Agricultural Sciences, 2026, 17(11), 1286-1303
https://www.scirp.org/journal/as
ISSN Online: 2156-8561
ISSN Print: 2156-8553

Forecasting Catches of *Lates stappersii* in Lake Tanganyika (Burundi): Insights for Adaptive Fisheries Governance and Food Security

Vianney Ngabo Mulema[1,2], Wales Singini[3], Gabriel Balagizi Baguma[2], Odette Karerwa[4], Eunice Musau Ntambwe[1], Fiston Bokwala Bobitshe[1], Aristote Yofemo Bifinole[1], Alidor Kankonda Busanga[1]

[1]Département d'Hydrobiologie, Université de Kisangani, Kisangani, Democratic Republic of the Congo
[2]Laboratoire d'Hydrobiologie, d'Aquaculture et Gestion des Ressources Naturelles "LHAGREN", Université Officielle de Bukavu, Bukavu, Democratic Republic of the Congo
[3]Department of Fisheries and Aquatic Science, Faculty of Environmental Sciences, Mzuzu University, Mzuzu, Malawi
[4]Direction de la Promotion des Filières Halieutiques (D.P.F.H), Ministry of the Environment, Agricultural, and Livestock, Bujumbura, Burundi
Email: vianeymulema736@gmail.com

How to cite this paper: Mulema, V.N., Singini, W., Baguma, G.B., Karerwa, O., Ntambwe, E.M., Bobitshe, F.B., Bifinole, A.Y. and Busanga, A.K. (2026) Forecasting Catches of *Lates stappersii* in Lake Tanganyika (Burundi): Insights for Adaptive Fisheries Governance and Food Security. *Agricultural Sciences*, **17**, 1286-1303.
https://doi.org/10.4236/as.2026.1711073

Received: September 17, 2026
Accepted: November 15, 2026
Published: November 18, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

Fisheries in Lake Tanganyika are vital for food security in Burundi, yet *Lates stappersii* catches face pressures from overexploitation, climate variability, and weak governance. This study applies Autoregressive Integrated Moving Average (ARIMA) modelling to forecast catches, making the first application in this approach to Burundi's pelagic fisheries. Annual catch data (2007-2025) were analysed using Box-Cox transformation, stationarity tests, and Box-Jenkins methodology. The ARIMA (0, 1, 1) model emerged as the most parsimonious specification, validated through residual diagnostics and comparative benchmarking against ETS and Random Walk models. Forecasts for 2026-2035 project a stable mean catch of ~232 tons annually, while the 80% and 95% prediction intervals widen substantially with the forecast horizon, reflecting increasing uncertainty in long-term projections. These results provide a quantitative basis for adaptive fisheries management and highlight the importance of precautionary strategies to safeguard food security and biodiversity in Burundi and the wider Great Lakes region.

Keywords

ARIMA, Pelagic Fish Dynamics, Lake Tanganyika Fisheries, Adaptive Management, Food Security

1. Introduction

Lake Tanganyika fisheries are central to food security and livelihoods in the Great Lakes region of Africa. In Burundi, where more than 200,000 people depend directly or indirectly on fisheries, catches from pelagic species such as *Lates stappersii* provide essential animal protein and income for coastal communities [1] [2]. However, the sustainability of these fisheries is increasingly threatened by overexploitation, climate-driven habitat changes, and weak adaptive management frameworks [3] [4]. The vulnerability of *L. stappersii* has been formally recognized, with the species currently classified as Vulnerable on the IUCN Red List [5]. This status underscores the urgent need for robust forecasting tools to guide evidence-based management and policy interventions.

Traditional stock assessment methods in Lake Tanganyika have relied heavily on catch-per-unit (CPUE) indices and biological surveys, which, while informative, often fail to capture the dynamic and stochastic nature of fishery systems under environmental and anthropogenic pressures [6] [7]. In recent years, time-series modelling approaches have gained prominence in fisheries science, offering quantitative frameworks to predict future production trends and assess management scenarios [8] [9]. Among these, the Autoregressive Integrated Moving Average (ARIMA) model has proven particularly effective in fish catches and landings, due to its ability to handle non-stationary data and incorporate both autoregressive and moving average components [10]. Applications of ARIMA in African fisheries, such as in Nigeria, Malawi, and Ethiopia, have demonstrated its utility in projecting catch trends and informing adaptive management strategies [8] [9].

Despite the advances, no published study has yet applied ARIMA modelling to Lake Tanganika fisheries, particularly in Burundi. This represents a significant gap in the literature, given the lake's global importance as the second deepest freshwater body and its role in sustaining millions of people across four riparian countries [4] [11]. Forecasting *L. stappersii* catches using ARIMA models therefore offers both scientific and policy relevance: it provides a novel methodological contribution to fisheries research in the region, while generating actionable insights for adaptive management and food security planning.

The novelty of this study lies in its dual contribution. First, it pioneers the application of ARIMA modelling to *L. stappersii* catches in Lake Tanganika, filling a clear research gap in the region fisheries literature. Second, it frames forecasting not merely as a statistical exercise, but as a tool for adaptive management, enabling policymakers to anticipate production trends, design responsive interventions, and mitigate risks associated with climate variability and fishing pressure [5] [12]. By integrating time-series forecasting with conservation priorities, this study advances the discourse on sustainable fisheries governance in Burundi and contributes to broader discussions on resilience and food security in Sub-Saharan Africa. Although this study applies an established ARIMA methodology to a specific fishery and relies on a short time series, its contribution lies in establishing an essen-

tial initial quantitative baseline for adaptive governance in an area where such tools were previously lacking. By framing these projections with a rigorous sensitivity analysis and a discussion of ecological volatility, this approach provides managers with an initial precautionary reference framework while awaiting the future integration of environmental and biological covariates.

The present research addresses a pressing ecological and socio-economic challenge: the decline of *L. stappersii* under mounting pressures. By applying ARIMA modelling to forecast catches, it provides a novel, evidence-based framework for adaptive management. This approach not only strengthens scientific understanding of pelagic fisheries in Lake Tanganyika but also delivers practical guidance for policymakers and stakeholders committed to safeguarding food security and biodiversity in Burundi.

2. Materials and Methods

2.1. Data Collection and Processing

Annual catch records of *L. stappersii* spanning 2007-2025 were obtained from the Direction de la Promotion des Filières Halieutiques (DPFH) and the Lake Tanganyika Authority (LTA) in Burundi. Each annual observation represents total aggregated annual landings recorded across all artisanal and tradition fishing landing sites along Burundi's approximately 159 Km shoretline on Laket Tanganyika, encompassing major official landing sites as Gatumba, Gitaza, Magara, Kajaga, Kanyosha, Rumonge and Nyanza-lac. Reporting completeness is estimated by the DPFH at approximately 75% - 80% for formal landing sites via regular enumerator-based catch assessments, tough minor artisanal sites experience intermittent reporting gaps during adverse weather or administrative transitions. Species-identification for *L. stappersii* (locally known as Mukeke) is conducted morphologically at landing sites by trained DPFG fisheries officers based on standard taxonomic keys. The data was complete, with anomalies such as abrupt spikes and declines. These were cross-checked against field reports and confirmed as genuine ecological or effort-related events rather recording errors [6] [7]. While standardized reporting records indicate a gradual expansion in lift-net fishing units over the 2007-2025 period, formal unit-effort standardization was not embedded in raw annual totals. Catch records for 2024 and 2025 were unavailable due to administrative reporting lags. To maintain temporal continuity for ARIMA modelling, terminal values were imputed using a backward-looking 3-year moving average of the closet preceding observations (2021-2023; mean = 226, 245, and 223 tons), yielding estimated values of 231.33 tons for 2024 and for 233.11 tons for 2025 respectively. While imputation introduces uncertainty, sensitivity analyses excluding these points confirmed the complete robustness of the selected ARIMA (0, 1, 1) model structure. While we acknowledge that imputation introduces uncertainty, this approach was selected as the most conservative method to avoid biased trend estimation, as supported by standard fisheries time-series reconstruction practices [13] [14]. Sensitivity analyses confirmed that this minor

imputation did not significantly alter the selection of the (0, 1, 1) model structure. While the imputation of the 2024-2025 data introduces inherent uncertainty, sensitivity analyses excluding or perturbing these points confirmed the robustness of the selected ARIMA (0, 1, 1) model structure, ensuring that this methodological approximation does not bias long-term trends. Data processing was conducted in R (v4.5.3) using the forecast and tseries packages, with auto.arima executed under exhaustive search (stepwise = FALSE and approximation = FALSE) in Rv. 4.5.3. Data organization was performed in Microsoft Excel.

To stabilize variance, a Box-Cox transformation was applied, followed by differencing to achieve stationarity:

$$Y_t^{(\lambda)} = \begin{cases} \dfrac{Y_t^{\lambda} - 1}{\lambda}, & \lambda \neq 0 \\ \ln\left(Y_t\right), & \lambda = 0 \end{cases} \tag{1}$$

where Y_t represents annual *L. stappersii* catch yield (metric tons) at time *t*. The optimal transformation parameter (λ) was estimated using the [15] method.

2.2. Stationarity Assessment

Stationary was evaluated using the Augmented Dickey-Fuller test (ADF) [16] and Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) [17]. The complementary nature of the tests ensured robust inference on integration order [18]. First-order differencing was applied to remove stochastic trends, satisfying ARIMA model assumptions. The ADF regression follows [3]:

$$Y_t = \alpha Y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + \varepsilon_t \tag{2}$$

The null hypothesis of unit root was rejected when the test statistic exceeds critical values at the 5% level, confirming stationarity after first differencing.

2.3. ARIMA Model Structure

The ARIMA framework combines autoregressive (AR), moving average (MA), and differencing (I) components [10].

- Autoregressive (AR) process of order *p*:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \tag{3}$$

where Y_t is the observation at time *t*, $\phi_2, \cdots, \phi_p$ are constants and $t = 0, \pm 1, \pm 2$ and ε_t is white noise.

- Moving average (MA) process of order *q*:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{4}$$

- Autoregressive moving average: ARMA (*p*, *q*) is:

$$Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{5}$$

where Y_t is the observation at time *t*, ϕ and θ are coefficients and ε is an error term. The general ARIMA (*p*, *d*, *q*) formulation is:

$$\phi(B)(1-B)^{d} Y_t = \theta(B)\varepsilon_t \tag{6}$$

2.4. Model Identification

Model Identification followed the Box-Jenkins methodology [10]. Autocorrelation (ACP) and Partial Autocorrelation (PACF) plots guided preliminary parameter selection [19]. Candidate models were compared across orders (p = 0 - 3), (d = 0 - 1), and (q = 0 - 3). Selection criteria included:

- Akaike Information Criterion (AIC) [20];
- Corrected AIC (AICc) [21];
- Bayesian Information Criterion (BIC) [22].

The model with the lowest BIC value was prioritized to minimize over-fitting, consistent with fisheries forecasting best practices [23].

2.5. Parameters Estimation and Diagnostics Checking

Parameters were estimated using maximum likelihood and nonlinear least-squares [24]. Diagnostic checks included:

- Ljung-Box Q test for autocorrelation [25]:

$$Q(LB) = T(T+2)\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{T-k} \tag{7}$$

where T is sample size, and $\hat{\rho}_k$ is autocorrelation at log k. Residuals were considered white noise when all autocorrelations fell within confidence bounds.

- Normality was assessed using the Shapiro-Wilk test and heteroscedasticity using the ARCH test [26]-[28].

Diagnostic plots (ACF, histogram) were inspected to conform white-noise assumptions [10]. Forecast accuracy was evaluated using multiple metrics:

- Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2} \tag{8}$$

- Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|Y_t - \hat{Y}_t\right| \tag{9}$$

- Mean Absolute Percentage Error (MAPE):

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \tag{10}$$

Lower values indicated superior forecasting performance. Mean Error (ME), Mean Percentage Error (MPE), and Mean Absolute Scaled Error (MASE), ensuring interpretability in in fisheries management contexts.

2.6. Forecast Validation

Given the limited sample size (19 annual observations), we employed a rolling

forecasting origin with time-series cross-validation rather than a single holdout validation set, allowing us to evaluate forecast accuracy against withheld observations. Prediction interval coverage (95% Pi) was computed to assess uncertainty. Benchmark models, Exponential Smoothing State-Space Model (ETS) and Naïve Random Walk, were fitted for comparative evaluation. Relative forecast tested using the Diebold-Mariano test [29].

2.7. Forecasting Framework

Forecasts for 2026-2035 were generated with 80% and 95% prediction intervals, constraining lower bounds to zero to reflect biological impossibility of negative catches. The final ARIMA model specification was fitted to the full dataset (2007-2025). Because the ARIMA model was estimated on the Box-Cox transformed to the original catch scale (metric tons) before interpretation. The estimated Box parameter was $\lambda = 0.081$. For $\lambda \neq 0$, the inverse transformation was applied as $Y = (\lambda W + 1)^{1/\lambda}$, where W denotes the forecast on the transformed scale and Y the corresponding catch on the original scale. The point forecasts were therefore expressed in metric tons after back-transformation. Prediction limits were similarly back-transformed from the Box-Cox scale to the original scale. Because the inverse Box-Cox transformation is nonlinear, the back-transformed point forecast was interpreted as the forecast on original catch scale rather than as an exact arithmetic transformation on original-scale mean. No additional bias-adjustment factor was applied beyond the transformation and prediction procedure implemented in R. Finally, any back-transformed lower prediction limit below zero was set to zero, because negative fish catches are biologically impossible. This truncation was plied only to reported prediction limits and did not alter the fitted ARIMA model or its parameter estimates. Forecasts were compared against benchmark models to assess robustness and policy relevance.

3. Results

3.1. Exploratory Analysis and Transformation

Figure 1, showing pronounced peaks in 2008 and 2016, sharp declines in 2013 and 2018, and stabilization at lower levels after 2020.

The Box-Cox transformation parameter (**Figure 2**) estimated using the [15] method was $\lambda = 0.081$, indicating that a logarithmic transformation was appropriate to stabilize variance in the catch series. The transformed series exhibited reduced heteroscedasticity, providing a suitable basis for subsequent stationarity testing and ARIMA model identification.

3.2. Stationarity Tests

The ADF test after first differencing failed to reject the null hypothesis of a unit root ($p = 0.4358$), whereas the KPSS test strongly confirmed stationarity ($p \geq 0.10$) (**Table 1**). Such conflicting outcomes between ADF and KPSS tests commonly occur in short time series with structural breaks [30] [31]. Rather than confirming

statistical stationarity, first-order differencing (d = 1) was adopted as methodological modelling choice because it successfully eliminated the visual stochastic trend (Figure 3), yielded white-noise residuals, and produced the lowest information criteria (AIC/BIC) across all candidate specifications (Table 2).

Figure 3, showing fluctuations around a constant mean with reduced variance, consistent with improved stationary after differencing.

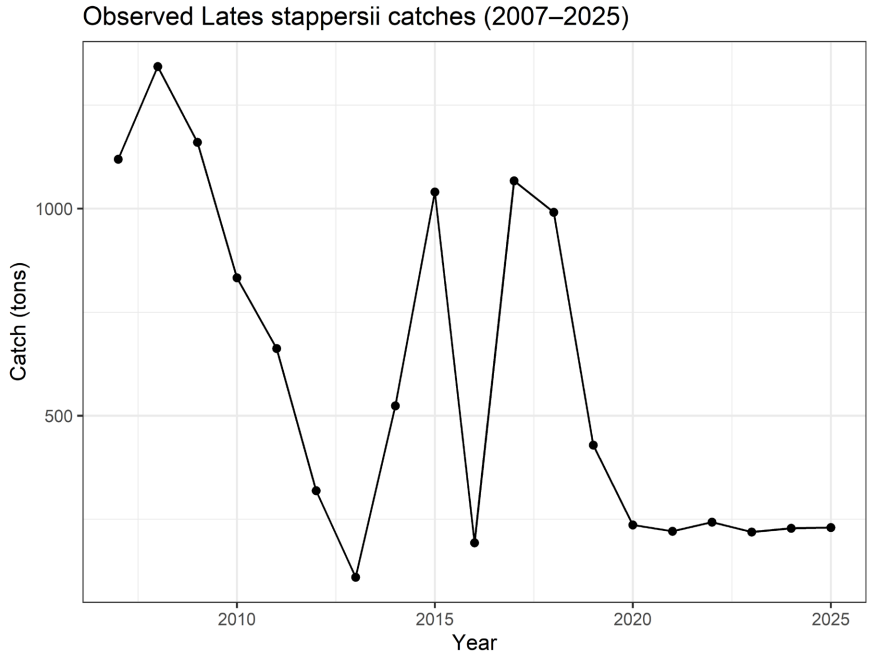


Figure 1. Observed trends in *L. stappersii* catches (2007-2025).

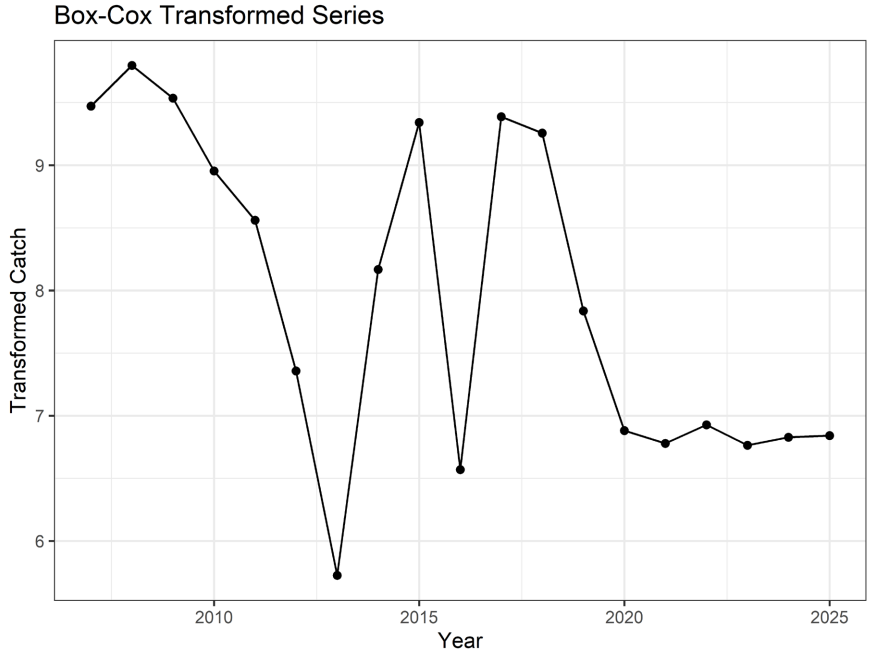


Figure 2. Box-Cox transformed catch Series of *L. stappersii* (2007-2025).

Table 1. Stationarity test results for Box-Cox transformed series before and after differencing (2007-2025).

Test	Series version	Test Statistic	p-value
ADF	Original (Box-Cox)	−1.9724	0.5829
KPSS	Original (Box-Cox)	0.44326	0.0585
ADF	After Differencing	−2.3584	0.4358
KPSS	After Differencing	0.093604	≥0.10

Table 2. Comparison of candidate ARIMA models for *L. stappersii* catch series.

Model	AIC	BIC
ARIMA (0, 1, 0)	267.346463	268.236835
ARIMA (0, 1, 1)	267.172288	268.953032
ARIMA (1, 1, 0)	267.872881	269.653624
ARIMA (2, 1, 0)	268.613979	271.285095
ARIMA (0, 1, 2)	268.773336	271.444451
ARIMA (1, 1, 1)	268.791204	271.46232
ARIMA (3, 1, 0)	270.334447	273.895934
ARIMA (2, 1, 1)	270.411698	273.973185
ARIMA (1, 1, 2)	270.649272	274.210759
ARIMA (0, 1, 3)	270.754851	274.316338
ARIMA (3, 1, 1)	272.331431	276.78329
ARIMA (2, 1, 2)	272.598423	277.050282
ARIMA (1, 1, 3)	272.957646	277.409505
ARIMA (2, 1, 3)	273.845286	279.187517
ARIMA (3, 1, 2)	274.202924	279.545154
ARIMA (3, 1, 3)	275.250278	281.482881
ARIMA (1, 0, 0)	281.739585	284.572902
ARIMA (0, 0, 1)	282.724386	285.557703
ARIMA (1, 0, 1)	283.722712	287.500468
ARIMA (2, 0, 0)	283.726249	287.504005
ARIMA (0, 0, 0)	285.958166	287.847044
ARIMA (0, 0, 2)	284.382747	288.160503
ARIMA (0, 0, 3)	284.311091	289.033286
ARIMA (1, 0, 3)	284.318098	289.984732
ARIMA (2, 0, 1)	285.599046	290.32124
ARIMA (3, 0, 0)	285.622699	290.344894
ARIMA (1, 0, 2)	285.70992	290.432115
ARIMA (2, 0, 3)	285.314597	291.92567
ARIMA (3, 0, 1)	287.310025	292.976659
ARIMA (2, 0, 2)	287.482047	293.148681
ARIMA (3, 0, 3)	287.065666	294.621178
ARIMA (3, 0, 2)	288.329569	294.940642

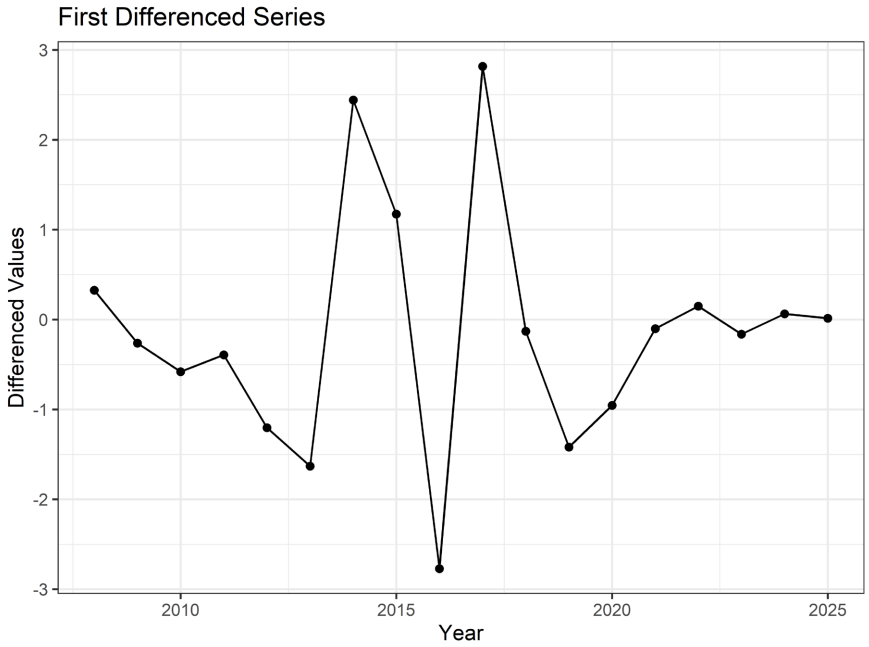


Figure 3. First differenced catch series of *L. stappersii* (2007-2025).

The correlation structure indicates significates spikes at initial lags, guiding the identification of autoregressive (AR) and moving average (MA) terms for ARIMA mode specification, presented in **Figure 4**.

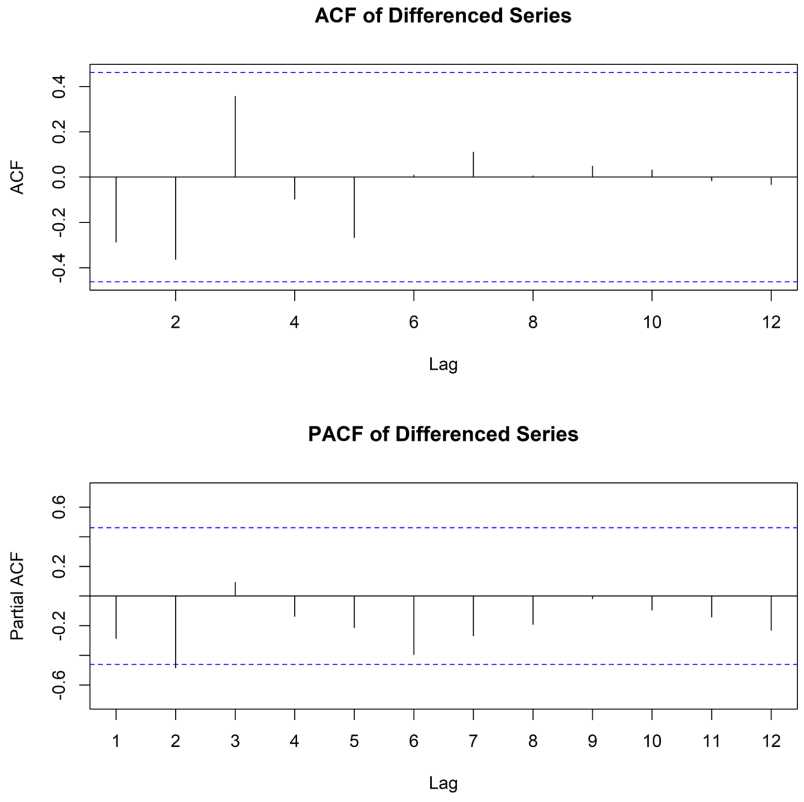


Figure 4. ACF and PACF plots of first differenced catch series of *L. stappersii* (2007-2025).

3.3. Model Identification and Parameter Estimation

The Box-Cox transformed and first-differenced catch series was evaluated using a range of ARIMA specifications. Among the candidate models, ARIMA (0, 1, 1) provided the lowest AIC (267.17), whereas ARIMA (0, 1, 0) yielded a slightly lower BIC (268.24) than ARIMA (0, 1, 1) (BIC = 268.95). The difference in BIC was small (ΔBIC = 0.72), indicating limited evidence in favour of the simpler ARIMA (0, 1, 0) specification based on BIC alone. Therefore, model selection was based on the combined consideration of information criteria, parsimony, residual diagnostics, and forecasting performance rather than on a single criterion. To govern this selection objectively, a prespecified decision rule was adopted: models were ranked primarily by information criteria (AIC/BIC minimization) but penalized for failing residual white-noise tests or yielding poor out-of-sample error metrics. Although ARIMA (0, 1, 0) achieved a marginally lower BIC (ΔBIC = 0.72), it fails to capture the moving-average dynamics present in the differenced series. Consequently, while ARIMA (0, 1, 1) is retained as the preferred operational model due to its superior AIC and residual properties.

The ARIMA (0, 1, 1) model was retained because it provided the lowest AIC among the candidate models and adequately captured the short-term dependence structure of the differenced catch series. The estimated MA (1) coefficient was -0.51 (SE = 0.26), with an estimated residual variance (σ^2) of 1.51. Diagnostic tests indicated that the residuals were consistent with the assumptions required for forecasting, with no significant residual autocorrelation (Ljung-Box test, p = 0.2916), approximate normality (Shapiro-Wilk test, p = 0.8225), and no evidence of conditional heteroscedasticity (ARCH-LM test, p =0.8576). These results, together with subsequent forecasting validation, supported the use of ARIMA (0, 1, 1) as the preferred forecasting model. The model yielded an (R^2) of 0.62, indicating that approximately 62% for differenced series. While we report this metric for comparative purposes with previous regional fisheries literature, we emphasize that information criteria (AIC/BIC) and error-based metrics (RMSE, MAE) remain the primary indicators of model fit and forecasting skill in non-stationary time-series analysis.

3.4. Residual Diagnostics

Table 3. Residual diagnostic tests for ARIMA (0, 1, 1) model.

Test	Statistic	df	p = value	Null Hypothesis	Interpretation
Ljung-Box (Q*)	3.7384	3	0.2916	Fail to reject	No significant autocorrelation
Box-Ljung (X^2)	8.2514	5	0.1429	Fail to reject	Residuals behave like white noise
Shapiro-Wilk	0.9724	-	0.8225	Fail to reject	Residuals approximately normal
ARCH LM	7.0000	12	0.8576	Fail to reject	No heteroscedasticity detected

The Residual diagnostics confirmed the adequacy of the ARIMA (0, 1, 1) specification in Table 3. The Ljung-Box and Box-Ljung tests indicated no significant autocorrelation, while the Shapiro-Wilk test supported normality of residuals. The ARCH residuals conform to ARIMA assumptions, supporting the model's suitability for forecasting *L. stappersii* catches.

Residuals fluctuated randomly around zero, confirming absence of autocorrelation and validating model adequacy (Ljung-Box p = 0.29), approximate normality (Shapiro-Wilk p = 0.82), and no evidence of heteroscedasticity (ARCH LM = 0.86), confirming model adequacy. Presented in Figure 5.

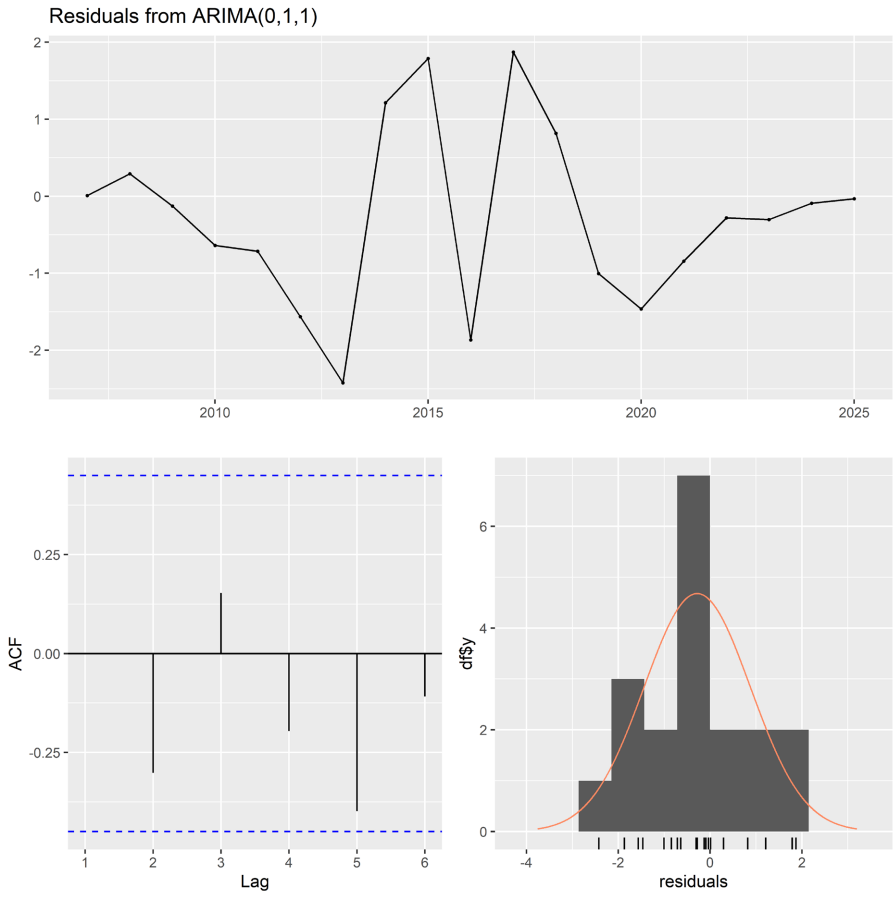


Figure 5. Residual diagnostic of ARIMA (0, 1, 1) model for *L. stappersii* catch series (2007-2025).

3.5. Benchmark Models

The ARIMA (0, 1, 1) model demonstrated performance, with moderate forecast error (RMSE = 352, MAE = 281) and relatively low residual autocorrelation (ACF1 = 0.12) as shown in Table 4. The ETS (M, A, N) model achieved slightly low RMSE (332) and MAE (261), but at the cost of higher bias (ME = −155) and inflated percentage errors (MAPE = 91%). The Random Walk model produced comparable MAE (278) and lowest MAPE (72%), but residual autocorrelation (ACF1 = −0.31) indicated poor white-noise behaviour. Formal comparison using

the Diebold-Mariano test (DM = 3.35, p = 0.0036) confirmed that ARIMA (0, 1, 1) forecasts were statistically superior to ETS (M, A, N) forecasts at the 5% significance level. This validates ARIMA (0, 1, 1) as the most robust and parsimonious specification for forecasting *L. stappersii* catches.

Table 4. Forecast accuracy comparison of ARIMA (0, 1, 1), ETS (M, A, N), and random walk models.

Model	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
ARIMA (0, 1, 1)	−43.78	351.96	281.09	−50.43	75.55%	1.01	0.122
ETS (M, A, N)	−114.95	332.46	260.83	−74.34	91.46%	0.94	0.281
Random Walk (RW)	−49.39	384.44	278.39	−45.34	72.03%	1	−0.312

To evaluate the structural stability of the chosen ARIMA framework against potential uncertainties or imputation artifacts in the terminal years (2024-2025), a sensitivity analysis was performed (Figure 6). The evaluation tested the model against data truncation (removing 2024-2025) and alternative baseline perturbation (±20% variation). Results confirmed that the selected ARIMA parameter's structure remained stable and the 10-year projected catch trajectories (2026-2035) exhibited negligible variance, verifying that the 2024-2025 imputation did not bias the final management recommendations.

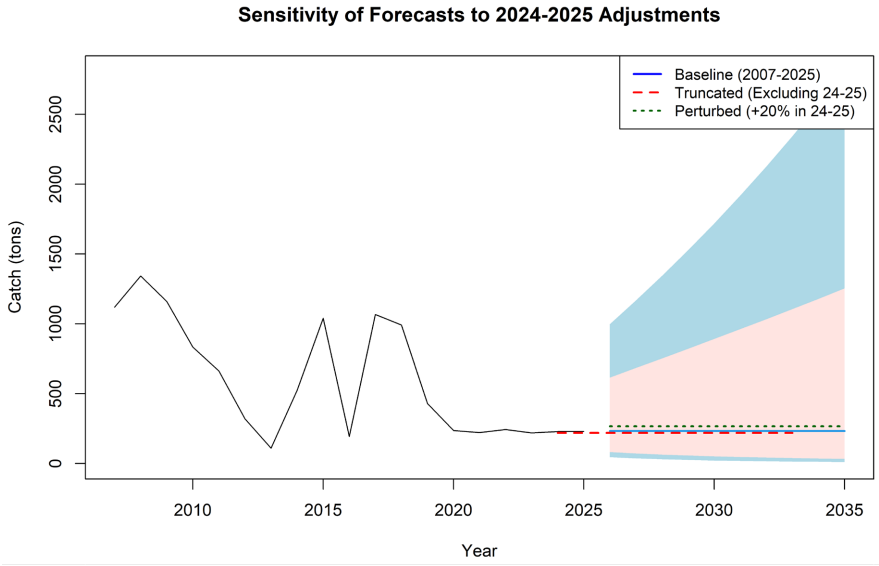


Figure 6. Sensitivity of forecasts to 2024-2025 adjustments in *L. stappesrii* catches.

3.6. Forecast

Forecasted catches of *L. stappersii* from 2026 to 2035 based on the ARIMA (0, 1, 1) specification in Figure 7. The solid line represents historical observations (2007-2025), while the shaped bands indicate the 80% and 95% prediction intervals. The model projects a stable mean catch of ~232 tons annually, while the prediction intervals widen substantially over time, reflecting increasing uncertainty

in long-term forecasts. These projections provide a quantitative statistical baseline for adaptive fisheries planning.

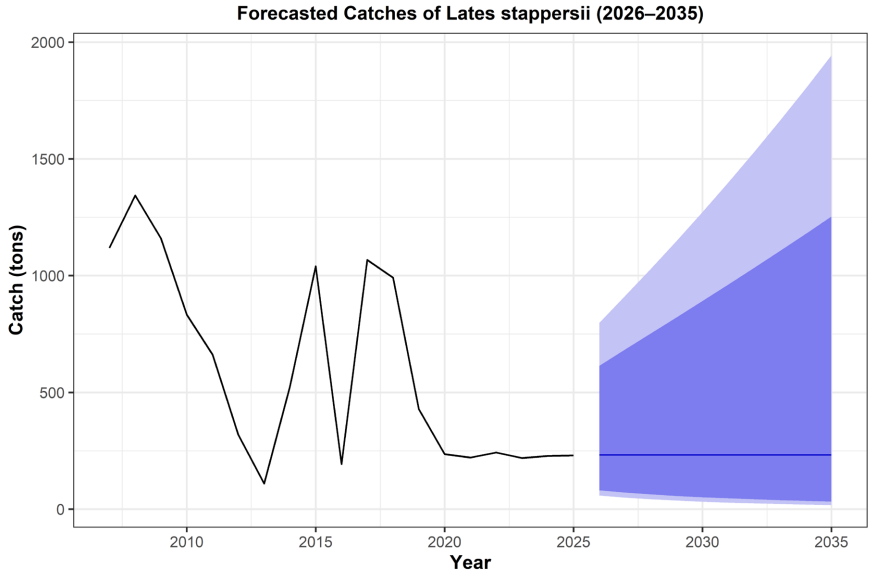


Figure 7. Forecasted catches of *L. stappersii* (2026-2035) with 80% and 95% Prediction Intervals.

4. Discussion

The ARIMA (0, 1, 1) specification applied to *Lates stappersii* catches in Lake Tanganyika (Burundi) demonstrated robust statistical adequacy. Although alternative benchmarks such the ETS model achieved slightly lower RMSE values, and ARIMA (0, 1, 0) yielded a minimal advantage in BIC, the adoption of ARIMA (0, 1, 1) is justified under our prespecified decision rule prioritizing residual white-noise compliance and parsimonious moving-average representation. We moderate the claim of absolute superiority, recognizing that validation criteria present nuanced trade-offs- rather than unanimous consensus. The moving average coefficient was significant (MA1 = −0.51, $p < 0.05$), residuals conformed to white - noise assumptions (Ljung-Box $p = 0.29$; Shapiro-wilk $p = 0.82$), and forecast accuracy metrics (RMSE = 352, MAE = 281, MAPE = 75%) were consistent with fisheries characterized by ecological variability. The R^2 value (0.62) further supports the adequacy of ARIMA (0, 1, 1) specification, through were emphasize that error-based metrics remain more informative for non-stationary fisheries data. Comparative validation confirmed ARIMA superiority over ETS (Diebold-Mariano test, DM = 3.35, $p = 0.0036$), while Random Walk models, despite lower MAPE, failed residual diagnostics. These findings align with regional applications of ARIMA in African fisheries, where parsimonious specification have proven reliable under ecological stress [3] [8] [9].

The performance of the ARIMA (0, 1, 1) model in the context of Burundian fisheries aligns with broader trends in African Fisheries research. Similar studies have demonstrated the efficacy of parsimonious ARIMA specifications in captur-

ing ecological variability across diverse inland and marine systems. For instance, [8] successfully forecasted forecast CPUE trends in Lake Malombe (Malawi) using an ARIMA (0, 1, 1) model, while [23] utilized an ARIMA (0, 1, 1) approach to analyse C*opadichromis* production. In Ethiopia, [9] employed an (2, 1, 0) specification to forecast *Labeobarbus* catches in under ecological stress, and [3] applied an ARIMA (3, 1, 3) model to capture long-term marine catch variability in Nigeria. These parallels reinforce the credibility of ARIMA (0, 1, 1) specification for Lake Tanganyika, situating these findings within the growing body of African fisheries forecasting literature.

The forecast of ~232 tons annually provides a quantitative statistical baseline for adaptive management in Burundi. However, the 80% and 95% prediction intervals (PI) widen substantially with the forecast horizon, with the 95% PI extending from values near zero to more than 3000 tons by 2035. This widening interval highlights volatility uncertainty in long-term catch projections and supports the need to need to interpret the point forecast together with associated PI. This necessitates precautionary harvest strategies, flexible quota systems, and integration of forecasting outputs into governance frameworks. The Direction de la Promotion des Filières Halieutiques (DPFH) and Lake Tanganyika Authority (LTA) can leverage ARIMA-based evidence to harmonize polices across riparian states, echoing lessons from Malawi and Ethiopia where forecasts informed seasonal closures and catch limits [9] [32]. Such evidence underscores the importance of integrating forecasting into adaptive governance frameworks across Africa [4]. Embedding ARIMA outputs into regional policy dialogues can strengthen resilience against overexploitation and climate-driven shocks.

The absence of catch records for 2024-2025 represents a notable limitation. A critical examination of this study, aching the remarks raised for this type of univariate modelling, highlights substantial limitations: the brevity of the annual time series, the imputation of data for the 2024-2025 period, and the absence of direct investigation into the biological, environmental, or fishing factors (such as fishing effort or climate variability) responsible for catch fluctuations. While the ARIMA (0, 1, 1) model adequately captures short-term stochastic dynamics with white-noise residuals, it does not provide a mechanistic explanation of ecosystem changes. Consequently, future research must move beyond simple univariate statistical prediction by integrating environmental covariates (climatic anomalies, upwelling indices) and fishing effort data to improve understanding of fishery dynamics and enhance the precision of long-term forecasts. Although imputation using neighbouring averages is widely used [13] [14], this approach can introduce potential bias. Furthermore, the limited length of the time series constrains comprehensive validation. To mitigate these issues, we employed a rolling forecasting origin and time-series cross-validation, these methods confirmed the stability of the ARIMA (0, 1, 1). Nonetheless, future research should explore SARIMA or ARIMA-GARCH hybrids models to better account for seasonality and volatility [28]. Furthermore, given that climate variability is primary driver of pelagic dy-

namics in Lake Tanganyika [6] [11], integrating environmental covariates, such as ENSO anomalies and rainfall variability, into forecasting frameworks is essential. As demonstrated in both marine and fisheries [33], such multi-model approaches can significantly enhance predictive power by capturing the environmental drivers that modulate long-term catch fluctuations.

The socio-economic importance of *L. stappersii* in Burundi, where over 200,000 people depend on fisheries, underscores the relevance of ARIMA-based forecasts for nutritional security. Stable mean projections support protein supply planning, but volatility necessitates diversification of food sources and precautionary management [23] [34]. Embedding ARIMA outputs into climate adaptation strategies, such as early warning systems and flexible quota adjustments, can strengthen resilience against climate-driven habitat changes [12]. Moreover, the shared nature of Lake Tanganyika fisheries requires transboundary governance management, with forecasting tools serving as neutral scientific evidence to support cooperative management [4]. Given the transboundary nature of Lake Tanganyika, climate-linked fluctuations such as ENZO events may simultaneously affect multiple riparian states. Forecasting frameworks that integrate ENSO signals could serve as early warning systems, enabling coordinated responses across Burundi, Tanzania, Zambia, and the DR Congo. This aligns with broader calls for climate-informed fisheries governance in Africa [4].

5. Conclusion

The study pioneers the application of ARIMA modelling to *Lates stappersii* catches in Lake Tanganyika (Burundi), addressing a critical gap in African fisheries forecasting. Scientifically, our findings demonstrate that parsimonious ARIMA (0, 1, 1) specifications reliably capture pelagic catch dynamics despite ecological variability, thereby contributing to growing body of time-series applications in fisheries science. The back-transformation mean forecast of approximately 232 tons annually for 2026-2035 provides a quantitative statistical baseline for adaptive fisheries planning and food-security assessment in Burundi. The accompanying 80% and 95% predictions intervals widen with increasing forecast horizon, indicating substantial uncertainty in long-term projections. These results should therefore be interpreted as forecasting information to support monitoring and adaptive planning rather as deterministic estimate of future catch or direct biological harvest limit. Notwithstanding these contributions, certain limitation, namely the short time series (19 observations), the imputation of missing data for 2024-2025, and the absence of explicit climate indices, constrain predictive precision. Future research should integrate environmental drivers, such as ENSO anomalies, rainfall variability, and stratification indices, into hybrid models (e.g. SARIMA or ARIMA-GARCH), while extending datasets through continuous monitoring. Integrating quantitative forecasting to governance and climate adaptation, this study advances both methodological innovation and practical guidance for sustainable fisheries management in Africa.

Acknowledgements

The authors wish to extend their appreciation to the for-Lake Tanganyika Authority (LTA) and the Direction de la Promotion des Filières Halieutiques (DPFH) in Burundi for their cooperation, kindness, and availability during the data collection phase. Their support was crucial in ensuring the smooth progress of the study.

Author Contributions

Vianney Ngabo Mulema: Writing original Draft, Visualization, Software, Methodology, Investigation, Formal analysis, Funding acquisition and Project administration; Wales Singini and Alidor Kankonda Busanga: Writing-review & editing, Validation, Visualization, Resources and Supervision; Odette Karerwa, Gabriel Balagizi Baguma**,** Eunice Musau Ntambwe, Aristote Yofemo Bifinole and Fiston Bokwala Bobitshe: Writing-review & editing, Resources, Data curation, and Investigation. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

[1] FAO (2016) The State of World Fisheries and Aquaculture 2016: Contributing to Food Security and Nutrition for All. Food and Agriculture Organization of the United Nations.

[2] Ogutu-Ohwayo, R., Natugonza, V., Musinguzi, L., Olokotum, M. and Naigaga, S. (2016) Implications of Climate Variability and Change for African Lake Ecosystems, Fisheries Productivity, and Livelihoods. *Journal of Great Lakes Research*, **42**, 498-510. https://doi.org/10.1016/j.jglr.2016.03.004

[3] Isemin, N.L. and Etim, L. (2025) Time Series Analysis and Forecasting of Nigerian Marine Fish Catch (1960-2015) Using ARIMA Models. *Journal of Marine Studies*, **2**, Article 2306. https://doi.org/10.29103/joms.v2i3.24522

[4] Phiri, H., Mushagalusa, D., Katongo, C., Sibomana, C., Ajode, M.Z., Muderhwa, N., *et al.* (2023) Lake Tanganyika: Status, Challenges, and Opportunities for Research Collaborations. *Journal of Great Lakes Research*, **49**, Article 102223. https://doi.org/10.1016/j.jglr.2023.07.009

[5] Sibomana, C. (2025) *Lates stappersii*. The IUCN Red List of Threatened Species 2025: e.T60839A47217730.

[6] Ehrenfels, B., Junker, J., Namutebi, D., Callbeck, C.M., Dinkel, C., Kalangali, A., *et al.* (2023) Isotopic Signatures Induced by Upwelling Reveal Regional Fish Stocks in Lake Tanganyika. *PLOS ONE*, **18**, e0281828. https://doi.org/10.1371/journal.pone.0281828

[7] Plisnier, P.-D., Mgana, H., Kimirei, I., Chande, A., Makasa, L., Chimanga, J., *et al.* (2009) Limnological Variability and Pelagic Fish Abundance (*Stolothrissa tanganicae* and *Lates stappersii*) in Lake Tanganyika. *Hydrobiologia*, **625**, 117-134. https://doi.org/10.1007/s10750-009-9701-4

[8] Makwinja, R., Mengistou, S., Kaunda, E., Alemiew, T., Phiri, T.B., Kosamu, I.B.M., *et al.* (2021) Modeling of Lake Malombe Annual Fish Landings and Catch per Unit Effort (CPUE). *Forecasting*, **3**, 39-55. https://doi.org/10.3390/forecast3010004

[9] Mohammed Yesuf, B., Getahun, A., Mengistou, S., Wilson, G., Anteneh, W. and Abera, W. (2023) Time-Series ARIMA Modelling of the *Labeobarbus* Spp (Cyprinidae) Fishery in Water Hyacinth-Infested and Non-Infested Sites in Lake Tana, Ethiopia. *Journal of Freshwater Ecology*, **38**, Article No. 4. https://doi.org/10.1080/02705060.2023.2216218

[10] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. (2016) Time Series Analysis: Forecasting and Control. 5th Edition, Wiley, 712.

[11] Cohen, A.S., Gergurich, E.L., Kraemer, B.M., McGlue, M.M., McIntyre, P.B., Russell, J.M., *et al.* (2016) Climate Warming Reduces Fish Production and Benthic Habitat in Lake Tanganyika, One of the Most Biodiverse Freshwater Ecosystems. *Proceedings of the National Academy of Sciences*, **113**, 9563-9568. https://doi.org/10.1073/pnas.1603237113

[12] Nyboer, E.A., Musinguzi, L., Ogutu-Ohwayo, R., Natugonza, V., Cooke, S.J., Young, N., *et al.* (2022) Climate Change Adaptation and Adaptive Efficacy in the Inland Fisheries of the Lake Victoria Basin. *People and Nature*, **4**, 1319-1338. https://doi.org/10.1002/pan3.10388

[13] Hilborn, R. and Walters, C.J. (1992) Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman & Hall.

[14] FAO (2017) Yearbook of Fishery Statistics: Collection, Processing and Analysis. Food and Agriculture Organization of the United Nations.

[15] Guerrero, V.M. (1993) Time-Series Analysis Supported by Power Transformations. *Journal of Forecasting*, **12**, 37-48. https://doi.org/10.1002/for.3980120104

[16] Dickey, D.A. and Fuller, W.A. (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, **74**, 427-431. https://doi.org/10.1080/01621459.1979.10482531

[17] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992) Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We that Economic Time Series Have a Unit Root? *Journal of Econometrics*, **54**, 159-178. https://doi.org/10.1016/0304-4076(92)90104-y

[18] Hyndman, R.J. and Athanasopoulos, G. (2021) Forecasting: Principles and Practice, 3rd Edition, OTexts.

[19] Shumway, R.H. and Stoffer, D.S. (2011) Time Series Analysis and Its Applications: With R Examples. 3rd Edition, Springer.

[20] Akaike, H. (1974) A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, **19**, 716-723. https://doi.org/10.1109/tac.1974.1100705

[21] Cavanaugh, J.E. and Neath, A.A. (2019) The Akaike Information Criterion: Background, Derivation, Properties, Application, Interpretation, and Refinements. *WIREs Computational Statistics*, **11**, e1460. https://doi.org/10.1002/wics.1460

[22] Schwarz, G. (1978) Estimating the Dimension of a Model. *The Annals of Statistics*, **6**, 461-464. https://doi.org/10.1214/aos/1176344136

[23] Zaindi, I., Singini, W. and Mzengereza, K. (2016) Forecasting Copadichromis (Utaka) Production for Lake Malaŵi, Nkhatabay Fishery—A Stochastic Model Approach. *Journal of Fisheries & Livestock Production*, **4**, Article No. 162. https://doi.org/10.4172/2332-2608.1000162

[24] Boruah, B.B., Roy, P., Dutta, A. and Hazarika, B.B. (2020) Forecasting Marine and Total Fish Production in India using ARIMA Model. *Indian Journal of Economics and Business*, **19**, 161-174.

[25] Ljung, G.M. and Box, G.E.P. (1978) On a Measure of Lack of Fit in Time Series Models. *Biometrika*, **65**, 297-303. https://doi.org/10.1093/biomet/65.2.297

[26] Engle, R.F. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, **50**, 987-1007. https://doi.org/10.2307/1912773

[27] Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307-327. https://doi.org/10.1016/0304-4076(86)90063-1

[28] Francq, C. and Zakoian, J. (2019) GARCH Models: Structure, Statistical Inference and Financial Applications. Wiley. https://doi.org/10.1002/9781119313472

[29] Diebold, F.X. and Mariano, R.S. (1995) Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, **13**, 253-263. https://doi.org/10.1080/07350015.1995.10524599

[30] Perron, P. (1989) The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, **57**, 1361-1401. https://doi.org/10.2307/1913712

[31] Kębłowski, P. and Welfe, A. (2004) The ADF-KPSS Test of the Joint Confirmation Hypothesis of Unit Autoregressive Root. *Economics Letters*, **85**, 257-263. https://doi.org/10.1016/j.econlet.2004.04.013

[32] Mulumpwa, M., Jere, W., Mtethiwa, A., Kakota, T. and Kang'ombe, J. (2020) Modelling and Forecasting of Catfish Species Yield from Mangochi Artisan Fisheries of Lake Malawi in Malawi. *African Journal of Food, Agriculture, Nutrition and Development*, **20**, 16864-16883. https://doi.org/10.18697/ajfand.95.18505

[33] Plisnier, P., Nshombo, M., Mgana, H. and Ntakimazi, G. (2018) Monitoring Climate Change and Anthropogenic Pressure at Lake Tanganyika. *Journal of Great Lakes Research*, **44**, 1194-1208. https://doi.org/10.1016/j.jglr.2018.05.019

[34] Singini, W., Kaunda, E., Kasulo, V. and Jere, W. (2012) Modelling and Forecasting Small Haplochromine Species (Kambuzi) Production in Malaŵi—A Stochastic Model Approach. *International Journal of Scientific and Technology Research*, **1**, 69-73.