UNIABUJA Journal of Engineering and Technology

https://ujet.uniabuja.edu.ng/

ISSN: 2714-3236 (Online); 2714-3228 (Print)

Volume 3, Issue 1, 2025; 1-11

Modelling of Selected River Water Quality Indicators using Autoregressive Integrated Moving Average (ARIMA) Techniques

Ibrahim VANKE[1], Salihu A. ABDULKADIR[2*], Felix A. MODI[3]

[1,3]*Department of Agric & Bioenvironmental Engineering, Adamawa State Polytechnic, Yola, Adamawa State, Nigeria*
[2*]*Department of Agricultural Engineering, University of Abuja, Abuja, Nigeria*

[1]*vankeibrahim4@gmail.com*, [2*]*salihuabdu110@gmail.com*, [3]*fmodiahmed@gmail.com*

Abstract

This study applies the Autoregressive Integrated Moving Average (ARIMA) modelling technique to Predict key water quality indicators of the river Benue in Jimeta, Yola, Nigeria. Utilizing a ten-year dataset (2011–2021) obtained from the Adamawa State Ministry of Water Resources. The research focused on three essential parameters: pH, calcium (mg/L), and iron (mg/L). Following the Box-Jenkins methodology, the data were analysed for stationarity for test using the Augmented Dickey-Fuller test, model identification via ACF/PACF analysis, parameter estimation, diagnostic checking, and forecasting. Results indicated that the ARIMA (0,0,1) model best fits the pH and iron data showing values are stable, primarily influenced by short-term random shocks (MA process)., while the ARIMA (1,0,0) model suits calcium, indicating values fluctuate more, influenced by their immediate past value (AR process). Forecasting results showed a stable average pH of 7.23, fluctuating calcium levels averaging around 61.85 mg/L, and a consistent iron concentration of approximately 0.1523 mg/L over the projected ten-year period (2023–2032). Diagnostic checks confirmed that all selected models were stable, stationary, and invertible, with no unit roots. These findings demonstrate ARIMA's effectiveness in capturing temporal dynamics in water quality and provide a reliable foundation for proactive environmental monitoring, planning, and decision-making.

Keywords: *Modelling, River Benue, Water, Forecasting, Time Series, Concentration, Stationarity Test, Management.*

1.0 Introduction

Water pollution is a growing global concern, driven by rapid population growth, industrialization, and agricultural expansion. As water is a critical natural resource and a core element of ecosystems, monitoring its quality is essential for sustaining human health and environmental balance. Degradation in water quality leads to increased costs in treatment and heightened health risks for communities that rely on untreated or poorly monitored sources (Ejiohuo, *et al.*, 2024). These concerns necessitate continuous monitoring and robust assessment systems to ensure the safe and sustainable use of water resources.

Water quality is indeed influenced by both natural factors like climate, geology, and biological processes, and anthropogenic (human-caused) factors such as industrial and agricultural activities, land use changes. Since no single parameter can fully describe water quality, comprehensive assessment involves monitoring a wide range of physical, chemical, and biological indicators (Ayob *et al.*, 2018). Limited water quality data, particularly in smaller streams and rural areas, can hinder efforts to understand spatial patterns and implement effective water management programs (XiaoYing *et al.*, 2017). Modeling approaches that correlate watershed characteristics with observed water quality can help fill data gaps and inform pollution control strategies (Aho *et al.*, 2018).

Predictive modelling plays a vital role in water quality management by forecasting changes in key parameters and aiding early intervention. Time series models, especially the Autoregressive Integrated Moving Average (ARIMA), have been successfully used to forecast variables such as pH, turbidity, ammonia, and heavy metals in river systems (Ayob *et al.*, 2018; Sentas *et al.*, 2018); These models are particularly valuable due to their ability to handle non-stationary data and capture dynamic environmental changes. They serve as essential tools for policy-makers and environmental agencies to simulate future scenarios and design effective control measures (Twinomuhangi *et al.*, 2025; Ahamad *et al.*, 2015).

ARIMA models have been extensively applied in river water-quality forecasting because they capture time-dependent structures in environmental data through the Box–Jenkins procedure of model identification, estimation, diagnostic checking and forecasting. Shahid and Mohsenipour (2017) demonstrated this capability on the Johor River in Malaysia, successfully predicting short-term fluctuations in pH and other hydrological variables. Similarly, Sentas *et al.* (2018) showed that ARIMA effectively described the temporal dynamics of both water quality and quantity in Greece's River Pinios. Hardiyanti (2020) confirmed the method's predictive strength for river water quality in Indonesia, provided the time series is stationary. Where river systems exhibit strong seasonal cycles, seasonal ARIMA extensions have been shown to improve forecast accuracy, as reported in Water Practice & Technology (2022). More recently, Wang, Chen and Tang (2024) proposed ARIMA–MLP and ARIMA–SSA–LSTM hybrid models, which combine the linear strengths of ARIMA with machine-learning techniques to better capture nonlinear dynamics, achieving higher predictive accuracy than classical ARIMA. Despite these advances, ARIMA remains best suited to short-term forecasts and is limited in its ability to incorporate exogenous environmental drivers (Ayob *et al.,* 2017; Sentas *et al.,* 2018).

ARIMA, as defined by Stellwagen, and Tashman (2013), means autoregression (AR), Integrated (I), and moving average (MA) components to model time-dependent processes. The strength of ARIMA lies in its ability to use historical data trends and residual errors to forecast future values, making it ideal for applications where environmental variables fluctuate unpredictably over time. When used with high-quality data, ARIMA models can forecast water quality trends with reasonable accuracy and support adaptive water management strategies (Chakraborty *et al.,* 2019; Zhu *et al.,* 2021). As such, applying ARIMA to river water quality analysis contributes significantly to ensuring the long-term sustainability of freshwater resources.

The River Benue, a vital water source for Jimeta–Yola, faces increasing contamination risks, necessitating ARIMA-based predictive modeling to forecast key quality indicators (pH, calcium, iron) for sustainable management and public health protection. ARIMA modeling provides a reliable tool for forecasting key water quality parameters, enabling timely interventions and supporting sustainable management of the River Benue. The water quality of River Benue requires thorough investigation, as it is a vital drinking water source for communities like Jimeta Yola. Ongoing human activities are contributing to environmental degradation and potential contamination, posing risks to public health. Therefore, regular monitoring and classification of water quality are essential for effective management and protection of the river. This study aims to develop and evaluate ARIMA models for forecasting selected river water quality indicators, thereby supporting effective environmental monitoring and decision-making for Jimeta Yola, Adamawa State Nigeria.

2.0 Materials and Methods

2.1 The Study Area

The study was conducted along the River Benue in Jimeta, located in Yola North Local Government Area of Adamawa State, Nigeria, within latitudes 9˚7'30'' and 10˚50''N and longitudes 11˚40'' and 13˚20''E. The area lies within the sub-Sudan and northern Guinea savannah vegetation zones, characterized by sparse vegetation comprising grasses, shrubs, and a mix of indigenous and exotic trees. The topography is generally flat and is intersected by the River Benue (Adaeze *et al.,* 2017). The region receives less than 1000 mm of mean annual rainfall in its central and north-western parts, with rainfall distribution influenced by altitude. Rainfall typically begins in April in the south and May in the north, ending between September and November, with the rainy season lasting between 120 and 210 days. Temperature patterns follow typical West African savannah trends, with high solar radiation causing a gradual increase in temperatures from January to April. Temperatures drop at the onset of rains due to cloudiness, rise again slightly after the rains, and fall further during harmattan in December. Minimum temperatures can fall to 18˚C (December–January), while maximums can reach 40˚C in April, with average monthly temperatures ranging from 26˚C to 27.8˚C (Adebayo & Tukur, 1999; Tukur *et al.,* 2004).

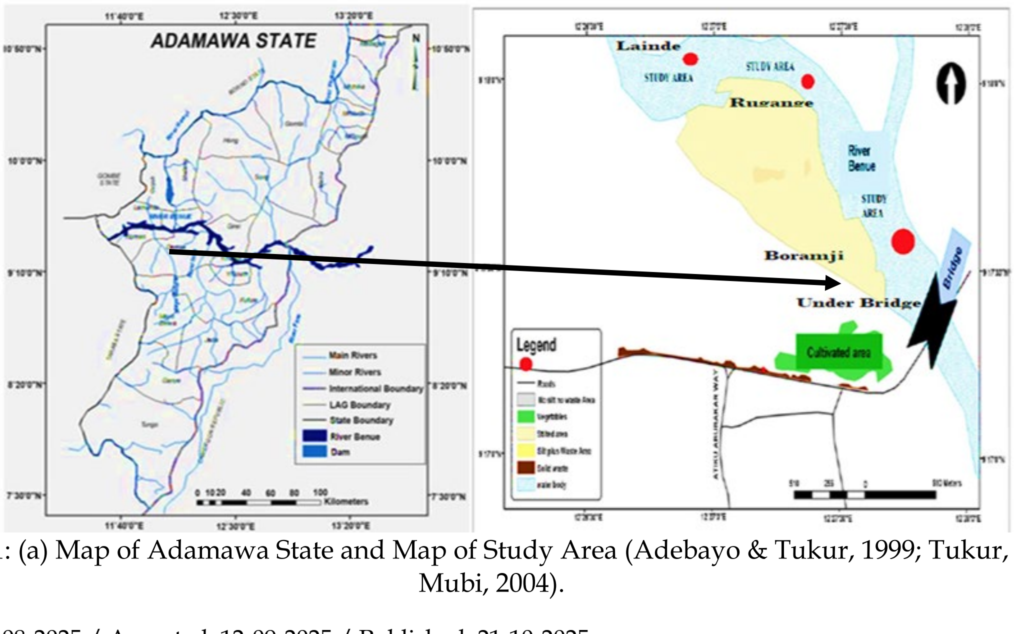

Figure 1: (a) Map of Adamawa State and Map of Study Area (Adebayo & Tukur, 1999; Tukur, Bashir, & Mubi, 2004).

Received: 28-08-2025 / Accepted: 12-09-2025 / Published: 21-10-2025

https://doi.org/10.70118/ujet.2025.0301.01

2.2 ARIMA Model Equations

The ARIMA (Autoregressive Integrated Moving Average) model is a widely used statistical method for analysing and forecasting time series data. It is denoted as ARIMA(p, d, q), where p is the order of autoregression (AR), d is the order of differencing (I), and q is the order of moving average (MA).

2.2.1 Autoregressive (AR) Part

The AR(p) process is represented as:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \quad (1)$$

2.2.1. Integrated (I) Part

The I(d) process applies differencing to make the series stationary:

$$Y'_t = (1 - B)^d Y_t \quad (2)$$

For example, with $d = 1$:

$$Y'_t = Y_t - Y_{t-1} \quad (3)$$

2.2.2 Moving Average (MA) Part

The MA(q) process is:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta\_q \varepsilon_{t-}q \quad (4)$$

2.2.3 Full ARIMA Model

Combining AR, I, and MA gives the ARIMA(p, d, q) equation:

$$\phi(B)(1 - B)^d Y_t = c + \theta(B)\varepsilon_t \quad (5)$$

where:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \ (AR\ polynomial)$$
$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta\_qB\_q \ (MA\ polynomial)$$

2.3 Implementation of ARIMA model for water properties

The flowchart in Figure. 2 depicts the technique for estimating the ARIMA model, which includes the following steps (Wang. 2018).

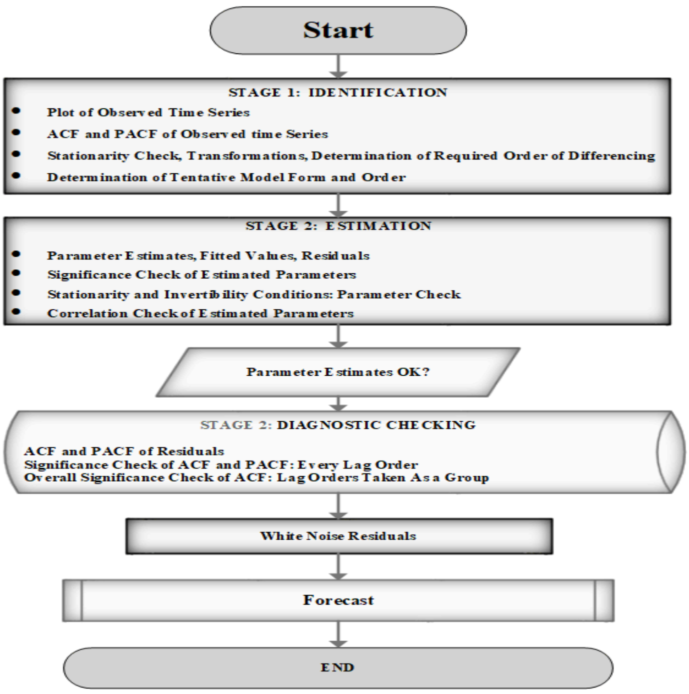

Figure 2: The Iterative Box-Jenkins (ARIMA) Modelling Strategy

Received: 28-08-2025 / Accepted: 12-09-2025 / Published: 21-10-2025
https://doi.org/10.70118/ujet.2025.0301.01

2.3.1. Identification

In this stage, the observed time series data were plotted to visually assess patterns and trends. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the observed series were examined to understand the correlation structures at different lags. A stationarity check was conducted, and if the series was non-stationary, transformations (e.g., differencing or logarithmic transformation) were applied to stabilize the mean and variance over time. The required order of differencing was determined, and a tentative model form and order (i.e., values of p, d, q) were identified based on visual and statistical interpretations of the ACF and PACF plots (AbdulWahid, & Arunbabu, 2022; Wang, 2018).

2.3.2 Estimation

In the estimation phase, model parameters were estimated using statistical techniques such as maximum likelihood estimation. The fitted values and residuals were computed and assessed. The significance of the estimated parameters was evaluated to ensure that each contributed meaningfully to the model. Conditions for stationarity and invertibility were verified to confirm that the model adhered to theoretical assumptions. The correlation among estimated parameters was also checked to detect multicollinearity or redundancy. If the parameter estimates were found to be satisfactory, the process proceeded to the next stage. If not, the model form was revisited and modified (AbdulWahid, & Arunbabu, 2022; Wang, 2018).

2.3.3. Diagnostic Checking

The residuals from the estimated model were analyzed using ACF and PACF plots to ensure they behaved like white noise—that is, having no significant autocorrelations. The significance of the residual autocorrelations was checked at every lag order, and an overall test was conducted to determine whether the group of lag orders was collectively insignificant. If the residuals passed these checks, they were considered white noise (AbdulWahid, & Arunbabu, 2022; Wang, 2018).

2.3.4. Forecasting and Conclusion

Once white noise residuals were confirmed, the model was deemed valid, and forecasting was carried out based on the final ARIMA model. The process concluded at this point, having developed a statistically sound model for time series forecasting (AbdulWahid, & Arunbabu, 2022; Wang, 2018).

2.4 Data collection and Analysis

This study utilized a ten-year dataset (2011–2021) consisting of monthly monitoring records of three key water quality parameters—pH, calcium (mg/L), and iron (mg/L)—from the River Benue. The dataset was obtained from the Adamawa State Ministry of Water Resources through the State Water Board, which operates designated monitoring stations along the river. Measurements were systematically recorded to capture spatial and temporal variations in water quality across the study period.

The study focused on predictive modeling of the three selected parameters to assess long-term trends and variability. Time series methods were employed using EViews 13 (S&P Global) statistical software, which facilitated data preprocessing, stationarity testing, model estimation, and forecast evaluation. The structured approach ensured reliability of results and provided insights into the temporal dynamics of water quality indicators in the River Benue.

3.0 Results and Discussion

3.1 ARIMA Model Stationarity Test

Table 1 presents the results of the Augmented Dickey-Fuller (ADF) test conducted on the time series data for pH, calcium, and iron. The test statistics for all three parameters were considerably more negative than the 5% critical value of –2.8859, confirming that the null hypothesis of a unit root can be rejected.

The stationarity of selected water quality parameters was assessed using the Augmented Dickey-Fuller (ADF) test to determine the presence of unit roots in the time series data. Parameters tested included pH, calcium, and iron. The ADF test statistics for all parameters were substantially more negative than the 5% critical value threshold of approximately -2.8859, indicating a strong rejection of the null hypothesis that the series contains a unit root. Furthermore, the associated p-values were extremely low (0.00001), reinforcing the conclusion of statistical significance. These results confirm that the datasets for all parameters are stationary, implying that their fluctuations over the ten-year period were random and not influenced by persistent trends or long-term systematic changes.

Received: 28-08-2025 / Accepted: 12-09-2025 / Published: 21-10-2025
https://doi.org/10.70118/ujet.2025.0301.01

Table 1: Augmented Dickey–Fuller Stationarity Test Results for Key Water-Quality Parameters

Parameter	ADF Test (Level)	Test Critical (Level)	Stationarity (Level)	5% significant (Level)
pH	-9.49251	-2.885863	Stationary Data	0.00001
Calcium (mg/L)	-8.34001	-2.885863	Stationary Data	0.00001
Iron (mg/L)	-7.52078	-2.885863	Stationary Data	0.00001

3.2 ARIMA Model Identification

Table 2 presents the model selection results for ARIMA models fitted to pH, Calcium (mg/L), and Iron (mg/L) using the Akaike Information Criterion (AIC) and Schwarz Criterion (SC) as performance metrics. The table shows that for the pH parameter, the ARIMA (0,0,1) model recorded the lowest AIC (-3.9984) and SC (-3.9484), indicating the best model fit with minimal complexity. Similarly, for Calcium, the ARIMA (1,0,0) model achieved the lowest AIC and SC values (0.52223), suggesting it provides the optimal balance between fit and simplicity. In the case of Iron, the ARIMA (0,0,1) model was found to be the most suitable, with the lowest AIC (2.194081) and SC (2.263768) values among the evaluated models. These showed that ARIMA (0,0,1) for pH and Iron, and ARIMA (1,0,0) for Calcium, as the most appropriate models for forecasting and further analysis of water quality parameters. These findings align with related studies where ARIMA models were shown to effectively capture temporal dynamics of water quality indicators, providing reliable forecasts for pH, dissolved solids, and heavy metals (Wang, Chen and Tang 2024; Shrestha & Kazama, 2007). The reliance on AIC and SC as selection criteria is consistent with best practices in time series modeling, ensuring models are both parsimonious and robust for environmental data analysis.

Table 2: AIC and SC Values for pH, Calcium, and Iron

Model (ARIMA) ph	Akaike info criterion (AIC)	Schwarz criterion (SC)
(0,0,1)	-3.9984	-3.9484
(1,0,0)	-3.9961	-3.9287
(0,0,0)	-3.9949	-3.9264
(1,0,1)	-3.9832	-3.8903
Model (ARIMA) Calcium (mg/L)	Akaike info criterion (AIC)	Schwarz criterion (SC)
(1,0,0)	0.52223	0.52223
(0,0,1)	0.534675	0.534675
(0,0,0)	0.568687	0.536694
(1,0,1)	0.536694	0.568687
Model (ARIMA) Iron (mg/L)	Akaike info criterion (AIC)	Schwarz criterion (SC)
(0,0,1)	2.194081	2.263768
(1,0,1)	2.209595	2.302512
(1,0,0)	2.23749	2.307177
(0,0,0)	2.346735	2.393193

3.3 ARIMA Model Estimation

Table 3 presents the estimated ARIMA model coefficients for pH, Calcium (mg/L), and Iron (mg/L) parameters, including the intercept (C), autoregressive (AR), and moving average (MA) components, along with the standard error (S.E) of regression. Results indicated that for pH, the model includes all three components with an intercept of 7.23544, a weak AR term (0.15308), a minor MA effect (0.05233), and an S.E of 0.23167, indicating limited unexplained variability. The Calcium model features an intercept of 60.687578 and a small AR term (0.254311), with no MA component, and a relatively high S.E of 18.04508, suggesting significant variation likely due to natural or measurement factors. For Iron, the model includes an intercept of 0.1786807 and a moderately negative MA term (-0.4134143), with no AR component, and a low S.E of 0.055342, reflecting limited unexplained variation. Overall, the ARIMA models indicate differing dynamics for each parameter, with pH and Iron influenced more by moving average components, while Calcium exhibits autoregressive behavior without MA influence. These findings are in line with previous studies, where pH and calcium often display weak autoregressive behavior due to buffering and gradual changes, whereas iron concentrations tend to be more responsive to abrupt external disturbances, showing stronger moving average effects (Shamsuddin & Johari, 2016; Zafra-Mejía *et al.*, 2024). Such consistency reinforces the suitability of ARIMA models for capturing the temporal behavior of different water quality indicators.

Received: 28-08-2025 / Accepted: 12-09-2025 / Published: 21-10-2025

https://doi.org/10.70118/ujet.2025.0301.01

Table 3: ARIMA Model Coefficients for pH, Calcium, and Iron

Parameter	Intercept C	Autoregressive (AR)	Moving Average (MA)	S.E of Regression
pH	7.23544	0.15308	0.05233	0.23167
Calcium (mg/L)	60.687578	0.254311	-	18.04508
Iron (mg/L)	0.1786807	-	0.4134143	0.055342

3.4 ARIMA Model Equations

The ARIMA model equations for pH, Calcium (mg/L), and Iron (mg/L) are presented, with model identification carried out using ACF and PACF plots, followed by estimation and validation through diagnostic checks.

3.4.1 pH – ARIMA (0,0,1)

The ARIMA (0,0,1) model for pH characterizes the series as a stationary moving average process of order one. This means the current value depends mainly on the overall mean and the influence of present and past random shocks, rather than lagged pH values. Equation 6 shows the intercept establishes a long-term mean of about 7.24, which represents stable water quality conditions near neutrality. The moving average coefficient is small and positive, indicating that roughly 5.2% of the previous error contributes to the current value. This suggests the series responds primarily to current random disturbances rather than prolonged effects of past fluctuations. The model shows that river pH remains generally stable around its mean, with minor short-term variations caused by random shocks. Past errors have minimal impact and fade quickly. Tizro *et al.* (2016) demonstrated that ARIMA models can effectively represent water quality parameters, noting that pH typically remains stationary and can be well described using low-order models, which reflects its inherent stability over time. In a similar context, Viccione *et al.* (2023) highlighted the usefulness of ARIMA models in water quality forecasting, as they efficiently account for short-term random fluctuations without the need for extensive differencing, making them well-suited for stable parameters such as pH. Consequently, the ARIMA (0,0,1) model confirms that river pH is generally stable around its mean value of approximately 7.24, with only slight short-term variations driven by transient shocks and minimal carryover from past errors.

$$Intercept\ (c)\ =\ 7.23544;\ MA\ (1)\ coefficient\ \theta_1\ =\ 0.05233.$$

Equation:

$$Y_t\ =\ 7.23544\ +\ \varepsilon_t\ +\ 0.05233\varepsilon_{t-1} \tag{6}$$

3.4.2. Calcium – ARIMA (1,0,0)

The ARIMA (1,0,0) model for Calcium concentration represents a first-order autoregressive process, where the current value depends on its long-term mean, the preceding observation, and a random error term. Equation 7 shows an intercept of 60.69 mg/L, which reflects the baseline or average Calcium level around which the series oscillates. The autoregressive coefficient of 0.25431 means that about 25.4% of the previous value contributes to the current observation, indicating moderate persistence in the data. Since the coefficient is positive, higher Calcium concentrations in one period tend to produce slightly higher values in the next. The model reveals that Calcium concentrations stay relatively steady around 60.69 mg/L, with moderate dependence on past levels. Additional short-term variability comes through random shocks that affect the system. Previous studies have shown that water quality variables, particularly chemical parameters such as Calcium, often exhibit autoregressive patterns. Tizro *et al.* (2016) demonstrated that ARIMA models effectively capture the persistence of chemical water quality indicators, with values depending on past observations. Similarly, Katimon *et al.* (2018) applied ARIMA models to the Johor River Basin and confirmed their suitability for stationary water quality parameters influenced by preceding values. More recently, Zafra-Mejía, *et al.* (2024) emphasized the effectiveness of ARIMA models in forecasting geochemical parameters in groundwater, noting that such parameters remain stable around long-term means while displaying moderate dependence on historical concentrations.

$$Intercept\ (c)\ =\ 60.68758;\ AR\ (1)\ coefficient\ \varphi_1\ =\ 0.25431.$$

Equation:

$$Y_t\ =\ 60.68758\ +\ 0.25431Y_{t-1}\ +\ \varepsilon_t \tag{7}$$

3.43. Iron (mg/L) – ARIMA (0,0,1)

The ARIMA (0,0,1) model for Iron concentration describes the series as a first-order moving average process, where the current value depends on a constant mean, the present random disturbance, and the effect of the previous error term. Equation 8 shows that the intercept of 0.17868 mg/L represents the long-term average Iron concentration, while the negative moving average coefficient of –0.41341 indicates that about

41% of the previous shock influences the current observation in the opposite direction. When Iron concentration runs higher than expected in one period, it tends to push the following value down. When it's lower than expected, it tends to pull the next value up. The parameters q=1 confirm that the series stays stationary and responds mainly to short-term shocks rather than past values, creating fluctuations around a stable mean of roughly 0.18 mg/L.

Both pH and Iron follow MA (1) processes and stay stable with random short-term variations. Calcium behaves differently, following an AR (1) process that shows greater dependence on its own past values and moderate persistence around an average of about 61.85 mg/L. This means short-term shocks, such as runoff or localized discharges, drive fluctuations while the series remains stationary with no long-term persistence (Ayob *et al.*, 2017; Zafra-Mejía *et al.*, 2024). The estimated mean is below the WHO guideline of 0.3 mg/L (Annem, 2017), though episodic events may temporarily increase concentrations above safe thresholds (Ekström *et al.*, 2016). Compared with calcium, which follows an AR (1) process with stronger dependence on past values, both pH and iron follow MA (1) processes, remaining generally stable but sensitive to short-lived shocks.

$$Intercept\ (c) = 0.17868;\ MA\ (1)\ coefficient\ \theta_1 = -0.41341.$$

Equation:

$$Y_t = 0.17868 + \varepsilon_t - 0.41341\varepsilon_{t-1} \tag{8}$$

pH & Iron → ARIMA (0,0,1): values were stable, and primarily influenced by short-term random shocks (MA process). pH and iron gave a random fluctuation with consistent averages in the forecasts (pH ~7.23; Iron ~0.152 mg/L), while Calcium concentration depends more on its own past values than on random shocks, having autoregressive structure fluctuations forecasts of average ~61.85 mg/L). That predicted a sharp cutoff at lag 1, indicating AR (1) was the best fit.

3.5 ARIMA Model Diagnosis

Table 3 shows the ARIMA model diagnostic analysis for pH, Calcium (mg/L), and Iron (mg/L) parameters and it confirms the stability and reliability of the respective models based on the inverse roots of their AR and MA components. For pH, the presence of a single MA (1) root with a modulus of 0.15308 indicates that the ARMA model is invertible, ensuring a stable prediction structure. The Calcium model, which includes AR (1), MA (1), and MA (2) components, also shows all root moduli within the unit circle—0.9928, 0.90018, and 0.46665 respectively—confirming both stationarity and invertibility. This reflects a well-fitting and stable model capable of accurately capturing the temporal behavior of Calcium levels. Similarly, the Iron model includes an MA (1) component with a root modulus of 0.41341, which also lies within the unit circle, indicating invertibility. Collectively, the diagnostic results show that all ARIMA models are structurally sound, with no roots outside the unit circle, validating their use for reliable forecasting and time series analysis of the respective water quality parameters. Similar diagnostics are widely used in time series modelling of water quality: Hardiyanti *et al.* (2020) employed diagnostic checking of ARIMA residuals and root-moduli in their analysis of pH, COD, and BOD, confirming invertibility and stationarity of selected models; Zafra-Mejía *et al.* (2024) in their study of drinking water supply systems likewise report that ensuring all AR and MA roots lie within the unit circle is essential to validate model forecasts. These consistent findings uphold the reliability of your models for forecasting and trend analysis of water quality indicators. ARIMA model diagnostics rely on checking that all AR and MA roots lie inside the unit circle, which ensures stationarity and invertibility—essential for reliable forecasts (Brockwell & Davis, 2016; Ragavan & Fernandez, 2006). Water-quality studies commonly apply this test and often find short-memory MA behaviour in parameters influenced by episodic events (Ayob *et al.*; Hardiyanti, 2020; Zafra-Mejía *et al.*, 2024). While classical ARIMA and seasonal variants remain strong baseline models, recent work shows that SARIMAX, transfer-function, and hybrid ARIMA–machine-learning approaches can outperform them when nonlinear dynamics or external drivers are important (Costa *et al.*, 2023; Wang et al., 2024; García-Guerrero *et al.*, 2025). For monitoring, iron and pH respond mainly to short-term shocks, so real-time sensors and short-lead forecasts are recommended, whereas calcium's more persistent behavior allows less frequent sampling but requires attention to its longer autocorrelation when planning management actions (WHO, 2019).

Table 3: ARIMA Diagnostic – Inverse Roots of AR/MA Polynomials for pH, Calcium, and Iron

Parameter	**Component**	**Root(s)**	**Modulus**	**Cycle**	**Model Property**
pH	MA (1)	-0.15308	0.15308	–	ARMA model is invertible
Calcium (mg/L)	AR (1)	-0.9928	0.9928	–	ARMA model is stationary
	MA (1)	-0.90018	0.90018	–	ARMA model is invertible
	MA (2)	-0.46665	0.46665	–	ARMA model is invertible
Iron (mg/L)	MA (1)	-0.41341	0.41341	–	ARMA model is invertible

No root lies outside the unit circle for all components, indicating that all ARIMA models are stable (stationary and/or invertible as applicable).

3.5 ARIMA Model Forecasting

The Autoregressive Integrated Moving Average (ARIMA) model is a powerful statistical tool commonly used for forecasting time series data. It combines autoregressive (AR), differencing (I), and moving average (MA) components to capture both short-term fluctuations and long-term patterns in sequential datasets. In water quality studies, ARIMA provides reliable forecasts of parameters such as pH, calcium, and iron, offering valuable insights for sustainable resource planning and management.

3.5.1 Ph Model forecasting

The ten-year ARMA forecast for pH (Jan. 2023–Dec. 2032) as shown in Figure 2 projects a relatively stable trend with mean predicted pH ~ 7.23, indicating little expected drift in the environmental acidity/alkalinity of the River Benue under current conditions. The stability in pH forecasts is supported by analogous studies; for example, Hardiyanti *et al.* (2020) forecast pH using ARIMA for an Indonesian river and likewise found forecasted pH values staying near neutral/stable ranges (7.40) over multi-year horizons, suggesting buffering capacities in similar freshwater systems. The stability in your forecast accords with findings that in many river water quality time series, pH tends to revert and exhibits minimal systematic trend unless impacted by external perturbations (e.g. pollution, land use change). Nonetheless, as with others (Hardiyanti *et al.*, 2020), you should consider that unexpected changes (anthropogenic discharge, climate shifts) may alter future dynamics, so forecasts are best used as baseline expectations rather than guarantees. Comparable stability has been documented in rivers with strong natural buffering capacity in Malaysia and Indonesia (Ayob *et al.*, 2017; Hardiyanti, 2020), in European catchments (Ragavan & Fernandez, 2006) and in large North American rivers (Stets et al., 2015). pH is a key regulator of aquatic chemistry: even modest departures from the recommended 6.5–8.5 range can increase the solubility and toxicity of metals such as iron and aluminum, alter nutrient cycling, and affect drinking-water aesthetics and corrosivity (WHO, 2019; Wang, 2016; Ekström *et al.*, 2016). Similar impacts on aquatic life and infrastructure have been highlighted by Saalidong *et al.* (2022) and García-Guerrero *et al.* (2025). This study provides one of the few decade-scale ARMA forecasts for a major West African river and explicitly links rigorous ARIMA diagnostics to actionable monitoring guidance, bridging statistical modelling and water-resources management. There is need to add exogenous drivers like rainfall and discharge using SARIMAX or transfer-function models (Costa *et al.*, 2023; Zafra-Mejía *et al.*, 2024). Hybrid ARIMA–machine-learning models are needed to capture nonlinear responses (Wang *et al.*, 2024). Short-lead forecasts should be integrated into automated early-warning systems with real-time sensors (Bownik, 2021).

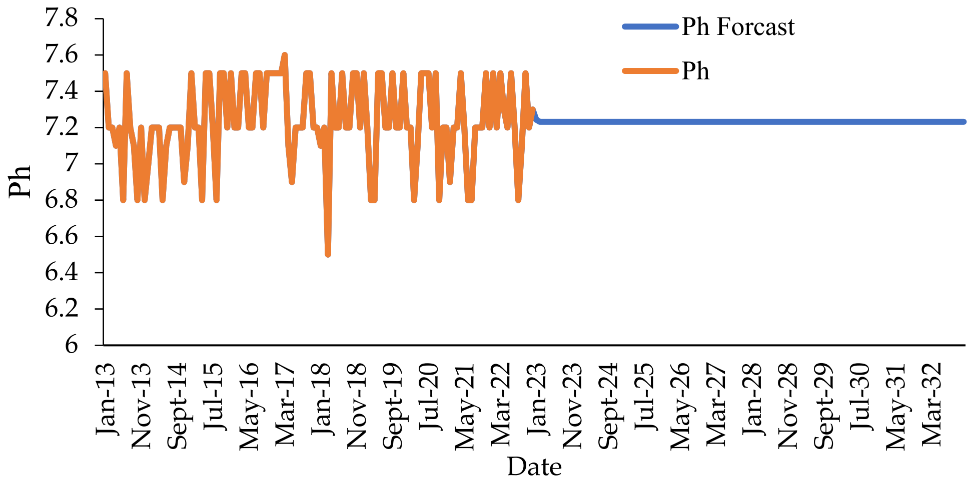

Figure 2: Ph ARMA Model forecast for Ten (10) Years

3.5.2 Calcium Model forecasting

Figure 3 shows a ten-year ARMA forecast for calcium, with an average concentration of about 61.85 mg L^{-1}. The model points to regular seasonal ups and downs but no long-term drift. Similar behaviour has been observed elsewhere: in Bogotá, Zafra-Mejía *et al.* (2024) found that calcium and related ions follow a moderate, seasonally predictable pattern, and Veerendra *et al.* (2023) reported comparable cycles in Indian surface waters linked to rainfall and agricultural runoff. Long-term studies from Europe and North America also show that natural carbonate weathering helps keep calcium levels steady over decades (Neal & Kirchner, 2000; Stets *et al.*, 2015). Calcium is more than just a number on a chart: it is a key hardness ion that helps buffer pH, limits the solubility of toxic metals, and affects both aquatic life and water-supply infrastructure. Levels around 60 mg L^{-1} fit comfortably within World Health Organization guidelines, providing some protection against pipe corrosion—although very high hardness can lead to scaling (WHO, 2019; McGowan *et al.*, 2021). This study adds one of the few decade-long ARMA forecasts for a major West African river, giving managers a new evidence base for hardness control and treatment planning. For now, calcium levels look likely to stay broadly stable while responding to seasonal rainfall and runoff. But forecasts built only from past patterns can miss sudden changes from land-use shifts, extreme events, or unmonitored pollution. To keep predictions robust, future work should update the models regularly, include external drivers such as rainfall and discharge through SARIMAX or transfer-function approaches (Costa *et al.*, 2023; Zafra-Mejía *et al.*, 2024), and test hybrid ARIMA–machine-learning models to capture more complex, nonlinear responses (Wang *et al.*, 2024; García-Guerrero *et al.*, 2025).

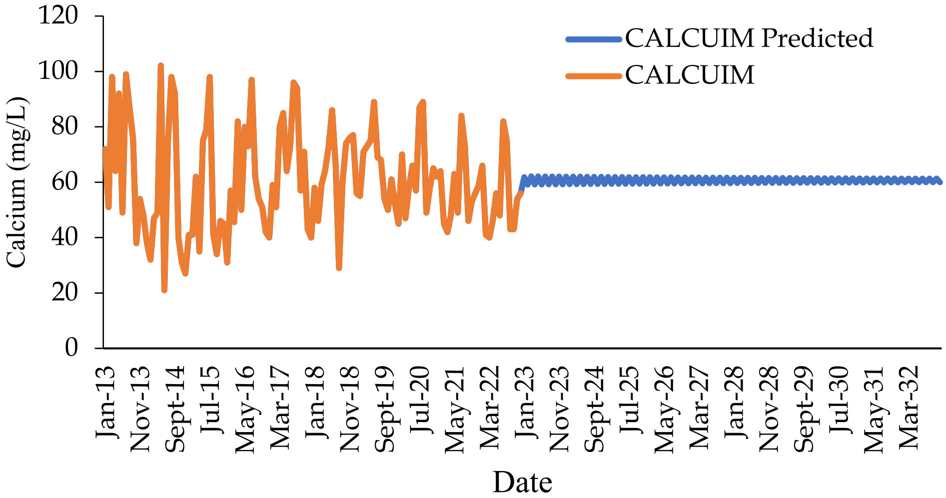

Figure 3: Calcium (mg/l) ARMA Model forecast for Ten (10) Years

3.5.1 Iron Model forecasting

Figure 4 shows the ten-year ARMA forecast for dissolved iron, with an average of about **0.15 mg L^{-1}** and only small short-term ups and downs. This steady pattern is very similar to what other long-term water-quality studies have found. For example, in Bogotá's drinking-water system, Zafra-Mejía *et al.* (2024) observed that iron behaved like a typical trace metal, with much smoother seasonal cycles than more variable indicators such as turbidity. Likewise, monitoring at India's Krishnagiri Reservoir showed that iron levels changed only slightly from month to month compared with parameters like chlorophyll-a (Abdul Wahid & Arunbabu, 2022). Long-term data from North American rivers point to the same explanation: slow geochemical processes such as redox cycling and carbonate buffering tend to keep dissolved iron stable over time (Stets *et al.*, 2015). Even though the forecasted levels are well below World Health Organization aesthetic guidelines, iron still matters for both water treatment and the ecosystem. Sudden spikes often triggered by heavy rains, dredging or other disturbances can stain pipes, clog treatment systems and even mobilize other metals (WHO, 2019; Hu *et al.*, 2019; Krueger *et al.*, 2020).

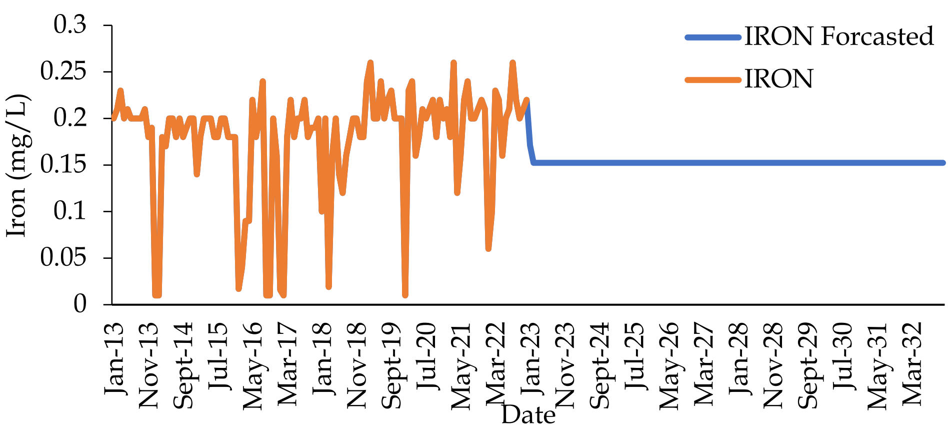

Figure 4: Iron (mg/l) ARMA Model forecast for Ten (10) Years

Received: 28-08-2025 / Accepted: 12-09-2025 / Published: 21-10-2025

https://doi.org/10.70118/ujet.2025.0301.01

This work is one of the first decade-long ARMA forecasts for iron in a major West African river, giving water managers a valuable baseline for planning treatment and monitoring. Looking ahead, the models could be made more powerful by adding external drivers such as rainfall and river discharge using SARIMAX or transfer-function approaches (Costa *et al.,* 2023; Zafra-Mejía *et al.,* 2024), by testing hybrid ARIMA–machine-learning methods to capture sudden or nonlinear changes (Wang *et al.,* 2024), and by linking short-lead forecasts to automated sensor networks for real-time early warning of iron spikes (Bownik, 2021). Regularly updating and re-validating the models will also be essential as land use and climate continue to change.

4.0 Conclusion

This study concludes that ARIMA time-series modelling can be a practical, easy-to-use tool for keeping track of the River Benue's water quality. By analysing ten years of monthly data (2011–2021) on pH, calcium and iron, we found simple but reliable models: a moving-average type model (ARIMA 0,0,1) works best for pH and iron, while calcium follows a first-order autoregressive model (ARIMA 1,0,0). All three passed the usual statistical checks, meaning their forecasts can be trusted. Looking ahead to 2023–2032, the river is expected to remain chemically stable: pH should stay close to neutral at about 7.2, calcium will fluctuate mildly around 62 mg L^{-1}, and iron will hold near 0.15 mg L^{-1}—comfortably within World Health Organization guidelines. In simple terms, pH and iron mainly react to short-term disturbances such as brief runoff events, while calcium shows a slightly stronger "memory" of its own past levels. For water managers and local authorities, these forecasts offer a solid baseline for planning treatment and monitoring. But it's important to remember that ARIMA models only look at past patterns. Sudden changes from land-use shifts, pollution events or climate extremes can quickly upset these trends. To keep forecasts useful, the models should be updated regularly and, where possible, improved by including outside drivers like rainfall or river discharge, or by blending ARIMA with modern machine-learning methods. Coupled with real-time sensors, such an approach can provide early warnings and support sustainable, long-term stewardship of the River Benue's water resources.

References

Abdul Wahid, A., & Arunbabu, E. (2022). Forecasting water quality using seasonal ARIMA model by integrating in-situ measurements and remote sensing techniques in Krishnagiri reservoir, India. *Water Practice & Technology*, *17*(5), 1230-1252.

Adaeze, J., Dishan, E., & Tella, I. (2017). Plants species diversity along river benue bank under the influence of siltation and solid waste effluents, Adamawa state, Nigeria. *Open Access Library Journal*, *4*(12), 1-13.

Adebayo, A., & Tukur, A. (1999). Climate, sunshine, temperature, evaporation and relative humidity. *Adamawa state in maps*, 20-22.

Ahamad, K., Nazary, B., Mudoi, P., Rani, R., Bharati, V., & Singh, H. (2015). Application of regression analysis for surface water quality modeling. *IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)*, 65-70.

Aho, M., Akpen, G., & Ekwule, O. (2018). Predictive regression models of water quality parameters for river Amba in Nasarawa State, Nigeria. *Intl. J. Innov. Eng. Sci. Res*, *2*, 24-33.

Annem, S. (2017). *Determination of iron content in water*. [Capstone Project]. Minnesota State University.

Ayob, K., Shahid, S., & Mohsenipour, M. (2018). Modeling water quality and hydrological variables using ARIMA: a case study of Johor River, Malaysia. *Sustainable Water Resources Management*, *4*, 991-998. https://doi.org/10.1007/s40899-017-0202-8

Brockwell, P. J., & Davis, R. A. (2016). Introduction to Time Series and Forecasting (3rd ed.). Springer.

Bownik, A., & Wlodkowic, D. (2021). Advances in real-time monitoring of water quality using automated analysis of animal behaviour. *Science of the Total Environment*, *789*, 147796.

Chakraborty, T., Chakraborty, A. K., & Mansoor, Z. (2019). A hybrid regression model for water quality prediction. *Opsearch*, *56*, 1167-1178.

Costa, G. E. D. M. E., Menezes Filho, F. C. M. D., Canales, F. A., Fava, M. C., Brandão, A. R. A., & de Paes, R. P. (2023). Assessment of Time Series models for Mean Discharge modeling and forecasting in a Sub-basin of the Paranaíba River, Brazil. *Hydrology*, *10*(11), 208.

Ejiohuo, O., Onyeaka, H., Akinsemolu, A., Nwabor, O. F., Siyanbola, K. F., Tamasiga, P., & Al-Sharify, Z. T. (2025). Ensuring water purity: Mitigating environmental risks and safeguarding human health. *Water Biology and Security*, *4*(2), 100341.

García-Guerrero, J. C., Álvarez-Alvarado, J. M., Carrillo-Serrano, R. V., Palos-Barba, V., & Rodríguez-Reséndiz, J. (2025). A Review of Water Quality Forecasting Models for Freshwater Lentic Ecosystems. *Water*, *17*(15), 2312.

Hardiyanti, D. Y. (2020). Forecasting river water quality using Autoregressive Integrated Moving Average (ARIMA). In Proceedings of the International Conferences on Information System and Technology (CONRIST 2019), 158–163. https://www.scitepress.org/Papers/2019/99072/99072.pdf

Received: 28-08-2025 / Accepted: 12-09-2025 / Published: 21-10-2025
https://doi.org/10.70118/ujet.2025.0301.01

Hu, R., Gwenzi, W., Sipowo-Tala, V. R., & Noubactep, C. (2019). Water treatment using metallic iron: A tutorial review. *Processes*, *7*(9), 622.

Jaffar, O., Najah, A., Elshafie, A., & Karim, O. A. (2009). Prediction of Johor River water quality parameters using artificial neural networks. *European journal of scientific research*, *28*(3), 422-435.

Jeziorski, A., & Smol, J. P. (2017). The ecological impacts of lakewater calcium decline on softwater boreal ecosystems. *Environmental Reviews*, *25*(2), 245-253.

Katimon, A., Shahid, S., & Mohsenipour, M. (2018). Modeling water quality and hydrological variables using ARIMA: a case study of Johor River, Malaysia. *Sustainable Water Resources Management*, *4*(4), 991-998.

Krueger, K. M., Vavrus, C. E., Lofton, M. E., McClure, R. P., Gantzer, P., Carey, C. C., & Schreiber, M. E. (2020). Iron and manganese fluxes across the sediment-water interface in a drinking water reservoir. *Water research*, *182*, 116003.

Neal, C., & Kirchner, J. W. (2015). Sodium and chloride levels in rainfall, mist, streamwater and groundwater at the Plynlimon catchments, mid-Wales: inferences on hydrological and chemical controls. *Hydrology and Earth System Sciences*, *4*(2), 295-310.

Ragavan, A. J., & Fernandez, G. C. (2006). Modeling water quality trend in long-term time series. *Proceedings of SAS Users Group Meeting (SUGI 31)*, Paper 205.

Saalidong, B. M., Aram, S. A., Otu, S., & Lartey, P. O. (2022). Examining the dynamics of the relationship between water pH and other water quality parameters in ground and surface water systems. *PloS one*, *17*(1), e0262117.

Sentas, A., Psilovikos, A., Karamoutsou, L., & Charizopoulos, N. (2018). Monitoring, modeling, and assessment of water quality and quantity in River Pinios, using ARIMA models. *Desalination and Water Treatment*, *133*, 336-347.https://doi.org/10.5004/dwt.2018.23239

Shrestha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environmental Modelling & Software, 22*(4), 464–475. https://doi.org/10.1016/j.envsoft.2006.02.001

Stellwagen, E., & Tashman, L. (2013). ARIMA: The Models of Box and Jenkins. *Foresight: The International Journal of Applied Forecasting*, (30).

Stets, E. G., et al. (2015). Long-term trends in alkalinity in large rivers. *Science of the Total Environment, 488–489*, 280–289.

Tukur, A., Bashir, B., & Mubi, A. (2004). Agricultural land use and land degradation in Adamawa State. *Global Journal of Environmental Sciences*, *3*(1), 27-31.

Twinomuhangi, M. B., Bamutaze, Y., Kabenge, I., Wanyama, J., Kizza, M., Gabiri, G., & Egli, P. E. (2025). Analysis of stationary and non-stationary hydrological extremes under a changing environment: A systematic review. *HydroResearch*.

UNESCO. (2005). *Water Resources Systems Planning and Management.* S. a. C. O. United Nations Educational.

Veerendra, G. T. N., Kumaravel, B., Rao, P. K. R., Dey, S., & Manoj, A. P. (2024). Forecasting models for surface water quality using predictive analytics. *Environment, Development and Sustainability*, *26*(6), 15931-15951.

Water Practice & Technology. (2022). Forecasting water quality using seasonal ARIMA. Water Practice & Technology, 17(5), 1230–1241. https://doi.org/10.2166/wpt.2022.075

Wang, T., Chen, W., & Tang, B. (2024). Water quality prediction using ARIMA–SSA–LSTM hybrid models. Water Science, 24(4), 1282–1296. https://iwaponline.com/ws/article/24/4/1282/101339/Water-quality-prediction-using-ARIMA-SSA-LSTM

Wang, G. C. (2018). A guide to Box-Jenkins modeling. *The Journal of Business Forecasting*, *27*(1), 19.

Wang, Z. (2016). Metal toxicity to freshwater organisms as a function of pH. *Environmental Research, 151*, 356–365.

World Health Organization. (2019). *Guidelines for drinking-water quality* (Vol. 1). World health organization.

XiaoYing, Y., Liu Qun, L. Q., Luo XingZhang, L. X., & Zheng Zheng, Z. Z. (2017). Spatial regression and prediction of water quality in a watershed with complex pollution sources.

Zafra-Mejía, C. A., Rondón-Quintana, H. A., Urazán-Bonells, C. F. (2024). ARIMA and TFARIMA analysis of the main water quality parameters in the initial components of a drinking water supply system. *Hydrology, 11*(1), Article 10

Zhu, G., Zhang, S., Bian, Y., & Hursthouse, A. S. (2021). Multi-linear regression model for chlorine consumption by waters. *Environmental Engineering Research*, *26*(4).

Received: 28-08-2025 / Accepted: 12-09-2025 / Published: 21-10-2025
https://doi.org/10.70118/ujet.2025.0301.01