2023

Fitting Time Series Models to Fisheries Data to Ascertain Age

Kathleen S. Kirch
Old Dominion University, kkirch@odu.edu

Norou Diawara
Old Dominion University, ndiawara@odu.edu

Cynthia M. Jones
Old Dominion University, cjones@odu.edu

Original Publication Citation

Kirch, K. S., Diawara, N., & Jones, C. M. (2023). Fitting time series models to fisheries data to ascertain age. *Journal of Probability and Statistics*, *2023*, 1-10, Article 9991872. https://doi.org/10.1155/2023/9991872

Hindawi
Journal of Probability and Statistics
Volume 2023, Article ID 9991872, 10 pages
https://doi.org/10.1155/2023/9991872

Research Article

Fitting Time Series Models to Fisheries Data to Ascertain Age

Kathleen S. Kirch,[1] **Norou Diawara**,[2] **and Cynthia M. Jones**[1]

[1]*Department of Ocean & Earth Sciences, Old Dominion University, Norfolk, VA 23529, USA*
[2]*Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA 23529, USA*

Correspondence should be addressed to Norou Diawara; ndiawara@odu.edu

Received 21 August 2023; Revised 22 September 2023; Accepted 27 September 2023; Published 7 October 2023

Academic Editor: Zacharias Psaradakis

The ability of government agencies to assign accurate ages of fish is important to fisheries management. Accurate ageing allows for most reliable age-based models to be used to support sustainability and maximize economic benefit. Assigning age relies on validating putative annual marks by evaluating accretional material laid down in patterns in fish ear bones, typically by marginal increment analysis. These patterns often take the shape of a sawtooth wave with an abrupt drop in accretion yearly to form an annual band and are typically validated qualitatively. Researchers have shown key interest in modeling marginal increments to verify the marks do, in fact, occur yearly. However, it has been challenging in finding the best model to predict this sawtooth wave pattern. We propose three new applications of time series models to validate the existence of the yearly sawtooth wave patterned data: autoregressive integrated moving average (ARIMA), unobserved component, and copula. These methods are expected to enable the identification of yearly patterns in accretion. ARIMA and unobserved components account for the dependence of observations and error, while copula incorporates a variety of marginal distributions and dependence structures. The unobserved component model produced the best results (AIC: −123.7, MSE 0.00626), followed by the time series model (AIC: −117.292, MSE: 0.0081), and then the copula model (AIC: −96.62, Kendall's tau: −0.5503). The unobserved component model performed best due to the completeness of the dataset. In conclusion, all three models are effective tools to validate yearly accretional patterns in fish ear bones despite their differences in constraints and assumptions.

1. Introduction

Sustainable fisheries rely on the use of age-based demographic models that account for reproduction and growth to maximum economic benefit. To achieve this goal, it is required that fish be ascribed correct ages [1]. Ascribing correct ages relies on validating putative annual marks on fish bones [2]. The most commonly used technique to validate marks is marginal increment analysis (MIA) [2, 3]. MIA evaluates the width of newly accreted material laid down on the edge of fish ear bones measured over a year compared to material accreted over the previous year [2] (Figure 1). The ratio is typically averaged by month over all age groups and plotted over a single year. The plot shows a single dip if one band is produced yearly (Figure 2). This dip validates the pattern of a putative annual ring, allowing for accurate ageing of the fish to the nearest year.

For this pattern verification, few instances exist of statistically validated quantitative approaches because of irregular patterns observed due to changing environmental conditions. One such irregular pattern often looks like a sawtooth wave with an abrupt drop in accretion yearly to make an annual band [5–7]. Such sawtooth wave patterns provide challenges in fitting data to statistical models. Attempts have been made to overcome the difficulty of modeling the sawtooth wave pattern. Okamura et al. [6] used circular statistics and conversion of the month of capture to the median of the month and then another conversion of the median of the month to radians with accretion ratios as the linear vector. The circular method provides a short-term forecast that assumes the data are uniformly distributed around one preferred direction [8]. In contrast, Phelps et al. [7] used analysis of variance to identify significant differences in increment widths between months, but not in the

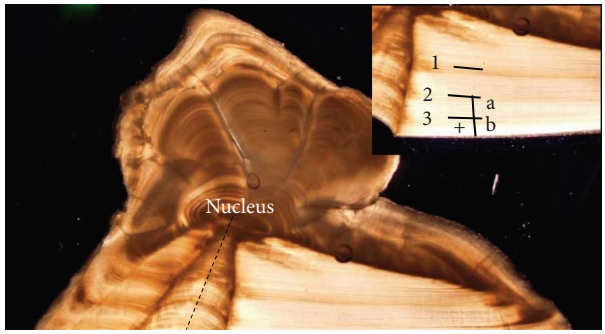

Figure 1: This is the microscopic cross section of the ear bone of a 3-year-old Atlantic croaker. The age is evidenced by the three numbered dark bands. Year 1 begins at the nucleus and continues down the succal groove (dashed line) to band 1. Year 2 begins at band 1 and continues to band 2. Year 3 begins at band 2 and ends at band 3. Accretional growth during year 3 extends to the edge of the ear bone as indicated by "+." Inset: the marginal increment ratio is b/a.

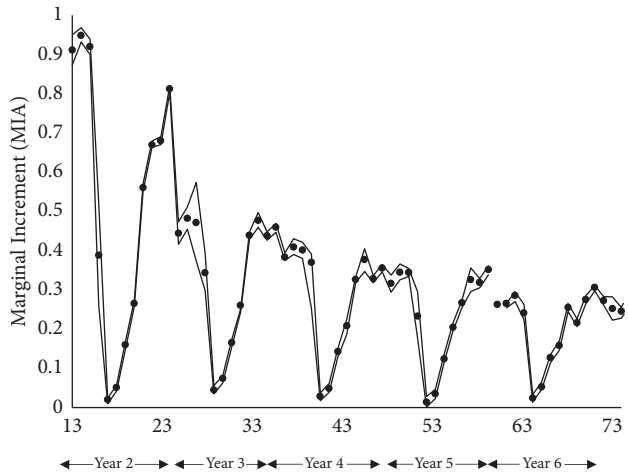

Figure 2: Time series of marginal increment analysis (MIA) data with one standard error interval from [4]. Notice the dip in May of each year. By fisheries convention, age in months is arbitrarily assigned a birthdate of January 1 to all fishes born in the northern hemisphere. The x axis has two legends. The first shows age in months, and the second shows the distribution of months as years.

years' trend. Analysis with a time series model is appropriate for yearly patterned data because time series models identify seasonal patterns and should detect a yearly accretion pattern.

The advantage of time series analysis is that it accounts for the dependence of the observations and errors in the dataset [9]. Time series analysis of marginal increments can allow for the objective determination of the temporal, seasonal, and cyclic nature of growth accretion. Time series models, such as autoregressive integrated moving average (ARIMA), unobserved component models (UCMs), and Gaussian copula, can pinpoint seasonal changes through time.

A time series approach works especially well when data are complete and abundant. Other assumptions of ARIMA include the independence and identical distribution under the Gaussian nature of the error terms in the model statement, as well as the constant variance assumption or stationarity. Unfortunately, the sawtooth wave pattern in MIA data complicates finding a well-fitted time series model.

The UCM can break a time series of data into trend, cyclical, seasonal, autoregressive, and lagged components [10, 11]. Bian et al. [12] gave a summary of UCM applications but did not include analysis in fisheries research. To our knowledge, this paper is the first to use UCM to model MIAs. UCMs use smoothing models that allow for improved analysis and put more weight onto observations that are closer to each other due to the nonignorable correlation of time series measurements [13]. The single-error model is in fact broken up into its component calibration.

The Gaussian copula construction emanates from Joe [14] and other authors [15, 16]. However, just as in the UCM case, its use in fisheries research is still in its infancy [17, 18]. Copula models easily incorporate a variety of marginal distributions and different dependence structures, unlike ARIMA whose assumptions rely on identical joint distributions at all-time values and identical Gaussian marginal distributions [9]. As a first advantage, the copula approach for parameter estimation in multistage models captures dependence from the temporal samples. A second advantage is that the copula circumvents the nontrivial computations in the variances of estimates. Gaussian copulas are very common in economics, epidemiology, time series analysis, and climate change [19–23].

We propose the comparison of the ARIMA, UCM, and copula-based statistical time-series methods while incorporating serial dependence. The goal is to evaluate the feasibility and preliminary efficacy of the models on fish ear bone accretion data. We are fortunate in this first test to have an unusually complete dataset without often-missing seasonal components due to conditions that prohibit fishing such as adverse weather or lack of fish availability.

The paper is organized as follows. In Section 2, the data are described with the ARIMA, UCM, and copula models and the statistical analysis methods are presented. In Section 3, the solutions are presented. In Section 4, we end with a discussion of the performance of these models using our complete dataset.

2. Materials and Methods

2.1. Data Collection. The data for this study were extracted from the data included by Barbieri et al. [4] using DataThief III, Version 1.7 (2015) to extract the data points from a graph. According to the authors' methods, fish were collected each month between June 1988 and June 1991 from commercial fisheries in the Chesapeake Bay and brought to their laboratory. Next, individual biometrics were recorded, and ear bones were excised.

Fish ear bones (sagittal otoliths) were sectioned and then mounted on microscope slides and magnified to a range from 0.8x to 8x and photographed using an Olympus DP71 camera and program Cells Sense (Figure 1). The images were then uploaded to program Image-Pro Plus v. 6.2.0.424

(Media Cybernetics Inc.) for marginal increment measurement, in micrometers. The first marginal increment was measured from the last dark band to the outer edge of the ear bone. The second is measured from the last dark band to the previous dark band. These measurements are then combined as a ratio of the last partial accretion to the last full accretion band (Figure 1). The average ratio was then calculated for each age class by its month of capture. The average was plotted by month and visually inspected for the drop in the sawtooth wave. Fish were categorized by the age at capture, regardless of year of capture. The youngest age caught was 1 year.

2.2. Model Diagnostics and Goodness of Fit Measurements. Consider the data as n observations of type $y_1, \ldots, y_n$, the likelihood function can be written as follows:

$$L(\Theta; y_1, \ldots, y_n) = \prod_{i=1}^{n} f_y(y_i; \Theta), \tag{1}$$

where

$$\Theta = \left[\theta_1, \ldots, \theta_p\right]^T, \tag{2}$$

is the vector of p parameters associated with the model equation.

The log-likelihood equation can then be formulated as follows:

$$\log(L) = \sum_{i=1}^{n} \log\ \left\{f_y(y_i; \Theta)\right\}. \tag{3}$$

Akaike's information criterion is then calculated by the following equation:

$$\mathrm{AIC} = -2\log(L) + 2K, \tag{4}$$

where $K = p$ is the number of parameters in the model [24, 25].

The mean squared error is formulated as

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2, \tag{5}$$

and Kendall's tau that measures the difference in the measure of concordance and discordance between the marginal CDFs is

$$\begin{aligned}\tau = {} & \Pr\{(X_1 - X_2)(Y_1 - Y_2) > 0\} \\ & - \Pr\{(X_1 - X_2)(Y_1 - Y_2) < 0\},\end{aligned} \tag{6}$$

for pairs of observations (X_1, Y_1) and (X_2, Y_2), and can also be formulated as follows:

$$\tau = 4\int_0^1\int_0^1 C(u, v)\mathrm{d}C(u, v) - 1, \tag{7}$$

as by Sun et al. [16], where C represents the copula function that will be described in Section 2.5, with u and v capturing the marginal distribution of X and Y, respectively.

2.3. Time Series Model. The data from the time series process X_t are described as an ARIMA (p, d, q) model and are defined as follows:

$$\phi(B)X_t = \theta(B)Z_t, \tag{8}$$

where p, d, and q are non-negative integers, based on a sample size of n, with

$$\begin{aligned}X_t &= (1 - B)^d Y_t, \\ Z_t &\sim \mathrm{WN}\left(0, \sigma^2\right), \\ \phi(z) &= 1 - \phi_1 z - \ldots - \phi_p z^p, \\ \theta(z) &= 1 + \theta_1 z + \ldots + \theta_q z^q,\end{aligned} \tag{9}$$

where Y_t is the transformed X_t series with the mean retrieved and B is the backward shift operator [26].

The parameter p is associated with the autoregressive portion of the process, while q is associated with the moving average portion of the process. To help estimate the autoregressive parameter (p), the sample autocorrelation function can be used once graphed:

$$\widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma}(0)}, \quad -n < h < n, \tag{10}$$

where γ is the sample autocovariance function and is defined as follows:

$$\widehat{\gamma}(h) \coloneqq n^{-1}\sum_{t=1}^{n-|h|}\left(x_{t+|h|} - \overline{x}\right)\left(x_t - \overline{x}\right), \ -n < h < n. \tag{11}$$

To help estimate the moving average parameter (q), the sample partial autocorrelation function can be used once graphed:

$$\begin{aligned}\alpha(0) &= 1, \\ \alpha(h) &= \phi_{\mathrm{hh}}, \quad h \geq 1,\end{aligned} \tag{12}$$

where ϕ_{hh} is the last component of

$$\phi_h = \Gamma_h^{-1}\gamma_h, \tag{13}$$

where

$$\begin{aligned}&\Gamma_h = [\gamma(i - j)]^h, \quad i, j = 1 \ldots T, h \geq 1, \\ &\gamma_h = \gamma(h).\end{aligned} \tag{14}$$

The likelihood function is as follows [27]:

$$L(\theta|Y_t) = \prod_{n=2}^{n}\frac{1}{\left(2\pi^{n/2}\right)|\Sigma|^{1/2}}\exp\left(\frac{-1}{2}Y'\Sigma^{-1}Y\right), \tag{15}$$

where Σ is the variance/covariance matrix of the observed time series data.

R package "astsa" is used to loop through possible combinations of p and q ranging from zero to three. [28]. AIC, log likelihood, and MSE were used to decide the best model.

2.4. Unobserved Component Model. Components of the models can be assumed unobserved and must be estimated under a time series model. The unobserved component model (UCM) offers such flexibility. Unobserved components can be modeled using the following equation:

$$Y_t = \mu_t + \gamma_t + \psi_t + r_t + \sum \varphi_i Y_{t-i} + \varepsilon_t, \tag{16}$$

where μ_t represents the trend component, γ_t represents the seasonal component, ψ_t represents the cycle trend, r_t is the autoregressive term, $\sum \varphi_i Y_{t-i}$ is a regressive term involving the lagged dependent variables, ε_t is the error term assumed to be independent with an identical Gaussian distribution, and μ_t, γ_t, ψ_t, and r_t are assumed to be independent of each other [13].

The likelihood function is as follows [27]:

$$L(\theta|Y) = p_1(y_1)\prod_{t=2}^{n} p_t\left(y_t|y_{t-1};\theta\right), \tag{17}$$

where $Y = (y_1, \ldots, y_n)'$.

Analysis is conducted using SAS "proc ucm." Six models are tested using various combinations of the level, slope, cycle, and season and described in Table 1. Model 1 uses a stochastic level, slope, cycle, and season; model 2 uses a stochastic level and slope, no cycle, and stochastic season; model 3 uses a stochastic level, fixed slope, and stochastic cycle and season; model 4 uses a stochastic level, no slope, and stochastic cycle and season; model 5 uses a stochastic level, fixed slope, no cycle, and stochastic season; model 6 uses a stochastic level, no slope or cycle, and stochastic season.

2.5. Copula Model. A copula is defined as follows:

$$P(Y_t \le y_t, Y_{t-1} \le y_{t-1}) = C\{P(Y_t \le y_t), P(Y_{t-1} \le y_{t-1})\},$$
$$t = 1, 2, \ldots, T, C\colon [0,1]^2 \longrightarrow [0,1], \tag{18}$$

which satisfies the following requirement:

$$\begin{aligned} C(u,0) &= C(0,v) \\ &= 0, C(u,1) = u, C(1,v) = v, \text{for } 0 \le u, v \le 1, \\ &\quad C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1v_1) \\ &\ge \text{for } 0 \le u_1 \le u_2 \le 1 \wedge 0 \le v_1 \le v_2 \le 1. \end{aligned} \tag{19}$$

In addition, Sklar's theorem states that for the following equation,

$$\Pr(X \le x, Y \le y) = C\{F(x), G(y)\}, \tag{20}$$

a unique function of C can be found if F and G are continuous [16, 19].

This paper will use beta marginals described as follows:

Table 1: Combination of all possible methods used to test the unobserved component models (UCMs).

	Level	Slope	Cycle	Season
Model 1	Stochastic	Stochastic	Stochastic	Stochastic
Model 2	Stochastic	Stochastic	—	Stochastic
Model 3	Stochastic	Fixed	Stochastic	Stochastic
Model 4	Stochastic	—	Stochastic	Stochastic
Model 5	Stochastic	Fixed	—	Stochastic
Model 6	Stochastic	—	—	Stochastic

$$\begin{aligned} G(y) &= \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)}, \\ B(\alpha,\beta) &= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \\ \mu &= E[Y_t], \\ \sigma^2 &= \mathrm{Var}(Y_t), \end{aligned} \tag{21}$$

where α and β are the shape and scale parameters and Γ is the Gamma function. The likelihood function is as follows [29]:

$$L(\theta|Y) = p(y_1,\theta)\prod_{t=2}^{n} p_t\left(y_t|y_{t-1}, \ \ldots \ y_1; \ \theta\right). \tag{22}$$

The equation used in the copula analysis was

$$\begin{aligned} \mathrm{MIA}_t = \tan\left(\frac{2\pi t}{12} - 4\right)\left(\sin\left(\frac{2\pi t}{12} - 4\right)\right. \\ \left. + \frac{2}{\pi^2}\cos\left(\frac{2\pi t}{12} - 4\right)\right) + \varepsilon_t, \end{aligned} \tag{23}$$

where sin, cos, and tan are the trigonometric functions and $\varepsilon_t \sim \text{iid } N(0, \sigma^2)$ are the error terms.

R package "gcmr" was used to fit the copula models [30]. For the sake of interpretability of smoothness, ARIMA (1, 1, 0) was selected. Due to time dependency, trigonometric functions were used in the model equation [31–33]. The copula is more parsimonious than UCM or time series. It includes only the time series parameters in a more succinct form based on the marginal distributions of Y_t's. The copula equation was estimated and reported along with the log likelihood and AIC. Kendall's tau was calculated to estimate a correlation between marginal accretion width over consecutive time periods.

3. Results

3.1. Data. Ear bones from a total of 1185 fish were prepared, and increments were measured by Barbieri et al. [4]. The monthly average increment by age is graphed in Figure 2. A clear decrease in accretion width can be observed in May of each year, as well as a decreasing variance over time.

3.2. ARIMA Model. The model with the highest log likelihood (65.6458), lowest AIC (–117.292), and MSE (0.00805) had an autoregressive order of 2 and a moving average order

of 3 (Table 2). All the estimated parameters were significant except the third moving average term, and when this third term is removed, the constant term becomes insignificant (Table 3). The QQ plots of these two models show the majority of the data with a normal distribution (Figure 3).

3.3. Unobserved Component Model. The six models tested using various combinations are presented in Table 4. Model 4 provided the highest log likelihood (67.874) and lowest AIC (−123.7), while model 1 produced the lowest MSE (0.00626). The AIC and MSE were smaller than those of the ARIMA model, while the log likelihood is larger than the best of the ARIMA models. (Table 4). All models produced an identical graph with an extremely narrow confidence interval (Figure 4).

The residuals for models 2, 3, and 5 had a distinct sinusoidal pattern, while the others had a slightly more random appearance (Figure 5). The QQ plots of the residuals show most of the data follow a normal distribution (Figure 6). Models 2, 5, and 6 had a sinusoidal pattern in the ACF graphs (Figure 7).

3.4. Copula Model. Figure 8 describes the graph of the copula function. Tangent was added to the equation to properly model the drop of the sawtooth wave. The copula produced a log likelihood, AIC, and Kendall tau of −52.31, −188.9633, and −0.5503, respectively. Kendall's tau is negative since as time increases, the marginal increment decreases. Such a result has not been captured under ARIMA or UCM. The AIC and log likelihood were smaller than those of ARIMA or UCM. The variance of the copula model was larger than that of either ARIMA or UCM (Figure 9). The conditional residual plot showed dips corresponding to May of each year, and QQ plots showed the majority of the data within a normal distribution (Figure 9). The marginal residual plot showed dips corresponding to May of each year, and the QQ plot showed the majority of the data within a normal distribution (Figure 9). The pattern in the MIA residuals shows that alternative models are better fits.

Table 2: Model diagnostics for the ARIMA (p, 1, q) model.

$p\backslash q$	0	1	2	3
(a)				
0	−80.6663	−83.841	−81.8519	−88.0023
1	−83.211	−81.8484	−90.054	−89.9316
2	−82.1529	−96.3694	*−83.3204*	**−117.292**
3	−80.8888	−99.4612	−107.115	−104.951
(b)				
0	0.017768	0.016501	0.016498	0.014002
1	0.016651	0.016499	0.013748	0.013359
2	0.016426	0.012415	*0.014449*	**0.008054**
3	0.016247	0.011402	0.010011	0.010014
(c)				
0	42.33315	44.92048	44.92594	49.00115
1	44.60548	44.92422	50.02699	50.96579
2	45.07647	53.18471	*47.66019*	**65.64576**
3	45.44441	55.73059	60.5575	60.47542
(d)				
0	−84.66630	−85.56429	−81.29855	−85.17230
1	−84.93434	−81.29511	−87.22398	−84.82492
2	−81.59961	−93.53942	*−78.21372*	**−109.90819**
3	−78.05882	−94.35452	−99.73167	−95.29084

The bold values show the best model parameters, while the italic values show the second best model parameters. (a) The AICs for selected ARIMA models. (b) The MSE for selected ARIMA models. (c) The log likelihood for selected ARIMA models. (d) The BIC for selected ARIMA models.

Table 3: Parameter estimates for best ARIMA models.

	Estimate	Standard error	t value	p value
(a)				
AR 1	1.6169	0.0559	28.932	<0.0001
AR 2	−0.9253	0.0506	−18.2994	<0.0001
MA 1	−2.063	0.1885	−10.9446	<0.0001
MA 2	1.2442	0.3565	3.4899	0.0009
MA 3	−0.1262	0.1803	−0.7	0.4864
Constant	−0.0051	0.0019	−2.6825	0.0093
(b)				
AR 1	0.9625	0.0715	13.4632	<0.0001
AR 2	−0.9573	0.0561	−17.073	<0.0001
MA 1	−0.8789	0.0877	−10.0175	<0.0001
MA 2	1	0.0861	11.6125	<0.0001
Constant	−0.0083	0.0161	−0.5134	0.6094

(a) Parameter estimates for best time series model, ARIMA (2, 1, 3) (bold in Table 2). (b) Parameter estimates for second best time series model, ARIMA (2, 1, 2) (italic in Table 2).

4. Discussion

The sawtooth wave pattern of MIA data proved challenging to model, but when analyzed with ARIMA, UCM, and copula, these methods provided precise timing of accretional patterns in fish ear bones. The models demonstrated that, for Atlantic croaker, dark bands had formed by May and occurred only once during the year. Thus, providing a model to validate the formation of dark bands was evident after a sharp drop in accretion while also providing statistical metrics of model fit, attributes missing from qualitative measures.

The first step in our approach was to formulate more flexible assumptions about the dependence structure of the process. More precisely, the joint density of the accretion process could directly and conveniently describe the stochastic process. Using the dependence structure of ARIMA, the conditional probability of the sawtooth pattern was well described. The UCM model matched the results from ARIMA as the autocorrelation was captured under Gaussian white noise and is also extendable when one considers the copula types of distributions. In this paper, we applied the copula ideas of Salinas-Gutierrez et al. [34] and Alqawba and Diawara [35], with beta marginals.

ARIMA and UCM produced the best results. ARIMA is well established and widely used in analysis of complete datasets. The UCMs are valuable extensions of the time series with variance decreasing over time. The confidence intervals in both contain most of the recorded data and matched the seasonal pattern well. In both models, parsimony is obtained

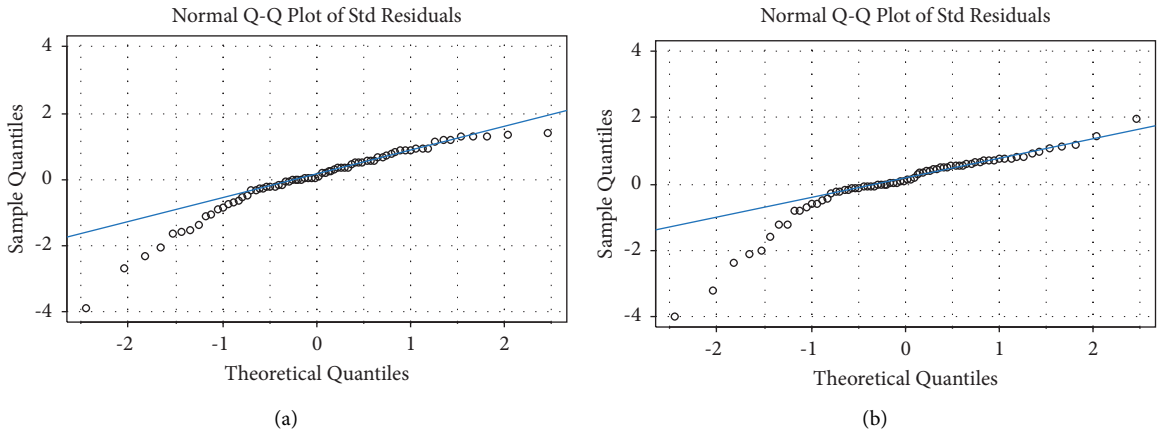

Figure 3: QQ plots of the best two ARIMA models: (a) (2, 1, 3) and (b) (2, 1, 2).

Table 4: Summary statistics for each of the six UCMs.

Models	AIC	MSE	Log likelihood
1	−121.2	**0.00626**	67.594
2	−91.53	0.01001	49.766
3	−118	0.00649	65.014
4	**−123.7**	0.00844	**67.874**
5	−93.53	0.01001	49.766
6	−99.68	0.01277	52.842

Model 4 has the lowest AIC and largest log likelihood. Model 1 has the lowest MSE as shown in bolded text.

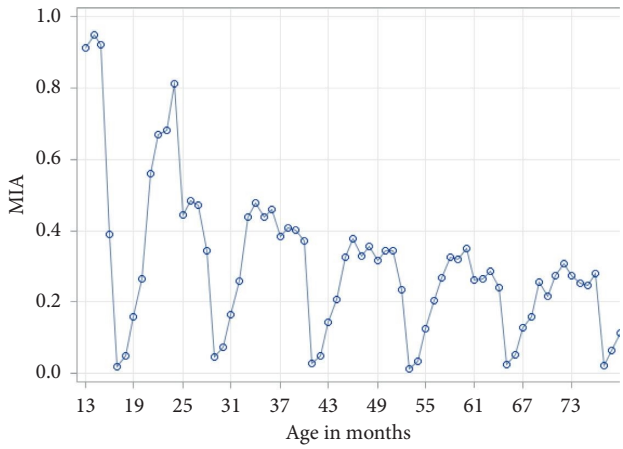

Figure 4: MIA data fit from each of the six unobserved component models (UCMs). The same graph was produced for all six models. The confidence intervals are very narrow.

mainly because of complete data. The cycle was automatically obtained, giving us close to perfect time-varying predictions.

Although the copula did not perform well, the results were still acceptable for its use in MIA. The variance estimate is higher in this case than in ARIMA and UCM. The challenge may be due to the choice of the marginal distribution not being as good fit to these data. The way ARIMA and UCM components were captured, as regularity of cycle and seasonality, provides less of an emphasis on sampling. The strength of copulas is in capturing a flexible correlation structure when additional variables, such as temperature or length, are measurable. In conclusion, all three models are effective tools to validate yearly accretional patterns in fish ear bone despite their differences in constraints and assumptions.

Overall, we validated the annual pattern of the MIA data with all three models (ARIMA, UCM, and copula). In this case, we had a full dataset and have shown we can use these three methods with quantitative results that validate the qualitative visual results shown by Barbieri et al. [4] and Foster [3]. We anticipate that copulas will outperform

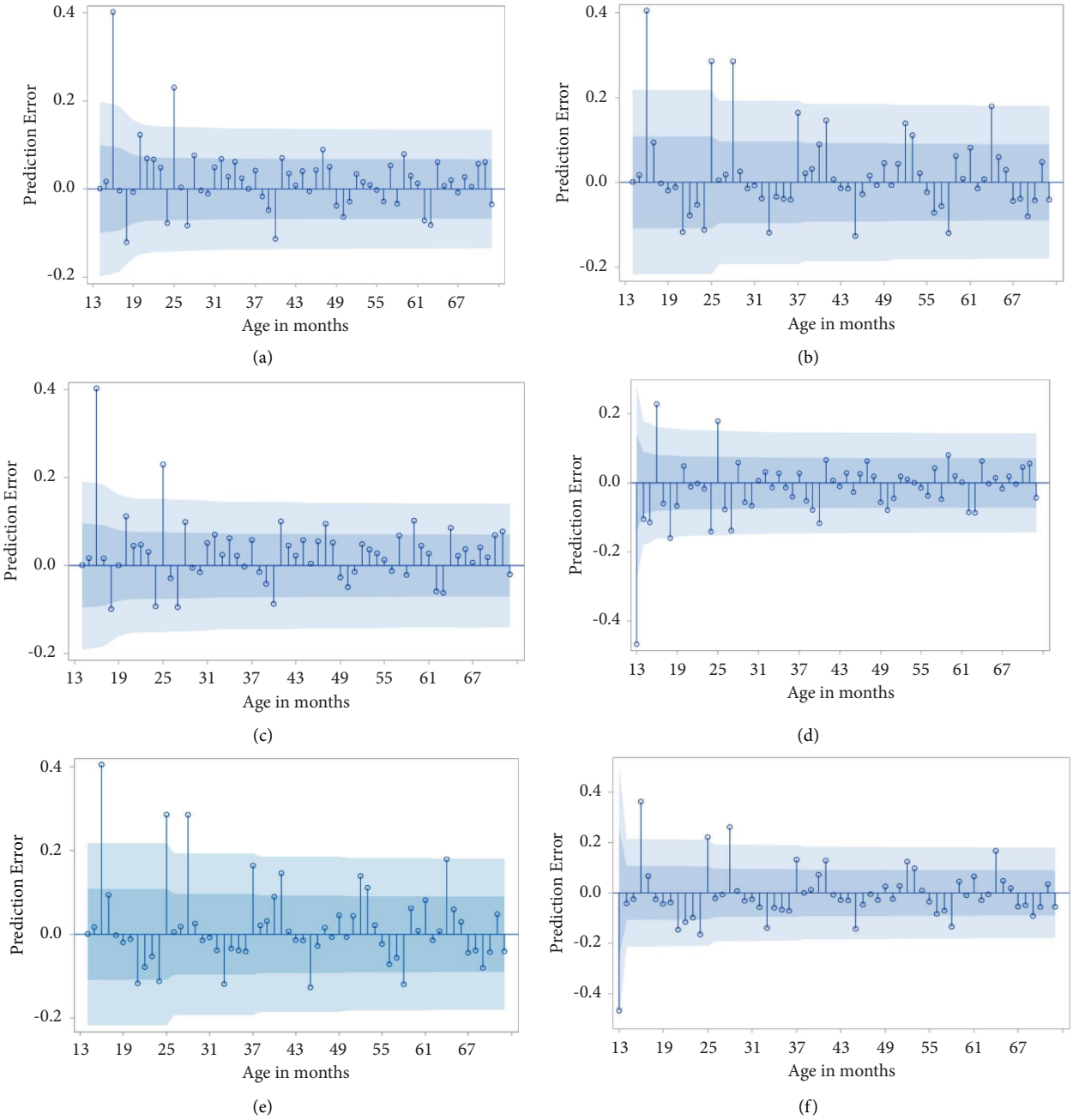

FIGURE 5: Residual plot of the six UCMs. Dark blue indicates one standard error, while light blue indicates two standard errors. Note: models (d, f) are displayed on a different y-scale than the other models to highlight the modeling of the MIA at early ages; also, the x-scale indicates months, i.e., 13 = January, 19 = July, etc. (a) Model 1 uses a stochastic level, slope, cycle, and season. (b) Model 2 uses a stochastic level and slope, no cycle, and stochastic season. (c) Model 3 uses a stochastic level, fixed slope, and stochastic cycle and season. (d) Model 4 uses a stochastic level, no slope, and stochastic cycle and season. (e) Model 5 uses a stochastic level, fixed slope, no cycle, and stochastic season. (f) Model 6 uses a stochastic level, no slope or cycle, and stochastic season.

ARIMA and UCMs when challenged with incomplete data where imputation is necessary, when some variable transformations are not recommended, when missingness cannot be avoided, and the effects of covariates are not removable. In the future, we will test the performance of copulas when challenged with both the sawtooth wave pattern and incomplete datasets (Kirch et al., in process). We hope to increase the use of copula in this field and generalize this

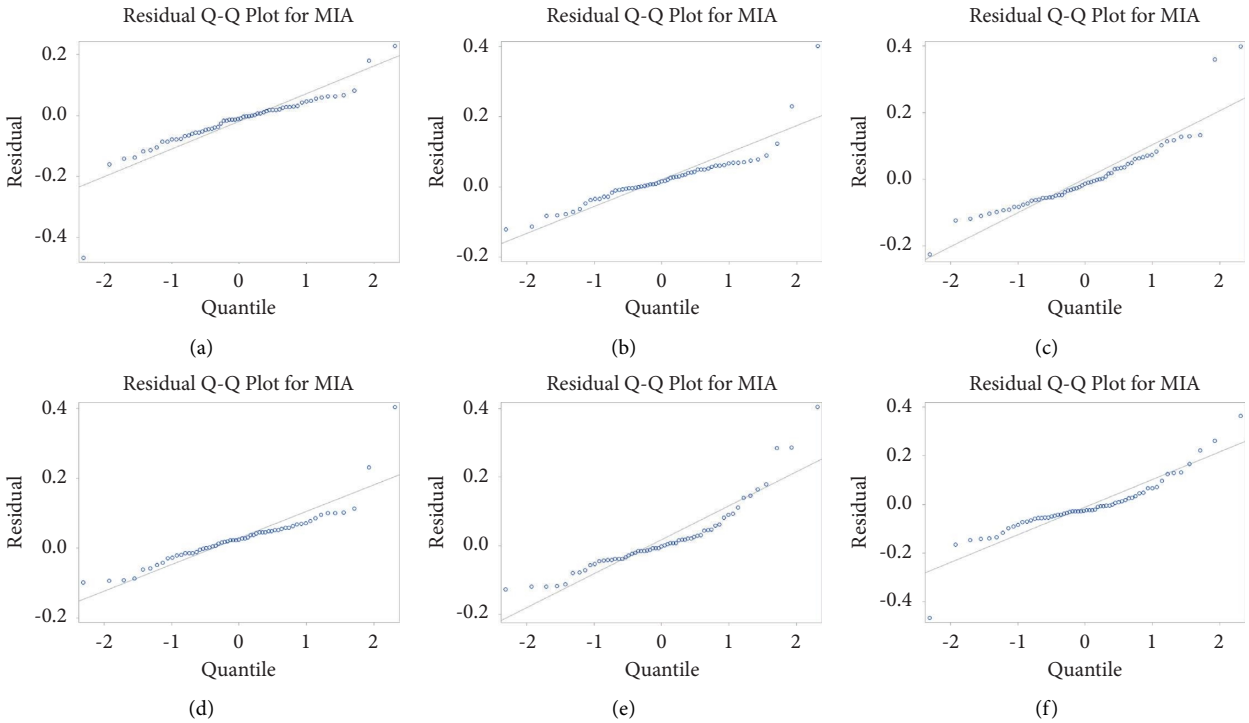

FIGURE 6: QQ plots of the six UCMs. (a) Model 1 uses a stochastic level, slope, cycle, and season. (b) Model 2 uses a stochastic level and slope, no cycle, and stochastic season. (c) Model 3 uses a stochastic level, fixed slope, and stochastic cycle and season. (d) Model 4 uses a stochastic level, no slope, and stochastic cycle and season. (e) Model 5 uses a stochastic level, fixed slope, no cycle, and stochastic season. (f) Model 6 uses a stochastic level, no slope or cycle, and stochastic season.

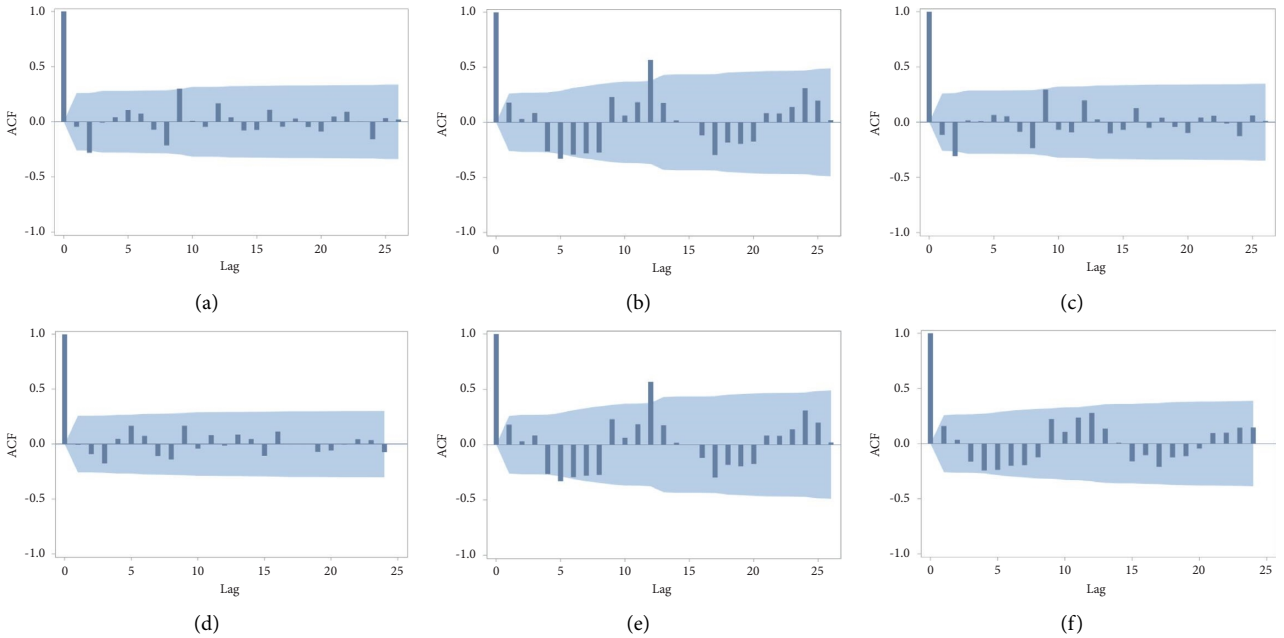

FIGURE 7: Autocorrelation function plot of the six UCMs. Dark blue indicates two standard errors from the mean. (a) Model 1 uses a stochastic level, slope, cycle, and season. (b) Model 2 uses a stochastic level and slope, no cycle, and stochastic season. (c) Model 3 uses a stochastic level, fixed slope, and stochastic cycle and season. (d) Model 4 uses a stochastic level, no slope, and stochastic cycle and season. (e) Model 5 uses a stochastic level, fixed slope, no cycle, and stochastic season. (f) Model 6 uses a stochastic level, no slope or cycle, and stochastic season.

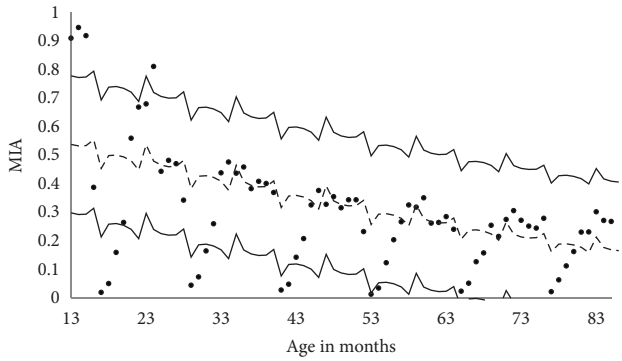

Figure 8: Graph of Gaussian copula with beta marginal fit to model MIA data. The copula estimate is indicated by the dashed line and shows the model fit; a 95% confidence interval is indicated by solid lines.

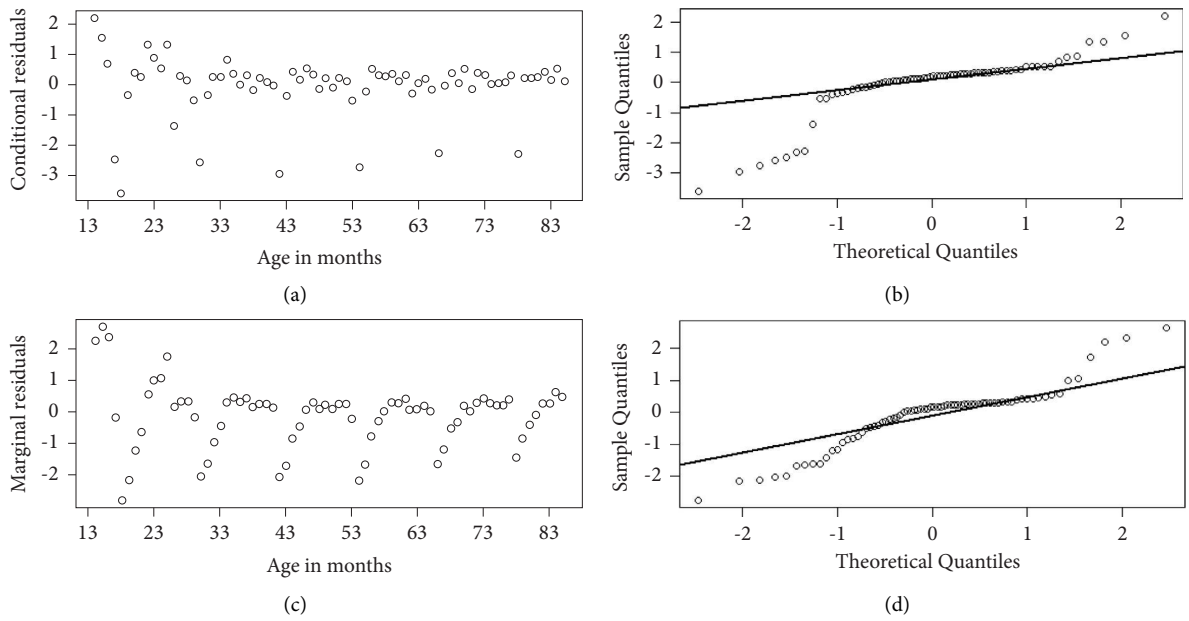

Figure 9: (a) Conditional residuals of the copula model. (b) QQ plot of the conditional residuals of the copula model. (c) Marginal residuals of the copula model. (d) QQ plot of the marginal residuals of the copula model.

method of estimation for higher dimension problems. In further research, the exploration of incomplete datasets and other copulas should bring in interesting results.

Data Availability

The data used to support the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] S. Gebremedhin, S. Bruneel, A. Getahun, W. Anteneh, and P. Goethals, "Scientific methods to understand fish population dynamics and support sustainable fisheries management," *Water*, vol. 13, no. 4, pp. 574–620, 2021.

[2] S. E. Campana, "Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods," *Journal of Fish Biology*, vol. 59, no. 2, pp. 197–242, 2001.

[3] J. R. Foster, "Age, growth, and mortality of Atlantic croaker," *Micropogonias undulatus*, vol. 8, 2001.

[4] L. Barbieri, M. E. Chittenden, and C. M. Jones, "Age, growth, and mortality of Atlantic croaker, *Micropogonias undulatus*, in the Chesapeake Bay Region, with a discussion of apparent geographic changes in population dynamics," *Fishery Bulletin*, vol. 92, no. 1, pp. 1–12, 1994.

[5] P. G. D. A. Martínez, I. Babushkin, L. Bergé et al., "Boosting terahertz generation in laser-field ionized gases using a saw-tooth wave shape," *Physical Review Letters*, vol. 114, no. 18, Article ID 183901, 2015.

[6] H. Okamura, A. E. Punt, Y. Semba, and M. Ichinokawa, "Marginal increment analysis: a new statistical approach of testing for temporal periodicity in fish age verification," *Journal of Fish Biology*, vol. 82, no. 4, pp. 1239–1249, 2013.

[7] Q. E. Phelps, J. H. Kim, C. Lim, and J. S. Odenkirk, "Marginal increment analysis of northern snakehead otoliths," *American Fisheries Society Symposium*, vol. 89, pp. 1–6, 2019.

[8] L. Landler, G. D. Ruxton, and E. P. Malkemper, "Circular data in biology: advice for effectively implementing statistical procedures," *Behavioral Ecology and Sociobiology*, vol. 72, no. 8, pp. 128–210, 2018.

[9] S. Makridakis, S. C. Wheelwright, and V. E. McGee, *Forecasting: Methods and Applications*, John Wiley & Sons, Hoboken, NJ, USA, 1983.

[10] S. J. Koopman and M. Ooms, *Forecasting daily time series using periodic unobserved components time series models*, Tinbergen Institute, Amsterdam, Netherlands, 2004.

[11] B. Zheyong, Z. Zhang, X. Liu, and X. Quin, "Unobserved component model for predicting monthly traffic volume," *Journal of Transpotation Engineering*, vol. 145, no. 12, pp. 1–10, 2019.

[12] Z. Bian, Z. Zhang, X. Liu, and X. Qin, "Unobserved component model for predicting monthly traffic volume," *Journal of Transportation Engineering Part A: Systems*, vol. 145, no. 12, pp. 1–10, 2019.

[13] Y. Yang and H. Zhang, "Spatial-temporal forecasting of tourism demand," *Annals of Tourism Research*, vol. 75, pp. 106–119, 2019.

[14] H. Joe, *Dependence Modeling with Copulas*, Chapman and Hall/CRC, Boca Raton, FL, USA, 2014.

[15] M. S. Alqawba, N. Diawara, and N. Rao Chaganty, "Zero-inflated count time series models using Gaussian copula," *Sequential Analysis*, vol. 38, no. 3, pp. 342–357, 2019.

[16] L. Sun, X. Huang, M. S. Alqawba, J. Kim, and T. Emura, *Copula-based Markov Models for Time Series: Parametric Inference and Process Control*, Springer Nature Singapore Pte Ltd, Singapore, 2020.

[17] G. R. Hosack, G. W. Peters, and S. A. Ludsin, "Interspecific relationships and environmentally driven catchabilities estimated from fisheries data," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 71, no. 3, pp. 447–463, 2014.

[18] C. Marsh, N. Sibanda, M. Dunn, and A. Dunn, "A copula-based habitat preference index in fish spatial population modelling," *Procedia Environmental Sciences*, vol. 27, pp. 2–5, 2015.

[19] F. Alanazi, "The spread of COVID-19 at hot-temperature places with different curfew situations using copula models," in *Proocedings of the 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, New York, NY, USA, August 2021.

[20] A. J. Patton, "A review of copula models for economic time series," *Journal of Multivariate Analysis*, vol. 110, pp. 4–18, 2012.

[21] M. S. Alqawba, D. Fernando, and N. Diawara, "A class of copula-based bivariate Poisson time series models with applications," *Computation*, vol. 9, no. 10, pp. 108–112, 2021.

[22] J. Yin, T. Ma, J. Li, G. Zhang, X. Cheng, and Y. Bai, "Global increases in lethal compound heat stress: hydrological drought hazards under climate change," *Journal of Affective Disorders*, vol. 312, no. 18, pp. 1–8, 2022.

[23] J. Yin, P. Gentine, L. Slater et al., "Future socio-ecosystem productivity threatened by compound drought-heatwave events," *Nature Sustainability*, vol. 6, no. 3, pp. 259–272, 2023.

[24] K. P. Burnham, D. R. Anderson, and K. P. Huyvaert, "AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons," *Behavioral Ecology and Sociobiology*, vol. 65, no. 1, pp. 23–35, 2011.

[25] W. W. Piegorsch and A. J. Bailer, *Analyzing Environmental Data*, John Wiley and Sons Ltd, West Sussex, UK, 2005.

[26] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, Springer Science+Business Media, Inc, New York, NY, USA, 2002.

[27] E. M. Almetwally, S. Dey, and S. Nadarajah, "An overview of discrete distributions in modelling COVID-19 data sets," *Sankhya*, vol. 85, no. 2, pp. 1403–1430, 2022.

[28] R. J. Hyndman and R. Killick, "CRAN task view: time series analysis," 2022, https://CRAN.R-project.org/view=TimeSeries.

[29] A. Guolo and C. Varin, "Beta regression for time series analysis of bounded data, with application to Canada Google® Flu Trends," *Annals of Applied Statistics*, vol. 8, no. 1, pp. 74–88, 2014.

[30] G. Masarotto and C. Varin, "Gaussian copula regression in R," *Journal of Statistical Software*, vol. 77, no. 8, pp. 1–26, 2017.

[31] C. Chesneau, "On new types of multivariate trigonometric copulas," *Applied Mathematics*, vol. 1, no. 1, pp. 3–17, 2021.

[32] C. Chesneau, "A study of the power-cosine copula," *Open Journal of Mathematical Analysis*, vol. 5, no. 1, pp. 85–97, 2021.

[33] C. Chesneau, "A note on a simple polynomial-sine copula," *Asian Journal of Mathematics and Applications*, vol. 2, pp. 1–14, 2022.

[34] R. Salinas-Gutierrez, A. Hernandez-Aguirre, and E. R. Villa-Diharce, "Using copulas in estimation of distribution algorithms," *MICAI 2009: Advances in Artificial Intelligence*, p. 5845, Springer, Berlin, Germany, 2009.

[35] M. Alqawba and N. Diawara, "Copula-based Markov zero-inflated count time series models with application," *Journal of Applied Statistics*, vol. 48, no. 5, pp. 786–803, 2020.