MJFAS
Malaysian Journal of Fundamental and Applied Sciences

Someetheram *et al.* | Malaysian Journal of Fundamental and Applied Sciences, Vol. 22 (2026) 36–50

RESEARCH ARTICLE

Forecasting Kemaman River Water Level Using Hybrid ARIMA-STL Model

Vikneswari Someetheram[a], Muhammad Fadhil Marsani[a,*], Muhammad Wafiy Adli[b], Mohd Radhie Mohd Salleh[c], Basri Badyalina[d]

[a]School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia; [b]School of Geography Section, School of Humanities, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia; [c]Hydraulics and Hydrology Research Group, Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 UTM, Johor Bahru, Johor Malaysia; [d]Faculty of Computer and Mathematical Sciences, University Teknologi Mara, Cawangan Johor Kampus Segamat, 85000, Johor, Malaysia

Abstract The rise in river water levels is a critical indicator for flood risk and early warning systems particularly in flood-prone areas such as the Kemaman River in Terengganu, Malaysia. This study aims to develop a reliable forecasting model to predict daily water level fluctuations and enhance flood preparedness. A total of 3,287 daily water level observations between 1 January 2001 and 31 December 2009 were used as the unit of analysis. The research addresses the limitation of traditional Autoregressive Integrated Moving Average (ARIMA) models in capturing non-linear and seasonal structures by proposing a hybrid forecasting model that integrates the ARIMA model with Seasonal and Trend Decomposition using Loess (STL). This hybrid ARIMA-STL model improves the ability to capture underlying seasonal patterns and long-term trends in water level data. The findings reveal that the hybrid model offers more accurate and stable predictions compared to the standalone ARIMA model that is effective for early warning systems and water resource management. This study fills a research gap by applying STL decomposition to enhance classical time series forecasting in hydrology that highlights the novelty of integrating statistical and decomposition techniques for improved daily river water level prediction.

Keywords: Machine learning, time series, predictive models, statistical method, flood forecasting.

Introduction

Flooding occurs when a water body accumulates excessive water beyond its capacity causing overflow into surrounding floodplain areas that often leads to widespread destruction of infrastructure, displacement of communities and significant economic losses. From a hydrological perspective, the time taken for river flow to return to normal levels differentiates monsoon floods which are prolonged from flash floods which are sudden and short-lived [1]. In Malaysia, states like Kelantan, Terengganu and Pahang frequently experience severe flooding during the monsoon season due to intense and prolonged rainfall [2]. Predicting water levels is crucial for successful flood control. Time-based data analysis allows for projections of future water heights [3]. Such projections are key components of flood alert systems, enabling quick responses and lowering potential harm [4]. Observing water heights and river flow is also necessary for judging the long-term viability of water resources. This becomes more important as global changes and human actions intensify flood occurrences and severity [5].

The Autoregressive Integrated Moving Average (ARIMA) model is a commonly used statistical approach for time series analysis and forecasting, capturing patterns through autoregression, differencing, and moving averages [6]. ARIMA is a well-established statistical approach that requires conditions such as stationarity and the presence of white noise to ensure accurate predictions. Despite these requirements,

***For correspondence:**
fadhilmarsani@usm.my

Received: 22 March 2025
Accepted: 13 August 2025

ARIMA is widely favored for forecasting due to its robustness in handling non-stationary and seasonal trends, often serving as a benchmark approach. [7] showcased its effectiveness by utilizing a seasonal ARIMA model based on the Box and Jenkins approach to forecast rainfall data in Sylhet while [8] highlighted its frequent use by hydrology experts for forecasting river flow. Its popularity extends to diverse applications including reservoir flood risk assessments under climate change scenarios [9] and groundwater level predictions to enhance early warning systems [10]. Recently, [11] investigates the predictive capabilities of the SARIMA model for forecasting Malaysia's mean sea level from three coastal stations in Malaysia. The findings highlight SARIMA's effectiveness in capturing seasonal patterns and short-term dependencies, demonstrating its potential as a reliable tool for coastal hazard management and climate change mitigation. Recent advancements in ARIMA modeling include incorporating machine learning models for enhanced accuracy and versatility particularly for nonlinear and high-dimensional datasets. Furthermore, its integration with big data technologies has expanded its applicability in real-time forecasting systems that makes it a vital tool in water resource management and disaster preparedness. [12] proposes a hybrid decomposition method with ARIMA model for accurate runoff prediction at the Lijin hydrological station on the lower Yellow River. The hybrid model has the ability to address modal mixing and noise with ARIMA's strength in handling linear trends, the model effectively forecasts complex hydrological data. The results demonstrate high predictive performance and confirms its potential for nonlinear and non-stationary runoff forecasting. However, the study focuses solely on runoff data and does not explore river water level prediction which limits its applicability to flood forecasting scenarios where water level data is a more direct and critical indicator. [13] proposes a flexible framework for forecasting hydrological time series to support water resource management that focuses on river water level prediction at the Vitia station on the Moravae Binçës River. By using monthly data from 2014 to 2022, the study compares the performance of ARIMA and Exponential Smoothing (ETS) models. The results demonstrate that both models are effective with predictive analyses identifying critical periods of high and low water levels relevant for flood risk management. Nevertheless, the study does not incorporate any data decomposition techniques to enhance model input quality which limits the model's ability to capture complex patterns in nonstationary and nonlinear hydrological data that is a key challenge in accurate river water level forecasting. [14] proposes the use of ARIMA models to forecast pre and post monsoon groundwater levels in observation wells across the Jammu Himalaya region from 2015 to 2034. Using historical data from 1994 to 2014 and validating with 2009–2014 data, the ARIMA model was selected for its best fit and strong agreement with observed values. The forecast reveals a declining groundwater trend, emphasizing the need for sustainable aquifer management and informed policy decisions. However, the study is limited to groundwater prediction and does not implement decomposition-enhanced methods. To address this limitation, this study proposes the integration of Seasonal-Trend decomposition using Loess (STL) with ARIMA. This hybrid approach is adopted to overcome ARIMA's inability to effectively capture complex seasonal fluctuations and non-linear patterns in the data to improve forecasting accuracy and robustness which limits the model's ability to capture complex patterns in nonstationary and nonlinear hydrological data

Time series data often exhibit intricate patterns such as trends, seasonality and randomness that can be challenging to analyze [15]. Decomposition techniques are instrumental in uncovering these hidden patterns by isolating the components of a time series [16]. Among these techniques, the Seasonal and Trend decomposition using Loess (STL) method is particularly advantageous due to its flexibility in handling various forms of seasonality, its robustness to outliers and its user-defined control over the smoothness of the trend-cycle relationship [17]. STL has been extensively utilized in hydrological research, particularly for streamflow and water level forecasting applications [13-20]. Recent advancements in time series analysis emphasize hybrid approaches that integrate STL with machine learning models to improve predictive accuracy. For instance, [21] highlighted the role of STL-LSTM models in achieving precise short-term water demand forecasting within intelligent water supply systems. In this study, STL is applied to break down water level data into trend, seasonal and random components, enabling a comprehensive analysis of the underlying patterns. To further enhance forecasting performance, STL is combined with ARIMA in a hybrid model, ARIMA-STL is benchmarked against the traditional ARIMA model. This comparative analysis seeks to assess the efficiency of both methods in predicting daily water levels, highlighting the potential of hybrid approaches to better capture intricate patterns in time series data.

The primary objective of this study is to develop an accurate and reliable forecasting approach for daily river water levels to support flood prediction and early warning systems. To achieve this, the study first applies the Autoregressive Integrated Moving Average (ARIMA) model to forecast water levels, leveraging its ability to capture temporal dependencies in time series data. Furthermore, a hybrid forecasting model is developed by integrating ARIMA with the Seasonal and Trend Decomposition using LOESS (STL) method which combines the strengths of both techniques to more effectively capture seasonal patterns, long-term trends and irregular fluctuations in the data. Finally, the performance of the

individual ARIMA model and the ARIMA-STL hybrid model is systematically evaluated using daily water level data. The evaluation focuses on comparing their accuracy, reliability and predictive effectiveness to identify the model that delivers superior forecasting capabilities for river water level prediction.

The following organization is employed in this paper: Section 2 presents a concise overview and preliminary explanation and the methodology of ARIMA, STL and ARIMA-STL. In Section 3, the detail of the performance metrics chosen in this study will be explained. Section 4 provides an in-depth discussion on the study area along with a detailed analysis of the proposed models' performance and results. In addition, a comparison with ARIMA and ARIMA-STL will be executed. Finally, conclusions are given in the final Section 5.

Methodology

Autoregression Integrated Moving-Average (ARIMA)

ARIMA models are widely recognized for their ability to accurately forecast future values within time series data [22]. These models utilize three key components: Autoregressive (AR), Moving Average (MA), and differencing. The AR component captures the correlation between current and past values, while the MA component accounts for the relationship between the current value and past forecast errors. Differencing is applied to achieve stationarity, a crucial requirement for effective time series modeling. An ARIMA model is expressed as ARIMA(p,d,q), where 'p' signifies the order of the AR component, 'd' represents the number of differencing steps needed for stationarity, and 'q' indicates the order of the MA component. The selection of these parameters is critical for optimal model performance, and is typically determined through analysis of autocorrelation and partial autocorrelation functions

AR Model

The ARIMA(p,0,0) model or the common approach of the AR model with order p is written in equation 1.

$$X_t = \delta + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \ldots + \emptyset_p X_{t-p} + \varepsilon_t \quad (1)$$

where X_t is observed value at time t, δ a constant value, $\emptyset_p$ is autoregressive parameter of order p and ε_t is the error value at time t.

MA Model

The ARIMA (0,0,q) model or the common approach of the MA model with order q is written in equation 2.

$$X_t = \delta - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (2)$$

where X_t is the observed value at time t, δ is a constant value, θ_q is the autoregressive parameter of order q, and ε_{t-q} is the error values at time t – q

ARMA Model

With order p = 1 and q = 1, the combination of the MA and AR models is frequently applied. The model ARMA (1,1) can be represented by the expression in equation 3 or equation 4.

$$(1 - \emptyset_1 B) X_t = \delta + (1 - \theta_1 B) \varepsilon_t \quad (3)$$

$$X_t = \delta + \emptyset_1 X_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (4)$$

where X_t is the observed value at time t, δ is a constant value, $\emptyset_1$ is the autoregressive parameter of order 1, θ_1 is the autoregressive parameter of order 1, ε_t is the error values at time t and B is the backshift operator,

ARIMA Model

Box and Jenkins outlined three phases, identification, estimation, and diagnostic checking in ARIMA modeling, applicable to both stationary and non-stationary data, especially when differencing is needed. In hydrology, experts frequently use ARIMA and ARMA models for forecasting river flow, given their effectiveness in handling non-stationary and seasonal trends [9]. The standard ARIMA(p,d,q) model is widely used due to its consistent high performance [23]. This study applies the standard ARIMA model, considering the absence of trends in the change of mean, variance, and autocorrelation, which indicates stationarity. The following are the suggested ARIMA model processes:

:

(a) Model Identification

Identification involves checking the stationarity of the data and applying differencing if needed. The number of differencing steps determines the ARIMA model's integration order. Suitable ARMA model calculated based on Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) with the

order ("p" and "q") of AR and MA terms identified through comparisons between the observed data's PACF and ACF. The calculation of the ACF and PACF plots are shown by the following equations 5 to 14.

i. The ACF at lag k:

$$r_k = \mathrm{Corr}(Z_t, Z_{t-k} \quad (5)$$

$$r_k = \mathrm{Corr}(Z_t, Z_{t-k}) \quad (6)$$

The standard error (SE) of autocorrelation at lag k is:

$$r_k = \frac{\sum_{t=1}^{n-k}(Z_t-\bar{Z})(Z_{t-k}-\bar{Z})}{\sum_{t=1}^{n}(Z_t-\bar{Z})^2} \quad (7)$$

$$SE_{rk} = \sqrt{\frac{1+2\sum_{j=1}^{k-1} rj^2}{n}} \quad (8)$$

$$r_k = \frac{\sum_{t=1}^{n-k}(Z_t-\bar{Z})(Z_{t-k}-\bar{Z})}{\sum_{t=1}^{n}(Z_t-\bar{Z})^2} \quad (9)$$

$$SE_{rk} = \sqrt{\frac{1+2\sum_{j=1}^{k-1} rj^2}{n}} \quad (10)$$

$$r_{kk} = \frac{r_k-\sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1-\sum_{j=1}^{k-1} r_{k-1,j} r_j} \quad (11)$$

$$r_{kk} = \frac{r_k-\sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1-\sum_{j=1}^{k-1} r_{k-1,j} r_j} \quad (12)$$

ii. The PACF:

with standard error (SE) of partial autocorrelation at lag k and t statistics presented respectively

where Z_t is stationarized data, n is the total data number, $j = 1,2,3,\ldots,k-1$ and

$$SE_{r_{kk}} = \sqrt{\frac{1}{n}} \quad (13)$$

$$t_{r_{kk}} = \frac{r_{kk}}{SE_{r_{kk}}} \quad (14)$$

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-1}$$

The possible outcome from this process can only have AR term, the MA term and combination of AR and MA terms that is (ARMA). There are a few points that we can use to find whether it is AR, MA or ARMA. The points shown below: -

i. If the ACF plot tails off exponentially and the PACF is suddenly cut off after p lag. The lag p is considered as the auto regressive term.
ii. If the PACF tails off exponentially and the ACF is suddenly cut off after lag q. The number of lag q is considered as the moving-average term.
iii. If both ACF and PACF tail off after a few lags, then it will be mixed model.

The model identification is used for both non-seasonal and seasonal conditions. The seasonal composition, the observation of ACF and PACF has n number of lag intervals. If lag-1 is set as the first point for non-seasonal the lag-1 + n is used for the next observation point.

(b) Parameter Estimation

Once a model has been identified, typically, several candidate models are put forward. These proposed models must undergo parameter estimation processes, which include the application of a statistical p-test. For each parameter, the model choice necessitates a p-value lower than the specified alpha (α). With a confidence level of 90% for this analysis, alpha is established at 0.1.

(c) Diagnostic Checking

Following parameter estimation for our initial model, we evaluated its forecasting potential by conducting a diagnostic check, specifically examining the ACF. The null hypothesis (H0) states residual ACF follows a white noise pattern, whereas the alternative hypothesis (H1) asserts that the residual ACF does not exhibit white noise characteristics. The flowchart of ARIMA method is illustrated in Figure 1.

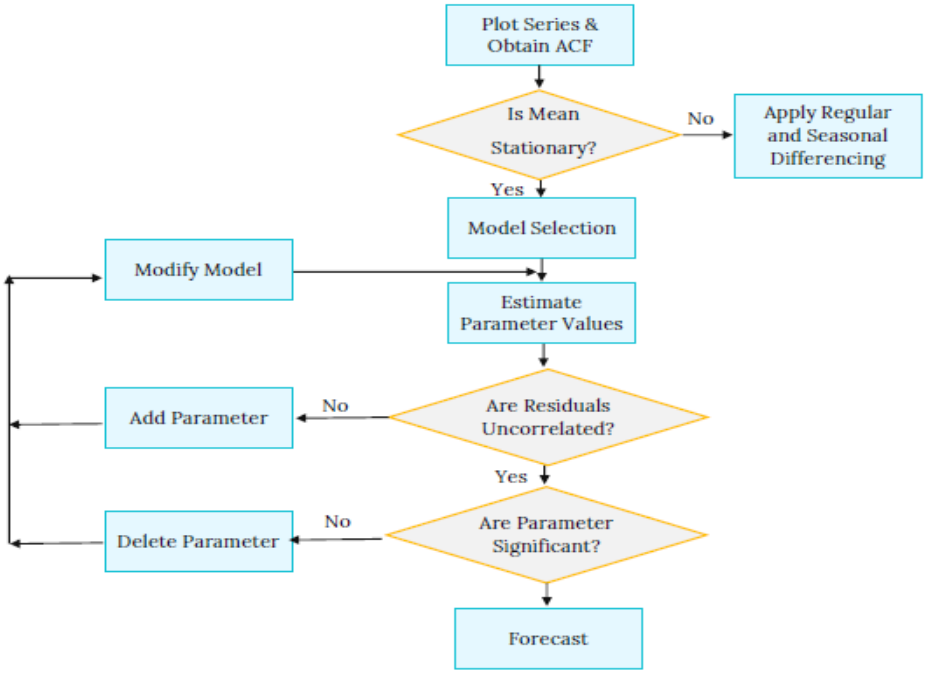

Figure 1. Flowchart of the ARIMA forecasting processes

Seasonal and Trend Decomposition of Time Series using Loess

The Seasonal-Trend decomposition procedure based on Loess (STL) is a powerful tool for analyzing time series data, separating it into its fundamental components: seasonal, trend, and remainder. This decomposition facilitates a clearer understanding of the underlying patterns within the data. STL employs locally weighted regression, a non-parametric approach, to estimate these components, offering flexibility in capturing complex, non-linear relationships. Specifically, it utilizes Locally Weighted Scatterplot Smoothing (Loess), as detailed in [18], to filter the time series. The Loess methodology allows for the robust estimation of non-linear trends and seasonal variations, even in the presence of outliers or abrupt changes in the data. This robust nature, combined with its ability to handle various types of seasonality, makes STL a valuable technique for time series analysis across diverse fields.

The formula for Loess can be represented as follows:

$$\hat{Y}(x) = \frac{\sum(w(x_i)*y_i)}{\sum w(x_i)} \quad (15)$$

where $\hat{Y}(x)$ represents the estimated value of the response variable at the point x, $w(x_i)$ represents the weight assigned to the data point at x_i, based on its distance from x and y_i denotes the observed response variable value corresponding to x_i.

The weights, $w(x_i)$, is determined by the kernel function and bandwidth parameter, controlling the neighborhood size for regression. The kernel assigns higher weights to nearby points, emphasizing their influence on the estimation point. As explained in [24], the algorithm includes seasonal, trend, and remainder components. The equation for STL is expressed as in equation 16:

$$Y(t) = S(t) + T(t) + R(t) \quad (16)$$

where Y(t) represents the observed value of the time series at time t, S(t) represents the estimated seasonal component at time t, T(t) represents the estimated trend component at time t and R(t) represents the estimated remainder component at time t.

ARIMA-STL Hybrid Model

A model is introduced to improve future water level predictions in this research. Combining various models, as suggested by [20], enhances prediction accuracy. ARIMA and STL are widely utilized in time series forecasting, where ARIMA models autoregressive and moving average components, while STL decomposes the data into few elements. The next step is forecasting for the ARIMA-STL Hybrid Model.

i. Individual Model Selection: To find the ARIMA model orders (p, d, q), and decompose the water level data, STL is used to separate the trend, seasonal, and residual components. The trend represents long-term variation, the seasonal component shows cyclical patterns, and the residual component captures random fluctuations. Next, an automated Exponential Smoothing Space State model (ESSSM) in R-Studio is applied, displaying the smoothing parameter and initial states of the model.
ii. Model Combination: After the best model for each method was defined, ARIMA and STL method were combined into the model. The weightage for the model will be set to 0.5 because the method uses the equally average weightage.
iii. Forecasting: Forecast future data after the model that satisfies the conditions.
iv. Performance Evaluation: Assess the performance of both models using RMSE, MAE, and MAPE values.

Data Normalization

Normalization occurs early in the process, before model training. A range of 0 to 1 is used for sigmoid activation, and -1 to 1 for hyperbolic activation [25]. However, normalization poses challenges, requiring retransformation of output values to the original scale for performance analysis. Linear transformation, developed by [26], is employed to normalize water level data within a range of 0 to 1. The equation given in equation 17 [27].

$$x_{\mathrm{n}} = \frac{(x_0 - x_{\min})}{(x_{\max} - x_{\min})} \tag{17}$$

The scaled water level data denoted by x_{n} is derived from the original measurements, x_0. This transformation is based on the range of the original data, utilizing the minimum,($x_{\min}$) and maximum values (, $x_{\max}$) [28].

Performance Metrics

Various methods can be applied for performance evaluation in hydrological forecasting models. Scale-dependent error measures typically express errors in the same or squared unit as the variable being analyzed. In this study, RMSE, MAE, and MAPE are utilized to assess model performance, as presented in Equations 18 to 20 [29].

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - \hat{z}_i)^2} \tag{18}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|z_i - \hat{z}_i|^2 \tag{19}$$

where n is the total number data, $\hat{z}_i$ the forecast value and $z_{i\,\mathrm{t}}$ is the actual water level value.

Mean Absolute Percentage Error (MAPE)

$$MAPE = \sum_{i=1}^{n}\frac{100}{n}\left|\frac{z_i - \hat{z}_i}{z_i}\right| \tag{20}$$

Results and Discussion

Data set and Study Area

Terengganu is in northern Malaysia, bordered by Kelantan and Pahang. Kemaman River with 167 km long as illustrated in Figure 2. The gauging station at Kemaman river at Air Puteh bridge was chosen for the water level of Kemaman river because the Kemaman river at the bridge usually will have floods when comes to rainy season. The data sets were obtained from Department of Irrigation and Drainage (DID) Terengganu, Malaysia. A total of 3287 daily water level data were taken at Kemaman River between 1 January 2001 and 31 December 2009. According to the raw water level data, the greatest value recorded is 285.85 m^3/s, and the least value recorded is 2.12 m^3/s. Water Level data on average is 3.04 m^3/s.

The quality of the water level dataset was carefully assessed prior to model development to ensure the reliability of the forecasting results. The dataset is automatically recorded from remote river gauging stations and transmitted to the National Water Management Centre via a telemetry system on a 24-hour basis, contained some missing values. These gaps are common in hydrological datasets and are typically caused by power interruptions, sensor malfunctions or transmission failures at the monitoring stations. However, no significant outliers were detected in the data during the quality check, which involved visual inspection and statistical verification of the time series. To address the issue of missing values, the affected entries were marked as 'NA' and handled appropriately in the preprocessing phase to ensure they did not bias the model training. Additionally, data normalization was applied by scaling the training data to a fixed range between 0 and 300. This step was performed to standardize the data input and enhance the stability and performance of the ARIMA and ARIMA-STL models. These procedures ensured that the dataset was clean, consistent, and suitable for reliable water level forecasting.

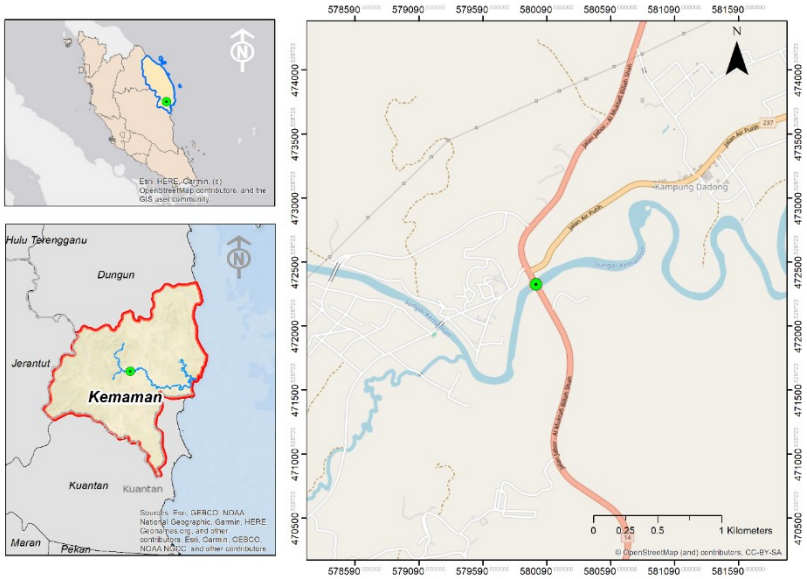

Figure 2. Kemaman river at Air Puteh bridge

ARIMA

Model Identification

Analysis of the ACF and PACF is required to identify the ARIMA model. The normalised water level data's ACF plot was analysed to determine the data's stationary in the first stage. These ACF and PACF plots were analysed to determine q and p terms. The nonseasonal ARIMA model's general theoretical model identifications are shown in Table 1.

The terms “tail off” and “cut off” refer to the characteristic patterns observed in the ACF and PACF plots. A “tail off” indicates a gradual decline in autocorrelation across multiple lags that suggests a non-zero AR or MA structure and depends on whether it appears in the PACF or ACF, respectively. In contrast, a “cut off” refers to a sharp drop in correlation after a certain lag with all subsequent lags showing insignificant values that indicates the order of the AR or MA component. For example, if the ACF cuts off after lag q while the PACF tails off, this supports the selection of an MA(q) model. Conversely, if the PACF cuts off after lag p and the ACF tails off, an AR(p) model is appropriate. These interpretations are based on whether the autocorrelations exceed the 95% confidence bounds around zero in the respective plots.

Table 1. General theoretical model identifications for the non-seasonal ARIMA model

Model	ACF	PACF
AR terms (p)	Tail off	Cut off
MA terms (q)	Cut off	Tail off
AR(p) or MA(q)	Cut off	Cut off
ARMA terms (p,q)	Tails off	Tails off
No non-seasonal terms	No spike	No spike

Figure 3 displays the water level data in Kemaman River from January 2001 to December 2009, along with the ACF plot of daily water level. The plot shows a repeating pattern, suggesting seasonal non-stationarity. However, there is no clear long-term trend. The ACF plot for the original series has a linear tailing off, indicating seasonal non-stationarity. The PACF interpretation is deemed meaningless due to the ACF's pattern. The ACF at non-seasonal lags (1 to 11, 13 to 23) gradually tails off, suggesting non-stationarity, requiring differencing with d=1.

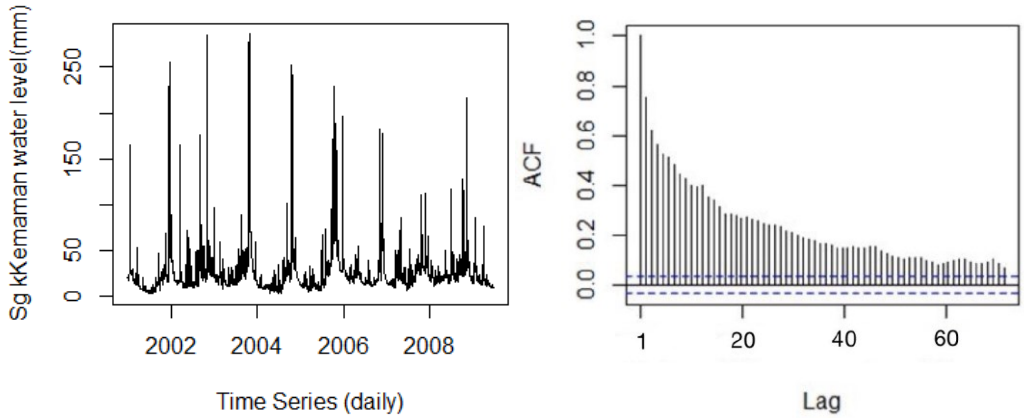

Figure 3. Time Series Plot and ACF Plot for Original Daily Water Level of Kemaman River

In Figure 4, the ACF shows a cutoff pattern at lag 1, and the PACF tails off non-seasonally for the daily water level in Kemaman River. Following simplicity principles, MA (1) is fitted for the non-seasonal component, leading to the choice of ARIMA (0,1,1) as the non-seasonal model.

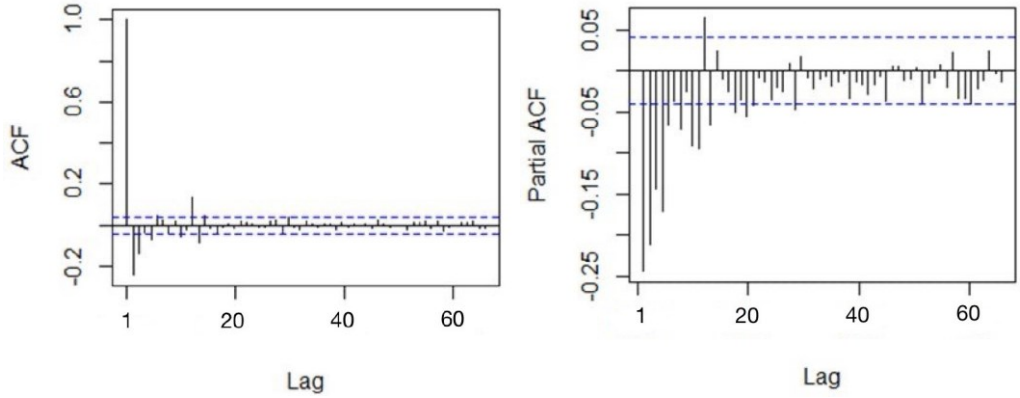

Figure 4. ACF and PACF Plot for Daily Water Level of Kemaman River after conducting first differencing of non-seasonal

Figure 5 shows a noticeable spike at lag-14, which diminishes by lag-21 that indicates potential seasonality in the data. Similarly, Figure 5 presents the PACF plot of the river water level time series. The plot shows a gradual tailing-off pattern rather than a sharp cutoff, with several significant spikes appearing at regular intervals. This slow decay in the partial autocorrelations suggests the presence of a seasonal structure within the data. In time series analysis, a tailing-off pattern in the PACF especially at specific seasonal lags often indicates that the series exhibits seasonal dependencies. Such patterns support the inclusion of seasonal components in the modeling process as they reflect repeated cycles or trends over time. Therefore, the observed behavior in Figure 5 aligns with the theoretical structure of a standard Seasonal Moving Average (SMA) (1) process, suggesting that an SMA (1) model is appropriate for capturing seasonal variations. Considering both the non-seasonal and seasonal components, the ARIMA (1,1,1)(0,1,1)14 model is selected for further analysis. This model effectively integrates differencing, autoregressive, and moving average components to enhance forecasting accuracy.

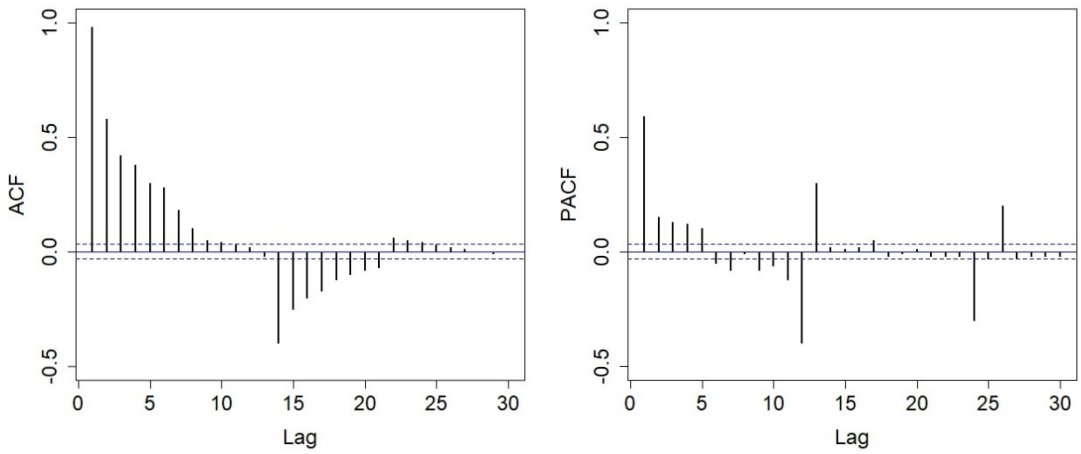

Figure 5. ACF and PACF plot of the seasonal differenced water level data

Parameter Estimation

The statistical t-test and p-test are being used in the parameter estimation approach. The value of standard error (SE) will be generated by the R Studio and the smaller value will be more preferrable. The p-value is used to obtain the probability of the t-test. The p-value is less than the significant level α (0.10) will be considered for the ARIMA models to be statistically significant.

Table 2. Parameter estimation of ARIMA Model for Kemaman River

Model	Coefficient	SE	p-value
ARIMA $(0,1,1)(0,1,1)^{14}$			
MA (1)	-0.456902	0.026577	$2.2e^{-16}$
SMA (1)	-0.999134	0.051165	$2.2e^{-16}$
ARIMA $(1,1,1)(0,1,1)^{14}$			
AR (1)	0.477869	0.033521	$2.2e^{-16}$
MA (1)	-0.861374	0.021508	$2.2e^{-16}$
SMA (1)	-0.999985	0.013116	$2.2e^{-16}$
ARIMA $(0,1,2)(0,1,1)^{14}$			
MA (1)	-0.393742	0.020025	$2.2e^{-16}$
MA (2)	-0.241423	0.021175	$2.2e^{-16}$
SMA (1)	-0.999999	0.011480	$2.2e^{-16}$
ARIMA $(0,1,1)(1,1,1)^{14}$			
MA (1)	-0.459411	0.026650	$2.2e^{-16}$
SAR (1)	-0.029972	0.021137	0.1562
SMA (1)	-0.994565	0.011998	$2.2e^{-16}$
ARIMA $(0,1,1)(0,1,2)^{14}$			
MA (1)	-0.45960	0.026648	$2e^{-16}$
SMA (1)	-1.025099	0.022656	$2e^{-16}$
SMA (2)	0.030462	1.4375	0.5216

Table 3. RMSE and AIC for ARIMA Model Kemaman River

Model	RMSE	AIC
ARIMA $(0,1,1)(0,1,1)^{14}$	18.07673	19984.94
ARIMA $(1,1,1)(0,1,1)^{14}$	17.50583	19841.43
ARIMA $(0,1,2)(0,1,1)^{14}$	17.59235	19863.84
ARIMA $(0,1,1)(1,1,1)^{14}$	18.09993	19984.92
ARIMA $(0,1,1)(0,1,2)^{14}$	18.10021	19984.88

Table 3 shows the fitting of ARIMA $(0,1,1)(0,1,1)^{14}$, ARIMA $(1,1,1)(0,1,1)^{14}$, ARIMA $(0,1,2)(0,1,1)^{14}$, ARIMA $(0,1,1)(1,1,1)^{14}$ and ARIMA $(0,1,1)(0,1,2)^{14}$ models respectively. Based on Table 2, the p-value for ARIMA $(0,1,1)(0,1,1)^{14}$, ARIMA $(1,1,1)(0,1,1)^{14}$ and ARIMA $(0,1,2)(0,1,1)^{14}$ are less than 0.1, implying that the models are significant. While the model for ARIMA $(0,1,1)(1,1,1)^{14}$ and ARIMA $(0,1,1)(0,1,2)^{14}$ are not significant because of the p-value that is more than 0.1.

Diagnostic Checking

The accuracy is determined by the Ljung-Box test of white noise residuals. The null hypothesis, H_0 is said that the residual ACF are white noise while the alternative hypothesis, H_1 said the residual ACF are not white. If the residuals are independently distributed and have a mean of zero, a time series is defined as white snoise. Table 4 shows the Ljung-Box test of ARIMA models.

Table 4. Ljung-Box test of ARIMA Model for Kemaman River

Model	2	7	14	30
ARIMA $(0,1,1)(0,1,1)^{14}$	0.6899437	0.03214573	$1.070746\ e^{-07}$	$1.749869\ e^{-07}$
ARIMA $(1,1,1)(0,1,1)^{14}$	0.7889057	0.03128219	$4.248283e^{-08}$	$1.605112e^{-07}$
ARIMA $(0,1,2)(0,1,1)^{14}$	0.041708	0.0008783	$2.021712\ e^{-10}$	$3.38805e^{-09}$
ARIMA $(0,1,1)(1,1,1)^{14}$	$1.332268e^{-15}$	0	0	0
ARIMA $(0,1,1)(0,1,2)^{14}$	$1.332268e^{-15}$	0	0	0

Table 5. Summary of the Simplest Model

Model	Coefficients	Independence	AIC
ARIMA $(0,1,1)(0,1,1)^{14}$	Significant	Independent	19984.94
ARIMA $(1,1,1)(0,1,1)^{14}$	Significant	Independent	19841.43
ARIMA $(0,1,2)(0,1,1)^{14}$	Significant	Independent	19863.84
ARIMA $(0,1,1)(1,1,1)^{14}$	Not Significant	Independent	19984.92
ARIMA $(0,1,1)(0,1,2)^{14}$	Not Significant	Independent	19984.88

Model selection criteria such as Akaike info criterion (AIC) is used to decide on the better model for further analysis between these three models that are significant. It is observed that the ARIMA $(1,1,1)(0,1,1)^{14}$ smallest AIC (19841.43) value as compared to ARIMA $(0,1,1)(0,1,1)^{14}$ and ARIMA $(0,1,2)(0,1,1)^{14}$ models indicating that ARIMA (1,1,1) $(0,1,1)^{14}$ better fits and employ for the further analysis. The MA component ARIMA $(1,1,1)(0,1,2)^{14}$ is increased to check the best model to be fitted.

Table 6 presents the summary of the overfitting process on the series. In time series analysis, the overfitting process involves progressively adding more parameters such as additional AR, MA or seasonal terms to an initially adequate model to examine whether these added terms significantly improve model performance.

Table 6. Summary of the Base Model and Overfitted Models on the Series

Model	Coefficient	RMSE	AIC
	First layer		
ARIMA $(0,1,1)(0,1,1)$	Significant	18.07673	19984.94
ARIMA $(1,1,1)(0,1,1)^{14}$	Significant	17.50583	19841.43
ARIMA $(0,1,2)(0,1,1)^{14}$	Significant	17.59235	19863.84
ARIMA $(0,1,1)(1,1,1)^{14}$	Not Significant	18.09993	19984.92
ARIMA $(0,1,1)(0,1,2)^{14}$	Not Significant	18.10021	19984.88
	Second layer		
ARIMA $(2,1,1)\ (0,1,1)^{14}$	Not Significant	17.5569	19889.1
ARIMA $(1,1,2)(0,1,1)^{14}$	Not Significant	17.50599	19843.42
ARIMA $(1,1,1)(1,1,1)^{14}$	Significant	17.49233	19840.61
ARIMA $(1,1,1)(0,1,2)^{14}$	**Significant**	**17.49088**	**19840.51**
	Third Layer		
ARIMA $(2,1,1)(0,1,2)^{14}$	Not Significant	-	-
ARIMA $(1,1,2)(0,1,2)^{14}$	Not Significant	-	-
ARIMA $(1,1,1)(1,1,2)^{14}$	Not Significant	-	-

(Note: the bold row represents the significant model which satisfies the white noise assumptions)

This process helps to ensure that the selected model is not only accurate but also parsimonious by avoiding unnecessary complexity that may fit noise instead of the true underlying pattern. In the first layer, several base ARIMA models were assessed. Among them, the ARIMA $(1,1,1)(0,1,1)^{14}$ model showed good performance with a low RMSE and AIC and statistically significant coefficients. The second layer introduced additional parameters to test for overfitting. The model ARIMA $(1,1,1)(0,1,2)^{14}$ yielded the lowest RMSE and AIC with all coefficients significant that indicates a slight improvement and suggesting a better model fit. In the third layer, further overfitting was applied by introducing even more parameters. However, none of these models showed improved performance, and their coefficients were statistically insignificant. This indicates that the added complexity did not capture meaningful patterns and may lead to overfitting. Therefore, the overfitting process confirms that ARIMA $(1,1,1)(0,1,2)^{14}$ is the most appropriate model for this time series as it achieves the best balance between accuracy and model simplicity.

ARIMA-STL Model

This study also discussed the Hybrid Model that consists of ARIMA and STL for the time series forecasting. [29] also said in the research that combining multiple forecast leads to better forecasting.

Individual Model Selection ARIMA Model

The model used in this research is implemented in R-Studio. The "HybridModel" package in R-Studio, along with the "auto.arima" and "stl" functions, aids in selecting the ARIMA (5,0,3) model with a non-zero mean. Coefficient tests show significance for most ARIMA components, except for MA(2). The automated ARIMA model, justified by the "forecastHybrid" package, is integrated with the STL model to create a hybrid model, promising enhanced forecasting accuracy.

STL Forecasting Model

For the STL model the “stl” function will give us the trend, seasonal and remainder value of the STL decomposition from the training data of Kemaman River. Figure 6 shows the plot of each Trend, Seasonal and Remainder of the decomposition shown below:

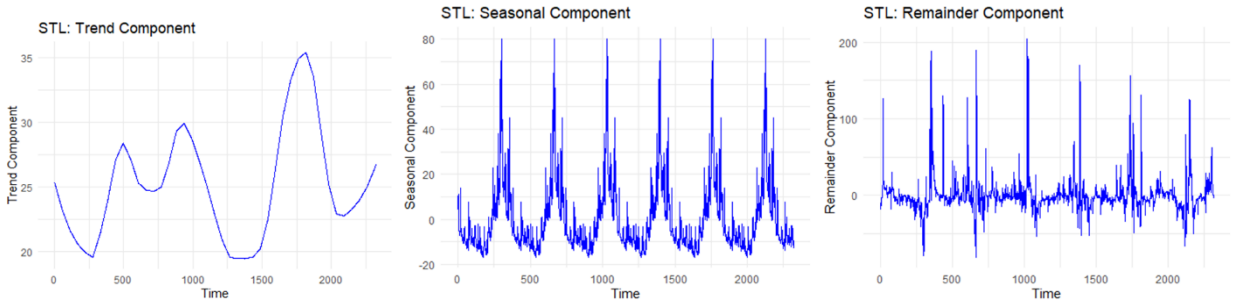

Figure 6. Trend, Seasonal and Remainder plot of trend component

The STL decomposition extracts trend, seasonal, and remainder components from the dataset. The subsequent step involves constructing a forecasting model using the Exponential Smoothing Space State Model (ESSSM). The output from the STL components (trend, seasonality, and remainder) assists in forming the ESSSM forecasting model, denoted as "(A, N, N)," specifying an additive error component, a non-seasonal trend, and no seasonal component. This model employs the Simple Exponential Smoothing method with a smoothing parameter (α) of 0.5806, determining the emphasis on recent data for adaptation to changes. The initial state of the level component is 8.6804, representing the starting point for updating and forecasting. The scale parameter (sigma) is 16.1804, indicating the estimated standard deviation of the error term, signifying greater volatility if the value is high. The study assigns equal weight 0.5 to each forecasting model (STL and ESSSM) in the hybrid model, recognizing both methods in the final prediction.

Forecasting Results and Discussion

75% of the data is allocated for training the ARIMA model while the remaining 25% is reserved for testing. The selected area has four water level categories: normal, alert, warning, and danger, with threshold levels set at 29.0 m, 30.0 m, 30.5 m, and 31.0 m, respectively. A flood warning is issued when the water level exceeds the normal threshold. Accurate water level prediction models are crucial for timely alerts to authorities and residents. After ensuring the statistical adequacy of the ARIMA (1,1,1)(0,1,2)14 model, the next step involves forecasting future water levels in the Kemaman River. Using the hybrid model with assigned weights, the forecasting process utilizes the "forecast()" function in R Studio. The analysis

includes two datasets: the training set for model building and the test set for evaluation. Figure 7 illustrates the forecasting plot for both ARIMA and ARIMA-STL, and Table 7 provides the forecasting errors for the test data.

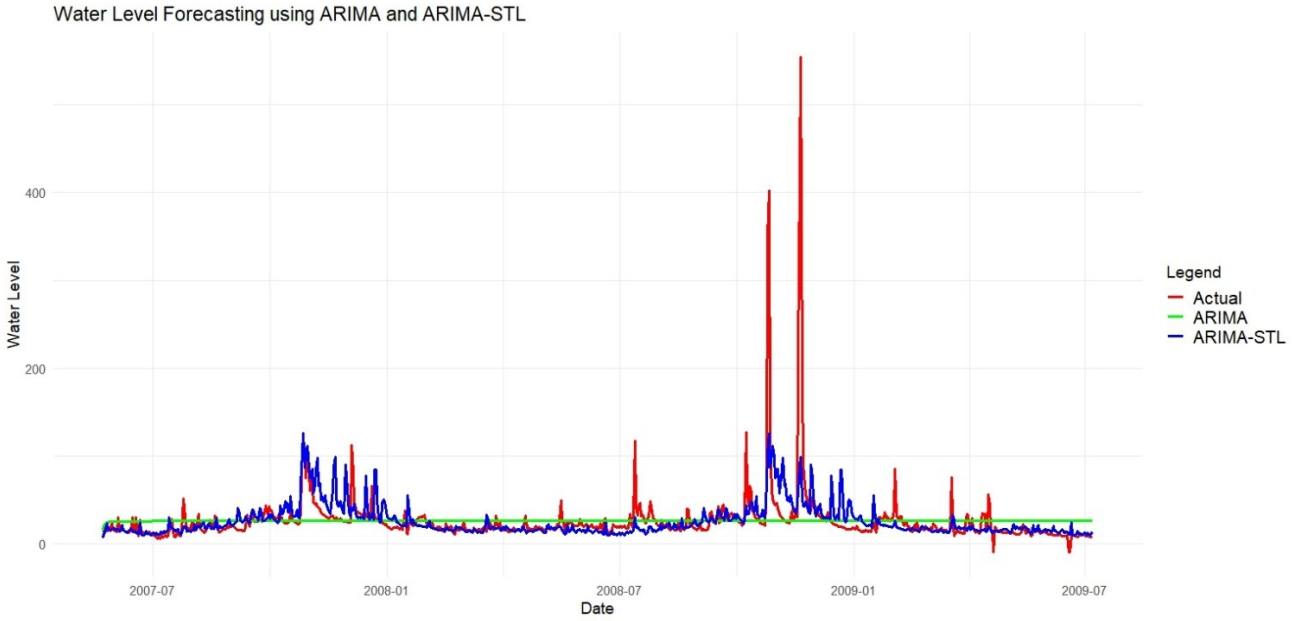

Figure 7. Actual plot vs. forecast plot of ARIMA and ARIMA-STL Hybrid model

Table 7. Performance evaluation of ARIMA $(1,1,1)(0,1,2)^{14}$ and ARIMA-STL

Method	Train Data			Test Data		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
ARIMA $(1,1,1)(0,1,2)^{14}$	17.4909	7.0653	25.9389	17.9096	9.2544	37.9499
Hybrid: 0.5 ARIMA-STL	16.2213	6.3947	23.4415	15.6480	8.7682	33.4702

This study evaluates model performance using RMSE, MAE, and MAPE, comparing two forecasting approaches (ARIMA and ARIMA-STL) for predicting the Kemaman River's water level. Figure 7 illustrates the forecast plots alongside the actual data. The comparison reveals that the ARIMA-STL hybrid model performs better, closely resembling the actual data plot. In contrast, the ARIMA $(1,1,1)(0,1,2)^{14}$ forecast plot shows a substantial deviation from the actual data plot.

The results suggest that both the ARIMA and ARIMA-STL models perform similarly in forecasting accuracy, making them both suitable for time series analysis. The metrics used for evaluation (RMSE, MAE, and MAPE) help assess the models' accuracy. Lower RMSE and MAE values indicate better accuracy, and comparing these metrics with other models is important for evaluation. MAPE, a relative measure, reflects the average percentage difference, with lower values indicating better accuracy. Essentially, these measures enable the evaluation and comparison of the models' predictive performance. After testing on the validation dataset, the ARIMA-STL model outperforms the ARIMA model, with lower RMSE and MAE values. The RMSE decreases by 12.63%, and the MAE decreases by 5.25% by using ARIMA-STL. Further analysis using MAPE shows a better performance, with ARIMA-STL having a 33.47% lower MAPE compared to the benchmark ARIMA model's 37.95%. In conclusion, the ARIMA-STL hybrid model achieves superior accuracy, forecasting the daily water level of Kemaman River.

These findings demonstrate the effectiveness of integrating decomposition techniques such as STL with ARIMA to enhance forecasting accuracy particularly for hydrological time series data. Importantly, this improved predictive capability has broader implications in the context of extreme climatic events and the growing impacts of global warming. Climate change has intensified the frequency and severity of extreme weather conditions such as intense rainfall and prolonged droughts which directly affect river water levels

and increase the risk of floods and water scarcity. Reliable water level forecasting as demonstrated in this study contributes significantly to early warning systems, flood risk management and sustainable water resource planning. The ability of the ARIMA-STL model to more accurately capture seasonal trends and fluctuations makes it a valuable tool in anticipating and responding to climate-induced hydrological variability. Therefore, the outcomes of this research not only advance methodological approaches in time series forecasting but also support broader climate resilience and adaptation strategies for river basin management under changing global conditions.

However, despite the improved accuracy, the proposed ARIMA-STL model has certain limitations that future researchers may consider addressing. One notable drawback is its inability to accurately capture sharp peaks or extreme fluctuations in the river water level series. This limitation arises because ARIMA and STL are inherently linear and additive models that primarily capture overall trend and seasonality but struggle with highly nonlinear and abrupt changes in the data such as those caused by sudden heavy rainfall, flash floods or structural changes in river flow dynamics. STL decomposes the series into seasonal, trend, and remainder components using a smoothing approach, which tends to suppress short-term spikes. Meanwhile, the ARIMA model is better suited to handle linear dependencies over time but does not have mechanisms to respond quickly to sudden anomalies or rare extreme events. As a result, while the ARIMA-STL model is effective for general pattern forecasting, it may underpredict the magnitude of peak flows, which are critical in flood risk assessments.

Conclusion

This research focused on comparing the methods of forecasting for daily water level forecasting between autoregressive integrated moving-averaged (ARIMA) and autoregressive integrated moving-averaged integrated with Seasonal and Trend Decomposition of Time Series using Loess (ARIMA-STL) model. The data of daily water level of Kemaman River was chosen for this reserach. Evaluation metrices used in this paper are RMSE, MAE and MAPE). Among the ARIMA models, The forecasted model by ARIMA $(1,1,1)(0,1,2)^{14}$ model at Kemaman River produced better results with RMSE, MAE and MAPE are 17.49088, 7.065331 and 25.93892 respectively compared to other ARIMA models. Next, the forecasted test data set used and compared by both ARIMA model and ARIMA-STL model. The comparing was evaluated by evaluation metrics. The results show that both the ARIMA model and ARIMA-STL model are practical in solving the time series forecasting problem. The forecasted model by ARIMA-STL model produces performances evaluation with RMSE, MAE and MAPE that are 15.64795, 8.768218 and 33.47029 respectively that concludes the proposed ARIMA-STL model shows the better accuracy in creating the forecasting model compared to its benchmark ARIMA.These findings are consistent with previous studies such as [12] that reported improved forecasting accuracy using decomposition-based hybrid models and [13] that found that incorporating seasonal and trend components enhances model performance in hydrological time series forecasting. The improved performance of the ARIMA-STL model confirms the benefit of decomposing complex time series before modeling as also supported by [14] that emphasized the need for robust techniques in predicting water levels under changing climatic conditions.

Future work should explore the application of this hybrid modeling framework across multiple river stations to incorporate spatial and temporal variability. Additionally, integrating other decomposition techniques or machine learning methods such as wavelet transform, Long Short Term Memory and Support Vector Machine could further enhance forecasting performance. Given the increasing occurrence of extreme hydrological events due to global warming, future studies should also focus on evaluating model robustness under extreme climatic scenarios and during monsoon periods, to support early warning systems and adaptive water resource management strategies.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgment

All the authors gratefully acknowledged the financial support from the “Ministry of Higher Education Malaysia for Fundamental Research Grant Scheme with Project Code: FRGS/1/2022/STG06/USM/03/1.
“

References

[1] Wohl, E., & Lininger, K. B. (2022). Hydrology and discharge. In A. Gupta (Ed.), *Large rivers: Geomorphology and management* (2nd ed., pp. 42–75). John Wiley & Sons.

[2] Shah, S. M. H., Mustaffa, Z., & Yusof, K. W. (2017). Disasters worldwide and floods in the Malaysian region: A brief review. *Indian Journal of Science and Technology, 10*(2), 1–9. https://doi.org/10.17485/ijst/2017/v10i2/110385.

[3] Brockwell, P. J., & Davis, R. A. (2009). *Time series: Theory and methods* (Springer Series in Statistics). Springer.

[4] Parvaze, S., Khan, J. N., Kumar, R., & Allaie, S. P. (2021). Temporal flood forecasting for trans-boundary Jhelum River of Greater Himalayas. *Theoretical and Applied Climatology, 144*, 493–506. https://doi.org/10.1007/s00704-021-03562-8.

[5] Makanda, K., Nzama, S., & Kanyerere, T. (2022). Assessing the role of water resources protection practice for sustainable water resources management: A review. *Water, 14*(19), 3153. https://doi.org/10.3390/w14193153.

[6] Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. Holden-Day.

[7] Bari, S. H., Rahman, M. T., Hussain, M. M., & Ray, S. (2015). Forecasting monthly precipitation in Sylhet City using ARIMA model. *Civil and Environmental Research, 7*(1), 69–77.

[8] Adnan, R. M., Yuan, X., Kisi, O., & Curtef, V. (2017). Application of time series models for streamflow forecasting. *Civil and Environmental Research, 9*(3), 56–63.

[9] Yan, B., Mu, R., Guo, J., Liu, Y., Tang, J., & Wang, H. (2022). Flood risk analysis of reservoirs based on full-series ARIMA model under climate change. *Journal of Hydrology, 610*, 127979. https://doi.org/10.1016/j.jhydrol.2022.127979.

[10] Rashid, A., Alamgir, M., Ahmed, M. T., Salam, R., Islam, A. R. M. T., & Islam, A. (2022). Assessing and forecasting groundwater level fluctuation in Joypurhat district, northwest Bangladesh, using wavelet analysis and ARIMA modeling. *Theoretical and Applied Climatology, 150*(1–2), 327–345. https://doi.org/10.1007/s00704-022-04027-y.

[11] Marsani, M. F., Someetheram, V., Mohd Kasihmuddin, M. S., Mohd Jamaludin, S. Z., Mansor, M., & Badyalina, B. (2024). The application of seasonal autoregressive integrated moving average (SARIMA) model in forecasting Malaysia mean sea level. In *AIP Conference Proceedings* (Vol. 3123, No. 1). AIP Publishing. https://doi.org/10.1063/5.0170840.

[12] Zhang, M., Zhang, X., Qiao, W., Lu, Y., & Chen, H. (2023). Forecasting of runoff in the lower Yellow River based on the CEEMDAN–ARIMA model. *Water Supply, 23*(3), 1434–1450. https://doi.org/10.2166/ws.2023.097.

[13] Agaj, T., Budka, A., Janicka, E., & Bytyqi, V. (2024). Using ARIMA and ETS models for forecasting water level changes for sustainable environmental management. *Scientific Reports, 14*(1), 22444. https://doi.org/10.1038/s41598-024-58839-9.

[14] Qadir, J., Sultan Bhat, M., & Meer, M. S. (2024). Forecasting pre- and post-monsoon depth to water levels using the autoregressive integrated moving average (ARIMA) model in the outer plains of Jammu Himalaya. *Journal of the Geological Society of India, 100*(11), 1557–1567.

[15] Cryer, J. D. (2008). *Time series analysis*. Springer.

[16] Schreiber, T. (1999). Interdisciplinary application of nonlinear time series methods. *Physics Reports, 308*(1), 1–64. https://doi.org/10.1016/S0370-1573(98)00073-5.

[17] Tebong, N. K., Simo, T., Takougang, A. N., & Ntanguen, P. H. (2023). STL-decomposition ensemble deep learning models for daily reservoir inflow forecast for hydroelectricity production. *Heliyon*. https://doi.org/10.1016/j.heliyon.2023.e20606.

[18] Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics, 6*(1), 3–73.

[19] Xu, Z., Mo, L., Zhou, J., Fang, W., & Qin, H. (2022). Stepwise decomposition–integration–prediction framework for runoff forecasting considering boundary correction. *Science of the Total Environment, 851*, 158342. https://doi.org/10.1016/j.scitotenv.2022.158342.

[20] Yang, H., & Li, W. (2023). Data decomposition, seasonal adjustment method and machine learning combined for runoff prediction: A case study. *Water Resources Management, 37*(1), 557–581. https://doi.org/10.1007/s11269-022-03385-3.

[21] Liu, J., Zhou, X. L., Zhang, L. Q., & Xu, Y. P. (2023). Forecasting short-term water demands with an ensemble deep learning model for a water supply system. *Water Resources Management*, 1–22. https://doi.org/10.1007/s11269-023-04089-w.

[22] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons.

[23] Lyu, Z., Ororbia, A., & Desell, T. (2023). Online evolutionary neural architecture search for multivariate non-stationary time series forecasting. *Applied Soft Computing, 130*, 110522. https://doi.org/10.1016/j.asoc.2023.110522.

[24] Zhang, J., Wei, Y. M., Li, D., Tan, Z., & Zhou, J. (2018). Short-term electricity load forecasting using a hybrid model. *Energy, 158*, 774–781. https://doi.org/10.1016/j.energy.2018.05.154.

[25] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0.

[26] Someetheram, V., Marsani, M. F., Kasihmuddin, M. S. M., Jamaludin, S. Z. M., & Mansor, M. A. (2024). Double decomposition with enhanced least-squares support vector machine to predict water level. *Journal of Water and Climate Change, 15*(6), 2582–2594. https://doi.org/10.2166/wcc.2024.486.

[27] Lapedes, A., & Farber, R. (1987). How neural nets work. In *Neural Information Processing Systems*.

[28] Someetheram, V., Marsani, M. F., Kasihmuddin, M. S. M., Jamaludin, S. Z. M., Mansor, M. A., & Zamri, N. E.

(2025). Hybrid double ensemble empirical mode decomposition and k-nearest neighbors model with improved particle swarm optimization for water level forecasting. *Alexandria Engineering Journal, 115*, 423–433. https://doi.org/10.1016/j.aej.2025.01.082.

[29] Rahayu, S. P., Prastyo, D. D., & Wijayanti, D. G. P. (2017). Hybrid model for forecasting time series with trend, seasonal and calendar variation patterns. In *Journal of Physics: Conference Series* (Vol. 890, No. 1, 012160). IOP Publishing. https://doi.org/10.1088/1742-6596/890/1/012160.