International Journal of Probability and Statistics 2021, 10(1): 1-8
DOI: 10.5923/j.ijps.20211001.01

# On the Estimation of k-Regimes Switching of Mixture Autoregressive Model via Weibull Distributional Random Noise

**Rasaki Olawale Olanrewaju[1,*], Anthony Gichuhi Waititu[2], Nafiu Lukman Abiodun[3]**

[1]Pan African University, Institute for Basic Sciences, Technology and Innovation (PAUSTI), Kenya
[2]Department of Mathematical Sciences, Jomo Kenyatta University of Agriculture and Technology, Kenya
[3]Department of Mathematics and Statistics, Kabale University, Kabale, Uganda

**Abstract** This paper describes regime-switching, full range of shape changing distributions (multimodalities), and cycles traits that were characterized by time-varying series via Weibull distributional noise for time series with fluctuations and long-memory. We developed and established a Weibull Mixture Autoregressive model of *k*-regimes via $WMAR(k; p_1, p_2,\ldots, p_k)$ with Expectation-Maximization (EM) algorithm adopted as parameter estimation technique. The ergodic process for the $WMAR(k; p_1, p_2,\ldots, p_k)$ model was ascertained via the maximized derivation of the absolute value of the subtraction of its likelihood from its expected likelihood.

**Keywords** Expectation-Maximization, *k*-regimes, Mixture Autoregressive model, Regime-switching, Weibull Distribution

## 1. Introduction

Most economic and financial time series possesses traits such as regime shifting, outburst, outliers and change point like behavior that have be addressed by some non-linear time series models, such as Autoregressive Moving Average (ARMA), Self-Exciting Threshold Auto-Regressive (SETAR) and Autoregressive Conditional Heteroscedasticity (GARCH), and its variants. Problems of overall-stationarity, Conditional Heteroscedasticity, excessive skewness and kurtosis, non-linearity violation, full range of shape changing predictive distributions (multimodality) and ability to handle cycles are yet to be addressed fully in modeling fluctuating time series data [9], [2]. Financial, climate and economic time series are often driven by unimodal innovation series. This often implies a unimodal marginal and/or a unimodal conditional distribution for the time series itself. In reality, many financial, climate and economic time series exhibits multimodality either in the marginal or the conditional distribution. However, this article will be adopting a more robust marginal distributional of Weibull distributional form in the build-up of specifying the Mixture Autoregressive model denoted by $MAR(k; p_1, p_2,\ldots, p_k)$. The Weibull marginal distribution for autoregressive regimes-switching model will be adopted as a substitute for widely used Gaussian marginal distribution because of its ability to capture and model contaminated series characterized with traits such as skewness, kurtosis, *k*-regimes, outburst, outliers and change point like behavior [6], [5]. The $MAR(k; p_1, p_2,\ldots, p_k)$ with Gaussian distributional form of the random noise lacks the ability to fully capture and represent distortion caused by contaminated series with above said traits. Therefore, this paper develops, ascertains and estimates the Mixture Autoregressive of *k*-regimes with Weibull marginal distributional form, denoted by $WMAR(k; p_1, p_2,\ldots, p_k)$.

* Corresponding author:
rasakiolawale@gmail.com (Rasaki Olawale Olanrewaju)
Received: Dec. 7, 2020; Accepted: Dec. 30, 2020; Published: Feb. 6, 2021
Published online at http://journal.sapub.org/ijps

## 2. Literature Review

Regime-switching generalization via Mixture Autoregressive (MAR) model was propounded by [10] to relax all these mentioned stylized properties of outburst, outliers and change-point like behavior characterized linear and non-linear time series model but yet to be addressed by ARMA, ARIMAX, SETAR and GARCH. Shortly afterwards, [1], In his work “Prediction with Mixture Autoregressive models” deduced Mixture Autoregressive (MAR) models have the attractive property that the shape of the conditional distribution of a forecast depends on the recent history of the process. It was also ascertained by [1] that if the original MAR model is a mixture of normal distributions, then, the multi-step distributions are also mixtures of normal distributions.

A $p^{th}$ ordered model and explicitly expressed dimensional stationary distribution giving a mixture of Normal distribution with constant mixing weights via a general formulation for a univariate nonlinear autoregressive model was presented by [6]. They presented an illustration via an empirical example of interest rates of ten (10) years. In advancement to [6], [5] proposed a novel nonlinear Vector Autoregressive (VAR) otherwise called Gaussian Mixture Vector autoregressive (GMVAR) model. They developed and explained its asymptotic theory of maximum likelihood whose usefulness is vital when dealing with bivariate time series settings. [3] adopted Student's *t*- distribution as the error term for mixture autoregressive model via. Their intention was catering for meet inadequate ergodicity and stationary properties that had been a problem by Gaussian error term.

Furthermore, [4] 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 make 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 of 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.

From the available and reviewed literature, it was glaring that the $MAR(k; p_1, p_2, \dots, p_k)$ has not been subjected to $MAR(K : p_1, \dots, p_k)$ any of the candidates of the Extreme Valued Distributions (EVDs). In support of the need to subject the MAR model to any of the EVDs in order to capture fluctuation caused by extreme values, this article will be adopting the Weibull as the marginal distribution (random noise) for the MAR model as well as the extreme valued distributional form. The mean and variance of the multimodal conditional distribution for the $WMAR(k; p_1, p_2, \dots, p_k)$ will be ascertained. The Expectation-Maximization (EM) estimation technique will be adopted via E-step and M-step leading to a system of equation and Newton-Raphson iterative technique for estimating AR coefficients, and standard errors attached to each regime. As well as the Ergodic Process for the model.

## 3. Detailed Description of the Gamma Mixture Autoregressive Model

Normal mixture transitional distribution (GMTD) models for conditional Normal distribution was firstly introduced by [7], it was splinted-out and viewed as a finite (countable) mixture distribution by [11] as;

$$p(x) = w_1 f_1(x) + w_2 f_2(x) + \cdots + w_k f_k(x) \tag{1}$$

where $p(x)$ is whole mixture regime-switching probability density function of identically distributed function and $f_i(x)(i = 1, \dots, n)$ are the probability density functions which may depend on certain parameters; mixing weight or weighted probability $w_i > 0$ *such that* $w_1 + w_2 + \cdots + w_k = 1$ for $i = 1, \dots, n$.

A k-component of Mixture Autoregressive (MAR) model was defined by [9], [11] and [8] to be

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

Extending the k-component of MAR model in (2) to Weibull Mixture Auto-Regressive $WMAR(k; p_1, p_2, \dots, p_k)$ gives

$$X_t = \sum_{i=1}^{k} (w_k, \alpha_k, \beta_k) \, \Phi\left(\frac{x_t - \phi_{k,0} - \phi_{k,1} x_{t-1} - \cdots - \phi_{k,pk} \, x_{t-pk}}{\sigma_k}\right) \tag{3}$$

where; $x_t = \phi_{k,0} + \phi_{k,1} x_{t-1} + \cdots + \phi_{k,p} x_{t-pk} + \varepsilon_t^{(k)}$ $\quad \varepsilon_t^{(k)} = x_t - \phi_{k,0} - \phi_{k,1} x_{t-1} - \cdots - \phi_{k,p} x_{t-pk}$

$\mu_{k,t} = \phi_{k,0} + \phi_{k,1} x_{t-1} + \cdots + \phi_{k,p} x_{t-pk}$

Otherwise,

$$Xt = \begin{cases} \phi 1,0 + \sum_{i=1}^{p1} \phi 1, i\, xt-1 + \varepsilon_t^{(1)} & \text{with probability } \omega 1 \quad \& \, (\alpha_1, \beta_1) \\ \phi 2,0 + \sum_{i=1}^{p2} \phi 2, i\, xt-2 + \varepsilon_t^{(2)} & \text{with probability } \omega 2 \quad \& \, (\alpha_2, \beta_2) \\ \vdots & \vdots \qquad\qquad \vdots \\ \phi_k, 0 + \sum_{i=1}^{pk} \phi_k, i\, xt-p + \varepsilon_t^{(k)} & \text{with probability } \omega_k \quad \& \, (\alpha_k, \beta_k) \end{cases} \tag{4}$$

such that the model is denoted by $WMAR(k; p_1, p_2, \ldots, p_k)$, where $F\left(x_{(t)} / f_{t-1}\right)$ is the conditional cumulative distribution function of $X_t$ given the immediate past information, evaluated at $x_t$, $\phi_{pk}, \theta_{qk} \in (0,1)$, $0 < \phi_{pk}, \theta_{qk} < 1$, for $k = 1, \ldots\ldots, K$, $p, q \geq 1$. For mixing weights (weighted probabilities) $\omega_1 + \ldots + \omega_k \approx 1$, $\omega_i > 0$, for $k = 1 \cdots K$. $\Phi(.)$ is the Cumulative Distribution Function of the standard Weibull, where $\varepsilon_t^{(k)}$ & $X_t \approx g\left(x_t; \alpha, \beta\right) = \frac{\alpha}{\beta}\left(\frac{x_t}{\beta}\right)^{\alpha-1} \exp\left(-\frac{x_t}{\beta}\right)^{\alpha}$, with mean $E(X) = \beta\Gamma\left(1 + \frac{1}{\alpha}\right)$ and variance $\sigma^2 = \beta^2\left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \left(\Gamma\left(1 + \frac{1}{\alpha}\right)\right)\right]^2$.

### 3.1. The Conditional mean and Variance for WMAR Model

The conditional mean and variance for WMA model of $x_t$ given the immediate past information is as follow:

$$
\begin{aligned}
E\left(x_t / g(\alpha, \beta)_{t-1}\right) &= \sum_{k=1}^{K} \left(\omega_k, \alpha_k, \beta_k\right)\left(\phi_{k0} + \phi_{k1} x_{t-1} + \cdots + \phi_{kpk} x_{t-pk}\right) \\
&= \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k)\, \mu_{k,t} \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k)\, \mu_{k,t}
\end{aligned} \tag{5}
$$

since, $\mu_{k,t} = \sum_{k=1}^{K}\left(\phi_{i0} + \sum_{p=1}^{pk} \phi_{i1} x_{t-pk}\right) = \phi_{k0} + \phi_{k1} x_{t-1} + \cdots + \phi_{kpk} x_{t-pk}$

Since the mean of Weibull is $E(X) = \beta\Gamma\left(1 + \frac{1}{\alpha}\right)$,

$$
E\left(x_t / g(\alpha, \beta)_{t-1}\right) = \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k)\, \mu_{k,t} = \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k)\, \beta_k\, \Gamma\left(1 + \frac{1}{\alpha_k}\right) \tag{6}
$$

which depends on immediate past values of the time series and $g(\alpha, \beta)$ is the PDF of Weibull.

$$
E(x_t^2 / g(\alpha, \beta) x_{t-1}) = \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \sigma_k^2 + \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}^2
$$

So,

$$
Var(x_t / g(\alpha, \beta) x_{t-1}) = \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \sigma_k^2 + \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}^2 - \left(\sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}\right)^2 \tag{7}
$$

But, $\left(\sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}\right)^2 = \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k)\left(\mu_{k,t} - \sum_{s=1}^{K} (\omega_s, \alpha_s, \beta_s) \mu_{s,t}\right)^2$ for $s = t-1$ (immediate past values of the time series that depends on "$t$" the present values).

$$
\sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k)\left(\mu_{k,t} - \sum_{s=1}^{K} (\omega_s, \alpha_s, \beta_s) \mu_{s,t}\right)^2 = \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}^2 + \left(\sum_{s=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}\right)^2 \tag{8}
$$

Therefore,

$$
\begin{aligned}
Var(x_t / g(\alpha, \beta) x_{t-1}) &= \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \sigma_k^2 + \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}^2 - \left(\sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}\right)^2 \\
&= \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \sigma_k^2 + \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}^2 - \sum_{k=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}^2 - \left(\sum_{s=1}^{K} (\omega_k, \alpha_k, \beta_k) \mu_{k,t}\right)^2
\end{aligned}
$$

$$Var(x_t / g(\alpha,\beta)x_{t-1}) = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\sigma_k^2 - \left(\sum_{s=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}\right)^2 \quad (9)$$

The expression $\sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\sigma_k^2 - \left(\sum_{s=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}\right)^2$ is a positive (non-negative) and would be tantamount to zero if and only if the each regime mean is equal to each other, that is, $\mu_{1,t} = \mu_{2,t} = \mu_{3,t},\cdots,\mu_{K,t}$. The expression satisfies the non-negativity property of variance, so the variance **for WMAR Model is non-negative.**

Since the mean and variance of Weibull PDF are;

$$\beta\Gamma\left(1+\frac{1}{\alpha}\right) \text{ and } \sigma^2 = \beta^2\left[\Gamma\left(1+\frac{2}{\alpha}\right) - \left(\Gamma\left(1+\frac{1}{\alpha}\right)\right)\right]^2 \quad \textit{for } \alpha > 0 \text{ respectively}$$

Alternatively, $Var(x_t / g(\alpha,\beta)x_{t-1})$ can be rewritten as

$$Var(x_t / g(\alpha,\beta)x_{t-1}) = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\beta_k^2\left[\Gamma\left(1+\frac{2}{\alpha_k}\right) - \left(\Gamma\left(1+\frac{1}{\alpha_k}\right)\right)\right]^2 - \left(\sum_{s=1}^{K}(\omega_k,\alpha_k,\beta_k)\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)\right)^2 \quad (10)$$

### 3.2. Parameter Estimation for WMAR via EM Algorithm

Adopting the Expectation-Maximization (EM) algorithm Let $X = \{X_1, X_2,\cdots,X_n\}$ ; $\alpha_k = \{\alpha_0,\alpha_1,\cdots,\alpha_k\}^T$ ; $\beta_k = \{\beta_0,\beta_1,\cdots,\beta_k\}^T$ ; $\phi_k = \{\phi_{k0},\phi_{k1},\cdots,\phi_{k\,pk}\}^T$ $\omega_k = \{\omega_0,\omega_1,\cdots,\omega_k\}^T$ for $k = 1,\cdots,K$ .

Furthermore, let "S" be the unobserved random variable where $S_t$ is a $k-$ dimensional vector such that $S = \{S_1,S_2,\cdots,S_n\}$

$S_t = \{S_1,S_2,\cdots,S_t\}^T$ whose component is

$$S_{i,t} = \begin{cases} 1 & \textit{if } X_t \textit{ originated from the } j^{th} \textit{ weight (state)} \\ 0, & \textit{otherwise} \end{cases}$$

For $1 \le j \le K$ , that is,

$$P\left(S_t = (1,0,\cdots,0)^T\right) = \omega_1$$

$$P\left(S_t = (0,1,\cdots,0)^T\right) = \omega_2$$

$$\vdots$$

$$P\left(S_t = (0,0,\cdots,1)^T\right) = \omega_k$$

Let $\Theta = \{\phi_k,\omega_k,\alpha_k,\beta_k\}^T$ be the parameter space.

Given $S_t$ , the Weibull distribution of the complete data $(X_t,S_t)$ is given by

$$L_t(\Theta) = \prod_{k=1}^{K}\left[\omega_k\frac{\alpha_k}{\beta_k}\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1}\exp\left(-\frac{x_t}{\beta_k}\right)^{\alpha_k}\right]^{S_{k\,t}} \quad (11)$$

Allowing $L_t(\Theta)$ to be the conditional log-likelihood function at time "$t$". The log-likelihood is then $L(\Theta) = \sum_{t=1}^{n} L_t$ . Let

$L(\Theta)=\sum_{t=L+1}^{n} L_t$ be the joint conditional log-likelihood function for large sample size (n) that makes the effect of $\sum_{t=L}^{n} L_t$ trifling. So, the maximizing the conditional log-likelihood function $L(\Theta)$ gives

$$L_t(\Theta)=\log\prod_{k=1}^{K}\left[\omega_k\frac{\alpha_k}{\beta_k}\left(\frac{X_t}{\beta_k}\right)^{\alpha_k-1}\exp\left(-\frac{X_t}{\beta_k}\right)^{\alpha_k}\right]^{S_{kt}} \tag{12}$$

$$L(\Theta)=\sum_{t=L+1}^{n}\left[\sum_{k=1}^{K}S_{kt}\log(\omega_k)+\sum_{k=1}^{K}S_{kt}\log\alpha_k+\sum_{k=1}^{K}S_{kt}\alpha_k\log X_t-\sum_{k=1}^{K}S_{kt}\log X_t-\right.$$
$$\left.\sum_{k=1}^{K}S_{kt}\alpha_k\log\beta_k-\sum_{k=1}^{K}S_{kt}\alpha_k X_t+\sum_{k=1}^{K}S_{kt}\alpha_k\beta_k\right] \tag{13}$$

Where $X_t=\phi_{k0}+\phi_{k1}x_{t-1}+\cdots\cdots+\phi_k x_{t-pk}=\phi_{k0}+\sum_{i=1}^{Pk}\phi_{ki}x_{t-ki}$

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

$$\nabla_{w_k}L(\Theta)=\sum_{t=L+1}^{n}\left(\frac{S_{kt}}{w_k}-\frac{S_{Kt}}{w_K}\right)\text{for } k=1,\cdots,K-1 \tag{14}$$

$$\nabla_{\phi_{k0}}L(\Theta)=\sum_{t=L+1}^{n}\left(\frac{\alpha_k S_{Kt}}{X_t}-\frac{S_{Kt}}{X_t}+S_{Kt}\alpha_k\right)$$
$$=\sum_{t=L+1}^{n}S_{Kt}\left(\frac{\alpha_k}{X_t}-\frac{1}{X_t}+\alpha_k\right)$$
$$=\sum_{t=L+1}^{n}S_{Kt}\left(\frac{\alpha_k-1}{X_t}+\alpha_k\right)\text{for } k=1,\cdots,K-1 \tag{15}$$

$$\nabla_{\phi_{ki}}L(\Theta)=\sum_{t=L+1}^{n}\left(\frac{\alpha_k S_{Kt}-S_{Kt}}{X_t X_{t-i}}+S_{Kt}\alpha_k X_{t-i}\right)$$
$$=\sum_{t=L+1}^{n}S_{Kt}\left(\frac{\alpha_k-1}{X_t X_{t-i}}+\alpha_k X_{t-i}\right)\text{for } k=1,\cdots,K-1 \tag{16}$$

$$\nabla_{\alpha_k}L(\Theta)=\sum_{t=L+1}^{n}\left(\frac{S_{kt}}{\alpha_k}+S_{kt}\log X_t-S_{kt}\log\beta_k+S_{kt}X_t+S_{kt}\beta_k\right)$$
$$=\sum_{t=L+1}^{n}S_{kt}\left(\frac{1}{\alpha_k}+\log\frac{X_t}{\beta_k}+X_t+\beta_k\right)$$

$$\nabla_{\beta_k}L(\Theta)=\sum_{t=L+1}^{n}\left(-\frac{S_{kt}\alpha_k}{\beta_k}+\alpha_k S_{kt}\right)$$
$$=\sum_{t=L+1}^{n}\alpha_k S_{kt}\left(1-\frac{1}{\beta_k}\right) \tag{17}$$

For $k=1,2,\cdots,K;\quad i=0,\cdots,pk$

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

Let the function $x_t$ be a function of a random variable at time "$t$" and counter "$j$"

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

So,

$$\nabla^2_{\omega_k} L(\Theta) = -\sum_{t=L+1}^{n} \left( \frac{S_{K,t}}{\omega_K^2} - \frac{S_{K,t}}{\omega_K^2} \right) \tag{18}$$

$$\nabla_{\omega_k} L(\Theta) \nabla_{\omega_j} L(\Theta) = -\sum_{t=L+1}^{n} \left( \frac{S_{k,t}}{\omega_k^2} \right) \quad for \quad k \neq j \tag{19}$$

$$\nabla^2_{\phi_{k0}} L(\Theta) = \sum_{t=L+1}^{n} \left( \frac{\alpha_k S_{Kt}}{m(x_{t-i})^2} - \frac{S_{Kt}}{m(x_{t-i})^2} \right) = \sum_{t=L+1}^{n} \frac{S_{Kt}}{m(x_{t-i})^2} (\alpha_k - 1) \tag{20}$$

$$\nabla_{\phi_{koi}} L(\Theta) \nabla_{\phi_{koj}} L(\Theta) = \sum_{t=L+1}^{n} \frac{S_{Kt}}{m(x_{t-i}) m(x_{t-j})} (\alpha_k - 1) \quad for \quad i \neq j \tag{21}$$

$$\nabla^2_{\phi_{ki}} L(\Theta) = \sum_{t=L+1}^{n} S_{Kt} \left( \frac{\alpha_k - 1}{X_t m(x_{t-i})^2} + \alpha_k m(x_{t-i})^2 \right) \tag{22}$$

$$\nabla_{\phi_{ki}} L(\Theta) \nabla_{\phi_{kj}} L(\Theta) = \sum_{t=L+1}^{n} S_{Kt} \left( \frac{\alpha_k - 1}{X_t m(x_{t-i})^2 m(x_{t-j})^2} + \alpha_k m(x_{t-i})^2 m(x_{t-j})^2 \right) \quad for\ \ i \neq j \tag{23}$$

$$\nabla_{\alpha_k \alpha_k} L(\Theta) = \sum_{t=L+1}^{n} -S_{kt} \left( \frac{1}{\alpha_k^2} \right) \tag{24}$$

$$\nabla_{\alpha_k} L(\Theta) \nabla_{\alpha_j} L(\Theta) = \sum_{t=L+1}^{n} -S_{kt} \left( \frac{1}{\alpha_k^2 \alpha_j^2} \right) \quad for \quad k \neq j \tag{25}$$

$$\nabla_{\beta_k} L(\Theta) = \sum_{t=L+1}^{n} \left( -\frac{S_{kt}\alpha_k}{\beta_k} + \alpha_k S_{kt} \right) \tag{26}$$

$$\nabla_{\beta_k} L(\Theta) \nabla_{\beta_k} L(\Theta) = \sum_{t=L+1}^{n} \left( \frac{S_{kt}\alpha_k}{\beta_k^2} \right) \tag{27}$$

$$\nabla_{\beta_k} L(\Theta) \nabla_{\beta_j} L(\Theta) = \sum_{t=L+1}^{n} \sum_{t=L+1}^{n} \left( \frac{S_{kt}\alpha_k}{\beta_k^2 \beta_j^2} \right) \quad for \quad \cancel{k} \tag{28}$$

Employing the EM algorithm procedure for estimating $\Theta$, the parameter space via the $L(\Theta)$ in equation (13).

The first step from the acronym EM algorithm, that is the E-step, suppose the parameter space $\Theta$ is available, then the missing values for the unobserved data $(S_{L,t})$ is then replaced by impose means imposed on each parameter on the observed data X. Allowing $\gamma_{k,t}$ to be the imposed mean of $S_{k,t}$, then $\gamma_{k,t}$ is individual transition probability of the imposed mean all over the totality of the transition probability imposed expectation (that is, Bayes' theorem)

$$\gamma_{k,t} = \frac{\omega_k \frac{\alpha_k}{\beta_k} \left( \frac{X_t}{\beta_k} \right)^{\alpha_k - 1} \exp\left( -\frac{X_t}{\beta_k} \right)^{\alpha_k}}{\sum_{k=1}^{K} \omega_k \frac{\alpha_k}{\beta_k} \left( \frac{X_t}{\beta_k} \right)^{\alpha_k - 1} \exp\left( -\frac{X_t}{\beta_k} \right)^{\alpha_k}} \tag{29}$$

For $k = 1, \cdots, K; \quad t = L+1, \cdots, n; \quad S_{k,t} = \gamma_{k,t}$

The second case (M-step) where the missing data S is assumed to be guessed and to be replaced by their imposed means on the parameters. The estimates of the parameters of the parameter space, $\Theta$. The estimates of the parameters $\Theta$ can be obtained via $L(\Theta)$ by subtracting $\gamma_{k,t}$ from $L(\Theta)$ to give,

$$\omega_k = \sum_{t=L+1}^{n} \frac{\gamma_{k,t}}{n-L} \tag{30}$$

Such that $\phi_{k\,pk}$ "$k$" for that runs from $1,\cdots,p_k$ that could also be estimated via a system of equations (that is, the 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.

$$\Theta_k^{r+1} = \Theta_k^{r} + \left[E\left(-n\nabla^2_{\Theta_k} L(\Theta)/_{\left(\phi_{k0},\phi_{ki},\alpha_k,\beta_k\right)}\right)\right]^{-1} \times \nabla_{\Theta_k} L(\Theta)/_{\left(\phi_{k0},\phi_{ki},\alpha_k,\beta_k\right)} \tag{31}$$

### 3.3. Ergodic Process for WMAR Model

Strictly stationary and ergodic process WMAR model would be ascertained via

$$\lim_{l\to\infty} \mathrm{P}\left(\max\left|\frac{1}{n}\sum_{i=1}^{n}\left[l_t(\theta)-E\left(l_t(\theta)\right)\right]\right| \ge \varepsilon\right) = 0 \tag{32}$$

Where, $l_t(\theta)\ \&\ E\left(l_t(\theta)\right)$ are likelihoods and expectations of likelihoods respectively.

$$\mathrm{E}\left(L(\Theta)\right) = \mathrm{E}\left[\sum_{t=L+1}^{n}\left[\sum_{k=1}^{K} S_{kt}\log(\omega_k) + \sum_{k=1}^{K} S_{kt}\log\alpha_k + \sum_{k=1}^{K} S_{kt}\alpha_k \log X_t - \right.\right.$$
$$\left.\left.\sum_{k=1}^{K} S_{kt}\log X_t - \sum_{k=1}^{K} S_{kt}\alpha_k\log\beta_k - \sum_{k=1}^{K} S_{kt}\alpha_k X_t + \sum_{k=1}^{K} S_{kt}\alpha_k\beta_k\right]\right] \tag{33}$$

From Jensen's inequality, $\mathrm{E}\left(\ln(X)\right) \le \ln\mathrm{E}(X)$ for every positively integrable X.

$$\mathrm{E}\left(L(\Theta)\right) \le \sum_{t=L+1}^{n}\left[\sum_{k=1}^{K} S_{kt}\log(\omega_k) + \sum_{k=1}^{K} S_{kt}\log\alpha_k + \sum_{k=1}^{K} S_{kt}\alpha_k\log\mathrm{E}(X_t) - \right.$$
$$\left.\sum_{k=1}^{K} S_{kt}\log\mathrm{E}(X_t) - \sum_{k=1}^{K} S_{kt}\alpha_k\log\beta_k - \sum_{k=1}^{K} S_{kt}\alpha_k\mathrm{E}(X_t) + \sum_{k=1}^{K} S_{kt}\alpha_k\beta_k\right] \tag{34}$$

But, $E(X) = \beta\Gamma\left(1+\frac{1}{\alpha}\right)$

$$\mathrm{E}\left(L(\Theta)\right) \approx \sum_{t=L+1}^{n}\left[\sum_{k=1}^{K} S_{kt}\log(\omega_k) + \sum_{k=1}^{K} S_{kt}\log\alpha_k + \sum_{k=1}^{K} S_{kt}\alpha_k\log\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right) - \right.$$
$$\left.\sum_{k=1}^{K} S_{kt}\log\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right) - \sum_{k=1}^{K} S_{kt}\alpha_k\log\beta_k - \sum_{k=1}^{K} S_{kt}\alpha_k\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right) + \sum_{k=1}^{K} S_{kt}\alpha_k\beta_k\right] \tag{35}$$

So,

$$\left[l_t(\theta)-E\left(l_t(\theta)\right)\right] \approx \sum_{t=L+1}^{n} S_{kt}\alpha_k\log X_t - \sum_{t=L+1}^{n} S_{kt}\log X_t - \sum_{t=L+1}^{n} S_{kt}\alpha_k X_t -$$
$$\sum_{t=L+1}^{n} S_{kt}\alpha_k\log\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right) + \sum_{t=L+1}^{n} S_{kt}\log\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right) + \sum_{t=L+1}^{n} S_{kt}\alpha_k\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right) \tag{36}$$

$$\left|\frac{1}{n}\sum_{i=1}^{n}\left[l_t(\theta)-E\left(l_t(\theta)\right)\right]\right| = \approx \sum_{t=L+1}^{n} S_{kt}\log X_t\left(\alpha_k-1\right) - \sum_{t=L+1}^{n} S_{kt}\log\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)\left(\alpha_k-1\right)$$

$$-\sum_{t=L+1}^{n} S_{kt}\alpha_k X_t + \sum_{t=L+1}^{n} S_{kt}\alpha_k\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)$$

$$\approx(\alpha_k-1)\sum_{t=L+1}^{n} S_{kt}\left(\log X_t-\log\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)\right)-\sum_{t=L+1}^{n} S_{kt}\alpha_k\left(X_t-\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)\right) \quad (37)$$

$$\max\left|\frac{1}{n}\sum_{i=1}^{n}\left[l_t(\theta)-E\left(l_t(\theta)\right)\right]\right|\ge\varepsilon$$

$$\approx(\alpha_k-1)\sum_{t=L+1}^{n} S_{kt}\left(\log X_t-\log\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)\right)-\frac{\sum_{i=1}^{n}\sum_{t=L+1}^{n} S_{kt}\alpha_k\left(X_t-\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)\right)}{n}\ge\varepsilon \quad (38)$$

$$\lim_{l\to\infty}\mathrm{P}\left(\max\left|\frac{1}{n}\sum_{i=1}^{n}\left[l_t(\theta)-E\left(l_t(\theta)\right)\right]\right|\ge\varepsilon\right)=0$$

$$\approx\lim_{l\to\infty}P\left((\alpha_k-1)\sum_{t=L+1}^{n} S_{kt}\left(\log X_t-\log\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)\right)-\frac{\sum_{i=1}^{n}\sum_{t=L+1}^{n} S_{kt}\alpha_k\left(X_t-\beta_k\,\Gamma\left(1+\frac{1}{\alpha_k}\right)\right)}{n}\ge\varepsilon\right)<\infty \quad (39)$$

## 4. Conclusions

The proposed and formulated model of $WMAR(k; p_1, p_2, \ldots, p_k)$ being one of the candidates models for extreme valued mixture autoregressive was designed in order to capture, and correct stylized properties; entire outrange of switching-patterns with Weibull marginal distribution(multimodalities), change points like behavior, regime-switching (capable of handling recurring periodical sequence), and time-varying volatilities (conditional means-variances). The model was also formulated to absolve heavy-tailed, long-memory and non-Gaussian MAR model characterized by positive and strictly count (discrete) valued random noises (error terms) series.

## ACKNOWLEDGEMENTS

We wish to thank the reviewers.

## REFERENCES

[1] G.N. Boshankov, "Prediction with Mixture Autoregressive Models" Research Report No.6, 2006, Probability and Statistics Group School of Mathematics. The University of Manchester, 2006.

[2] G. Behrooz, S.M.N. Abu Hassan, and S. Tamat, "Application of the Threshold Model for Modelling and Forecasting of Ex change Rate in Selected ASEAN Countries" Sains Malaysiana, Vol.43(10), pp. 1609-1622, 2014.

[3] M. Meitz, D. Preve, and P. Saikkonen, "A mixture autoregres sive model based on Student's t- distribution", 2018.

[4] M. Maleki, and A. R. Nematollahi, "Autoregressive Models with Mixture of Scale Mixtures of Gaussian Innovations" Iran Journal of Science and Technology: Transactions A Science. Vol. 41(4), pp. 1099–1107, 2017. https://doi.org/10.1007/s40995-017-0237-6.

[5] L. Kalliovirta, M. Meitz, and P. Saikkonen," A Guassian Mix ture Vector Autoregressive". Journal of econometrics, Vol. 192, pp. 485-498, 2016.

[6] L. Kalliovirta, M. Meitz, and P. Saikkonen, "A Guassian Mix ture autoregresive model for univariate series" Journal of Time Series Analysis, Vol. 36, pp. 247-266, 2012.

[7] M. Li and M. Raftery, "Modeling flat stretches, burst and outli ers in time series using Mixture Transition Distribution Mod els" J.Am Statist. Ass., Vol. 91, pp. 1504-1514, 1966.

[8] N. Nguyen, "Hidden Markov Model for Stock Trading" Inter national Journal Financial *Studies*, Vol 6 (36), 2018. Doi:10.3390/ijfs6020036.

[9] C. S. Wong, and W. K. Li, "On a Mixture Autoregressive Mod el" Journal of Royal Statistics Society. B 62, pp. 95–115, 2000.

[10] C. S. Wong, W. S. Chan, and P. L. Kam," A Student t-mixture autoregressive Model with applications to heavy-tailed financial data" Singapore Economic Review Conference, pp. 1–10, 2009.

[11] C. S. Wong, Statistical inference for some nonlinear time series models, Ph.D thesis, University of Hong Kong, Hong Kong, 1998.