RT&A, No 4 (71)
Volume 17, December 2022

MLE OF A 3-PARAMETER GAMMA DISTRIBUTION ANALYSIS OF RAINFALL INTENSITY DATA SETS

[1]David.I.J., [2]Adubisi.D.O., [3]Ogbaji.O.E., [4]Adehi.U.M., and [5]Ikwuoche.O.P.

•

[1, 2, 3, 5]Department of Mathematics and Statistics, Federal University Wukari, Nigeria
[4]Department of Statistics, Nasarawa State University, Keffi, Nigeria
davidij@fuwukari.edu.ng[1], adubisiobinna@fuwukari.edu.ng[2], ogbajieka@yahoo.com[3], maryadehi@yahoo.com[4], philonye10@gmail.com[5]

Abstract

This research presents the maximum likelihood estimation of a three-parameter Gamma distribution with application to four types of average rainfall intensities in Nigeria. These data sets are average half-yearly, yearly, quarterly and monthly rainfall intensities. The fitted three-parameter Gamma is compared to a two-parameter Gamma distribution using empirical distribution function (EDF) tests. The tests used are Cramér-von Mises, Anderson-Darling and Kolmogorov-Smirnov statistics. Based on the results obtained at 10% significance level both the two-parameter and three-parameter Gamma distributions are of good fit to only the average yearly rainfall intensity data. A kernel density plot revealed that the average half-yearly, quarterly and monthly rainfall intensity data sets are multi-modal in nature hence a reason for both Gamma distributions poor fit to the data sets. Also, the PDF, CDF and Q-Q plots are presented which supported the outcome of the analysis.

Keywords: Gamma distribution, Anderson-Darling, Cramér-von Mises, Kernel density, Kolmogorov-Smirnov, Maximum likelihood estimation

1. Introduction

Classical analysis of statistical data in most fields including meteorology and hydrology has assumed that the data being analyzed may be reasonably modeled by distribution with somewhat *light tailed* where the tail of the density function approaches zero like some kind of exponential function (Arshad, Rasool & Ahmad, [1]). One of the most difficult problems in rainfall modeling is often the fitting of theoretical models to rainfall data (Richard, [10]). According to Hughes [8], the primary objective of modeling is frequently to generate a long representative time series of stream flow volumes from which water supply schemes can be designed. Wolfram [14] stated that Gamma distribution is a general type of statistical distribution that is related to the Beta distribution and arises naturally in processes for which the waiting times between Poisson distributed events are relevant. According to Alghazali & Alawadi [2], the two-parameter Gamma distribution is widely known and used in hydrological analysis. However, Chow *et al.*, [4] stated that the two-parameter Gamma distribution has a lower bound at zero, this condition handicaps its application to hydrological variables with lower bound larger than zero.

In the theory of probability and statistics, the gamma distribution is a two-parameter family of continuous probability distributions. It has a shape and scale parameters, say α and β respectively. If β is an integer, then the distribution represents the sum of β independent exponentially distributed random variables, each of which has a mean of α [which is equivalent to a rate parameter of α^{-1}] (Wackerly *et al.*, [12]). It often appear as solution to problems in Statistical Physics, for example, the energy density of classical ideal gas or the Wien (Vienna) distribution is an approximation to the relative intensity of black radiation as a function of the frequency (Crooks, [5]). The disadvantage of Gamma distribution is that the cumulative distribution function cannot be plotted. The 1-parameter gamma distribution is very limited in hydrological analysis due to its relative inflexibility in fitting to frequency distributions of hydrologic variation (Aksoy, [5]). Gamma distribution is widely used in many fields like reliability, survival analysis, hydrology, ecology, etc. (Dikko, *et al.*, [7]) Many variant of the gamma distribution exist and different estimation techniques have been used for estimating the gamma distribution parameters. These estimation techniques include methods of moment (MOM), percentile method, graphical estimation technique, maximum likelihood estimation (MLE), etc, with different modifications of the estimating techniques. The objective of this research is to present the estimation of the three parameters Gamma distribution using MLE and its application to four average rainfall intensity data sets for Nigeria.

2. Methods

In this section, the Gamma distribution assumptions for its applicability are presented. The probability density function (PDF) for the Gamma distribution is presented and its parameter estimation is presented using the maximum likelihood estimation technique. Four average rainfall intensity data sets which span for 115 years (1901 – 2015) are fitted for this research. The first data set is a quarterly data while the second data set used was obtained by collapsing the quarterly data to first half (FH) of the year and second half (SH) of the year, that is, average of first and second quarters to produce FH and average of third and fourth quarters to produce SH. The yearly rainfall intensity was used as the third data set and the monthly rainfall intensity data was used as the fourth. Data used was obtained from climate knowledge portal, https://climateknowledgeportal.worldbank.org.

2.1. PDF for A 3-Parameter Gamma Distribution

According to Aksoy [5], the Gamma distribution function is of three different types, 1-parameter, 2-parameters and 3-parameters Gamma distributions. If the continuous random variable x fits to the probability density function of:

$$f(x) = \frac{1}{\Gamma(k)} x^{k-1} e^{-x} ; x \geq 0 \tag{1}$$

it is said that the variable x is 1-parameter Gamma distributed, with the shape parameter k. The Gamma function $\Gamma(k)$ in equation (1) is generally expressed as:

$$\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx \tag{2}$$

when k = 1, equation (1) becomes a simple exponential distribution function. If x is replaced by x/β in equation (1) the 2-parameter Gamma distribution (2-PGD) with k being the shape parameter and β being the scale parameter is obtained as:

$$g(x;k,\beta) = \frac{1}{\beta^k \Gamma(k)} x^{k-1} e^{-x/\beta} \quad ; \ x \geq 0 \tag{3}$$

which can easily return to 1-parameter Gamma distribution for β =1. Gamma distribution with two parameters k and β denoting the shape and the scale parameters respectively are commonly used in hydrological studies (Alghazali & Alawadi, [2]). The shape of the rainfall distribution is regulated by the shape parameter and the scale parameter controls the variation of rainfall intensity series which is specified in the same unit as the random variable x (Suhaila, & Jemain, [11]). If x is replaced by $(x-\lambda)/\beta$ in equation (1) the 3-parameter Gamma distribution (3-PGD) with k, β and λ being the shape, scale and location parameters respectively is obtained as:

$$g(x,\theta)=\frac{1}{\beta^{k}\Gamma(k)}(x-\lambda)^{k-1}e^{-(x-\lambda)/\beta} \quad ;\ \theta=(\beta,\lambda,\mathrm{k})>0 \tag{4}$$

2.2. Parameter Estimation with Maximum Likelihood Estimation Technique

The likelihood function:

$$Lg(x,\theta)=\prod_{i=1}^{n}g(x,\theta) \tag{5}$$

Applying the likelihood function of equation (5) to equation (4) we have

$$\prod_{i=1}^{n}g(x,\theta)=\frac{1}{\beta^{nk}\left[\Gamma(k)\right]^{n}}\sum_{i=1}^{n}(x-\lambda)^{n(k-1)}e^{-\sum_{i=1}^{n}(x-\lambda)/n\beta} \tag{6}$$

Taking the logarithm (ln) of equation (6) we get

$$\begin{aligned} ln\prod_{i=1}^{n}g(x,\theta)&=ln\left[\beta^{k}\Gamma(k)\right]^{-n}+n(k-1)ln\sum_{i=1}^{n}(x-\lambda)-n\beta^{-1}\sum_{i=1}^{n}(x-\lambda)\\ &=-nk\,ln\,\beta-n\,ln\,\Gamma(k)+n(k-1)ln\sum_{i=1}^{n}(x-\lambda)-n\beta^{-1}\sum_{i=1}^{n}(x-\lambda)\\ &=-nk\,ln\,\beta-n\,ln\,\Gamma(k)+n(k-1)ln\sum_{i=1}^{n}x_i-n^{2}(k-1)ln\,\lambda\\ &\quad-n\beta^{-1}\sum_{i=1}^{n}x_i+n^{2}\beta^{-1}\lambda \end{aligned} \tag{7}$$

Differentiating equation (7) with respect to β and setting the derivative to zero, we have

$$\frac{d\,ln\,Lg(x,\theta)}{d\beta}=-nk\beta^{-1}+n\sum_{i=1}^{n}x_i\left(n\beta\right)^{-2}-\lambda\beta^{-1}$$

$$-nk\beta^{-1}+n\sum_{i=1}^{n}x_i\left(n\beta\right)^{-2}-\lambda\beta^{-2}=0$$

$$-nk\beta^{-1}+\overline{x}\beta^{-2}-\lambda\beta^{-2}=0$$

$$\overline{x}\beta^{-2}-\lambda\beta^{-2}=nk\beta^{-1} \tag{8}$$

Multiply both sides of equation (8) by $(nk)^{-1}\beta^{2}$ we have

$$\beta=\overline{x}\left(nk\right)^{-1}-\lambda\left(nk\right)^{-1}$$

$$\hat{\beta}=\left(nk\right)^{-1}\left[\overline{x}-\lambda\right] \tag{9}$$

Differentiating equation (7) with respect to λ and setting the derivative to zero, we have

$$\frac{d\,ln\,Lg(x,\theta)}{d\lambda}=-n^{2}(k-1)\lambda^{-1}+n^{2}\beta^{-1}$$

$$-n^{2}(k-1)\lambda^{-1}+n^{2}\beta^{-1}=0$$

$$n^{2}\beta^{-1}=n^{2}(1-k)\lambda^{-1}$$

(10)

Multiply both sides of equation (10) by $\left(n^{-2}\right)\lambda\beta$ we have

$$\hat{\lambda}=\hat{\beta}\left(1-k\right) \tag{11}$$

Differentiating equation (7) with respect to k and setting the derivative to zero, we have

$$\frac{d\,ln\,Lg(x,\theta)}{dk}=-n\,ln\,\beta-n\mathrm{D}\left[ln\boldsymbol{\Gamma}\left(k\right)\right]+n\,ln\sum_{i=1}^{n}x_i-n^2\,ln\lambda$$

$$-n\,ln\,\beta-n\mathrm{D}\left[ln\boldsymbol{\Gamma}\left(k\right)\right]+n\,ln\sum_{i=1}^{n}x_i-n^2\,ln\lambda=0$$

$$-n\,ln\,\beta+n\,ln\sum_{i=1}^{n}x_i-n^2\,ln\lambda=n\mathrm{D}\left[ln\boldsymbol{\Gamma}\left(k\right)\right]$$

$$\mathrm{D}\left[ln\boldsymbol{\Gamma}\left(k\right)\right]=ln\sum_{i=1}^{n}x_i-n\,ln\lambda-ln\,\beta$$

$$\mathrm{D}\left[ln\boldsymbol{\Gamma}\left(k\right)\right]=ln\left[\sum_{i=1}^{n}x_i-n\hat{\lambda}-\hat{\beta}\right] \tag{12}$$

where D in equation (12) is the derivative, this implies

$$\mathrm{D}\left[ln\boldsymbol{\Gamma}\left(k\right)\right]=\frac{d}{dk}ln\boldsymbol{\Gamma}\left(k\right)=\frac{\boldsymbol{\Gamma}'\left(k\right)}{\boldsymbol{\Gamma}\left(k\right)}$$

$$\frac{\boldsymbol{\Gamma}'\left(k\right)}{\boldsymbol{\Gamma}\left(k\right)}=-\gamma+\sum_{i=1}^{\infty}\left(\frac{1}{i}-\frac{1}{i+k-1}\right) \tag{13}$$

where γ is the Euler-Mascheroni constant and it is given as

$$\gamma=\lim_{n\to\infty}\left(-ln\left(n\right)+\sum_{i=1}^{n}\frac{1}{i}\right)\approx0.58 \tag{14}$$

Substituting the value of γ in equation (14) into equation (13) we have

$$-0.58+\sum_{i=1}^{\infty}\left(\frac{1}{i}-\frac{1}{i+k-1}\right) \tag{15}$$

Substituting equation (15) into equation (12) and inserting the estimates of $\hat{\beta}$ and $\hat{\lambda}$ we have

$$-0.58+\sum_{i=1}^{\infty}\left(\frac{1}{i}-\frac{1}{i+k-1}\right)=ln\left[\sum_{i=1}^{n}x_i-n\hat{\beta}k+n\hat{\beta}-\left(nk\right)^{-1}\bar{x}_i+\left(nk\right)^{-1}\hat{\lambda}\right] \tag{16}$$

Equation (16) does not exist in a closed form hence the estimation of k can only be obtained through numerical solution. This can be accomplished using any statistical software. In this research, Statistical Analytical System (SAS) version 9.4 is used to fit both the 2-PGD and 3-PGD.

2.3. Goodness of Fit Test

The goodness-of-fit tests based on empirical distribution function (EDF) are used in this research work. The EDF tests offer advantages over traditional chi-square goodness-of-fit test, including improved power and invariance with respect to the histogram midpoints (D'Agostino and Stephens, [6]). The empirical distribution function is defined for a set of n independent observations $X_1, \ldots, X_n$ with a common distribution function $F(x)$. If we Denote the observations ordered from smallest to largest as $X_{(1)}, \ldots, X_{(n)}$. The empirical distribution function, $F_n(x)$, is defined as:

$$F_n(x) = \begin{cases} 0; x < X_{(1)} \\ \dfrac{i}{n}, X_{(i)} \leq x \leq X_{(i+1)}; i = 1,2,\ldots,n \\ 1; x \geq X_{(n)} \end{cases} \tag{17}$$

Note that $F_n(x)$ is a step function that jump [$1/n$] in height at each observation, but in the case where two observations or more are equal, that is, when there are n_j observations at x_j, *then* $F_n(x)$ becomes a step function that jump [n_j/n] in height at each observation x_j. This function estimates the distribution function $F(x)$. At any value x, $F_n(x)$ is the proportion or fraction of observations less than or equal to x, while $F(x)$ is the probability of an observation less than or equal to x. EDF statistics measure the discrepancy between $F_n(x)$ and $F(x)$ which are used to conclude whether the empirical distribution $F_n(x)$ fit the hypothesize distribution $F(x)$. In this research, three EDF tests are used in testing the goodness of fit of each distribution fitted to the average monthly, quarterly, half-yearly and yearly rainfall intensity data. The EDF are Kolmogorov-Smirnov, Anderson-Darling and Cramer-von Mises. These GOF tests are presented below as follows.

2.3.1. Kolmogrov-Smirnov (D) Statistic

According to Wilks [13], the Kolmogorov-Smirnov (D) Statistic is defined as

$$D = Sup_x \left| F_n(x) - F(x) \right| \tag{18}$$

The Kolmogorov-Smirnov statistic belongs to the supremum class of empirical distribution function (EDF) statistics. This class of statistics is based on the largest vertical difference between $F(x)$ and $F_n(x)$. The Kolmogorov-Smirnov statistic is computed as the maximum of D^+ and D^-, where D^+ is the largest vertical distance between the EDF and the distribution function when the EDF is greater than the distribution function, and D^- is the largest vertical distance when the EDF is less than the distribution function.

$$D = \max(D^+, D^-) \tag{19}$$

D represents the maximum difference between the empirical and theoretical distributions over all real numbers x, and is referred to as the Kolmogorov-Smirnov value. $F_n(x)$ is the empirical cumulative probability of observing a value less than or equal to y and $1/n_p$ is added for each observation (x_i) that is greater than zero and less than or equal to y. $F(x)$ is the theoretical cumulative probability at x described by the estimated gamma distribution parameters $(\beta, \lambda, \mathrm{k})$. $F_n(x)$ and $F(x)$ are given as (Husak *et al.*, [9])

$$F_n(x) = \frac{\left(\left\{ i \in \{1,\ 2,\ \ldots,\ n\} : y_i \leq y \right\}\right)}{n} \tag{20}$$

$$F(x) = \int_0^x f(y)dy = \frac{1}{\hat{\beta}\hat{\alpha}\Gamma(\hat{\alpha})} \int_0^x y^{\hat{\alpha}-1} e^{-y/\hat{\beta}} dy \tag{21}$$

A smaller value of D implies a better fit between the observed and theoretical distributions for a fixed number of observations, n.

2.3.2. Anderson-Darling Statistic

The Anderson-Darling statistic and the Cramér-von Mises statistic belong to the quadratic class of EDF statistics. This class of statistics is based on the squared difference $\left(F_n(x) - F(x)\right)^2$. Quadratic statistics have the following general form:

$$Q = n\int_{-\infty}^{\infty} \left(F_n(x) - F(x)\right)^2 \psi(x) dF(x) \tag{22}$$

where, $\psi(x)$ is the weight function for the squared differences $\left(F_n(x) - F(x)\right)^2$.

When the weight function $\psi(x) = \left[F(x)(1-F(x)\right]^{-1}$, then the Anderson-Darling Statistic denoted by A^2 is defined as:

$$A^2 = n\int_{-\infty}^{\infty} \left(F_n(x) - F(x)\right)^2 \left[F(x)(1-F(x)\right]^{-1} dF(x) \tag{23}$$

The Anderson-Darling statistic (A^2) is computed as follows.

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}\left[(2i-1)\log U_{(i)} + (2n+1-2i)\log(1-U_{(i)})\right] \tag{24}$$

where, $U_{(i)}$ is the *ith* order Statistic.

2.3.3. Cramer-von Mises Statistic

Explained the Cramér-von Mises statistic as similar to Anderson-Darling Statistic, but in the case of Cramér-von Mises statistic, the weights function $\psi(x) = 1$. The Cramér-von Mises statistic denoted by (W^2) is defined by:

$$W^2 = n\int_{-\infty}^{\infty} \left(F_n(x) - F(x)\right)^2 dF(x) \tag{25}$$

The Cramér-von Mises Statistic (W^2) is computed as:

$$W^2 = \sum_{i=1}^{n}\left(U_{(i)} - \frac{2i-1}{2n}\right)^2 + \frac{1}{12n} \tag{26}$$

where, $U_{(i)}$ is the *ith* order Statistic.

3. Results

Results from the fitted distributions are presented below. Table 1 presents the empirical 2-PGD and 3-PGD mean and standard deviation (Std. Dev) values for the average half-yearly rainfall intensity, average yearly rainfall intensity, average quarterly rainfall intensity, and average monthly rainfall intensity data sets, that is, **AHYRI**, **AYRI**, **AQRI**, and **AMRI** respectively. It is observed that for all the data sets, the 2-PGD and 3-PGD estimates for the mean is the same as the empirical mean estimate. However, both fitted distributions estimates for the standard deviation are different from the empirical standard deviation for each data set except for the **AYRI** data set. Therefore, both the 2-PGD and 3-PGD estimated equivalent mean and standard deviation values to the that of the empirical mean and standard deviation values of 96.4014 and 7.7945 respectively.

Table 1: *Summary Statistics for the Rainfall Data*

Data Type	Statistic	Observed	2-Gamma Estimate	3-Gamma Estimate
AHYRI	Mean	96.401398	96.4014	96.4014
	Std. Dev	32.162844	32.74474	35.61137
AYRI	Mean	96.401398	96.4014	96.4014
	Std. Dev	7.7944973	7.87561	7.860996
AQRI	Mean	96.401398	96.4014	96.4014
	Std. Dev	78.321223	86.39379	90.5222
AMRI	Mean	96.401398	96.4014	96.4014
	Std. Dev	85.00559	110.8792	113.5311

The results from the summary statistics clearly give a clue that both the 2-PGD and 3-PGD will fit the average yearly rainfall intensity data better. However, such conclusion cannot be for certain until the fitted distributions are subjected to goodness of fit tests described earlier in section 2.3. The results for the parameter estimates from the 2-PGD and 3-PGD are presented in Table 2 and Figure 1, 2, 3, and 4 shows the histogram plots, the 2-PGD, and 3-PGD curves with the kernel density curve as well for the **AHRI**, **AYRI**, **AQRI**, and **AMRI** data sets.

Table 2: *Maximum Likelihood Parameter Estimates Results*

Data Type	Parameter	2-PGD Estimate	3-PGD Estimate
AHYRI	**Location**	****	41.0887
	Scale	11.12243	22.92728
	Shape	8.667296	2.412528
AYRI	**Location**	****	-10.213
	Scale	0.643406	0.579615
	Shape	149.8298	183.9402
AQRI	**Location**	****	4.125437
	Scale	77.42508	88.80176
	Shape	1.245093	1.039123
AMRI	**Location**	****	0.4245
	Scale	127.5312	134.296
	Shape	0.755904	0.714667

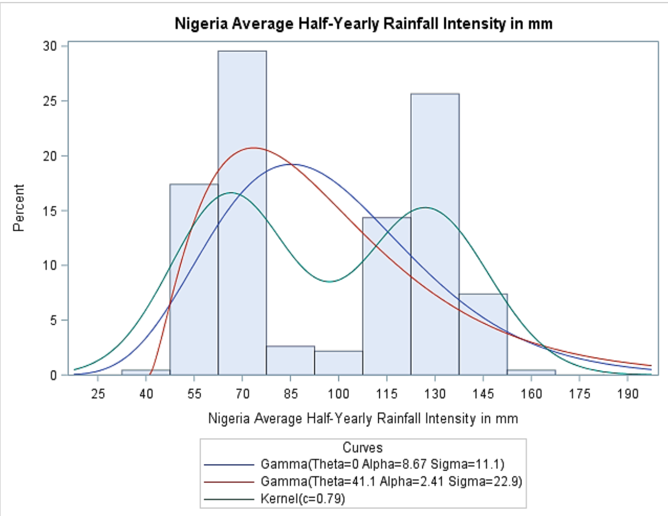

Figure 1: *Fitted Curve for AHYRI Data set*

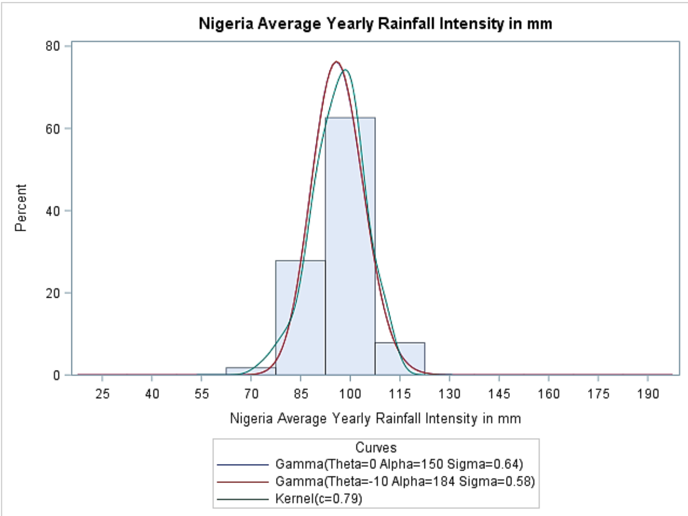

Figure 2: *Fitted Curve for AYRI Data set*

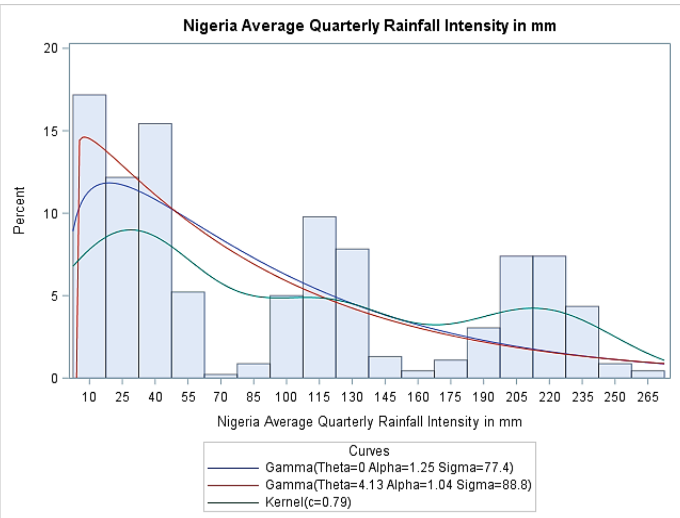

Figure 3: *Fitted Curve for AQRI Data set*

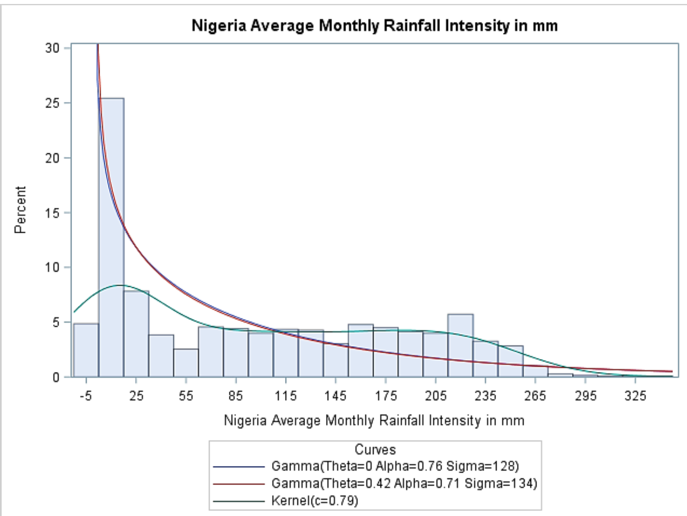

Figure 4: *Fitted Curve for AMRI Data set*

RT&A, No 4 (71)
Volume 17, December 2022

From the figures displayed, it can be seen that the two and three parameter Gamma distributions fits the **AYRI** data set (Figure 2) better compared to the **AHYRI**, **AQRI**, and **AMRI** data sets. Figure 2 shows a peaked shape with one mode compared to Figure 1, 2, and 3 with two modes, three modes and two modes respectively as depicted by the kernel density curve. To ascertain the 2-PGD and 3-PGD goodness of fit for all data sets, Table 3 presents Cramér-von Mises (***W***²), Anderson-Darling (***A***²), and Kolmogorov-Smirnov (***D***) statistics results for assessing the fitted distributions.

Table 3: *Criterion for Assessing Goodness of Fit*

Data Type and GOF Methods		Goodness of Fit Estimate (P-Values) 2-PGD	3-PGD
AHYRI	*D*	0.1809663(<*0.001*)	0.1972426(<*0.001*)
	*W*²	2.3407958(<*0.001*)	2.0138826(<*0.001*)
	*A*²	12.7906188(<*0.001*)	11.0297705(<*0.001*)
AYRI	*D*	**0.06071233(>*0.250*)**	**0.05959971(>*0.250*)**
	*W*²	**0.08224762(*0.194*)**	**0.07871282(*0.217*)**
	*A*²	**0.54769583(*0.161*)**	**0.52514845(*0.184*)**
AQRI	*D*	0.1095454(<*0.001*)	0.1179290(<*0.001*)
	*W*²	1.9423611(<*0.001*)	1.6980158(<*0.001*)
	*A*²	12.1566899(<*0.001*)	10.5830877(<*0.001*)
AMRI	*D*	4.18050(<*0.001*)	4.78650(<*0.001*)
	*W*²	6.45624(<*0.001*)	6.00624(<*0.001*)
	*A*²	38.67804(<*0.001*)	37.57614(<*0.001*)

Bold p-values imply good fit

From Table 3 above, it is clearly seen that the 2-PGD and 3-PGD are poor fit to Nigeria average half-yearly, quarterly, and monthly rainfall intensity data sets. The reason is that D, W² and A² statistic values produced p-values less than 0.01 but they produced p-values greater than 10% significance level for average yearly rainfall intensity. Therefore, it is clear from the goodness of fit statistics p-values that both the 2-PGD and 3-PGD are good fit to only the average yearly rainfall intensity data. To buttress the results discussed thus far, the cumulative density function (CDF), quantile estimates, and quantile plots (Q-Q plots) are presented. The CDF plots presented in Figure 5, 6, 7, 8, 9, 10, 11 and 12 clearly shows that only the 2-PGD and 3-PGD CDF plots for the AYRI data has a well fitted *S*-shape as seen in figure 7 and 8 respectively.

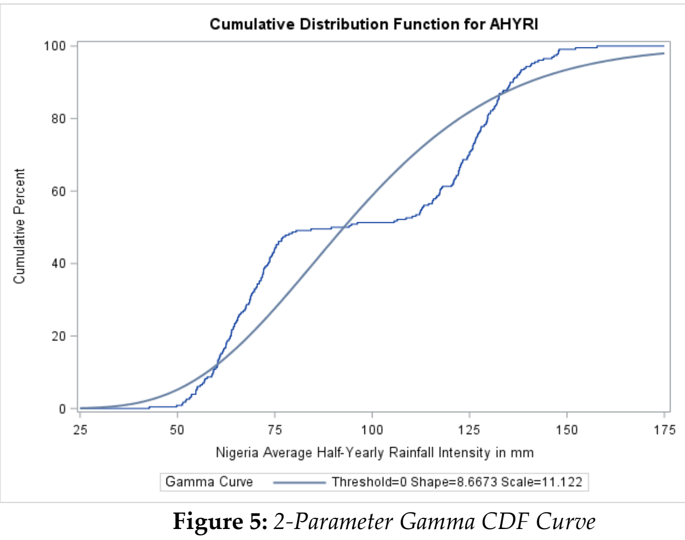

Figure 5: *2-Parameter Gamma CDF Curve*

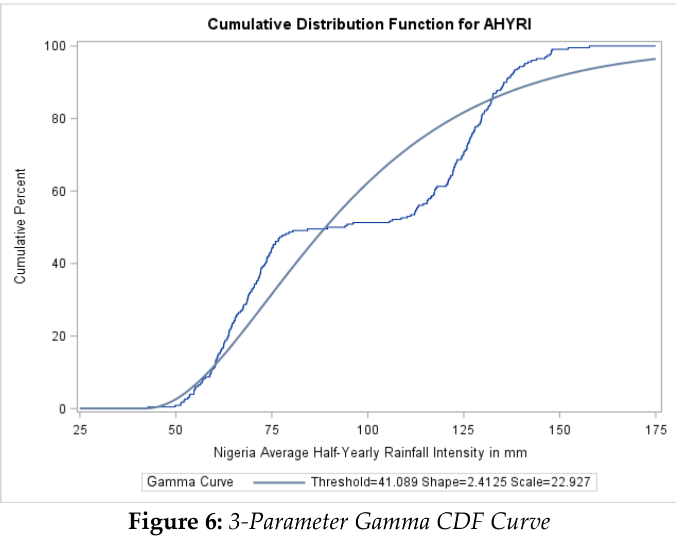

Figure 6: *3-Parameter Gamma CDF Curve*

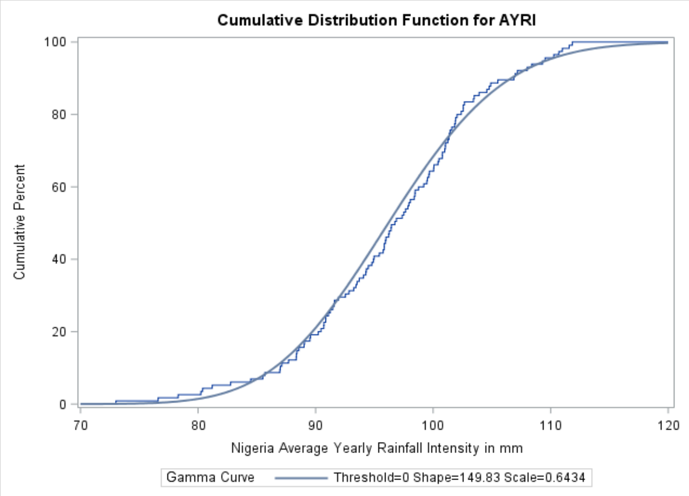

Figure 7: *2-Parameter Gamma CDF Curve*

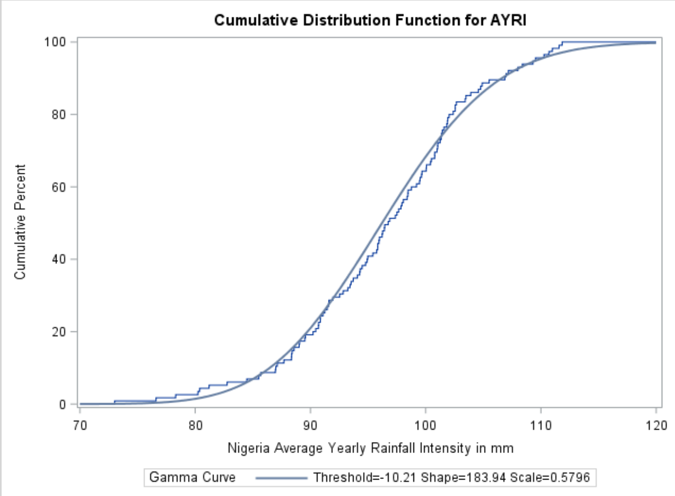

Figure 8: *3-Parameter Gamma CDF Curve*

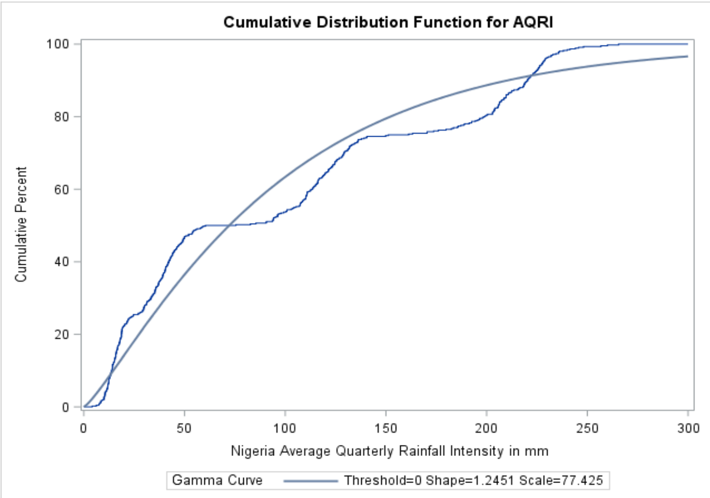

Figure 9: *2-Parameter Gamma CDF Curve*

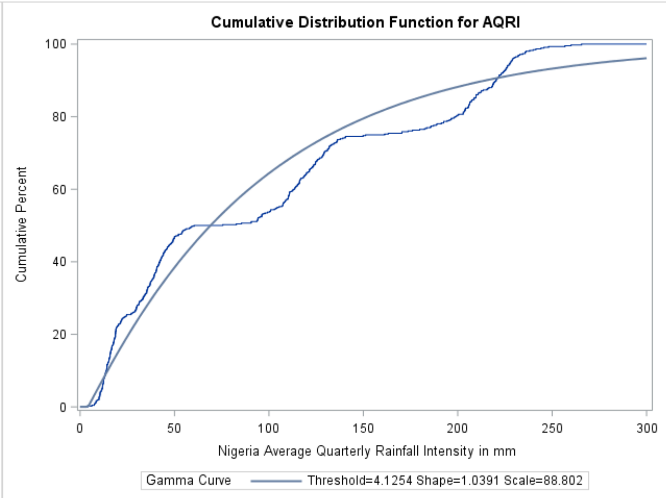

Figure 10: *3-Parameter Gamma CDF Curve*

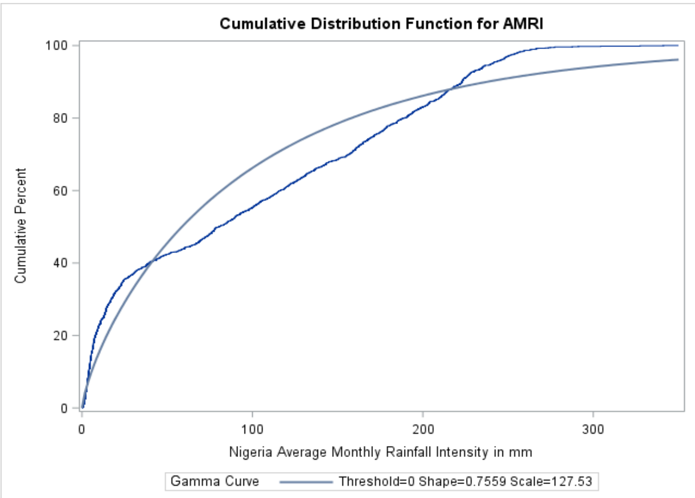

Figure 11: *2-Parameter Gamma CDF Curve*

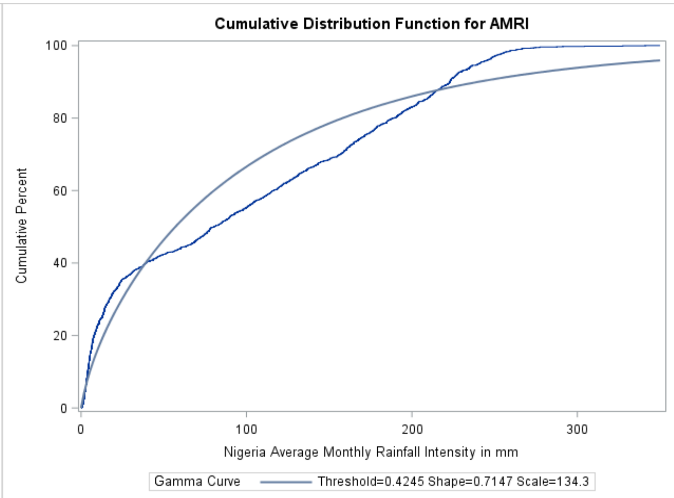

Figure 12: *3-Parameter Gamma CDF Curve*

The estimated quantile presented in Table 4 shows that the 2-PGD and 3-PGD estimated quantiles are similar to the empirical quantiles for AYRI compared to AHRI, AQRI and AMRI data sets.

Table 4: *Quantile Estimates from the three Distributions for the four Quarters*

	PERCENTAGE	OBSERVED	2-Gamma	3-Gamma
	1.0	51.2246	36.7578	46.8728
	5.0	54.7701	49.5577	53.2972
	10.0	59.2066	57.5328	58.4083
	25.0	65.4879	72.7802	70.2148
AHYRI	50.0	91.8013	92.7203	88.9753
	75.0	126.9564	116.0211	114.5791
	90.0	135.8106	140.0208	144.0947
	95.0	140.9320	155.8089	164.8805
	99.0	147.9636	188.3987	210.4818
	1.0	76.5994	79.0310	78.9703
	5.0	81.2038	83.8231	83.8090
	10.0	86.9932	86.4562	86.4594
	25.0	91.0947	90.9790	90.9992
AYRI	50.0	96.7386	96.1870	96.2083
	75.0	101.4315	101.5901	101.5931
	90.0	106.8712	106.6221	106.5916
	95.0	109.5241	109.7110	109.6526
	99.0	111.5782	115.6631	115.5364
	1.0	7.98870	2.14038	5.20545
	5.0	11.63405	8.06373	9.32423
	10.0	13.95009	14.59845	14.54737
	25.0	24.26035	33.88568	31.70857
AQRI	50.0	67.92154	72.16883	69.04598
	75.0	156.24420	133.05217	132.00585
	90.0	220.33728	210.27723	214.59713
	95.0	228.02246	267.51252	276.84191
	99.0	245.61700	398.36593	420.96745
	1.0	0.97407	0.25860	0.61226
	5.0	2.58519	2.19311	2.22199
	10.0	3.91610	5.56936	5.22762
	25.0	12.62992	19.93095	18.76462
AMRI	50.0	80.51468	58.62833	56.97441
	75.0	171.20598	132.96517	132.42207
	90.0	223.01040	237.61879	240.17013
	95.0	240.39757	319.16261	324.67099
	99.0	266.14557	512.60240	526.04011

The quantile plots for the 2-PGD and 3-PGD are presented in Figure 13, 14, 15, 16, 17, 18, 19 and 20. The 2-PGD and 3-PGD Q-Q plots for the AYRI data set showed almost all points fall on the reference straight line. This implies that the quantiles of the theoretical and data distribution agree for AYRI data set only.

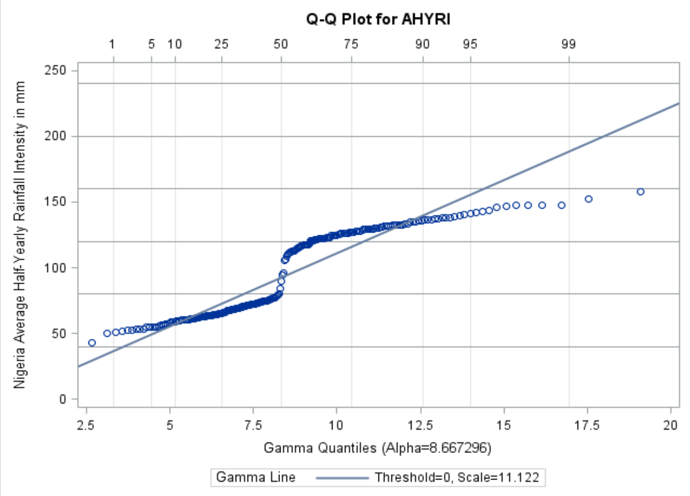

Figure 13: *AHYRI 2-P-Gamma Q-Q Plot*

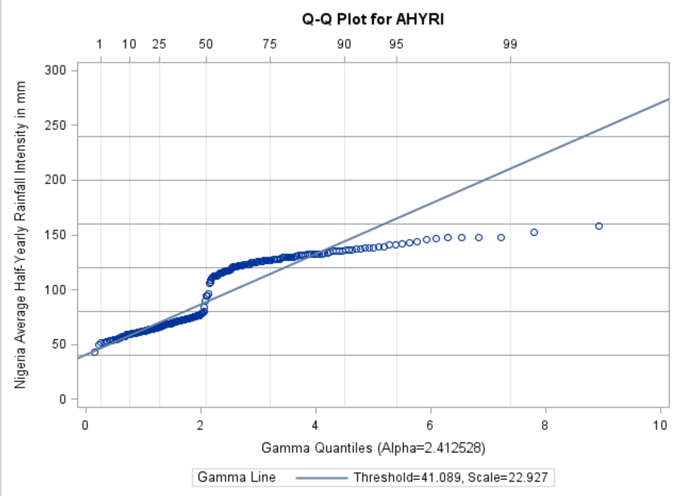

Figure 14: *AHYRI 3-P-Gamma Q-Q Plot*

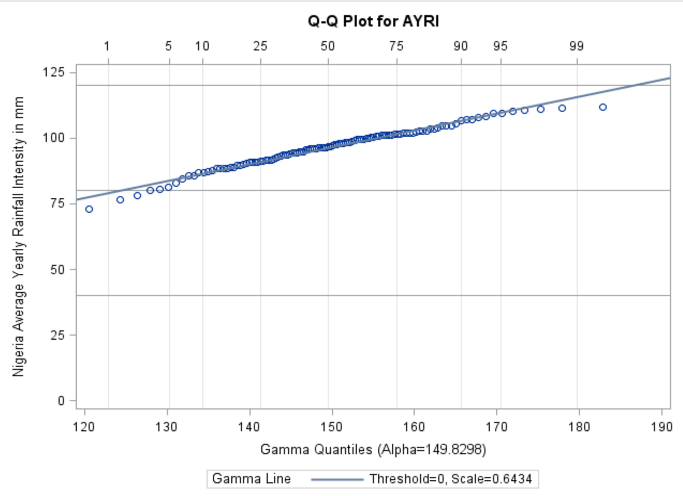

Figure 15: *AYRI 2-P-Gamma Q-Q Plot*

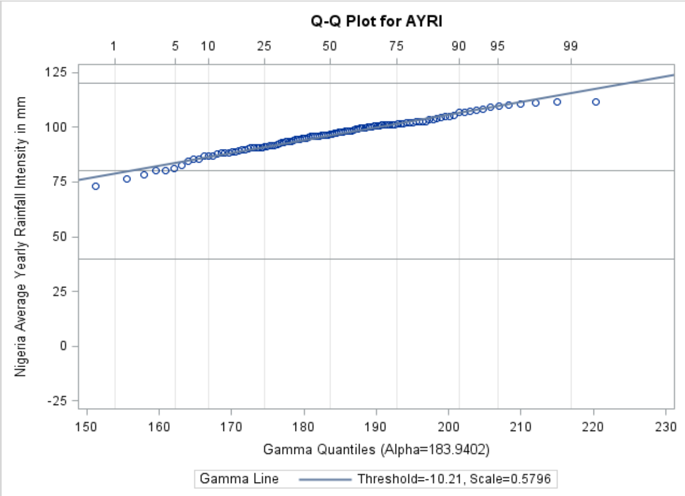

Figure 16: *AYRI 3-P-Gamma Q-Q Plot*

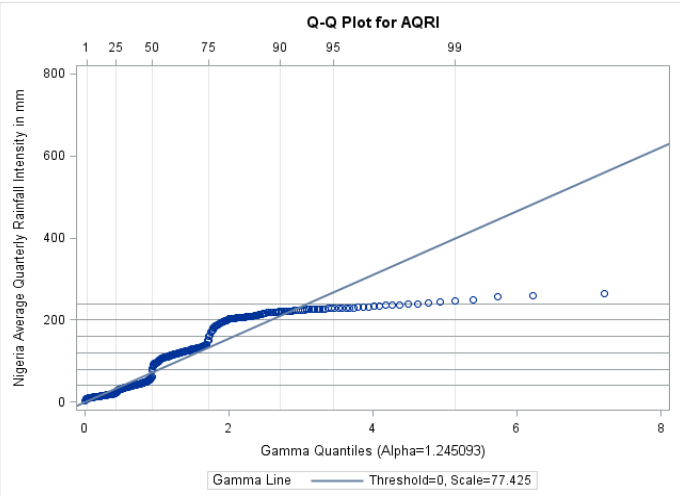

Figure 17: *AQRI 2-P-Gamma Q-Q Plot*

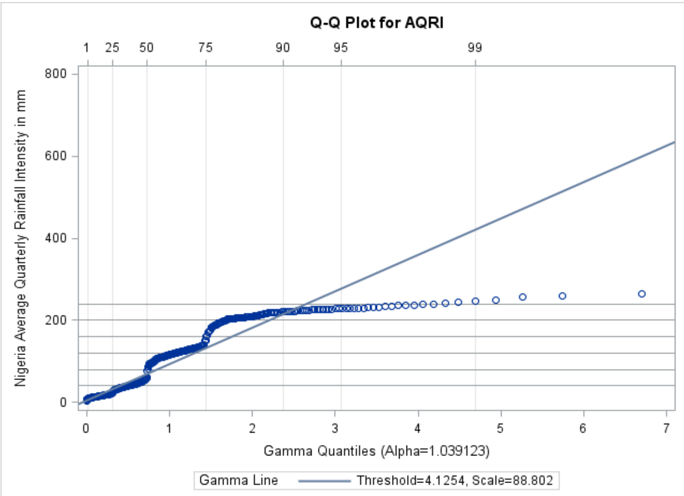

Figure 18: *AQRI 3-P-Gamma Q-Q Plot*

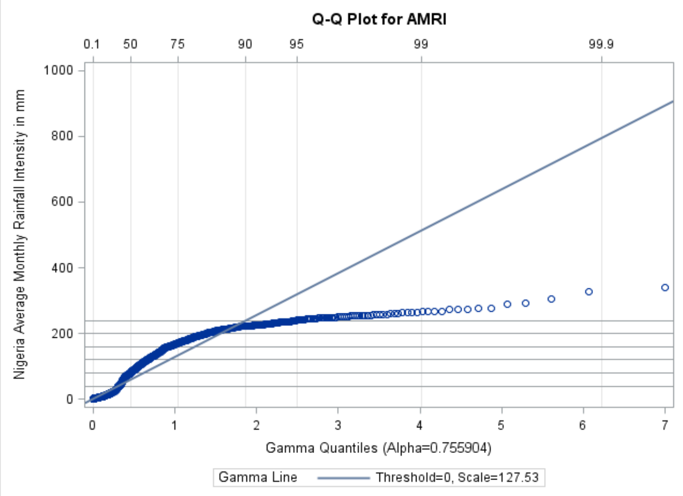

Figure 19: *AMRI 2-P-Gamma Q-Q Plot*

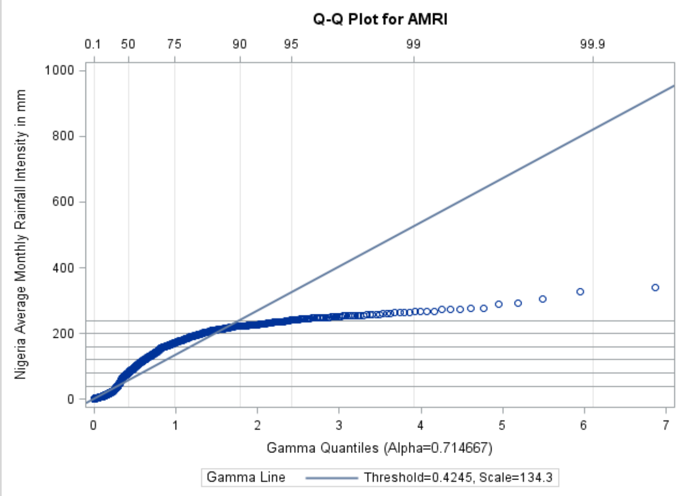

Figure 20: *AMRI 3-P-Gamma Q-Q Plot*

4. Conclusion

In this research, the maximum likelihood parameter estimation of a 3-PGD is presented. Also, its application to four different average rainfall intensity data sets was performed and compared to a 2-PGD. A goodness of fit test was performed using three criterions, that is, Cramér-von Mises (W^2), Anderson-Darling (A^2) and Kolmogorov-Smirnov (D) statistics. Based on the results obtained it is concluded that among the four data sets fitted, the 2-PGD and 3-PGD are good fit to Nigeria yearly rainfall intensity data set only. The PDF curves with kernel density curves, CDF curves and Q-Q plots showed supporting evidence as the goodness of fit statistics (W^2, A^2 and D) results. The kernel density curves showed that AHYRI, AQRI and AMRI data sets are multi-modal data sets and it is a major reason both the 2-PGD and 3-PGD fitted the data sets poorly. Hence, distributions that handle multi-modal data will be more suitable for fitting the AHYRI, AQRI and AMRI data sets.

REFERENCES

[1] Arshad, M., Rasool, M. T. & Ahmad, M. I. (2002). Rain Fall Intensity Estimate by Generalized Pareto Distribution. *Journal of Applied Sciences*, 2(7), 774 – 776.

[2] Alghazali, N. O. & Alawadi, D. A. (2014). Fitting Statistical Distributions of Monthly Rainfall for Some Iraqi Stations. *Civil and Environmental Research*, 6(6), 40-46.

[3] Aksoy, H. (2000). Use of Gamma Distribution in Hydrological Analysis. *Turkey Journal of Engineering and Environmental Science*, 24, 419-428.

[4] Chow, V. T., Maidment, D. R. & Mays, L. W. (1988). *Applied Hydrology*. McGraw Hill: Singapore.

[5] Crooks, G. E. (2014). *Survey of Simple Continuous Univariate Probability Distributions,* V0.6 Beta. Accessed on 13th January, 2022 from http://threepluseone.com/gud

[6] D'Agostino, R. B., & Stephens, M., eds. (1986). *Goodness-of-Fit Techniques*. New York: Marcel Dekker.

[7] Dikko, H. G., David, I. J., & Bakari, H. R. (2013). Modeling the Distribution of Rainfall Intensity using Quarterly Data. *IOSR Journal of Mathematics*, 9(1), 11 – 16.

[8] Hughes, D. A. (2016). Hydrological Model Comparison and Assessment: Criteria from Catchment Scales and Temporal Resolution. *Journal of Hydrological Sciences*, 61(10), 123-132.

[9] Husak, G. J., Michaelsen, J. & Funk, C. (2007). Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *International Journal of Climatology*, 27, 935-944.

[10] Richard, F. (2016). Modeling clouds and precipitation. Retrived on 12th of March, 2022 from http://www.ecmwf.int/en/about/media-centre/news/2016/modelling-clouds-and-precipitation

[11] Suhalia, J. & Jemain, A. A. (2007). Fitting Daily Rainfall Amount in Malaysia Using the Normal Transform Distribution. *Journal of Applied Sciences*, 7(14), 1880-1886.

[12] Wackerly, D. D., Mendenhall, W. & Scheaffer, L. R. (1996). *Mathematical Statistics*, 7th Edition. Thomson Higher Education: USA.

[13] Wilks, D. S. (1995). *Statistical Methods in the Atmospheric Sciences*. Academic Press, San Diego.

[14] Wolfram R. (2016). Mathematica: Version 12.0, Champaign, IL.