

REGIME SWITCHING OF MIXTURE AUTOREGRESSIVE PROCESS FOR DISCRETIZED TIME EVENTS

R. O. OLANREWAJU^{1*}, S. A. OLANREWAJU², §

ABSTRACT. This paper establishes the need for Poisson random noise for mixture autoregressive process for strictly discretized time events (count series), and as well possessed traits of regime switching, and multimodalities that are usually caused by jumps, fluctuations, and outliers. Consequently, a Poisson mixture autoregressive (PMAR) model with k -regimes, denoted by $PMAR(k : p_1, p_2, \dots, p_k)$ was established and developed, such that, the embedded associated k -regime autoregressive and Poisson coefficients were estimated via Expectation-Maximization (EM) algorithm. The limiting distribution (asymptotic property) of the $PMAR(k : p_1, p_2, \dots, p_k)$ process was ascertained via the Central Limit Theorem (CLT) as well as the lower bound variance estimator of the PMAR process. The model was applied to the significant wave height of the Belmullet Inner (Berth B) and Belmullet Outer (Berth A) of the Atlantic Ocean. The discretized time events of the berths gave a realization of two regimes switching for Berth A and B respectively. The second regime produced a minimum Mean Square Error (MSE) of 12.02 compared to 12.32 produced by first regime.

Keywords: Expectation-Maximization, Limiting Distribution, Multimodalities, Regime-Switching, Poisson.

AMS Subject Classification: 37M10, 60F05, 37A25, 62M10.

1. INTRODUCTION

Linear time series processes, such as Autoregressive (AR), Moving Average (MA), and Autoregressive Moving Average (ARMA) processes were developed to generalize both uniformly and non-uniformly time-variant events, and timely stochastic processes, but vast majority of these processes lacked the ability to capture traits, such as conditional distributions with time-varying dependency variance (volatility), tenure-changing traits, multimodalities ([4], [12], [14]). However, some non-linear time-variant processes,

¹ Africa Business School (ABS), Department of Business Analytics and Value Networks (BAVNs), X4JH+QJR, Avenue Mohamed Ben Abdellah Regragui, Rabat 10112, Morocco.
e-mail: olanrewaju_rasaq@yahoo.com; ORCID: <https://orcid.org/0009-0006-4494-2421>.

² Department of Statistics, University of Ibadan, Ibadan, 900001, Nigeria.
e-mail: sodiqadejare19@gmail.com; ORCID: <https://orcid.org/0009-0002-3852-7361>.

* Corresponding author.

§ Manuscript received: August 27, 2023; accepted: December 20, 2023.

TWMS Journal of Applied and Engineering Mathematics, Vol.15, No.3; © Işık University, Department of Mathematics, 2025; all rights reserved.

such as Threshold Auto-Regressive (TAR), Self-Exciting Threshold Autoregressive (SETAR), Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and its variants (such as EGARCH, APARCH, GJR-GARCH); and fractional integrated processes were able to bridge the time-varying dependency variance (volatility) with conditional distributions ([3], [22]).

By and large, the mentioned non-linearity processes still left out stylized traits, traits like change points like behavior, regime-switching (ability to handle cycles), time-varying volatilities (conditional variances), time-varying mixing weights (transitional weights), and full range of shape changing predictive distributions (otherwise known as multimodalities) ([21], [20], [13]). Ref.[19] proposed a regime-switching model called Mixture Autoregressive (MAR) $MAR(k : p_1, p_2, \dots, p_k)$ process that relaxes all the left-out stylized properties. Consequently, each of the regime has its own marginal conditional distribution (f_t) which is a replica of the immediate past distribution, such that, the conditional mean and variance in each regime (component) depends on the immediate past time.

Moreover, no effort has been made about the time-varying stylized regime switching in accommodating and capturing discretized (count) time events that are affected by large fluctuations (outliers). In advancement of the MAR model with Gaussian random noise, this article will be proposing and developing a Poisson mixture autoregressive (PMAR) process, mathematically denoted by $PMAR(k : p_1, p_2, \dots, p_k)$. The $PMAR(k : p_1, p_2, \dots, p_k)$ process will be in terms of the Poisson random noise distributional form of the MAR process. In line with this scope, the novelty of the $PMAR(k : p_1, p_2, \dots, p_k)$ process will be developed with its associated k -regime-switching, p_k (optimal lag) and transitional weight per each regime, such that the Expectation-Maximization (EM) algorithm will be adopted in estimating these mentioned parameters [22].

2. LITERATURE REVIEW

Ref.[20] unfolded that multiple steps conditional distribution of estimates of futuristic observations from MAR processes were tractable analytically. They claimed that MAR process was autoregressive processes of changing tenures/states of chronological successions of unobserved variables (otherwise called latent variable) that will be used to select varying states or components. However, they found out that if stationary process (es) was/were mixed with non-stationary process (es), the overall process becomes stationary. The problem of ascertaining associated number of components (transitional weights) in MAR and model selection criteria (AIC, BIC, etc.) was worked upon by [9]. Ref.[9] noted that the AIC and its variants were inadequate for the objective function because of their likelihood of overestimating the number of the transitional weights, which usually arise to unreliable estimable coefficients for retention of the MAR model. Due to this lacuna, they formulated a newfangled process performance criterion known as mixture regression criterion (MRC), such that its parameter estimation was based on quasi-likelihood with extended pertinence to non-normal processes, dependent observations, and count explanatory time events. They further proved the efficiency of the limiting distribution of the MRC and showed that it could be performed using MCMC inspections for one or more responses throughout components with unequal or equal sample sizes.

Moreover, Ref.[2] introduced a point mixture model called Location-Mixture Autoregressive (LMAR) process, which captured multiple steps conditional densities. The parameter estimation technique adopted was the Expectation-Maximization (EM) algorithm and precise multiple steps ahead predictions were forecast. They reported that the LMAR model outperformed related and similar models regarding out-of-sample forecasting accuracy. The LMAR model was reported to be superior in predictive performances of real-time

computation when streamlined to location (mean-cluster) parameter estimation alone. Ref.[7] propounded a supposition and verifiable study of chance illation for AR processes with limited (m-mixing weights) mixture of Scale Mixtures of Normal (SMN) innovation. The process involved AR processes with overall mean and mixture-mingling proportions' innovation. They adopted the EM parameter estimation technique, and reported that the derivations needed for Hessian matrix were also developed when preparing the algorithm. Ref.[6] improved and examined Finite Mixture (FM) model, accompanied with flexibility of classification of two parts of distributions based on scale mixtures of normal (TP-SMN) constitutive members. They claimed that the family makes room for robust estimation of FM models development with the ability to capture and absolve asymmetric and symmetric, and heavy and fat-tailed distributions. They further maintained that TP-SMN provides an alternative family member to scale mingling skewed normal (SMSN) family and vital traits of well-hierarchical expression of the family to obtain ML estimates of the model coefficients via an EM parameter estimation technique.

Ref.[16] theoretically assorted autoregressive random processes via a special type of distributional error term, transmuted Gamma distributed noise. After ascertaining that the transmuted Gamma distribution is a proper distribution function (pdf), they theoretically re-parameterized the Gamma parameters in terms of μ and σ^2 into the autoregressive random processes. The mean and variance of the mingling transmuted Gamma autoregressive process was ascertained coupled with its first and second-order stationarity. However, the ingrained k -components' of the mingling transmuted Gamma autoregressive coefficients were estimated via Expectation-Maximization (EM) algorithm, such that the Levinson-Durbin recursive technique was used to derive some step ahead predictions as well as the model sub-setting estimation. Ref.[15] described cyclical-like tenure changing of number of the stocks sold via mingling autoregressive random processes with Poisson and Extreme-Value-Distributions (Fréchet, Gumbel, and Weibull) error terms. They expounded the periodical market behavioral demand and transitory of the stock returns via the associated Markov transitional mixing weights, autoregressive processes, and multimodalities that usually cause the large fluctuations and long-memory. Application wise, it was deduced that Gumbel distributional error outstripped the Poisson and other Extreme-Value-Distributional stochastic errors to give a fitted generalization of Gumbel-MAR (2:1, 1). In extension, Ref.[12] subjected the Poisson and Extreme-Value-Distributions autoregressive random processes to Kullback-Leibler divergence measure in order to ascertain the proximity between the finite/delimited and infinite mixture density of each of the mingling processes.

[17] provided a flexible way to model MAR via a predictive distribution that depends on the history of the process, which accommodate asymmetry and multimodality with the use of Bayesian method. The merit of process was that it incorporates the uncertainty in the estimated processes into the prediction such that all the parameters were cover in the parameter space introduced. Unlike the known approaches, the MCMC approach adopted also introduced a re-labelling algorithm that deals with a posteriori label switching. Relatively, [18] extended the Gaussian Mixture Autoregressive (GMAR) to Fisher's z Mixture Autoregressive (ZMAR) process via the introduction of a four parameter Fisher's z random noise and adopted Bayesian paradigm as a way of quantifying the uncertainty in the modified process. The process extension was based on the mode as a stable location parameter of mixture of k -component Fisher's z autoregressive models with mixing weights that changes overtime. Markov Chain Monte Carlo (MCMC) algorithm was adopted as the parameter estimation technique such that, the process's model evaluation was juxtaposed with GMAR model and Student- t Mixture Autoregressive (TMAR) model.

[11] applied the full-scale real-time observations of the wave climate significant wave height (in metre), peak wave (in $^{\circ}\text{C}$) and sea temperature (in $^{\circ}\text{C}$) of the latitude and longitude of the wave buoys' Belmulletts of the Atlantic Ocean to wave-signal-amplitude cosine and sine regression as an extension to wave signal Fourier function and Wave-Shape Function (WSF) model. The associated regression coefficients were estimated via the Ordinary Least Square (OLS) technique, such that, the model wave signal, frequency, and phase were carved-out in order to ascertain the directional of the wave signal, time in seconds to complete a wave cycle, and its phase respectively.

Based on the review, it was obvious that $MAR(k : p_1, p_2, \dots, p_k)$ has not been treated for strictly uniformly count (discretized) time observations. In support of the need to subject the MAR process to Poisson random noise for one of the candidates of discretized distributions, the mean and variance of the $PMAR(k : p_1, p_2, \dots, p_k)$ process for k -changing tenures (multimodalities) will be ascertained. Additionally, the parameter estimation of the $PMAR(k : p_1, p_2, \dots, p_k)$ process will be via the EM-algorithm, and the limiting distribution (asymptotic property) of the process will be put to check via the Central Limit Theorem (CLT).

Contributively, no effort about time varying MAR process to accommodate, combine, and capture distributional fluctuations, flat stretches, jumps, fluctuations, and full range shape changing predictive distributions (multimodalities) that arises from rare events, specifically series emanated from count (discrete) series has been made. Similarly to the idea of Poisson regression model that was designed to cater for count regressors and dependent variable in order to avoid wasteful and unmatched random noise, the modification of the MAR process towards capturing and accommodating strictly discretized uniformly jump time series was contributively proposed to nullify the possibility of unmatched random noise (that is nullifying the possibility of adopting non-count random noise).

Significantly, the discretized MAR process with Poisson random noise will be applied to count series of the significant wave height observations (in metric) of the Belmullet Inner (Berth B) and Belmullet Outer (Berth A) of the Atlantic Ocean to approximately deduce the number of regimes associated to the 10-years recorded data from May 2012 to April 2021. The associated transitional weight per each regime will not only be deduced, but also how skewedness affected each regime via regime switching lambda will be ascertained for inferential deductions that will be of benefit to Sustainable Energy Authority of Ireland (SEAI). The inferential deduction will make it possible for SEAI to further study other influential climatic factors that must have drove significant wave height observations of the Belmullet Inner (Berth B) and Belmullet Outer (Berth A) of the Atlantic Ocean over the years of study.

3. SPECIFICATION OF POISSON MIXTURE AUTOREGRESSIVE (PMA) MODEL

Ref.[5] & Ref.[10] introduced mixture transition distribution as finite (countable) mixture of transition probability with hidden Markov traits as

$$f(x) = \eta_1 g_1(x) + \eta_2 g_2(x) + \dots + \eta_k g_k(x) \quad (1)$$

Where $f(x)$ is the whole mixture model function and $g_i(x)$ ($i = 1, \dots, k$) are the uniform probability distributions which rely on embedded coefficients; weighted transition probability of $\eta_i > 0$; $\eta_1 + \eta_2 + \dots + \eta_k \approx 1$ for $i = 1, \dots, k$. Ref.[1] & Ref.[8] defined a k -component of Mixture Autoregressive (MAR) model of $MAR(k : p_1, \dots, p_k)$ to be.

$$F(x_{(t)}/f_{t-1}) = P(X_{(t)} \leq x/F_{t-1}) = \sum_{i=1}^k w_k \Phi \left(\frac{x_t - \phi_{k,0} - \phi_{k,1}x_{t-1} - \dots - \phi_{k,p_k} x_{t-p_k}}{\sigma_k} \right) \tag{2}$$

Extending the k -component of MAR model in equation (2) to Poisson Mixture Autoregressive $PMAR(k : p_1, p_2, \dots, p_k)$ gives

$$X_t = \sum_{i=1}^k (\eta_k, \lambda_k) \Phi \left(\frac{x_t - \phi_{k,0} - \phi_{k,1}x_{t-1} - \dots - \phi_{k,p_k} x_{t-p_k}}{\sigma_k} \right) \tag{3}$$

Where,

$$\begin{aligned} x_t &= \phi_{k,0} + \phi_{k,1}x_{t-1} + \dots + \phi_{k,p_k}x_{t-p_k} + \varepsilon_t^{(k)} \\ \varepsilon_t^{(k)} &= x_t - \phi_{k,0} - \phi_{k,1}x_{t-1} - \dots - \phi_{k,p_k}x_{t-p_k} \\ \mu_{k,t} &= \phi_{k,0} + \phi_{k,1}x_{t-1} + \dots + \phi_{k,p_k}x_{t-p_k} \end{aligned}$$

Alternatively,

$$X_t = \begin{cases} \phi_{1,0} + \sum_{i=1}^{p_1} \phi_{1,i} x_{t-1} + \varepsilon_t^{(1)} & \text{with transitional weight } \eta_1 \text{ \& } \lambda_1 \\ \phi_{2,0} + \sum_{i=1}^{p_2} \phi_{2,i} x_{t-2} + \varepsilon_t^{(2)} & \text{with transitional weight } \eta_2 \text{ \& } \lambda_2 \\ \vdots & \vdots \\ \phi_{k,0} + \sum_{i=1}^{p_k} \phi_{k,i} x_{t-p} + \varepsilon_t^{(k)} & \text{with transitional weight } \eta_k \text{ \& } \lambda_k \end{cases} \tag{4}$$

Where, ϕ_{p_k} ranges from $[0, 1]$, for $k = 1, \dots, K$, $p_k \geq 1$. For mixing weights $\eta_1 + \dots + \eta_k \approx 1$, $\eta_i > 0$; for $k = 1 \dots K$; $\Phi(\cdot)$ is the Cumulative Distribution Function (CDF) of the Poisson distribution, where $\varepsilon_t^{(k)}$ & $X_t \approx g(x_t; \lambda) = \frac{\lambda^{x_t} e^{-\lambda}}{x_t!}$; $x_t = 0, 1, 2, 3, \dots$, with mean and variance λ , where ! is the factorial symbol.

3.1. Parameter Estimation for PMA via EM Algorithm. .

Let $X_t = \{X_1, X_2, \dots, X_n\}$; $\lambda_k = \{\lambda_0, \lambda_1, \dots, \lambda_k\}^T$; $\phi_k = \{\phi_{k0}, \phi_{k1}, \dots, \phi_{kp_k}\}^T$; $\eta_k = \{\eta_0, \eta_1, \dots, \eta_k\}^T$ for $k = 1, \dots, K$

Assuming that “Z” is an unobserved random variable, where Z_t is a k -dimensional vector such that $Z = \{Z_1, Z_2, \dots, Z_n\}$

$Z_t = \{Z_1, Z_2, \dots, Z_t\}^T$ Whose component is

$$Z_{i,t} = \begin{cases} 1 & \text{if } X_t \text{ spring up from the } j^{\text{th}} \text{ weight} \\ 0, & \text{otherwise} \end{cases}$$

For $1 \leq j \leq K$, that is, $P(Z_t = (1, 0, \dots, 0)^T) = \eta_1$, $P(Z_t = (0, 1, \dots, 0)^T) = \eta_2$, $\dots, P(Z_t = (0, 0, \dots, 1)^T) = \eta_k$

Let $\Theta = \{\phi_k, \eta_k, \lambda_k\}^T$ be the universal parameter space.

Given Z_t , the Poisson distribution of the complete data (X_t, Z_t) is

$$L_t(\Theta) = \prod_{k=1}^K \left[\eta_k \frac{\lambda_k^{X_t} e^{-\lambda_k}}{X_t!} \right]^{Z_{k,t}} \quad (5)$$

Supposing that the complete data function in equation (5) is at time “t”. The observed data log-likelihood function is $L(\Theta) = \sum_{t=1}^n L_t$. Supposing $L(\Theta) = \sum_{t=L+1}^n L_t$ be the complete data log-likelihood function for sample size (n). So, the directly-maximized log-likelihood function of $L(\Theta)$ gives

$$L_t(\Theta) = \log \prod_{k=1}^K \left[\eta_k \frac{\lambda_k^{X_t} e^{-\lambda_k}}{X_t!} \right]^{Z_{k,t}} \quad (6)$$

$$L(\Theta) = \sum_{t=L+1}^n \left[\sum_{k=1}^K Z_{k,t} \log(\eta_k) + \sum_{k=1}^K Z_{k,t} X_t \log \lambda_k - \sum_{k=1}^K Z_{k,t} \lambda_k - \sum_{k=1}^K Z_{k,t} \log X_t! \right] \quad (7)$$

where,

$$X_t = \phi_{k,0} + \phi_{k,1}x_{t-1} + \cdots + \phi_{k,p_k}x_{t-p_k} = \phi_{k,0} \sum_{i=1}^{p_k} \phi_{k,i}x_{t-k_i}$$

First-order derivatives of $L(\Theta)$ with respect to each of the parameter gives,

$$\frac{\partial L(\Theta)}{\partial \eta_k} = \sum_{t=L+1}^n \left(\frac{Z_{k,t}}{\eta_k} - \frac{Z_{K,t}}{\eta_K} \right) \text{ for } k = 1, \dots, K-1 \quad (8)$$

$$\begin{aligned} \frac{\partial L(\Theta)}{\partial \phi_{k0}} &= \sum_{t=L+1}^n \left(Z_{k,t} \log \lambda_k + \frac{Z_{K,t} \partial X_t!}{X_t!} \right) \\ &= \sum_{t=L+1}^n Z_{k,t} \left(\log \lambda_k + \frac{\partial X_t!}{X_t!} \right) \text{ for } k = 1, \dots, K-1 \end{aligned} \quad (9)$$

$$\frac{\partial L(\Theta)}{\partial \phi_{ki}} = \sum_{t=L+1}^n Z_{k,t} \left(\log \lambda_k + \frac{\partial X_t! \partial X_{t-1}!}{X_t! X_{t-1}!} \right) \text{ for } k = 1, \dots, K-1 \quad (10)$$

$$\begin{aligned} \frac{\partial L(\Theta)}{\partial \lambda_k} &= \sum_{t=L+1}^n \left(\frac{Z_{k,t} X_t}{\lambda_k} - Z_{k,t} \right) \\ &= \sum_{t=L+1}^n Z_{k,t} \left(\frac{X_t}{\lambda_k} - 1 \right) \end{aligned} \quad (11)$$

$$k = 1, 2, \dots, K; \quad i = 0, \dots, p_k$$

Second derivatives of $L_t(\Theta)$ with respect to each of the parameter will be derived with respect to letting a function of x_t be a function of a random variable at time “t” and counter “j” such that

$$\tau(x_t, j) = \begin{cases} 1 & \text{for } i = 0 \\ x_{t-i} & j > 0 \end{cases}$$

So,

$$\frac{\partial^2 L(\Theta)}{\partial \eta_k^2} = - \sum_{t=L+1}^n \left(\frac{Z_{k,t}}{\eta_K^2} - \frac{Z_{k,t}}{\eta_k^2} \right) \tag{12}$$

$$\frac{\partial^2 L(\Theta)}{\partial \eta_k \partial \eta_j} = - \sum_{t=L+1}^n \left(\frac{Z_{k,t}}{\omega_k^2} \right) \text{for } k \neq j \tag{13}$$

$$\frac{\partial^2 L(\Theta)}{\partial \phi_{k0}^2} = \sum_{t=L+1}^n Z_{k,t} \left(\log \lambda_k + \frac{\partial \tau(x_{t-i})!}{\tau(x_{t-i})!} \right) \text{for } k = 1, \dots, K-1 \tag{14}$$

$$\frac{\partial^2 L(\Theta)}{\partial \phi_{k0i} \partial \phi_{k0j}} = \sum_{t=L+1}^n Z_{k,t} \left(\log \lambda_k + \frac{\partial \tau(x_{t-i})! \partial \tau(x_{t-j})!}{\tau(x_{t-i})! \tau(x_{t-j})!} \right) \text{for } i \neq j \tag{15}$$

$$\frac{\partial^2 L(\Theta)}{\partial \phi_{ki}^2} = \sum_{t=L+1}^n Z_{k,t} \left(\log \lambda_k + \frac{\partial X_t! \partial \tau(x_{t-1})!}{X_t! \tau(x_{t-1})!} \right) \tag{16}$$

$$\frac{\partial^2 L(\Theta)}{\partial \phi_{ki} \partial \phi_{kj}} = \sum_{t=L+1}^n Z_{k,t} \left(\log \lambda_k + \frac{\partial X_t! \partial \tau(x_{t-i})! \partial \tau(x_{t-j})!}{X_t! \tau(x_{t-i})! \tau(x_{t-j})!} \right) \text{for } i \neq j \tag{17}$$

$$\frac{\partial^2 L(\Theta)}{\partial \lambda_k^2} = \sum_{t=L+1}^n \left(- \frac{Z_{k,t} X}{\lambda_k^2} \right) \tag{18}$$

$$\frac{\partial^2 L(\Theta)}{\partial \lambda_k \partial \lambda_j} = \sum_{t=L+1}^n \left(- \frac{Z_{k,t} X}{\lambda_k^2 \lambda_j^2} \right) \text{for } k \neq j \tag{19}$$

Applying the EM-algorithm procedure for estimating the universal parameter Θ , through the $L_t(\Theta)$ in equation (7).

E-step: The E-step of the EM-algorithm, we firstly assume that the universal parameter space Θ is known, then the missing values of the unknown data ($Z_{L,t}$) is replaced by the expected value of the observed data “X” of each parameter. Letting $\Omega_{k,t}$ to be the expected value of $Z_{k,t}$ makes $\Omega_{k,t}$ be the transition probability of the expected value in totality divided by the individual transition probability of each expected value (Bayes’ theorem). That is,

$$\Omega_{k,t} = \frac{\eta_k \frac{\lambda_k^{X_t} e^{-\lambda_k}}{X_t!}}{\sum_{k=1}^K \eta_k \frac{\lambda_k^{X_t} e^{-\lambda_k}}{X_t!}} \tag{20}$$

for $k = 1, \dots, K; t = L + 1, \dots, n; Z_{k,t} = \Omega_{k,t}$

M-step: The M-step of the EM algorithm is to maximize the expectations computed in the E-step. The missing data “Z” is assumed to be guessed and filled by their expected values of the parameters. The mixing weight of each regime can then be obtained via $L_t(\Theta)$ by subtracting $\Omega_{k,t}$ from $L_t(\Theta)$ to give,

$$\hat{\eta}_k = \sum_{t=L+1}^n \frac{\Omega_{k,t}}{n - L} \tag{21}$$

Such that $\hat{\phi}_{k p_k}$ for $k = 1, \dots, p_k$ that could also be estimated via a system of equations (that is, estimates of the parameters are then obtained by iterating these two steps until

convergence) or alternatively via Newton-Raphson iterative procedure of all the parameter space at once.

$$\Theta_k^{r+1} = \Theta_k^r + [E(-n\nabla_{\Theta_k}^2 L(\Theta)/(\phi_{k0}, \phi_{ki}, \lambda_k))]^{-1} \times \nabla_{\Theta_k} L(\Theta)/(\phi_{k0}, \phi_{ki}, \lambda_k) \quad (22)$$

3.2. Limiting Behavior (Asymptotic Property) for PMA Model. The limit distribution (Asymptotic property) for PMA model will be based on the Central Limit Theorem (CLT) for providing alternate approximation for probability of sum of independent random variable as the sample size approaches infinity.

The Central Limit Theorem: Let $\{X_i\}_{i=1}^n$ be a sequence of IID random variables, each having mean (μ) and variance (σ). Define $S_n = X_1 + X_2 + \dots + X_n$, $\sigma(S_n)$ as the standard deviation of S_n .

Then the distribution of

$$Z_n = \frac{S_n - E(S_n)}{\sigma(S_n)} = \frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma\sqrt{n}} \quad (23)$$

what intends to the standard normal as $n \rightarrow \infty$. That is,

$$P\left[\frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right] \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{x^2}{2}} dx \quad (24)$$

Letting $S_n = \sum X$, $E(S_n) = n\lambda$, $\sigma(S_n) = \sqrt{n\lambda}$ as the mean and variance of Poisson. The Moment Generating Function (MGF) of Z-variate as known from the normal distribution is

$$Z_n = \frac{\bar{X} - \mu}{\sqrt{\sigma}} \quad (25)$$

From the normal variate, and transforming into PMA model, MGF given by Z_n

$$M_{z(t)} = Z_n = \frac{\bar{X}_{S_n(t)} - n\lambda_k}{\sqrt{n\lambda_k}} = \frac{M(t) - n\lambda_k}{\sqrt{n\lambda_k}} \quad (26)$$

$$M_{z(t)} = \frac{M(t)}{\sqrt{n\lambda_k}} - \frac{n\lambda_k}{\sqrt{n\lambda_k}} \quad (27)$$

$$= M\left(\frac{t}{\sqrt{n\lambda_k}}\right) - t\sqrt{n\lambda_k} \quad (28)$$

Introducing $M_z(t) = e^{tx_t}$

where, $X_t = \phi_{k,0} + \phi_{k,1}x_{t-1} + \dots + \phi_{k,p_k}x_{t-p_k} = \sum_{i=1}^k \left(\phi_{i,0} + \sum_{k=1}^{p_k} x_{t-p_k} \right)$

$$M_{z(t)} = e^{-t\sqrt{n\lambda_k}} M\left(\frac{t}{\sqrt{n\lambda_k}}\right) \quad (29)$$

Recall that $M_{x_t}(t)$ for Poisson distribution is

$$M_{x_t}(t) = \exp \lambda [e^t - 1] = \exp -\lambda [1 - e^t] \quad (30)$$

So,

$$M\left(\frac{t}{\sqrt{n\lambda_k}}\right) = \exp \left[-n\lambda_k \left(1 - e^{-t/\sqrt{n\lambda_k}} \right) \right] \quad (31)$$

$$\Rightarrow M_{Zn}(t) = e^{-t\sqrt{n\lambda_k}} \exp \left[-n\lambda_k \left(1 - e^{-t/\sqrt{n\lambda_k}} \right) \right] \tag{32}$$

$$\log M_{zn}(t) = -t\sqrt{n\lambda_k} - n\lambda_k \left[1 - e^{-t/\sqrt{n\lambda_k}} \right] \tag{33}$$

Recall,

$$e^{-1} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \tag{34}$$

$$e^{X_t} = 1 - \frac{X_t}{1!} + \frac{X_t^2}{2!} - \frac{X_t^3}{3!} + \dots$$

$$e^{-X_t} = 1 + \frac{X_t}{1!} + \frac{X_t^2}{2!} + \frac{X_t^3}{3!} + \dots, \quad X_t = \sum_{i=1}^k \left(\phi_{i,0} + \sum_{k=1}^{p_k} x_{t-p_k} \right)$$

So,

$$e^{-t/\sqrt{n\lambda_k}} = 1 + \frac{t^1}{(\sqrt{n\lambda_k})^1 1!} + \frac{t^2}{(\sqrt{n\lambda_k})^2 2!} + \frac{t^3}{(\sqrt{n\lambda_k})^3 3!} + \dots \tag{35}$$

So,

$$\log M_{zn}(t) = -t\sqrt{n\lambda_k} - n\lambda_k \left(e^{-t/\sqrt{n\lambda_k}} \right) \tag{36}$$

$$\log M_{zn}(t) = -t\sqrt{n\lambda_k} - n\lambda_k \left[1 - \left(1 + \frac{t^1}{(\sqrt{n\lambda_k})^1 1!} + \frac{t^2}{(\sqrt{n\lambda_k})^2 2!} + \frac{t^3}{(\sqrt{n\lambda_k})^3 3!} + \dots \right) \right] \tag{37}$$

$$= -t\sqrt{n\lambda_k} - n\lambda_k \left[1 - 1 - \frac{t^1}{(\sqrt{n\lambda_k})^1 1!} - \frac{t^2}{(\sqrt{n\lambda_k})^2 2!} - \frac{t^3}{(\sqrt{n\lambda_k})^3 3!} - \dots \right] \tag{38}$$

$$= -t\sqrt{n\lambda_k} - n\lambda_k \left[-\frac{t^1}{(\sqrt{n\lambda_k})^1 1!} - \frac{t^2}{(\sqrt{n\lambda_k})^2 2!} - \frac{t^3}{(\sqrt{n\lambda_k})^3 3!} - \dots \right] \tag{39}$$

$$= -t\sqrt{n\lambda_k} + n\lambda_k + \frac{n\lambda_k t}{(\sqrt{n\lambda_k})} + \frac{n\lambda_k t^2}{2(\sqrt{n\lambda_k})^2} + \frac{n\lambda_k t^3}{(\sqrt{n\lambda_k})^3 3!} + \dots \tag{40}$$

$$= -t\sqrt{n\lambda_k} + t\sqrt{n\lambda_k} + \frac{t^2}{2!} + \frac{t^3}{3!\sqrt{n\lambda_k}} + \dots \tag{41}$$

$$= \frac{t^2}{2!} + \frac{t^3}{3!\sqrt{n\lambda_k}} \tag{42}$$

$$\Rightarrow \log M_{zn}(t) = \frac{t^2}{2} + \frac{t^3}{6\sqrt{n\lambda_k}} \tag{43}$$

$$\lim_{n \rightarrow \infty} \log M_{zn}(t) = \frac{t^2}{2} \quad (\text{Continuity from above}) \tag{44}$$

Taking exponential of both sides

$$M_{zn}(t) = \exp \left(\frac{t^2}{2} \right) \tag{45}$$

Comparing $M_{zn}(t) = \exp\left(\frac{t^2}{2}\right)$ with $\exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$, it means $\mu = 0$, $\sigma^2 = 1$

So, as Z_n converges to $N(0, 1)$

Which means that as $n \rightarrow \infty$, the limiting distribution of the PMA asymptotically approaches $N(0, 1)$. It is to be noted that the Poisson distribution is a limiting case of Negative-Binomial distribution.

4. NUMERICAL RESULTS

This research presents an assessment of the significant wave height of the met-ocean observational measurements of the Atlantic Ocean. Real-time observations from the Atlantic Marine Energy Test Site Data Dashboard wave buoys known as Belmullet Inner (Berth B) and Belmullet Outer (Berth A) on the west coast of Ireland of wave energy assessments based on 10-years recorded data from May 2012 to April 2021. The two berths A and B display real-time performance indicators related to wave climate for the Belmullet Inner (Berth B) and Belmullet Outer (Berth A) site respectively. The significant wave height observations gathered for the two berths were measured in metre (m) as whole number (natural; discretized; count numbers). The primary aim is to provide an assessment of wave traits for the two deployment berths based on regime switching of Poisson mixture autoregressive related process.

TABLE 1. Parametric Bootstrap Medians and 95% Percentile C.I of the Atlantic Ocean of Belmullet Inner (Berth B) and Belmullet Outer (Berth A).

Berth	lambda	Median	2.5%	97.5%	Goodness of fitness test
A	282.0723	282.0639	281.9815	282.1536	535.6261
B	282.0723	282.0736	281.9969	282.1322	535.6261

Keys: C.I=Confidence Interval

Interestingly, the mean and median for the significant wave height of berth (A) and (B) approximately coincide in terms of magnitude, this sufficiently suggest that the entire probability distribution for both of the berths are perfectly symmetrical in distribution with approximately zero skewed distribution. It is to be noted that lambda (λ_k) is referred to as the mean of each of the regime, such that, $k = 1, 2$. The estimated magnitude of 282.1322 and 282.1536 for the 97.5 percentile connotes that 97.5% of the significant wave height observations for berths A and B are below 282.1536 and 282.1322 respectively. In a similar manner, significant wave height observations for berths A and B are above 281.9815 and 281.9969. However, since the estimated χ^2 -value of 535.6261 for berth A and B coincide and greater than the critical value of 20.476, we reject the null hypothesis that the berths datasets do not follow Poisson distribution. From Fig.1 below, from the two histograms for the berths, the highest proportion was at approximately 0.12 for each of Belmullet Inner (Berth B) and Belmullet Outer (Berth A). This connote that the most frequent observations of the wave heights are around 300 metre. This shows how frequently the wave height observations fall into the bin of interval 0-500. The bin divided the shaded part into two parts; this suggests two regimes switching for the distribution.

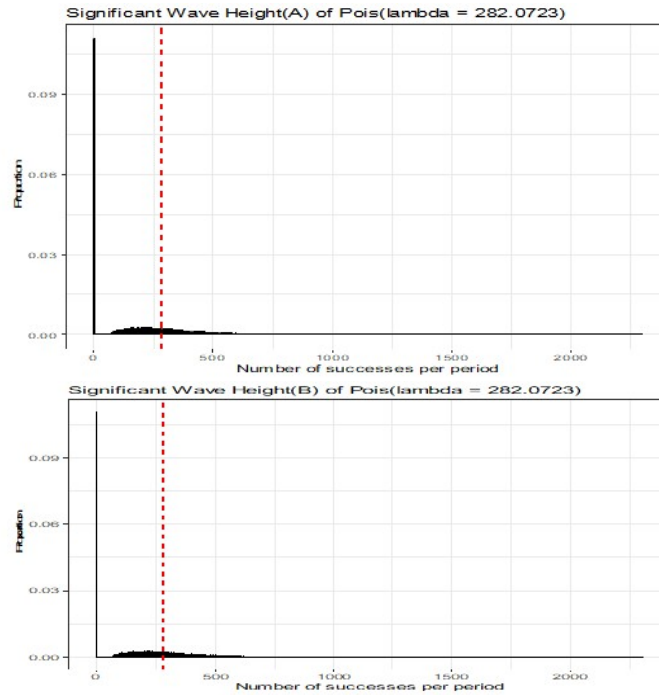


FIGURE 1. Histogram of Significant Wave Height of Belmullet Inner (Berth B) and Belmullet Outer (Berth A).

TABLE 2. Poisson Mixture Autoregressive Coefficients for Significant Wave Height of Belmullet Outer (A).

Berth	MSE	MAIC	MHQIC	MBIC	Moduli of AR
Model Performance	12.32	1275978.56	1276023.09	1276128.46	—
Regime 1	—	—	—	—	(1.55, 1.01, 1.55, 10.84)
Regime 2	—	—	—	—	(1.00, 2.08, 2.08, 4.69)
Mixing Weight	Reg. Mean	Var. Param	D.F. Param	Reg. Var	M.S.E(λ_k)
—	—	—	—	—	—
0.84	13.44	235.27	6.56	5082.61	0.0560
0.16	0.001	0.04	2.00	5.12	0.00076

Keys: Reg. Mean=Regime Mean; Var. Param= Variance Parameter; D.F. Param= Degrees of Freedom Parameters; Reg.Var= Regime Variance; Moduli of AR= Moduli of AR Poly Roots

The Mathematically Model of PMAR(2:4,4) =:

$$\hat{X}_{(A)} = \begin{cases} 36.12 - 0.90x_{t-1} - 0.31x_{t-2} + 0.43x_{t-3} - 0.04x_{t-4} & (\eta_1, \lambda_1) = (0.84, 13.44) \\ 0.001 + 0.58x_{t-1} + 0.31x_{t-2} + 0.15x_{t-3} + 0.05x_{t-4} & (\eta_2, \lambda_2) = (0.16, 0.01) \end{cases} \quad (46)$$

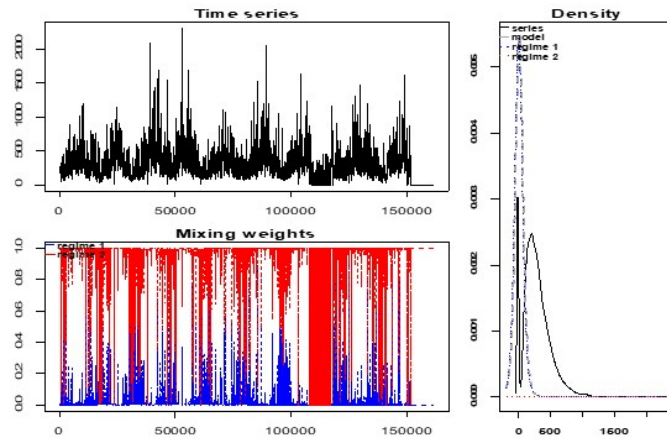


FIGURE 2. Time Plot and Regime Density of the Significant Wave Height of Belmullet Outer (A).

TABLE 3. Poisson Mixture Autoregressive Coefficients for Significant Wave Height of Belmullet Inner (B).

Berth	MSE	MAIC	MHQIC	MBIC	Moduli of AR
Model Performance	12.02	1275978.56	1276023.09	1276128.46	—
Regime 1	—	—	—	—	(1.55, 1.01, 1.55, 10.84)
Regime 2	—	—	—	—	(1.00, 2.08, 2.08, 4.69)
Mixing Weight	Reg. Mean	Var.Param	D.F. Param	Reg. Var	M.S.E(λ_k)
—	—	—	—	—	—
0.84	13.44	235.27	6.56	5082.61	0.0560
0.16	0.001	0.04	2.00	5.12	0.00076

The Mathematically Model of PMAR(2:4,4) =:

$$\hat{X}_{(B)} = \begin{cases} 36.12 - 0.90x_{t-1} - 0.31x_{t-2} + 0.43x_{t-3} - 0.04x_{t-4} & (\eta_1, \lambda_1) = (0.84, 13.44) \\ 0.001 + 0.58x_{t-1} + 0.31x_{t-2} + 0.15x_{t-3} + 0.05x_{t-4} & (\eta_2, \lambda_2) = (0.16, 0.01) \end{cases} \quad (47)$$

4.1. **Discussion of Results.** The Belmullet Inner (Berth B) and Belmullet Outer (Berth A) nearly possessed approximately the same PMAR model performance of AIC=1275978.56; BIC= 1276128.46; and HQIC=1276023.09 respectively. The two berths gave a realization of two regimes' switching of

$$\begin{aligned} \text{Reg}_1 &= 36.12 - 0.90x_{t-1} - 0.31x_{t-2} + 0.43x_{t-3} - 0.04x_{t-4} & (\eta_1, \lambda_1) &= (0.84, 13.44) \\ \text{Reg}_2 &= 0.001 + 0.58x_{t-1} + 0.31x_{t-2} + 0.15x_{t-3} + 0.05x_{t-4} & (\eta_2, \lambda_2) &= (0.16, 0.01) \end{aligned} \quad (48)$$

The transitional weight parameters are 0.84 and 0.003 for regime one (Reg_1) and regime two (Reg_2) respectively; the unconditional means (otherwise known as regime means) are 13.44 and 0.001 respectively. The degree of freedom and regime variance estimates are (6.56, 2.00) and (5082.61, 5.12) for regime one and two respectively. Second regime produced a small degree of freedom compared to the first regime, this connote that the second regime gave a reducing complexity and reduced redundant. Large degree of freedom by first regime might induce numerical problems. The transitional weight estimate of 0.84 by regime one indicates that in the long run, roughly 84% of the observations are generated

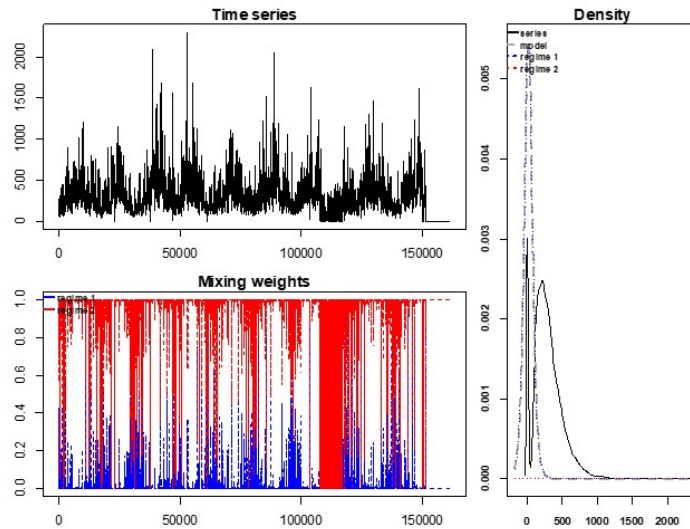


FIGURE 3. Time Plot and Regime Density of the Significant Wave Height of Belmullet Inner (Berth B).

from the first regime. Second regime seems to mostly account for the periods when the series takes smaller values and lesser volatility, because of its smaller variance parameter and regime variance of 0.16 and 5.12 respectively. The AR coefficients are somewhat similar in both berths, implying that it could be appropriate to restrict the berths to be identical despite observations from two separate test locations at various depths. Maybe due to the non-complexity of the second regime, it gave a minimum error forecast of Mean Square Error (MSE) of 12.02 compared to 12.32 produced by first regime.

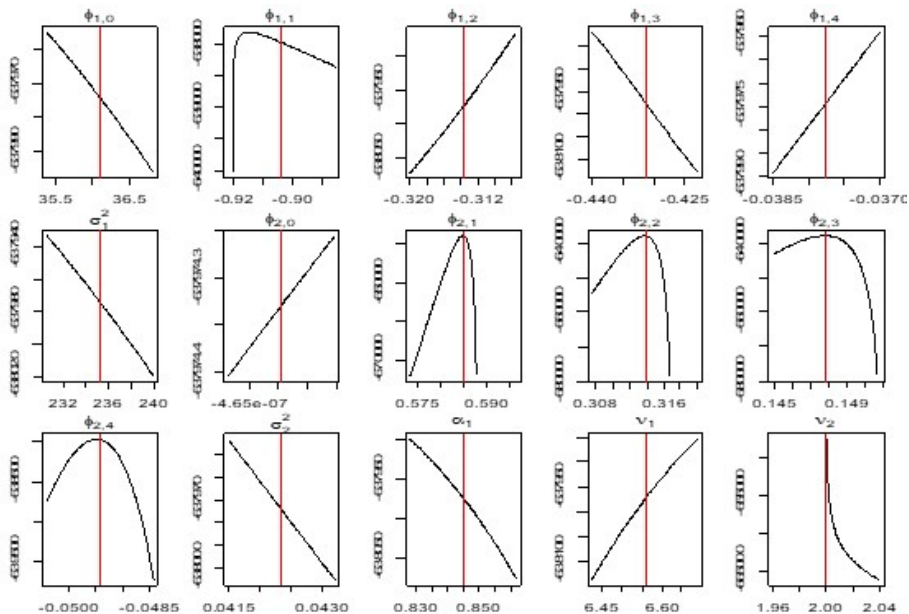


FIGURE 4. Graphical Representation of AR Coefficients Per the Two Regime for the Significant Wave Height of Belmullet Outer (A) and Inner (B).

5. CONCLUSION

The explicated model of $PMAR(k : p_1, p_2, \dots, p_k)$ with Poisson random noise was designed and expounded via mingling autoregressive processes for discretized (count) time events for transitional mixture of autoregressive with stylized properties — multimodalities, tenure-switching, time-varying volatilities (conditional variances), change points like behavior of recurring periodical series. The Poisson marginal distribution for mixture autoregressive model is the simplest counting distributional random noise for the process, though, other flexible counting Probability Mass Functions (PMFs), such as, Binomial, Negative-Binomial, Hyper-geometric, Multinomial, Polya-Aeppli, Poisson-Inverse Gaussian, Beta-Binomial, Compound Poisson distributions etc. could be a close substitute for Poisson random noise. Consequently, it will be ideal to use the $PMAR(k : p_1, p_2, \dots, p_k)$ model for count time events instead of the $MAR(k : p_1, p_2, \dots, p_k)$ with Gaussian random noise. It is to be noted that Gaussian takes values from any of the two sides of the real number line and Poisson for discretized time events is just a proper subset of the positive side of the real number line. This distinction and separation by $PMAR(k : p_1, p_2, \dots, p_k)$ process is needed in other to curb the infinitely many waste and redundancy values in the domain of the Gaussian and real number line. In conclusion, the positive integer of the significant wave height of Belmullet Inner (Berth B) and Belmullet Outer (Berth A) gave a realization of two regimes.

Conflict of Interest. Authors declared no conflict of interest of whatsoever.

Acknowledgement. We acknowledge the Sustainable Energy Authority of Ireland (SEAI) for releasing the Belmullet Inner (Berth B) and Belmullet Outer (Berth A) Atlantic Marine Energy datasets used in this project.

6. APPENDIX

PMA Lower Bound Variance of Estimator

$$g(x_t; \lambda_k) = \frac{\lambda_k^{x_t} e^{-\lambda_k}}{x_t!} \quad x_t = 0, 1, 2, 3, \dots$$

$$\frac{\delta \log g_x(x_t/\lambda_k)}{\delta \lambda_k} = \frac{\delta}{\delta \lambda_k} [-\lambda_k + x_t \log \lambda_k - \log x_t!] \quad (49)$$

$$= 1 + \frac{x_t}{\lambda_k} \quad (50)$$

$$I(\hat{\lambda}_k) = nE_x \left[\frac{\delta \log g_x(x_t/\lambda_k)}{\delta \lambda_k} \right]^2 = nE_x \left[\frac{x_t}{\lambda_k} - 1 \right]^2 \quad (51)$$

$$\frac{n}{\lambda_k^2} [x_t - \lambda_k]^2 = \frac{n}{\lambda_k} \quad (52)$$

$$\frac{\partial E(X_t)}{\lambda_k} = 1$$

Cramer Rao Lower Bound is

$$Var(\hat{\lambda}_k) \geq \frac{1}{n/\lambda_k} = \frac{\lambda_k}{n} \quad (53)$$

Efficiency of λ_k

$$M.S.E(\hat{\lambda}_k) = Var(\hat{\lambda}_k) + b^2(\lambda_k) = \frac{\lambda_k}{n} \quad (54)$$

$$\Rightarrow M.S.E(\hat{\lambda}_k) = Var(\hat{\lambda}_k) = \frac{\lambda_k}{n} \quad (55)$$

This implies that the lower bounded variance of each k -tenure-changing for the $PMAR(k : p_1, p_2, \dots, p_k)$ process is nothing but each regime-means divided by the sample size. In addition, it connotes that the Mean Square Error (MSE) of the k -tenure-changing $PMAR(k : p_1, p_2, \dots, p_k)$ process coincides with its lower bounded variance.

REFERENCES

- [1] Boshankov, G. N., (2006), Prediction with Mixture Autoregressive Models. Research Report No.6, 2006, Probability and Statistics Group School of Mathematics. The University of Manchester.
- [2] Cervone, D., Pillai, N. S., Pati, D., Berbeco, R., Lewis, J. H., (2015), A Location-Mixture Autoregressive Model for online Forecasting of Lung Tumor Motion, the Annals of Applied Statistics, 8(3), pp. 1341-1371. doi:10.1214/14-Aoas744 Institute of Mathematical Statistics.
- [3] Chawsheen, T. A., Broom, M., (2017), Seasonal Time-Series Modeling and Forecasting of Monthly Mean Temperature for Decision Making in the Kurdistan Region of Iraq, Journal of Statistical Theory and Practice, 11(4), pp. 604-633. <http://dx.doi.org/10.1080/15598608.2017.1292484>.
- [4] Graves, T., Gramacy, R., Watkins, N., Franzke, C., (2017), A Brief History of Long-Memory: Hurst, Mandelbrot, and the Road to ARFIMA 1951-1980, Entropy, 19(437). doi:10.3390/e19090437.4.
- [5] Le, N. D., Martin, R. D., Raftery. A. E., (1996), Modeling at Stretches, Burst and Outliers in Time Series using Mixture Transition Distribution Models, Journal of American Statistical Association, 91(436), pp. 1504-1515. doi:10.2307/2291576.
- [6] Maleki, M., Contreras-Reyes, J. E., Mahmoud, M. R., (2019), Robust Mixture Modeling Based on Two-Piece Scale Mixtures of Normal Family, Axioms, 8(38). doi:10.3390/axioms8020038.
- [7] Maleki, M., Nematollahi, M., (2017), Autoregressive Models with Mixture of Scale Mixtures of Gaussian Innovations, Iran Journal of Science and Technology: Transactions A Science, 41(4), 1099–1107. <https://doi.org/10.1007/s40995-017-0237-6>.
- [8] McLachlan, G. J., Lee, S. X., Suren, I. R., (2019), Finite Mixture Models, Annual Review of Statistics and Its Application Finite Mixture Models, 6, pp. 355-378. <https://doi.org/10.1146/annurev-statistics031017-100325>.
- [9] Naik, P. A., Shi, P., Tsai. C-L., (2007), Extending the Akaike Information Criterion to Mixture Regression Models, Journal of the American Statistical Association, 102 (477), pp. 244-254. doi:10.1198/016214506000000861.
- [10] Nguyen, N., (2018), Hidden Markov Model for Stock Trading. International Journal Financial Studies, 6(36). doi:10.3390/ijfs6020036.
- [11] Olanrewaju, R. O., Olanrewaju, S. A., Folorunsho, S., Dada, A. G., (2023). Time Variant Wave-Signal-Amplitude Trigonometry Regression of Latitudes and Longitudes of the Belmullets of the Atlantic Ocean, European Journal of Statistics, 3(12), pp.1-12. doi:10.28924/ada/stat.3.12.
- [12] Olanrewaju, R. O., Waititu, A. G., Nafiu, L. A., (2022), Kullback-Leibler Divergence of Mixture Autoregressive Random Processes via Extreme-Value-Distributions (EVDs) Noise with Application of the Processes to Climate Change. Transactions on Machine Learning and Artificial Intelligence, 10(1), pp. 1-18. doi:10.14738/tmlai.101.11544.
- [13] Olanrewaju, R. O., Waititu, A. G., Nafiu, L. A., (2021), Frechet Random Noise for k -Regime- Switching Mixture Autoregressive Model. American Journal of Mathematics and Statistics, 11(1), pp. 1-10. doi:10.5923/j.ajms.20211101.01.
- [14] Olanrewaju, R. O., Waititu, A. G., Nafiu, L. A., (2021), On the Estimation of k -Regimes-Switching of Mixture Autoregressive Model via Weibull Distributional Random Noise, International Journal of Probability and Statistics, 10(1), pp. 1-8. doi:10.5923/j.ijps.20211001.01.
- [15] Olanrewaju, R. O., Waititu, A. G., Nafiu, L. A., (2021), Bull and Bear Dynamics of the Nigeria Stock Returns Transitory via Mingled Autoregressive Random Processes, Open Journal of Statistics, 11(5), pp. 870-885. doi:10.4236/ojs.2021.115051.
- [16] Olanrewaju, R. O., Ojo, J. F., Adekola, L. O., (2021), Interswitching of Transmuted Gamma Autoregressive Random Processes, Journal of Mathematics and Statistical Science, (ISSN 2411-2518, USA), 7 (7), 183-202.

- [17] Ravagli, D., Boshnakov, G., (2022), Bayesian Analysis of Mixture Autoregressive Models Covering the Complete Parameter Space. *Computational Statistics*, 37, pp. 1399–1433. <https://doi.org/10.1007/s00180-021-01162-8>.
- [18] Solikhah, A., Heri, K., Nur, I., Kartika, F., (2021), Fisher's z Distribution-Based Mixture Autoregressive Model, *Econometrics*, 9(27). <https://doi.org/10.3390/econometrics9030027>.
- [19] Wong, C. S., Chan, W. S., Kam, P. L., (2009), A Student-*t*-Mixture Autoregressive Model with Applications to Heavy-Tailed Financial Data, *Singapore Economic Review Conference*, 1–10.
- [20] Wong, C. S., Li, W. K., (1998), On a Mixture Autoregressive Model, *Journal of Royal Statistical Society Series B, Statistical. Methodology* 2000, 62(1) pp. 95-115.
- [21] Wong, C. S., (1998), *Statistical inference for some Nonlinear Time Series Models*, Ph.D thesis, University of Hong Kong, Hong Kong.
- [22] Zeevi, A., Meir, R., Adler, R. J., (2015), Non-Linear Models for Time Series Using Mixtures of Autoregressive Model, *Journal of Econometric*, 192, 485-498.



Dr. R. O. Olanrewaju received his Ph.D degree in 2022 in Mathematics (Statistics option) from the Pan African University Institute for Basic Sciences, Technology, and Innovation (PAUSTI), Nairobi, Kenya. He is currently a Research Associate with Africa Business School (ABS), Mohammed VI Polytechnic University (UM6P), Rabat, Morocco. To his credit is more than forty-four (44) publications in peer-reviewed journals. His research interest is in the area of Econometric which includes Time Series, Bayesian Time Series, Bayesian Methods, Linear Models, Distribution Theory, and Advanced Probability Theory.



S. A. Olanrewaju holds Bachelor of Science in Statistics from the prestigious University of Ibadan. To his credit is more than ten (10) high-quality publications in peer-reviewed journals.