Transactions on Machine Learning and Artificial Intelligence - Vol. 10, No. 1
Publication Date: February, 25, 2022
DOI:10.14738/tmlai.101.11544.
Olanrewaju, R. O., Waititu, A. G., & Abiodun, N. L. (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). 09-26.

Kullback-Leibler Divergence of Mixture Autoregressive Random Processes via Extreme-Value-Distributions (EVDs) Noise with Application of the Processes to Climate Change

Rasaki Olawale Olanrewaju
Mathematical Sciences, Pan African
University Institute for Basic Sciences, Technology, and Innovation
Nairobi, Kenya

Anthony Gichuhi Waititu
Department of Mathematical Sciences
Jomo Kenyatta University of Agriculture and Technology
Nairobi, Kenya

Nafiu Lukman Abiodun
Department of Mathematics and Statistics
Kabale University, Kabale, Uganda

ABSTRACT

This paper designs inter-switch autoregressive random processes in a mixture manner with Extreme-Value-Distributions (EVDs) random noises to give EVDs-MAR model. The EVDs-MAR model comprises of Fréchet, Gumbel, and Weibull distributional error terms to form FMA, GMA, and WMA models with their embedded inter-switching transitional weights (ω_k), distributional parameters, and autoregressive coefficients (ϕ_{p_k}). The Kullback-Leibler divergence was used to measure the proximity (D) between finite/ delimited mixture density (g_k) and infinite mixture density (g) of the EVDs-MAR model with Expectation-Maximization (EM) algorithm adopted as the parameter estimation technique for the extreme mixture model. The FMA, GMA, and WMA models were subjected to monthly temperature in Celsius (°C) from 1900 to 2020 and annual rainfall in Millimeter (mm) from 1960 to 2020 datasets in Nigeria context.

Keywords: Extreme-Value-Distributions (EVDs); Kullback-Leibler; Autoregressive; Mixture.

MSC 2010 Classification: 37M10, 60G70, 40A20.

INTRODUCTION

Mixture models are models used when heterogeneous population of fractional parts consists of several subdivision of the population within each of which a relatively simpler time series models or other propounded models were applied. The source of the heterogeneity might be via trend, cyclical, regular or irregular fluctuations. Before the notable modeling of mixture

models; and finite mixture distributions around early 2000s by [1], the history of mixture models was initially originated and traced back to 1880s, when [2] used normal mixture to modeled outliers and employed the method of moments to decomposed normal mixtures. Ref. [2] convoluted two normal univariate data together and estimated the few parameters involved with the use of histogram to ascertain the number of components. Since then, especially starting from 1980s, few publications had been written on the theory as well as application of mixture models. For the past fifteen years, mixture models have started receiving great attention in the field of statistics. Mixture models have been reported to be vital in application to clustered analysis, spatial-temporal studies, changing behavioral studies, biometry, and psychological studies. The greater attention of mixture models was its extension to mixture Structural Equation Model (SEM), Bayesian mixture SEM, psychometrics, and missing data (incomplete data) modeling.

The ideology of mixture was based on the idea of Markov transition of states via unobserved latent random variables upon generalized distributional assumptions, mixing weights embedded with parameters to be estimated in each component. The parameter estimation of mixture models via Maximum Likelihood (ML) by the means of iterative technique received petty heed not until 1960s. Greater number of publications and technical works before and in the 1960s used ML approach. From the predating manuscript of Pearson's effort on mixture of models, Ref. [3] extracted and proposed an iterative procedure of estimation via reweighting algorithm that could be conceptualized as Expectation-Maximization (EM) algorithm as defined by Ref. [4], so as to estimate their shared means, variances, mixing weights, and other parameters embedded in the mixture or distributional modified mixture. On the other hand, it was when Ref. [4] formally conceptualized the iterative-reweighting algorithm into EM procedure that the ML solution converges.

SPECIFICATION MIXTURE AUTOREGRESSIVE MODEL

Ref. [5] splinted out and viewed as a finite (countable) mixture distribution;

$$f(x_t) = w_1 g_1(x_t) + w_2 g_2(x_t) + \cdots + w_k g_k(x_t) = \sum_{i=1}^{k} w_i g_i(x_t) \tag{1}$$

Such that $f(x)$ is the entire mixture regime switching and $g_i(x)(i = 1,\ldots,k)$ are the same probability density function per each regime depending on mixing transitional weights $w_i > 0 \ni \sum_{i=1}^{k} \omega_i \approx 1 \ \forall\ i = 1,\ldots,\mathrm{k}$.

Developing of the Extreme-Value-Distribution Mixture Autoregressive (EVDs-MAR) model

According to Ref. [6], Ref. [7] and Ref. [8] a K-component of Mixture Autoregressive (MAR) model can be defined as $MAR(K : p_1,\ldots,p_k)$ and express mathematically as:

$$\mathrm{X}_t = \sum_{\mathrm{i}=1}^{\mathrm{k}} \omega_k \Pi(\bullet)_k \ \Phi\left(\frac{x_t - \phi_{k,0} - \phi_{k,1} x_{t-1} - \cdots - \phi_{k,p_k} x_{t-p_k}}{\sigma_k}\right) \tag{2}$$

Where, $\mathrm{x}_t = \phi_{k,0} + \phi_{k,1} x_{t-1} + \cdots + \phi_{k,p} x_{t-p_k} + \varepsilon_t^{(k)}$

$$\varepsilon_t^{(k)} = x_t - \phi_{k,0} - \phi_{k,1} x_{t-1} - \cdots - \phi_{k,p} x_{t-p_k}$$

Where, $\phi_{pk} \in (0,1)$, for $k = 1,\ldots,K$, $p \geq 1$. For mixing transitional weights $\sum_{i=1}^{k} \omega_i \approx 1$, ə $\omega_i > 0$, for $k = 1,\cdots,K$; where $\Pi(\bullet)_k$ is the embedded distributional parameters of the random noise; $\Phi(.)$ is the Cumulative Distribution Function (CDF) of the standard distribution.
Then,

Extreme-Value-Distribution Mixture Autoregressive (EVDs-MAR) model holds when, $\varepsilon_t^{(k)} \approx$ **EVDs (Fréchet, Gumbel, and Weibull).**

Fréchet (EVD Type I) Distribution

$$\varepsilon_t^{(k)} \approx g\left(x_t;\alpha_k,\beta_k\right) = \frac{\alpha_k}{\beta_k}\left(\frac{\beta_k}{x_t}\right)^{\alpha_k+1} \exp\left(-\left(\frac{\beta_k}{x_t}\right)^{\alpha_k}\right) \tag{3}$$

Then, Fréchet Mixture Autoregressive (FMA) model holds since x_t of equation (2) has been embedded. Where α_k is the shape parameter, such that, $\alpha_k > 0$, and β_k is the scale parameter per each regime. It is to be noted that $(\alpha_k,\beta_k) = \Pi(\bullet)_k$ in FMA model.

Gumbel (EVD Type II) Distribution

$$\varepsilon_t^{(k)} \approx g\left(x_t;\beta_k,\mu_k\right) = \frac{1}{\beta_k}\exp\left(-z-\exp\left(-z\right)\right) \tag{4}$$

Then, Gumbel Mixture Autoregressive (GMA) model holds since x_t of equation (2) has been embedded. Where, $z = \frac{(x_t-\mu_k)}{\beta_k}$, μ_k is the location parameter, whereas β_k is the distribution scale, such that, $\beta_k > 0$ per each regime. It is to be noted that $(\mu_k,\beta_k) = \Pi(\bullet)_k$ in GMA model.

Weibull (EVD Type III) Distribution

$$\varepsilon_t^{(k)} \approx g\left(x_t;\alpha_k,\beta_k\right) = \frac{\alpha_k}{\beta_k}\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1} e^{-\left(\frac{x_t}{\beta_k}\right)^{\alpha_k}} \qquad x_t \geq 0 \tag{5}$$

Then, Weibull Mixture Autoregressive (WMA) model holds in this context. Where $\alpha_k > 0$ is called the shape parameter; $\beta_k > 0$ is the scale parameter of the distribution per each regime. It is to be noted that $(\mu_k,\beta_k) = \Pi(\bullet)_k$ in WMA model.

Specification Mixture Autoregressive Model

Suppose "Z" is a latent (unobserved) variable, where Z_t is a $k-$dimensional vector such that $Z_t = \{Z_1, Z_2, \cdots, Z_t\}^T$ whose component is

$$Z_{i,t} = \begin{cases} 1 & iff\ X_t\ is\ a\ member\ of\ the\ j^{th}\ regime \\ 0, & otherwise \end{cases} \quad \text{э } 1 \le j \le K$$

$$P\left[Z_t = (1,0,\cdots,0)^T\right] = \omega_1,\ P\left[Z_t = (0,1,\cdots,0)^T\right] = \omega_2, \cdots, P\left[Z_t = (0,0,\cdots,1)^T\right] = \omega_k$$

Fréchet Mixture Autoregressive (FMA) Model Parameter Estimation

Let $X = \{X_1, X_2, \cdots, X_n\}^T$; $\alpha_k = \{\alpha_0, \alpha_1, \cdots, \alpha_k\}^T$; $\beta_k = \{\beta_0, \beta_1, \cdots, \beta_k\}^T$; $\phi_k = \{\phi_{k0}, \phi_{k1}, \cdots, \phi_{k\,pk}\}^T$ $\omega_k = \{\omega_0, \omega_1, \cdots, \omega_k\}^T$ for $k = 1, \cdots, K$

Assuming Z_t is known, the Fréchet distribution of the complete data is given by (X_t, Z_t). The maximizing conditional log-likelihood function for FMA model gives

$$L_t(\Psi) = \log \prod_{k=1}^{K} \left[\omega_k \frac{\alpha_k}{\beta_k} \left(\frac{\beta_k}{X_t} \right)^{\alpha_k + 1} \exp\left(-\left(\frac{\beta_k}{X_t} \right)^{\alpha_k} \right) \right]^{Z_{kt}} \tag{6}$$

Where $\Psi = \{\phi_k, \omega_k, \alpha_k, \beta_k\}^T = \{\phi_k, \omega_k, \Pi(\bullet)\}^T$ is the universal parameter space.

Gumbel Mixture Autoregressive (GMA) Model Parameter Estimation

Let $X_t = \{X_1, X_2, \cdots, X_n\}^T$; $\mu_k = \{\mu_0, \mu_1, \cdots, \mu_k\}^T$; $\beta_k = \{\beta_0, \beta_1, \cdots, \beta_k\}^T$; $\phi_k = \{\phi_{k0}, \phi_{k1}, \cdots, \phi_{k\,pk}\}^T$ $\omega_k = \{\omega_0, \omega_1, \cdots, \omega_k\}^T$ for $k = 1, \cdots, K$

Assuming Z_t is known, the Gumbel distribution of the complete data is given by (X_t, Z_t). The maximizing conditional log-likelihood function for GMA model gives

$$L_t(\Xi) = \log \prod_{k=1}^{K} \left[\omega_k \frac{1}{\beta_k} \exp\left(-\frac{(X_t - \mu_k)}{\beta_k} - \exp\left(-\frac{(X_t - \mu_k)}{\beta_k} \right) \right) \right]^{Z_{kt}} \tag{7}$$

Where $\Xi = \{\phi_k, \omega_k, \mu_k, \beta_k\}^T = \{\phi_k, \omega_k, \Pi(\bullet)\}^T$ is the universal parameter space.

Weibull Mixture Autoregressive (WMA) Model Parameter Estimation

Let $X = \{X_1, X_2, \cdots, X_n\}^T$; $\alpha_k = \{\alpha_0, \alpha_1, \cdots, \alpha_k\}^T$; $\beta_k = \{\beta_0, \beta_1, \cdots, \beta_k\}^T$; $\phi_k = \{\phi_{k0}, \phi_{k1}, \cdots, \phi_{k\,pk}\}^T$; $\omega_k = \{\omega_0, \omega_1, \cdots, \omega_k\}^T$ for $k = 1, \cdots, K$

Assuming Z_t is known, the Weibull distribution of the complete data is given by (X_t, Z_t). The maximizing conditional log-likelihood function for the WMA model gives

$$L_t(\Omega) = \log \prod_{k=1}^{K} \left[\omega_k \frac{\alpha_k}{\beta_k} \left(\frac{X_t}{\beta_k} \right)^{\alpha_k - 1} \exp\left(-\frac{X_t}{\beta_k} \right)^{\alpha_k} \right]^{Z_{kt}} \tag{8}$$

Where $\Omega = \{\phi_k, \omega_k, \alpha_k, \beta_k\}^T = \{\phi_k, \omega_k, \Pi(\bullet)\}^T$ is the universal parameter space.

The embedded coefficients to be estimated for FMA, GMA and WMA models in equation (6), (7) and (8) respectively can be easily done using the Expectation-Maximization (EM) algorithm (via E-step and M-step). See Ref. [9]; Ref. [10] and Ref. [11].

KULLBACK-LEIBLER DIVERGENCE OF THE EVDS-MAR MODEL

We shall be using the Kullback-Leibler divergence to measure the proximity / distance (D) between finite/ delimited mixture density (g_k) and infinite mixture density (g) (see Ref. [12]). The Kullback-Leibler divergence is defined as follow:

$$D\left(g\|g_k\right)=\int g(x)\log\frac{g(x)}{g_k(x)}\omega_k\left(\partial x\right)$$

Where "D" is the distance; and the log-likelihood $\frac{1}{n}\sum_{i=1}^{n}g_k(x)$ is the independent sample of $\int g\log g_k$.

Kullback-Leibler Divergence as defined for continous density

$$\begin{aligned}D\left(g\,\|\,g_k\right)&=\int_{x_t\in X}g(x_t)\log\frac{g(x_t)}{g_k(x_t)}\omega_k\partial x_t\\&=\int_{x_t\in X}-\log\frac{g_k(x_t)}{g(x_t)}g(x_t)\omega_k\partial x_t\\&\geq-\log\left[\int_{x_t\in X}\log\frac{g_k(x_t)}{g(x_t)}g(x_t)\omega_k\partial x_t\right]\\&=-\log\left(\int_{x_t\in X}g_k(x_t)\omega_k\partial x_t\right)\\&=-\log(1)=0\end{aligned}\qquad(9)$$

This implies that the Kullback-Leibler divergence is a non-negative for the EVDs-MAR as well as the asymmetry property that the KL divergence is a asymmetrical measure of the two probability distributions of

$$D\left(g\,\|\,g_k\right)\neq D\left(g_k\,\|\,g\right)\qquad(10)$$

KL Divergence for Weibull Mixture Autoregressive (WMA) Model

$$g(x_t;\alpha,\beta)=\frac{\alpha}{\beta}\left(\frac{x_t}{\beta}\right)^{\alpha-1}e-\left(\frac{x_t}{\beta}\right)\quad x_t\in[0,\infty]$$

Where $g(x_t)$ & $g_k(x_t)$ denote the corresponding probability densities of Weibull without and with regime-switching for uniformly time series respectively.

$$D\left(g\,\|\,g_k\right)=\int_0^\infty g(x_t;\alpha,\beta)\log\frac{g(x_t;\alpha,\beta)}{g_k(x_t;\alpha_k,\beta_k)}\omega_k\partial x_t$$

$$= \int_0^{\infty} g(x_t;\alpha,\beta) \log \frac{\frac{\alpha}{\beta}\left(\frac{x_t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x_t}{\beta}\right)^{\alpha}}}{\frac{\alpha_k}{\beta_k}\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1} e^{-\left(\frac{x_t}{\beta_k}\right)^{\alpha_k}}} \omega_k \partial x_t \tag{11}$$

Factoring,

$$\log \frac{\frac{\alpha}{\beta}\left(\frac{x_t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x_t}{\beta}\right)^{\alpha}}}{\frac{\alpha_k}{\beta_k}\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1} e^{-\left(\frac{x_t}{\beta_k}\right)^{\alpha_k}}} = \log \frac{\alpha/\beta}{\alpha_k/\beta_k} + \log \frac{\left(\frac{x_t}{\beta}\right)^{\alpha-1}}{\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1}} + \log \frac{e^{-\left(\frac{x_t}{\beta}\right)^{\alpha-1}}}{e^{-\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1}}}$$

Let $A = \log \frac{\alpha/\beta}{\alpha_k/\beta_k}$; $B = \log \frac{\left(\frac{x_t}{\beta}\right)^{\alpha-1}}{\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1}}$; $C = \log \frac{e^{-\left(\frac{x_t}{\beta}\right)^{\alpha}}}{e^{-\left(\frac{x_t}{\beta_k}\right)^{\alpha_k}}}$

$$B = \log \frac{\left(\frac{x_t}{\beta}\right)^{\alpha-1}}{\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1}} = (\alpha-1)\log\frac{x_t}{\beta} - (\alpha_k-1)\log\frac{x_t}{\beta_k}$$

$$= (\alpha-\alpha_k)\log x_t + (\alpha_k-1)\log\beta_k - (\alpha-1)\log\beta$$

Let $P = (\alpha_k-1)\log\beta_k - (\alpha-1)\log\beta$ where P is a constant independent of x_t

$$C = \left(\frac{x_t}{\beta_k}\right)^{\alpha_k} - \left(\frac{x_t}{\beta}\right)^{\alpha}$$

Inserting A and P constants as well as the other terms into equation (11) gives

$$D(g \,||\, g_k) = \int_0^{\infty} g(x_t;\alpha,\beta)\left[A + (\alpha-\alpha_k)\log x_t + P + \left(\frac{x_t}{\beta_k}\right)^{\alpha_k} - \left(\frac{x_t}{\beta}\right)^{\alpha}\right]\omega_k \partial x_t$$

$$= (A+P)\int_0^{\infty} g(x_t;\alpha,\beta)\partial x_t + \int_0^{\infty} g(x_t;\alpha,\beta)(\alpha-\alpha_k)\omega_k \log x_t \partial x_t + \int_0^{\infty} g(x_t;\alpha,\beta)\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1}\partial x_t - \int_0^{\infty} g(x_t;\alpha,\beta)\left(\frac{x_t}{\beta}\right)^{\alpha-1}\partial x_t$$

$$= (A+P)\int_0^{\infty} g(x_t;\alpha,\beta)\partial x_t \tag{12}$$

$$+ \int_0^{\infty} g(x_t;\alpha,\beta)(\alpha-\alpha_k)\omega_k \log x_t \partial x_t \tag{13}$$

$$+ \int_0^{\infty} g(x_t;\alpha,\beta)\left(\frac{x_t}{\beta_k}\right)^{\alpha_k-1}\partial x_t \tag{14}$$

$$- \int_0^{\infty} g(x_t;\alpha,\beta)\left(\frac{x_t}{\beta}\right)^{\alpha-1}\partial x_t \tag{15}$$

The term $\int_0^\infty g(x_t;\alpha,\beta)\partial x_t = 1$ for a proper PDF of Gumbel distribution

The term $\int_0^\infty g(x_t;\alpha,\beta)\left(\frac{x_t}{\beta}\right)^\alpha \partial x_t$ can be solved by change of variable technique

Let $y = \left(\frac{x_t}{\beta}\right)^\alpha \Rightarrow \partial y = \frac{\alpha}{\beta}\left(\frac{x_t}{\beta}\right)^{\alpha-1}\partial x_t$

Also, $x_t = y^{1/\alpha\beta} \Rightarrow \log x_t = 1/\alpha \log y + \log\beta$ (16)

The term in equation (13) becomes

$$= (\alpha - \alpha_k)\int_0^\infty \frac{\alpha}{\beta}\left(\frac{x_t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x_t}{\beta}\right)^\alpha} \omega_k \log x_t \partial x_t$$

$$= (\alpha - \alpha_k)\int_0^\infty e^{-y}\left[\frac{1}{\alpha}\log y + \log\beta\right]\omega_k \partial y$$

$$= (\alpha - \alpha_k)\left[\frac{1}{\alpha}\int_0^\infty e^{-y}\log y dy \omega_k + \omega_k \log\beta \int_0^\infty e^{-y}dy\right]$$

$$= (\alpha - \alpha_k)\left[-\frac{\gamma\omega_k}{\alpha} + \omega_k \log\beta\right] \quad (17)$$

Where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant

Intergral of (14) can be considered as substitute in equation (16) and note that

$$x_t^{\alpha_k} = y^{\alpha_k/\alpha}\beta^{\alpha_k} \quad (18)$$

Then equation (16) becomes

$$= \frac{1}{\beta_k^{\alpha_k}}\int_0^\infty \frac{\alpha}{\beta}\left(\frac{x_t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x_t}{\beta}\right)^\alpha} x_t^{\alpha_k}\partial x_t$$

$$= \left(\frac{\beta}{\beta_k}\right)^{\alpha_k}\int_0^\infty e^{-y} y^{\alpha_k/\alpha}\partial y$$

$$= \left(\frac{\beta}{\beta_k}\right)^{\alpha_k}\Gamma\left(\frac{\alpha_k}{\alpha}+1\right)$$

Since $x_t^{\alpha_k} = y\beta^\alpha$ & solving the term in (18) gives

$$= -\frac{1}{\beta^\alpha}\int_0^\infty \frac{\alpha}{\beta}\left(\frac{x_t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x_t}{\beta}\right)^\alpha} x_t^{\alpha_k}\partial x_t$$

$$= -\left(\frac{\beta}{\beta}\right)^{\alpha} \int_0^{\infty} e^{-y} y \partial y = -1$$

Putting everything back in equation (11) gives

$$D(g \,||\, g_k) = \log \frac{\alpha/\beta}{\alpha_k/\beta_k} + \log \frac{\beta_k^{\alpha_k - 1}}{\beta^{\alpha - 1}} + (\alpha - \alpha_k)\left[\log \beta - \frac{\gamma}{\alpha}\right] + \left(\frac{\beta}{\beta_k}\right)^{\alpha_k} \Gamma\left(\frac{\alpha_k}{\alpha} + 1\right) - 1$$

$$= \log \frac{\alpha}{\beta^{\alpha}} - \log \frac{\alpha_k}{\beta_k^{\alpha_k}} + (\alpha - \alpha_k)\left[\omega_k \log \beta - \frac{\gamma \omega_k}{\alpha}\right] + \left(\frac{\beta}{\beta_k}\right)^{\alpha_k} \Gamma\left(\frac{\alpha_k}{\alpha} + 1\right) - 1 \tag{19}$$

Setting the shape parameters α and α_k to 1 gives two exponential distributions with inverse rate parameters $1/\beta$ & $1/\beta_k$. So, the close form equation for $D(g \,||\, g_k)$ gives

$$\omega_k \log(\beta_k) - \log(\beta) + \left(\frac{\beta}{\beta_k}\right) - 1 \tag{20}$$

KL Divergence for Fréchet Mixture Autoregressive (FMA) Model

$$\left(\omega_k \log(\beta_k) - \log(\beta) + \left(\frac{\beta}{\beta_k}\right) - 1\right)\left(\omega_k \log(\beta_k) - \log(\beta) + \left(\frac{\beta}{\beta_k}\right) - 1\right)^2 = \left(\omega_k \log(\beta_k) - \log(\beta) + \left(\frac{\beta}{\beta_k}\right) - 1\right)^3 \tag{21}$$

KL Divergence for Gumbel Mixture Autoregressive (GMA) Model

The Gumbel distribution,

$$g(x_t; \beta, \mu) = \frac{1}{\beta} \exp\left[-\frac{(x_t - \mu)}{\beta} - \exp\left(-\frac{x_t - \mu}{\beta}\right)\right]$$

The log-likelihood function

$$L(\mu, \beta) = -n \ln \beta - \sum_{i=1}^{n}\left[\frac{x_t - \mu}{\beta} + \exp\left(-\frac{x_t - \mu}{\beta}\right)\right]$$

Differentiating w.r.t to μ & β and equating to zero

$$\frac{\partial L(\mu, \beta)}{\partial \mu} = -\frac{n}{\beta} + \frac{1}{\beta}\sum_{i=1}^{n} \exp\left(-\frac{x_t - \mu}{\beta}\right) = 0 \tag{22}$$

$$\frac{\partial L(\mu, \beta)}{\partial \beta} = -\frac{n}{\beta} + \sum_{i=1}^{n}\left(-\frac{x_t - \mu}{\beta}\right) - \sum_{i=1}^{n}\left(-\frac{x_t - \mu}{\beta}\right)\exp\left(-\frac{x_t - \mu}{\beta}\right) = 0 \tag{23}$$

$$\Rightarrow \hat{\mu} = -\beta \ln\left[\frac{1}{n}\sum_{i=1}^{n} \exp\left(\frac{x_t}{\beta}\right)\right]; \quad \Rightarrow \hat{\beta} = \frac{1}{n}\sum_{i=1}^{n} x_t - \frac{\sum_{i=1}^{n} x_t \exp\left(\frac{x_t}{\beta}\right)}{\sum_{i=1}^{n} \exp\left(\frac{x_t}{\beta}\right)}$$

But, $D(g \,||\, g_k) = \int_{-\infty}^{\infty} g(x_t; \mu, \beta) \log \frac{g(x_t; \mu, \beta)}{g_k(x_t; \mu_k, \beta_k)} \omega_k \partial x_t$

$$= \int_{-\infty}^{\infty} g(x_t;\mu,\beta)\log g(x_t;\mu,\beta)\omega_k \partial x_t - \int_{-\infty}^{\infty} g(x_t;\mu,\beta)\log g(x_t;\mu_k,\beta_k)\omega_k \partial x_t \tag{24}$$

But,

$$\int_{-\infty}^{\infty} g(x_t;\mu,\beta)\log g(x_t;\mu_k,\beta_k)\omega_k \partial x_t$$

$$= -\omega_k \ln \beta_k + \frac{\mu_k - \mu}{\beta_k}\omega_k + \frac{\omega_k \beta}{\beta_k}\psi(1) - e^{\left(\frac{\mu_k-\mu}{\beta_k}\right)}\Gamma\left(\frac{\beta}{\beta_k}+1\right)\omega_k \tag{25}$$

$$\psi(z) = \frac{\partial \Gamma(z)}{\Gamma(z)};\quad \Gamma(z) = \int_{-\infty}^{\infty} e^{-y}y^z \partial y;\quad \partial\Gamma(z) = \frac{\partial}{\partial z}\int_{-\infty}^{\infty} e^{-y}y^z \partial y;\ \psi(z) = \frac{\int_{-\infty}^{\infty} e^{-y}y^z \partial y}{\frac{\partial}{\partial z}\int_{-\infty}^{\infty} e^{-y}y^z \partial y}$$

Where,

$$\int_{-\infty}^{\infty} g(x_t;\mu,\beta)\log g(x_t;\mu,\beta)\omega_k \partial x_t$$

In a similar vein,

$$= -\omega_k \ln \beta + \omega_k \psi(1) - \omega_k \tag{26}$$

$$D(g \,||\, g_k) = \omega_k \ln\frac{\beta}{\beta_k} + \omega_k \frac{\mu - \mu_k}{\beta_k} + \omega_k\left(1 - \frac{\beta}{\beta_k}\right)\psi(1) + \omega_k \frac{e^{(\mu_k-\mu)}}{\beta_k}\Gamma\left(\frac{\beta}{\beta_k}+1\right) - \omega_k \tag{27}$$

So,

Selection Criteria for the EVDs-MAR Model

According to Ref. [13], the selection criteria/model performance criteria for mixture autoregressive with any marginal distribution are:

$$BIC = -2L(\Theta) + \log(n - p_{\max})\left(3K - 1 + \sum_{k=1}^{K} p_k\right) \tag{28}$$

$$AIC = -2L(\Theta) + 2\left(3K - 1 + \sum_{k=1}^{K} p_k\right) \tag{29}$$

Where $L(\Theta)$ is the completed log-likelihood of each model; $p_{\max}$ is the maximum lag order of the entire regime; K is the number of regime.

NUMERICAL RESULTS OF APPLICATION

Data Set

In this section, equation (6), (7) and (8) for the EVDs-MAR (FMA, GMA and WMA) respectively will be applied to temperature and rainfall datasets. The monthly temperature in Celsius (^oC) from 1900 to 2020 was recorded and collected from the centralized Nigerian Meteorological Agency (NIMET), Abuja, Nigeria. In a similar version, Lagos annual level of rainfall gauged and recorded in Millimeter (mm) from 1960 to 2020 was collected from Ministry of Environment and Water Resources. The continuous-typed datasets for temperature and rainfall will be applied to the optimal and iterative solutions of the MAR model-types in equation (6), (7), and (8).

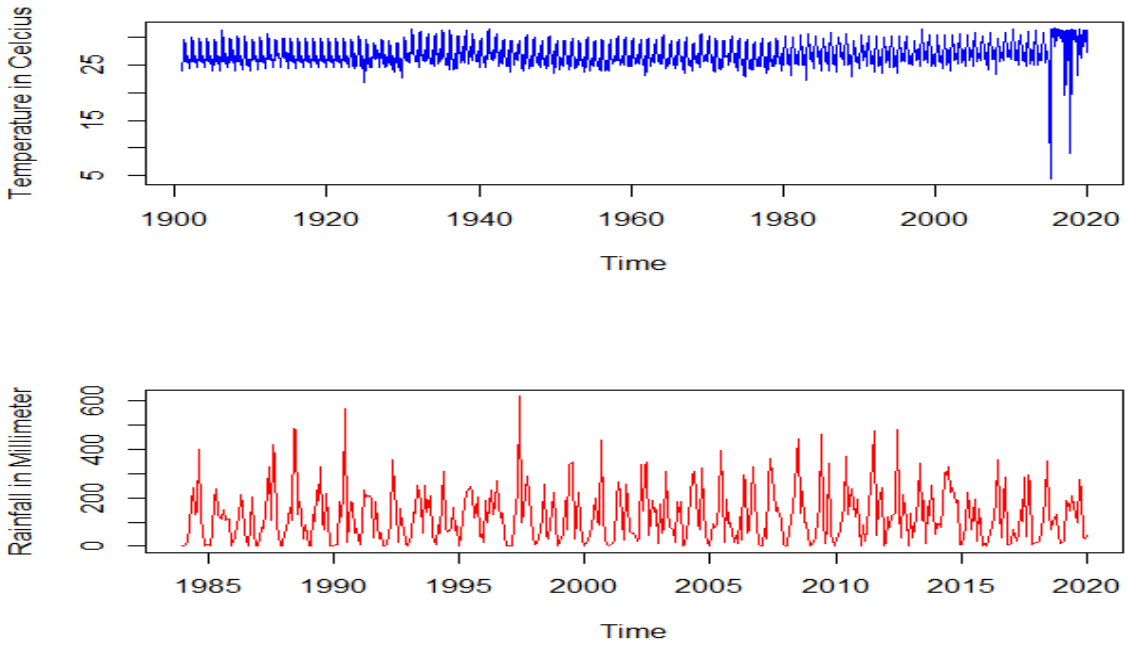

Figure 1. Time Plot of the Temperature and Rainfall Datasets in Celsius and Millimeter

Authors' Computation (2022)

From figure 1, it is glaring that the monthly temperature records in Celsius (^{o}C) from 1900 to1920 maintained a constant continual swinging temperature indexes between 25 ^{o}C and 28 ^{o}C. A miniature fall in the temperature indexes between 23 ^{o}C and 25 ^{o}C was recorded from 1922 to 1930. In the same manner, temperature records from 1935 till around 2016 experienced a drastic fall to 5 ^{o}C from 2017 to early 2019. The climate temperature change sporadically swings up to 20 ^{o}C in 2020. The level of rainfall gauge recorded in Millimeter (mm) ranges from 10(mm) and 400(mm) between 1961 and 1987. It skyrocketed to around 500(mm) towards the tail end of 1990, it maintained the normal statoscope that ranges from 10(mm), and 400(mm) from 1996 till 1997 before the highest downpour close to 600(mm) was recorded in the year 1988. From 1989 till 2020, the measured downpours swing between 10(mm) and 400(mm).

Table 1. Estimate of the Distributional Parameters and Goodness-of-Fit Test for the Fitted Climate Change on EVDs

Climate Change	Estimates	α	β	cor (α,β)	Loglik.	AIC	BIC	K-S Test Statistic
Temperature	Gumbel	25.8066 (0.1171)	4.2036 (0.0527)	0.3253	-3907.243	7818.486	7829.03	0.0193
	Fréchet	3.1214 (0.0357)	25.0793 (0.2237)	0.3225	-4850.682	9705.363	9715.908	0.0351
	Weibull	14.1645 (0.2779)	27.9061 (0.0549)	0.3279	-3116.16	6236.319	6246.864	0.004
	Normal	μ	σ	0	-3130.504	6265.007	6275.552	0.1366
		26.9736	2.1066					

		(0.0555)	(0.0393)					
Rainfall	Gumbel	77.3192 (3.9429)	79.1793 (3.0923)	0.3032	-2657.462	5318.925	5327.116	0.041
	Fréchet	0.3910 (0.0120)	19.4115 (2.5154)	-0.3504	-2846.929	5697.857	5706.049	0.001
	Weibull	0.8944 (0.0359)	120.6856 (6.6548)	0.2723	-2586.851	5177.703	5185.894	0.026
	Normal	μ	σ	0.0027	-2716.058	5436.116	5444.308	0.033
		125.8837 (5.2091)	109.7614 (3.6834)					

Keynotes: K-S= Kolmogorov Smirnov. In bracket are standard errors of estimates.
Authors' Computation (2022)

From table 1, literally fitting the temperature dataset to Gumbel, Fréchet, Weibull Normal distributions each, it is glaring that Weibull distribution is the true and best fit (ideal representation) with the most beneficial model performance of AIC=6236.319 and BIC=6246.864 compared to AIC=7818.486 and BIC=7829.03; AIC=9705.363 and BIC=9715.908; AIC=6265.007 and BIC=6275.552 for Gumbel, Fréchet, and Normal distributions respectively. There seems to be a weak positive correlation between the two parameters (scale and shape) in each of the distribution, except for normal distribution with no correlation between the shape and location parameter. The Kolmogorov Smirnov goodness-of-fit test for continuous distributions was adapted to test whether truly the climate change agents' datasets came from the EVDs or came from any of the EVDs. Since the K-S test statistic D= 0.00193, D=0.0351, D=0.004>0.04301 (critical value at α=5%) for Gumbel, Fréchet, and Weibull (EVDs) distributions, it implies the temperature sample came from EVDs while there is no sufficient evidence to attach sample to a Normal distribution. In a similar manner and like that of the temperature data, Weibull is the best fit for the rainfall gauges with ideal model performance AIC=5177.703 and BIC=5185.894. The K-S test statistic D= 0.041, D=0.001, D=0.026, D=0.033 >0.04301 (critical value at α=5%) for Gumbel, Fréchet, and Weibull (EVDs) and Normal distributions indicated that the rainfall sample readings can come from the four distributions.

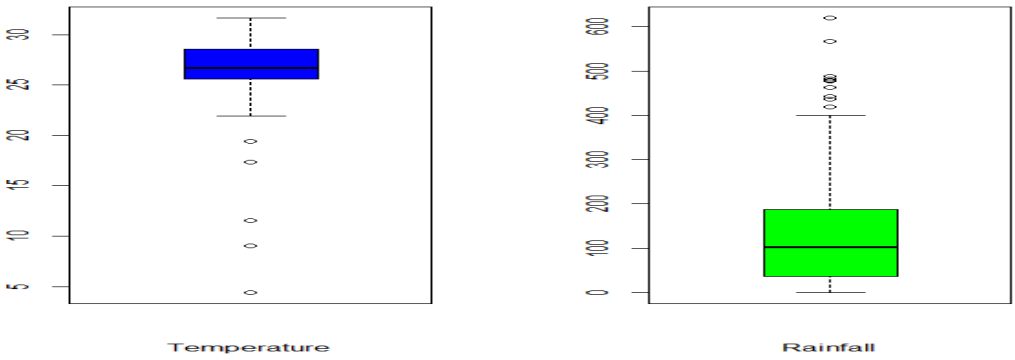

Figure 2. Boxplot of the Climate Change (Temperature and Rainfall).
Authors' Computation (2022)

From the boxplots (Box whisker plots) in figure 2 above that visually show the distribution of quantitative datasets and how skewed the datasets are through its quartiles and averages. Looking closely at the boxplot for temperature dataset, it is obvious that the median (the black line) is closer to the bottom box with the whisker about the same both sides of the box; it suggested possibly a slight left-skewed (negatively skewed) distribution for the temperature dataset. Moreover, since there are observations located outside the whisker of the boxplot, it means there are some numerical values that are far apart distanced from others. In other words, there are five different points (referred to as outliers or anomalies) to the outside of the whisker of the boxplot.

In a similar vein, the median (the black line) for the rainfall dataset is not at the middle of the box, closer to the bottom of the box and the whisker is shorter at the lower end of the box, this connote that the distribution of the dataset is obviously skewed to the right (positively skewed). There are 8 or 9 points of values to the top of the box, meaning obviously the rainfall dataset is contaminated with outliers. This indicated that the symmetrical distribution (Gaussian distribution) was not achieved, and would not be ideal and realistic for the random noise of the two datasets.

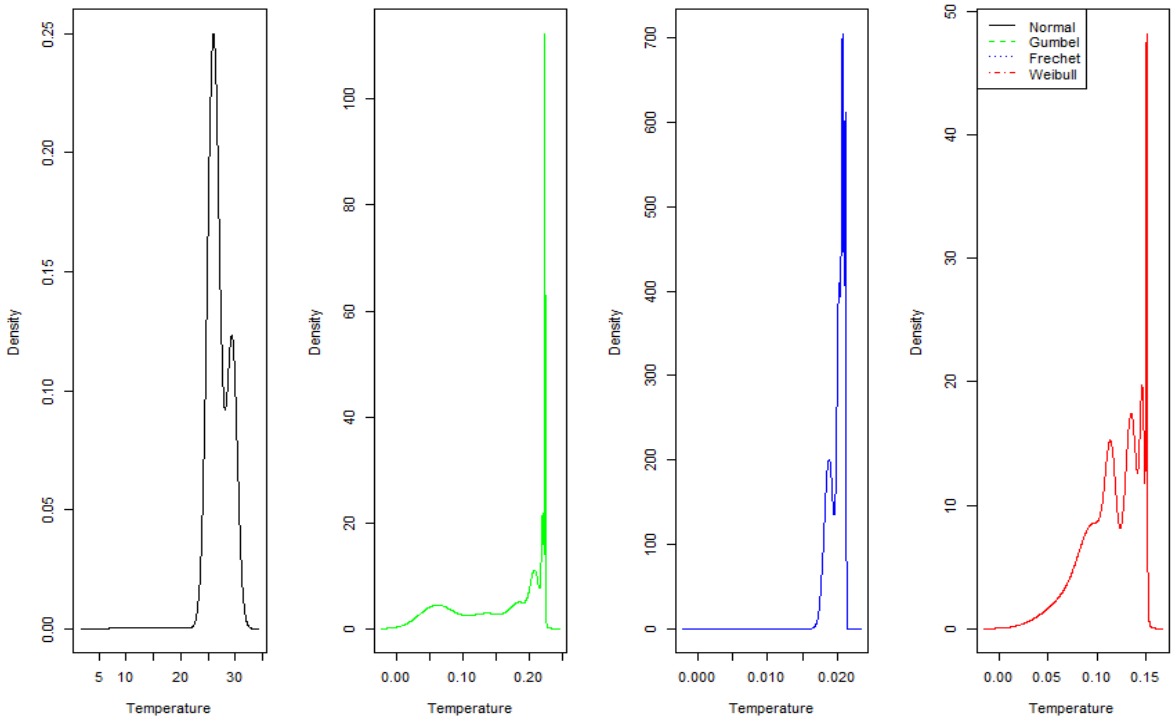

Figure 3. Distributional Density Plots of Normal and EVDs for Temperature. Authors' Computation (2022)

Figure 3 shows higher resolution details for the modes associated to temperature dataset as well as the location of outliers in detecting number of modes via density of distributions. By visualizing the distribution, it can be identified that bimodal was associated to Normal distribution and four (4)-modals was each associated to the EVDs (Gumbel, Fréchet, and Weibull) distributions.

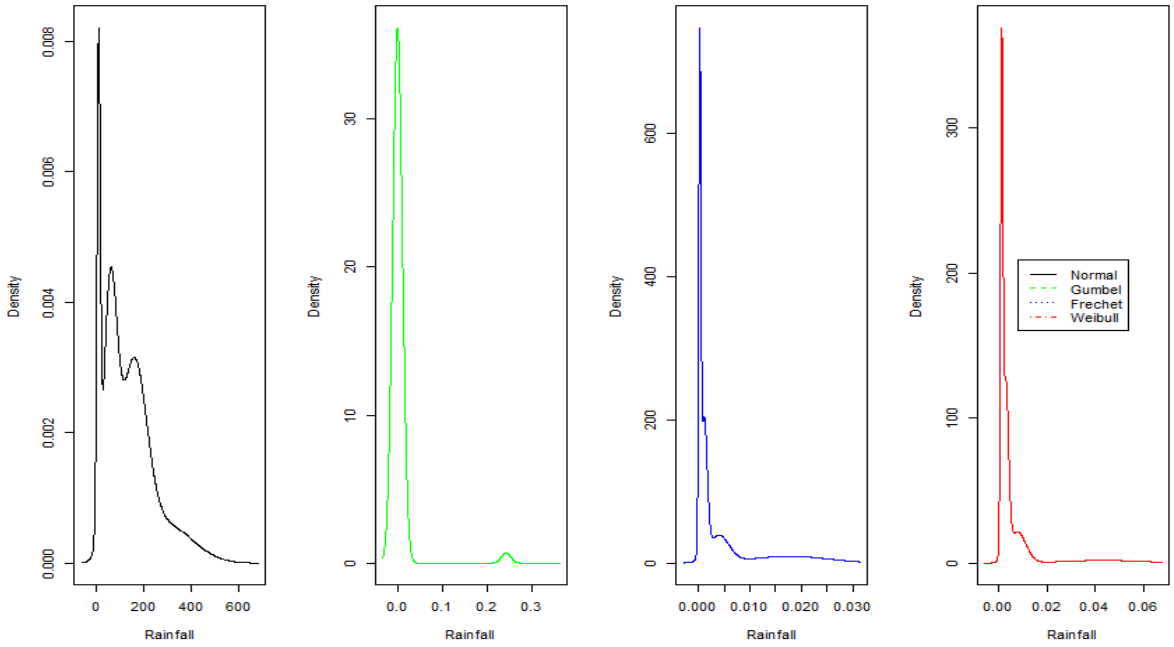

Figure 4. Distributional Density Plots of Normal and EVDs for Rainfall. Authors' Computation (2022)

Similarly, to interpretation of figure 3, figure 4 shows a tri-modal was associated to Normal distribution curve for rainfall dataset; bimodal was associated to Gumbel distribution curve for rainfall dataset; while tri-modal was associated to Fréchet, and Weibull distributions.

Table 2. Fitted of EVDs-MAR and SETAR Models to the Climate Change on EVDs.

Temperature	**% in low regime**	**% in high regime**	**AIC**	**BIC**	**Residuals variance**	**MAPE**	**Threshold Value**	
Norm-SETAR(2:2, 2)	84.98%	15.02%	1270	1279.3	2.393	4.112%	29.34	
Gumbel-SETAR(2:2, 1)	27.33%	72.67%	-8247	-8229	0.0032	31.408%	0.0937	
Fréchet-SETAR(2:2, 2)	15.09%	84.91%	-10577	-10564	0.0007	54.252%	0.0638	
Weibull-SETAR(2:2, 2)	15.09%	84.91%	-10381	-10362	0.0006	39.54%	0.0836	
MAR Models	Mixing Weight (ω_k) per regime	AIC	BIC	μ_k / α_k	β_k	σ_k	Log.Lik	EM Test for No. regime

Norm-MAR(2:2,1)	(0.911, 0.089)	4686.23	4694.418	(13.863, 18.145)	----	(0.953, 3.811)	-2318.125	1.9 ≈ 2 (0.0001)
Gumbel-MAR(4:1,4,1,1)	(0.1196,0.3533, 0.2632, 0.2637)	-20075.236	-20041.803	(0.2222, 0.0839, 0.0430, 0.1931)	(0.0007, 0.0255, 0.0347, 0.0111)	(0.056, 0.035, 0.026, 0.545)	2737.828	3.8 ≈ 4 (0.0003)
Fréchet-MAR(4:2,4,1,1)	(0.3472, 0.2606, 0.1715, 0.2207)	-21358.005	-21337.956	(0.0625, 0.0787, 0.0285, 0.1489)	(0.0536, 0.0456, 0.0628, 0.5638)	(0.010, 0.012, 0.018, 0.002)	3676.039	3.9 ≈ 4 (0.0017)
Weibull-MAR(4:2,4,2,2)	(0.3393, 0.3436, 0.0965, 0.2206)	-20114.272	-20097.65	(0.1436, 0.0199, 0.0010, 0.1448)	(0.4668, 0.8594, 0.5762, 0.4562)	(0.012, 0.015, 0.024, 0.005)	3504.656	3.99 ≈ 4 (0.0001)
Rainfall	**% in low regime**	**% in high regime**	**AIC**	**BIC**	**Residuals Variance**	**MAPE**	**Threshold Value**	
Norm-SETAR(2:1,2)	59.5%	40.5%	4100	4112	9959	3823%	133.6	
Gumbel-SETAR(2:1,1)	55.43%	44.57%	-6023	-6009	0.0031	142.9%	0.0033	
Fréchet-SETAR(2:1,1)	33.03%	66.97%	-4444	-4428	0.0296	789.4%	0.0007	
Weibull-SETAR(2:2,1)	77.38%	22.62%	-4010	-3997	0.0001	232%	0.0053	
MAR Models	**Mixing Weight (ω_k) per regime**	**AIC**	**BIC**	μ_k / α_k	β_k	σ_k	**Log.Lik**	**EM Test for No. regime**
Norm-MAR(3:2,1,1)	(0.2505,0.2453, 0.5042)	5241.93	5258.027	(154.616, 11.9954, 92.4006)	------	(105.2758, 16.3252, 57.2875)	-2592.466	2.7 ≈ 3 (0.0000)

Gumbel-MAR(2:1,1)	(0.6928,0.3072)	-11156.110	-11134.643	(0.0020, 0.0039)	(0.0010, 0.0002)	(0.009, 0.007)	2449.151	2.2 ≈ 2 (0.002)
Fréchet-MAR(3:1,1,1)	(0.1542,0.6110, 0.2348)	-11281.099	-11255.648	(0.0162, 0.0007, 0.0025)	(0.0095, 0.0005, 0.0020)	(0.001, 0.022, 0.040)	2102.836	3.1 ≈ 3 (0.0005)
Weibull-MAR(3:1,2,2)	(0.0929,0.5999, 0.3072)	-11171.006	-11154.738	(0.0321, 0.0014, 0.0048)	(0.0181, 0.0008, 0.0028)	(0.003, 0.004,0.502)	1966.962	2.9 ≈ 3 (0.0001)

Authors' Computation (2022)

It is stubbed and obvious from table 2 that the two regime SETAR model was at its best for the temperature dataset when Fréchet was the random noise with AIC=-10577; BIC=-10564, MAPE= 54.252% such that the series was partitioned into 15.09% for the low regime and 84.91% for the high regime at Fréchet-SETAR (2:2,2). Meaning, the optimal autoregressive order for low regime and high regime was at order 2 for each.

Narrowing down to MAR model that accommodates any number of regime associated to the series. It was captured that the MAR with Gaussian, that is, Norm-MAR(2:2,1) accrued a very high AIC and BIC of 4686.23 and 4694.418 respectively compared to negative AICs and BICs of -20075-236 and -20041.803; -21358.005 and -21337.956; -20114.272 and -20097.65 for Gumbel-MAR(4:1,4,1,1), Fréchet-MAR(4:2,4,1,1) and Weibull-MAR(4:2,4,2,2) respectively. All in all, the EVDs-MAR possessed smaller AIC and BIC compared to the linear models and other regime-switching non-linear models. It was also noted that Fréchet-MAR(4:2,4,1,1) outshined all the other models, as well as surmounted when used as random noise at each level of AR, ARMA, GARCH , SETAR and MAR model. This possibly indicated Fréchet as an ideal conditional error term for monthly time series when modelling contaminated series and mixture processes. Below is the coefficient table for the AR regime-switching processes as well as the fitted model for the ideal Fréchet-MAR(4:2,4,1,1).

Table 3. Components of the AR Processes for the Fréchet-MAR (4: 2, 4, 1, 1) on the Temperature Dataset

Coefficients	ω_k	ϕ_{11}	ϕ_{12}	ϕ_{21}	ϕ_{22}	ϕ_{23}	ϕ_{24}	ϕ_{31}	ϕ_{41}
Component 1	0.3472 (0.0269)	0.1021 (0.0214)	-0.6444 (0.0215)	----	----	----	----	----	----
Component 2	0.2606 (0.0249)	----	----	0.4112 (0.0354)	0.3602 (0.0529)	-0.5543 (0.0487)	0.1971 (0.0285)	----	----
Component 3	0.1715 (0.0129)	----	----	----	----	----	----	0.3285 (0.0135)	----
Component 4	0.2207 (0.0146)	----	----	----	----	----	----	----	0.0274 (0.0018)

In brackets are the standard errors of estimates
Authors' Computation (2022)

$$Fréchet-MAR\ (4:2,4,1,1) \approx (0.3472,0.0625,0.0536)\Phi\left(\frac{x_t-0.1021x_{t-1}+0.6444x_{t-2}}{0.010}\right)+$$
$$(0.2606,0.0787,0.0456)\Phi\left(\frac{x_t-0.4112x_{t-1}-0.3602x_{t-2}+0.5543x_{t-3}-0.1971x_{t-4}}{0.012}\right)+$$
$$(0.1715,0.0285,0.0628)\Phi\left(\frac{x_t-0.3285x_{t-1}}{0.018}\right)+(0.2207,0.1489,0.5638)\Phi\left(\frac{x_t-0.0273x_{t-1}}{0.002}\right) \quad (29)$$

Alternatively,

$$\widehat{X}_t=\begin{cases} x_t+0.1021x_{t-1}-0.6444x_{t-2} & \ni\ \omega_1=0.3472;\alpha_1=0.0625;\beta_1=0.0536 \\ x_t+0.4112x_{t-1}+0.3602x_{t-2}-0.5543x_{t-3}+0.1971x_{t-4} & \ni\ \omega_2=0.2606;\alpha_2=0.0787;\beta_2=0.0456 \\ x_t+0.3285x_{t-1} & \ni\ \omega_3=0.1715;\alpha_3=0.0285;\beta_3=0.0628 \\ x_t+0.0273x_{t-1} & \ni\ \omega_4=0.2207;\alpha_4=0.1489;\beta_4=0.5638 \end{cases}$$

$$\sum_{k=1}^{4}\omega_k=0.3472+0.2606+0.1715+0.2207\approx 1$$

$$AIC=-21358.005;BIC=-21337.956;Log.Lik=3676.039;RMSE=0.023;MAPE=0.3150$$

It is to be noted that all the AR coefficients in each regime are stationary, that is, $\phi_{k,i}\in(0,1)$. Similarly, from table 2, the two regime SETAR model was at its best for the annual rainfall dataset when Gumbel was the random noise with AIC=-6023; BIC=-6009; MAPE=142.9% such that the annual series partition was into 55.43% for the low regime and 44.57% for the high regime at Gumbel-SETAR (2:1,1). Meaning, optimal autoregressive order for low regime and high regime was at order 1 for each.

Narrowing down to MAR model, MAR with Gaussian, that is, Norm-MAR(3:2,1,1) accrued for a very high AIC and BIC of 5241.93 and 5258.027 respectively compared to negative AICs and BICs of -11156.110 and -11134.643; -11281.099 and -11255.648; -11171.006 and -11154.738 for Gumbel-MAR(2:1,1), Fréchet-MAR(3:1,1,1) and Weibull-MAR(3:1,2,2) respectively. Tout ensemble, the EVDs-MAR possessed smaller AICs and BICs compared to the linear models and

other regime-switching non-linear models. It was also noted that Fréchet-MAR(3:1,1,1) outperformed all the other models, as well as overcomed when subjected as random noise in the category of MAR models. This possibly indicated Fréchet as an ideal conditional error term for time series for MAR model when fitting annual time-varying contaminated series. Below is the coefficient table for the AR regime-switching processes as well as the fitted model for the better Fréchet-MAR(3:1,1,1).

Table 4. Components of the AR Processes for the Fréchet-MAR (4: 1, 1, 1) for the Rainfall

Coefficients	ω_k	ϕ_{11}	ϕ_{21}	ϕ_{31}
Component 1	0.1542 (0.0202)	0.099 (0.1072)	----	----
Component 2	0.6110 (0.3681)	----	0.0303 (0.2365)	----
Component 3	0.2348 (0.0270)	----	----	0.3913 (0.0460)

$$Fréchet-MAR\ (3:1,1,1) \approx (0.1542, 0.0162, 0.0095)\Phi\left(\frac{x_t - 0.0993x_{t-1}}{0.001}\right) + (0.6110, 0.0007, 0.0005)\Phi\left(\frac{x_t - 0.0303x_{t-1}}{0.022}\right) + (0.2348, 0.025, 0.0020)\Phi\left(\frac{x_t - 0.3914x_{t-1}}{0.040}\right) \quad (30)$$

$$\widehat{X}_t = \begin{cases} x_t + 0.0993x_{t-1} & \text{э } \omega_1 = 0.1542; \alpha_1 = 0.0162; \beta_1 = 0.0095 \\ x_t + 0.0303x_{t-1} & \text{э } \omega_2 = 0.6110; \alpha_2 = 0.0007; \beta_2 = 0.0005 \\ x_t + 0.3914x_{t-1} & \text{э } \omega_3 = 0.2348; \alpha_3 = 0.025; \beta_3 = 0.0020 \end{cases}$$

$$\sum_{k=1}^{3} \omega_k = 0.1542 + 0.6110 + 0.2348 \approx 1$$

$$AIC = -11281.099; BIC = -11255.648; Log.Lik = 2102.836; RMSE = 0.0024; MAPE = 0.112$$

It is to be noted that all the AR coefficients in each regime are stationary, that is, $\phi_{k,i} \in (0,1)$.

CONCLUSION

In this research, three new novel statistical, time series and regime-shifting EVDs-MAR (FMA, GMA and WMA) models were formulated to study the transitional shifting mechanisms of time-varying climate change of temperature and rainfall datasets. These non-linear time series models were formulated and analyzed qualitatively to climate change datasets. The three inter-switching and time-varying models contributed greatly to contaminated series of temperature and rainfall with outliers, excess kurtosis, and excess skewness. We employ the Kullback-Leibler divergence to derive the the proximity / distance (D) between finite/ delimited mixture density (g_k) and infinite mixture density (g). In conclusion**,** Fréchet gave the best fit among the regime switching for SETAR and EVDs-MAR models at Fréchet-MAR (3:1, 1, 1) with AIC=-11281.099, BIC=-11255.648 and Fréchet-MAR (4:2, 4, 1, 1) with AIC=-21358.005, BIC=-21337.956 for rainfall and temperature datasets respectively. In conclusion, it can be inferred that Fréchet random noise out of the three EVDs-MAR continous distributions is the best and ideal error term for mixture autoregressive model for modelling climate change.

References

McLachan, G..J. and D. Peel, Mixtures of factor analyzers. *In proceedings of the Seventeenth International Conference on Machine Learning*, Burlington, MA: Morgan Kaufman, 2000. 599 – 606.

Newcomb, S, A generalized theory of the combination of observations to obtain the best result. *American Journal of Mathematics*, 8: p. 343-366.

Pearson, K, Contributions to the mathematical theory of evolution. *Philosophy Transition of Royal Society London*, 1894, 185: p. 71-110.

Dempster, A.P., et al., Maximum likelihood from incomplete data via the EM-algorithm. *Journal of Royal Statistical Society B*, 1977, 39: p. 1-38.

Wong, C.S, *Statistical inference for some nonlinear time series models*, PhD thesis, University of Hong Kong, Hong Kong, 1998.

Wong, C.S. and W.K. Li, On a mixture autoregressive model. *Journal of Royal Statistical Society, Series B, Methodology*, 2000, 62(1): p. 95-115.

Boshnakov, G.N, *Prediction with mixture autoregressive models*. Research report No. 6, probability and Statistics Group Schools of Mathematics, the University of Mancheter, London, 2006.

Nguyen, N. Hidden markov model for stock trading. *International Journal Financial Studies*, 2018, 6(36): doi:3390/ijfs6020036.

Meitz, M., et al. A mixture autoregressive model based on student-t distribution. *SSRN Electronic Journal*, 2018, 2(4): p. 1-19. Doi:10/2139/ssrn.3177419.

McLachlan, G.J., et al. Finite mixture models. *Annual Review of Statistics and Its Application*, 2019, 6: p. 355-378. https://doi.org/10.1146/annurev-statistics031017-100325.

Olanrewaju, R.O., Frechet Random Noise for *k*-Regime-Switching Mixture Autoregressive Model. *American Journal of Mathematics and Statistics*, 2021, 11(1): p. 1-10. Doi:10.5923/j.ajms.20211101.01.

Li, L, et al, Stock Market Autoregressive Dynamics: A Multinational Comparative Study with Quantile Regression. *Mathematical Problems in Engineering*, 2016, 12(8): p. 1-15. https://dx.doi.org/10.1155/2016/1285768.

Maleki, M., et al., Robust Mixture Modeling Based on Two-Piece Scale Mixtures of Normal Family. *Axiom*, 2019, 8(38): doi: 10.3390/axioms8020038.