ANE.0000000000002864

[39] A. Pourzangbar, *et al*. "MACHINE LEARNING APPLICATION IN MODELLING MARINE AND COASTAL PHENOMENA: A CRITICAL REVIEW." *Frontiers in Environmental Eng*., vol.2 p.1235557, Sept. 2023, doi: https://doi.org/10.3389/fenve.2023.1235557