

# Estimation of Parameters in a Finite Mixture of Multivariate Gamma Distributions using Gaussian Approximation

V S Vaidyanathan\* and R Vani Lakshmi

Department of Statistics, Pondicherry University, Puducherry- 605 014, India

\*Corresponding Author: [vaidya.stats@gmail.com](mailto:vaidya.stats@gmail.com)

Received: 12<sup>th</sup> April 2015 / Revised: 6<sup>th</sup> January 2017 / Published: 11<sup>th</sup> March 2017

©IAppStat-SL2016

## ABSTRACT

*Finite mixture of multivariate gamma distributions is extensively used in the domains of stochastic modelling, reliability, hydrology and life testing. In this paper, we consider a multivariate gamma mixture model (MGMM) with independent marginals. A novel approach is proposed for estimating the parameters of this model. The approach makes use of Wilson-Hilferty approximation, MCLUST algorithm and the principle of maximum likelihood. Numerical illustrations based on simulated as well as real datasets have been implemented to assess the performance of the proposed approach. The results indicate that the proposed methodology provides reliable estimates for the model parameters.*

**Keywords:** Maximum Likelihood, MCLUST, Mean Square Error, Monte Carlo Simulation, Multivariate Gamma Mixture Model, Wilson-Hilferty Approximation

## 1. Introduction

Multivariate mixture models have received wide recognition in recent years owing to availability of high-dimensional data. Over the last few decades, several multivariate mixture models have been introduced to analyse multivariate events of interest in various spheres of research. Examples include multivariate skew-normal mixtures (Lin, 2009), multivariate skew mixtures (Lin, 2010) and multivariate Poisson mixtures (Karlis and Meligkotsidou, 2007). In particular, multivariate Gaussian mixtures have been extensively used to model data pertaining to genetics, astronomy and medical imaging. Kowalczyk and Tyrcha (1989) point out that “families of multivariate distributions with positively skewed marginals and positively dependent components are often needed as models for real data.” In this context,



Multivariate gamma distribution deserves considerable attention. Several forms of multivariate gamma distributions with gamma marginals are available in literature. A comprehensive study on various forms of multivariate gamma distributions can be found in Chapter 48 of Kotz et al. (2000).

It is pertinent to note that even though several forms of multivariate gamma distributions exist in literature, studies pertaining to mixture of multivariate gamma distributions have been scarce due to model complexities. Jones et al. (2000) which deals with a bivariate gamma mixture distribution where the scale parameters have a generalised Bernoulli distribution is an exception. This serves as a motivation to introduce a form of MGMM and propose a methodology to estimate its parameters.

We consider a finite mixture of  $k$ -variate gamma distributions with independent marginals.

Let  $Z_1, Z_2, \dots, Z_k$  denote independent random variates where  $Z_i \sim \text{Gamma}(\alpha_i, \beta_i), i = 1, 2, \dots, k$ . Here,  $\alpha_i$  and  $\beta_i$  represent the shape and scale parameters associated with  $Z_i$ . The joint distribution of  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_k)'$  is a multivariate gamma distribution with probability density function (pdf) given by

$$f(z_1, z_2, \dots, z_k) = \begin{cases} \frac{(z_1)^{\alpha_1-1} \cdot (z_2)^{\alpha_2-1} \dots (z_k)^{\alpha_k-1} \cdot e^{-\left(\frac{z_1}{\beta_1} + \dots + \frac{z_k}{\beta_k}\right)}}{\prod_{i=1}^k (\beta_i)^{\alpha_i} \prod_{i=1}^k \Gamma(\alpha_i)}, & 0 < z_1, \dots, z_k < \infty \\ 0, & \text{elsewhere} \end{cases} \quad (1)$$

where  $\alpha_i > 0$  and  $\beta_i > 0$ . The pdf of the resulting MGMM is defined as

$$f(\mathbf{z}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{i=1}^g \pi_i f_i(\mathbf{z}; \boldsymbol{\alpha}, \boldsymbol{\beta}) \quad (2)$$

where  $f_i(\mathbf{z}; \boldsymbol{\alpha}, \boldsymbol{\beta})$  denotes the pdf of  $i$ -th component which is a  $k$ -variate gamma distribution of the form given in equation (1) and  $g$  denotes the number of components.

Here,  $\pi_1, \pi_2, \dots, \pi_g$  represent mixture proportions or weights that satisfy the conditions (i)  $0 < \pi_i < 1 \forall i = 1, 2, \dots, g$  and (ii)  $\sum_{i=1}^g \pi_i = 1$ .

It can be seen that the total number of parameters involved in equation (2) is  $(2gk + g - 1)$ . In order to estimate the parameters, a novel methodology is introduced which makes use of Wilson-Hilferty (WH) approximation, MCLUST algorithm and the Maximum Likelihood (ML) principle.

WH approximation introduced by Wilson and Hilferty (1931) uses cube root transformation to a chi-square random variable to provide an approximation to normality. Krishnamoorthy et al. (2008) extended WH approximation to gamma distribution as follows. "Let  $X_{\alpha,\beta}$  denote a two-parameter gamma random variable, then  $X_{\alpha,\beta}^{1/3} \sim N(\mu, \sigma^2)$  where

$$\mu = \frac{\beta^{1/3} \Gamma(a + \frac{1}{3})}{\Gamma a} \text{ and } \sigma^2 = \frac{\beta^{2/3} \Gamma(a + \frac{2}{3})}{\Gamma a} - \mu^2.$$

MCLUST algorithm (Fraley and Raftery, 1998) provides a model-based clustering strategy that is used to estimate the parameters of Gaussian mixtures. It makes use of hierarchical clustering technique, Expectation-Maximization (EM) method and Bayesian Information Criterion (BIC) to identify cluster memberships and estimate the model parameters including mixture proportions. It provides reliable estimates even in the presence of noise, outliers and missing data. MCLUST is initiated by implementation of agglomerative hierarchical clustering method for the Gaussian mixture model to yield initial classifications for  $g$  mixture components. If  $g$  is not specified at the start of MCLUST algorithm, it assumes that the number of components lie between one and nine. Following this, EM algorithm is implemented corresponding to each initial classification and estimates for parameters are obtained. BIC is then computed for each resulting mixture model and the model having maximum BIC value is identified as the best model. An implementation of MCLUST in R can be found in Fraley et al. (2012).

The present article attempts to estimate the parameters of the model defined in equation (2) by combining the functionalities of WH approximation as well as MCLUST algorithm along with the principles of maximum likelihood. This is a multivariate extension to the methodology to estimate the parameters of a finite mixture of two-parameter gamma distributions proposed by Vani Lakshmi and Vaidyanathan (2016). Its potential uses involve scenarios where the marginals of the model are independent. Often, this is encountered in hydrology and reliability studies.

Rest of the article is divided into four sections. Section 2 describes the proposed methodology for parameter estimation in MGMM. Section 3 presents numerical illustration based on simulated data. In Section 4, an application pertaining to flood frequency analysis is presented and Section 5 concludes the paper with discussion.

## 2. Parameter estimation in MGMM

Consider a random sample of  $n$  observations from a  $g$ -component MGMM where component densities have the form defined in equation (1). The likelihood function corresponding to the mixture model is defined as

$$L(\mathbf{z}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{j=1}^n \sum_{i=1}^g \pi_{ij} f_i(z_1, z_2, \dots, z_k) \tag{3}$$

Following the definition in McLachlan and Peel (2004), the log-likelihood function for the complete data under the EM framework is written as

$$\log L(\mathbf{z}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{i=1}^g \sum_{j=1}^n \pi_{ij} \{ \log \pi_i + \log f_i(z_1, z_2, \dots, z_k) \} \tag{4}$$

where  $\pi_{ij} = \begin{cases} 1, & \text{if observation } j \text{ belongs to component } i \\ 0, & \text{otherwise} \end{cases}$

and  $f_i(z_1, z_2, \dots, z_k)$  denotes the  $k$ -variate gamma distribution corresponding to the  $i$ -th component with parameters  $\boldsymbol{\alpha}_i = (\alpha_{i1}, \dots, \alpha_{ik})'$  and  $\boldsymbol{\beta}_i = (\beta_{i1}, \dots, \beta_{ik})'$ . Substituting equation (1) in (4) and differentiating with respect to  $\alpha_{ip}, \beta_{ip}; p = 1, 2, \dots, k$ , we get the following log-likelihood equations.

$$\begin{aligned} \frac{\partial \log L}{\partial \alpha_{ip}} &= \sum_{i=1}^g \sum_{j=1}^n \pi_{ij} (\log(z_{ip})) - \sum_{i=1}^g \sum_{j=1}^n \pi_{ij} \log(\beta_{ip}) \\ &- \sum_{i=1}^g \sum_{j=1}^n \pi_{ij} \Psi(\alpha_{ip}) = 0 \end{aligned} \tag{5}$$

$$\frac{\partial \log L}{\partial \beta_{ip}} = \sum_{i=1}^g \sum_{j=1}^n \pi_{ij} \left( \frac{z_{ip}}{\beta_{ip}^2} \right) - \sum_{i=1}^g \sum_{j=1}^n \pi_{ij} \left( \frac{\alpha_{ip}}{\beta_{ip}} \right) = 0 \tag{6}$$

where  $\Psi(\cdot)$  denotes the digamma function. It can be observed that equations (5) and (6) do not provide closed-form solutions for  $\alpha_{ip}$  and  $\beta_{ip}; p = 1, 2, \dots, k$ . Therefore, one has to resort to iterative procedures or gradient-based search algorithms in order to estimate the parameters of the model. However, these approaches may result in estimates with large bias either because of sensitivity to starting values or complexity associated with the presence of digamma function coupled with interdependence of parameters. In the sequel, we introduce a novel approach to estimate the parameters of the model defined in equation (2).

To begin with, we transform  $Z_i \sim \text{Gamma}(\alpha_i, \beta_i), i = 1, 2, \dots, k$  as  $X_i = Z_i^{(1/3)}$ . From WH approximation, it can be seen that  $X_i \sim N(\mu_i, \sigma_i^2)$  where  $\mu_i = \frac{\beta_i^{1/3} \Gamma(\alpha_i + \frac{1}{3})}{\Gamma \alpha_i}$  and  $\sigma_i^2 = \frac{\beta_i^{2/3} \Gamma(\alpha_i + \frac{2}{3})}{\Gamma \alpha_i} - \mu_i^2, i = 1, 2, \dots, k$ . The independence of

$Z_i$ 's imply that  $X_i$ 's are independent and hence the pdf of  $\mathbf{X} = (X_1, \dots, X_k)'$  is given by

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-\left(\frac{k}{2}\right)} |\boldsymbol{\Sigma}|^{-\left(\frac{1}{2}\right)} e^{-\left(\mathbf{x}-\boldsymbol{\mu}\right)' |\boldsymbol{\Sigma}|^{-1} (\mathbf{x}-\boldsymbol{\mu}) / 2} \quad (7)$$

where  $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_k)'$  denote the mean vector and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_k^2 \end{bmatrix}_{k \times k}$$

denote the variance-covariance matrix. Substituting

equation(7) in (2), we obtain a multivariate Gaussian mixture model. Thus, WH approximation serves as an aid in transforming MGMM with independent marginals to multivariate Gaussian mixture model.

MCLUST is then implemented on the transformed model resulting in component memberships, estimates for  $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  as well as mixture proportions. However, our interest lies in estimating the parameters  $(\alpha_{ij}, \beta_{ij}), i = 1, 2, \dots, g; j = 1, 2, \dots, k$  of the gamma marginal densities. Towards this, the following steps are implemented.

1. Based on the estimates provided by MCLUST, compute  $100(1 - \delta)\%$  Confidence Interval (CI) for  $\mu_{ij}$  where  $\mu_{ij}$  denotes the mean corresponding to  $j$ -th marginal of the  $i$ -th component and  $\delta \in (0, 1)$  determines the coverage probability.

2. Using the relationship,  $\mu_{ij} = \frac{\beta_{ij}^{\frac{1}{3}} \Gamma(\alpha_{ij} + \frac{1}{3})}{\Gamma \alpha_{ij}}$  based on WH approximation, obtain all values of  $(\alpha_{ij}, \beta_{ij})$  that lies within the CI by an extensive two-dimensional search over the parameter space. Let this set be denoted by  $S$ .

3. Maximise the log-likelihood of the normal density corresponding to  $j$ -th marginal of the  $i$ -th component within  $S$  to obtain estimates for  $(\alpha_{ij}, \beta_{ij})$ , say  $(\hat{\alpha}_{ij}, \hat{\beta}_{ij})$ . The above steps are repeated for each component of the model.

In the following section, the performance of the proposed methodology is assessed in terms of the average value of the estimates (AVG), Bias, Standard Deviation (SD) and Mean Square Error (MSE) obtained through simulation studies.

### 3. Numerical Illustration

To assess the performance of the proposed approach, we consider a two-component MGMM of the form defined in equation (2) with  $k = 2$ . Datasets of sizes  $n_1$  and  $n_2$  are simulated from each component for the parameter choices given in Table 1.

Table 1: Choice of parameter values for Cases 1,2 and 3

Case	Component 1						Component 2					
	$\alpha_{11}$	$\beta_{11}$	$\alpha_{12}$	$\beta_{12}$	$n_1$	$\pi_1$	$\alpha_{21}$	$\beta_{21}$	$\alpha_{22}$	$\beta_{22}$	$n_2$	$\pi_2$
1	10.00	3.50	2.10	2.00	50	0.50	8.75	4.00	3.50	1.20	50	0.50
2	3.20	2.30	6.70	2.30	25	0.50	3.60	1.90	1.50	1.90	25	0.50
3	9.00	1.00	4.00	1.00	80	0.80	1.00	3.00	1.00	2.00	20	0.20

The above parameter choices ensure that the simulation study takes into account various levels of overlap among components, namely: low, medium and high. Figure 1 plots the mixture densities corresponding to the various cases considered in Table 1.

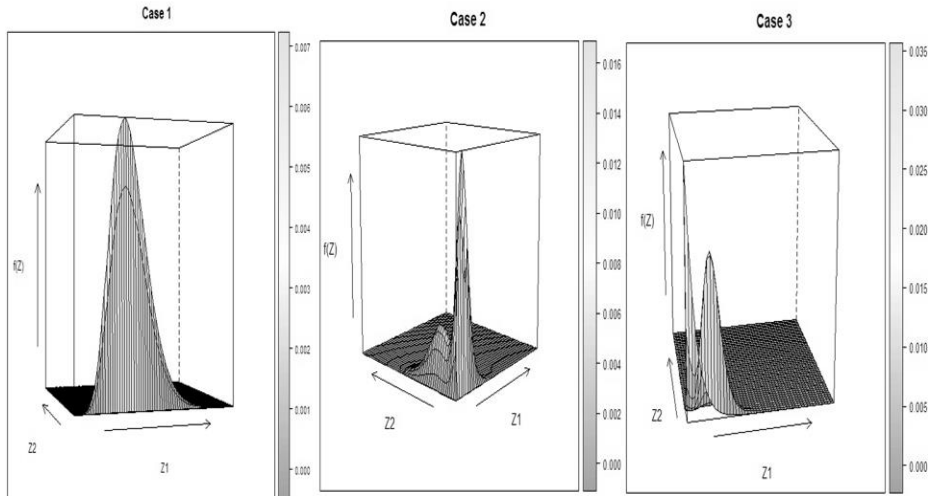


Figure 1: Mixture densities corresponding to the three different cases

From Figure 1, it can be observed that Case 1 is characterised by a high overlap among the components. In this case, the two components have same mean vector. In Case 2, there is a relatively low overlap between the components. The component-wise scale parameters are same in this case. In Case 3, there is a medium overlap among the components. In this case, the components have same variance-covariance matrix.

In order to assess the performance of the proposed approach for higher-order components, a simulation study involving three-component bivariate gamma

mixture model (Case 4) has also been carried out with parameter choice as given in Table 2.

Table 2: Choice of parameter values for Case 4

Component 1					
$\alpha_{11}$	$\beta_{11}$	$\alpha_{12}$	$\beta_{12}$	$n_1$	$\pi_1$
4.20	2.00	1.50	2.00	30	0.15
Component 2					
$\alpha_{21}$	$\beta_{21}$	$\alpha_{22}$	$\beta_{22}$	$n_2$	$\pi_2$
1.70	5.00	6.90	5.00	30	0.15
Component 3					
$\alpha_{31}$	$\beta_{31}$	$\alpha_{32}$	$\beta_{32}$	$n_3$	$\pi_3$
9.80	3.60	2.50	1.00	140	0.70

The proposed methodology is implemented for 500 simulation runs using a code developed in R version 3.1.1. The value of  $\delta$  is chosen as 0.001 so as to obtain a large confidence interval for means. The simulation results corresponding to the four cases along with misclassification rates are presented in Tables 3, 4 and 5 respectively.

From the simulation results, it is observed that the

- 1) proposed methodology provides estimates that are ‘close’ to the true parameter values for all the cases considered in the study as can be seen from the ‘small’ Bias values.
- 2) misclassification rate is low in cases where the components have low or medium overlap.
- 3) proposed methodology performs well for higher-order components.

Table 3: Simulation Results for Cases 1, 2 and 3

Case	Measure	Component 1				Component 2				Mixture Proportion	
		$\hat{\alpha}_{11}$	$\hat{\beta}_{11}$	$\hat{\alpha}_{12}$	$\hat{\beta}_{12}$	$\hat{\alpha}_{21}$	$\hat{\beta}_{21}$	$\hat{\alpha}_{22}$	$\hat{\beta}_{22}$	$\hat{\pi}_1$	$\hat{\pi}_2$
1	AVG	10.1078	3.6806	2.9538	1.6786	9.8804	3.9190	3.1426	1.6110	0.5540	0.4460
	Bias	0.1078	0.1806	0.8538	-0.3214	1.1304	-0.0810	-0.3574	0.4110	0.0540	-0.0540
	SD	1.9427	1.6741	1.2308	1.3565	2.3641	2.0036	1.5574	1.2188	0.2936	0.2936
	MSE	3.7782	2.8297	2.2407	1.9397	6.8554	4.0128	2.5484	1.6515	0.0889	0.0889
2	AVG	3.4770	2.2992	6.6680	2.2640	3.9472	1.8492	1.9618	1.5678	0.5289	0.4711
	Bias	0.2770	-0.0008	-0.0320	-0.0360	0.3472	-0.0508	0.4618	-0.3322	0.0289	-0.0289
	SD	1.0369	0.7849	1.9599	0.6492	0.9704	0.5675	1.3484	0.7242	0.0764	0.0764
	MSE	1.1497	0.6148	3.8346	0.4219	1.0603	0.3240	2.0277	0.6338	0.0067	0.0067
3	AVG	7.8438	1.1430	3.5922	1.1206	1.1358	1.7680	1.0200	1.7172	0.8262	0.1738
	Bias	-1.1562	0.1430	-0.4078	0.1206	0.1358	-1.2320	0.0200	-0.2828	0.0262	-0.0262
	SD	0.8439	0.1333	0.4619	0.1578	0.3161	0.6673	0.1279	0.5292	0.0500	0.0500
	MSE	2.0475	0.0382	0.3792	0.0394	0.1181	1.9623	0.0167	0.3595	0.0032	0.0032



Table 4: Simulation Results for Case 4

Measure	Component 1				Mixture Proportion
	$\hat{\alpha}_{11}$	$\hat{\beta}_{11}$	$\hat{\alpha}_{12}$	$\hat{\beta}_{12}$	$\hat{\pi}_1$
AVG	4.6600	1.5400	1.8000	1.5400	0.1377
Bias	0.4600	-0.4600	0.3000	-0.4600	0.0123
SD	0.3507	0.2966	0.2739	0.4159	0.0210
MSE	0.3100	0.2820	0.1500	0.3500	0.0005
Measure	Component 2				Mixture Proportion
	$\hat{\alpha}_{21}$	$\hat{\beta}_{21}$	$\hat{\alpha}_{22}$	$\hat{\beta}_{22}$	$\hat{\pi}_2$
AVG	1.8800	5.1400	7.1600	4.9000	0.1510
Bias	0.1800	0.1400	0.2600	-0.1000	-0.0010
SD	0.6221	1.4536	1.2239	1.0050	0.0069
MSE	0.3420	1.7100	1.2660	0.8180	0.0004
Measure	Component 3				Mixture Proportion
	$\hat{\alpha}_{31}$	$\hat{\beta}_{31}$	$\hat{\alpha}_{32}$	$\hat{\beta}_{32}$	$\hat{\pi}_3$
AVG	9.6400	3.8000	2.4400	2.0000	0.7113
Bias	-0.1600	0.2000	-0.0600	1.0000	-0.0113
SD	1.0502	0.5477	0.3050	1.2510	0.0225
MSE	0.9080	0.2800	0.0780	2.2520	0.0005

Table 5: Misclassification Rate

Case 1	Case 2	Case 3	Case 4
0.4988	0.1155	0.1584	0.0550

#### 4. Application

As a real-life application of the proposed methodology, we consider a dataset available in Yue (2001) that deals with stream flow measurements collected for a period of 77 years (1919-1995) from Madawaska basin in the Quebec province of Canada. Yue (2001) points out that the variables, flood duration

(expressed in number of days) and flood peak (expressed in  $m^3s^{-1}$ ) are independent ( $r = -0.0118, p \text{ value} = 0.9189$ ), each having gamma distribution. A scatter plot of the variables given in Figure 2 indicates the presence of a single cluster. Since the variables are independent with gamma marginals, the proposed approach is implemented using the joint density of flood duration and flood peak. The results indicate the presence of a single cluster with corresponding parameter estimates as provided in Table 6.

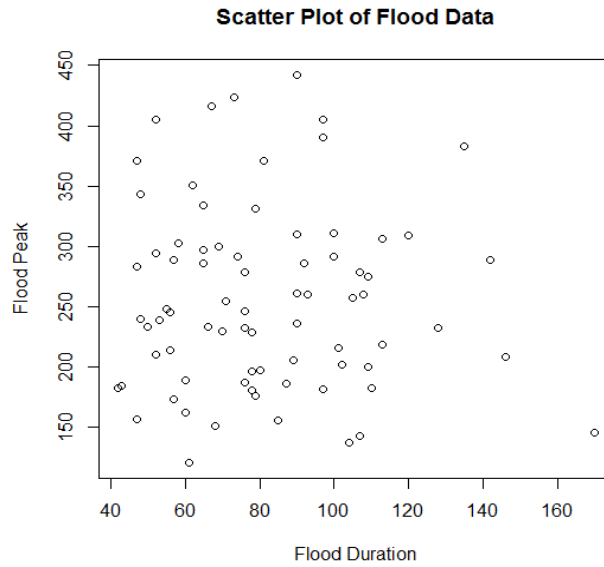


Figure 2: Scatter plot of flood duration and flood peak

Table 6: Parameter estimates of bivariate gamma model for flood data

$\hat{\alpha}_{11}$	$\hat{\beta}_{11}$	KS value	p-value	$\hat{\alpha}_{12}$	$\hat{\beta}_{12}$	KS value	p-value
9.6	8.4	0.0557	0.9707	13	19.7	0.0709	0.8337

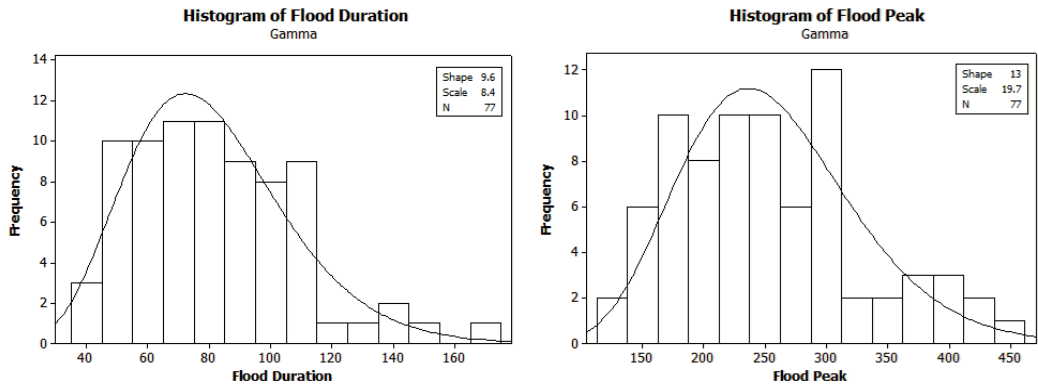


Figure 3: Histogram of flood duration and peak with their corresponding gamma fit curves

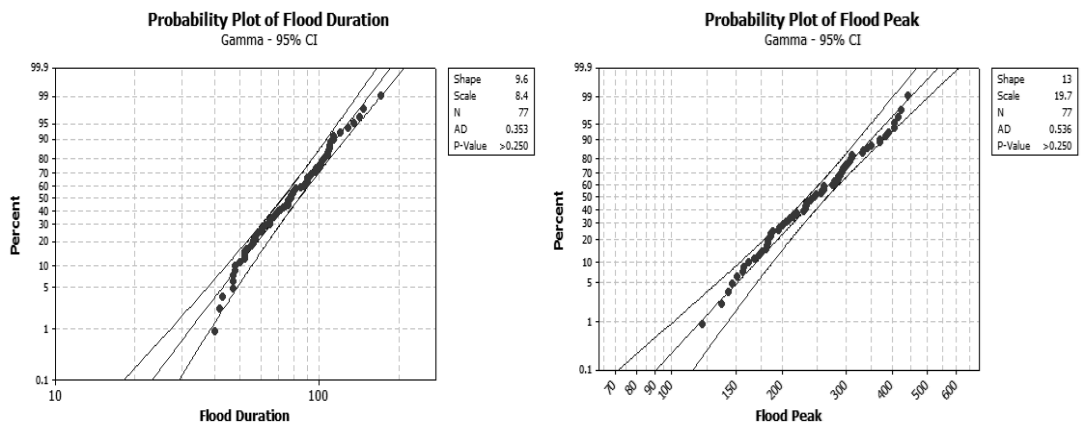


Figure 4: Probability plot of flood duration and peak using the estimates obtained by the proposed approach

### 5. Discussion

In this paper, we propose a novel approach to estimate the  $(2gk + g - 1)$  parameters of a  $k$ -variate gamma mixture model with  $g$  components assuming the variates to be independent. The proposed methodology incorporates the concepts of WH approximation, MCLUST algorithm and the principle of maximum likelihood.

The complexities associated with the estimation of parameters and identification of mixture proportions is considerably reduced due to the incorporation of WH approximation that transforms the MGMM to

multivariate Gaussian mixture setup. This enables the use of MCLUST which gives a framework to estimate the parameters of multivariate Gaussian mixture model with different covariance structures. MCLUST provides cluster memberships, estimates for component-wise  $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  and mixture proportions. The parameters of MGMM are then estimated by using the relationship between  $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  and  $(\alpha_{ij}, \beta_{ij}), i = 1, 2, \dots, g; j = 1, 2, \dots, k$  as given by WH approximation. Due to the non-linear nature of the relationship, closed form expression for the estimators do not exist. The proposed methodology overcomes this limitation by estimating the parameters through a confidence interval based search approach that involves  $\boldsymbol{\mu}$ . One may also use the estimates of  $\sigma_i^2, i = 1, 2, \dots, k$  to construct confidence interval in a similar manner and thereby obtain estimates for the parameters. However, this may be computationally more intensive than the present method.

The proposed method does not involve the use of calculus, complex iterative optimisation methods and re-parameterisation procedures. A key advantage of the method is its capability to determine the number of components present in the mixture model. However, computing time taken by the proposed method to obtain estimates is more owing to the fact that it searches the parameter space separately for each component.

From simulation results, it is clear that the proposed methodology results in estimates with 'small' Bias. Also, application of the proposed approach to real-life data indicates that the estimates obtained are good. This can be seen from large p-values associated with KS test. To conclude, the proposed methodology is able to identify the group memberships and estimate the parameters of MGMM for low, medium and high overlap between the components.

Further research is being carried out to extend the proposed methodology to estimate the parameters of MGMM in the presence of location parameters and dependent marginals.

### **Acknowledgement**

The authors express their sincere gratitude to the referees for their valuable comments and suggestions which helped in improving the article.

## References

1. Fraley C. and Raftery A.E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. *The Computer Journal*, 41(8): 578-588.
2. Fraley C., Raftery A.E., Brendan Murphy T. and Luca Scrucca. (2012). *mclust* Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. Technical Report: 597, Department of Statistics, University of Washington.
3. Jones G., Lai C.D. and Rayner J.C.W. (2000). A bivariate gamma mixture distribution. *Communications in Statistics: Theory and Methods*, 29(12): 2775-2790.  
DOI: 10.1080/03610920008832636
4. Karlis D. and Meligkotsidou L. (2007). Finite mixtures of multivariate Poisson distributions with application. *Journal of Statistical Planning and Inference*, 137:1942-1960.  
DOI: 10.1016/j.jspi.2006.07.001
5. Kotz S., Balakrishnan N. and Johnson N.L. (2000). *Continuous Multivariate Distributions, Volume 1: Models and Applications*. Wiley-Interscience Publications.
6. Krishnamoorthy K., Mathew T. and Mukherjee S. (2008). Normal-based methods for a gamma distribution. *Technometrics*, 50(1):69-78.  
DOI: 10.1198/004017007000000353
7. Kowalczyk T. and Tyrcha J. (1989). Multivariate Gamma Distributions-Properties and Shape Estimation. *Statistics: A Journal of Theoretical and Applied Statistics*, 20(3): 465-474.  
DOI: 10.1080/02331888908802197
8. Lin T. (2009), Maximum likelihood estimation for multivariate skew normal mixture models. *Journal of Multivariate Analysis*, 100:257-265. DOI: 10.1016/j.jmva.2008.04.010
9. Lin T. (2010), Robust mixture modeling using multivariate skew t distributions. *Statistical Computing*, 20:343-356.  
DOI: 10.1007/s11222-009-9128-9
10. McLachlan G. and Peel D. (2004), *Finite mixture models*, USA: John Wiley and Sons.
11. Vani Lakshmi R. and Vaidyanathan V.S. (2016). Parameter Estimation in Gamma Mixture Model Using Normal-based Approximation. *Journal of Statistical Theory and Applications*, 15(1):25-35.  
DOI: 10.2991/jsta.2016.15.1.3

12. Wilson E.B.andHilferty M.M. (1931).The distribution of chi-square. *Proceedings of the National Academy of Sciences of the United States of America*,17: 684-688.
13. Yue S. (2001). A bivariate gamma distribution for use in multivariate flood frequency analysis. *Hydrological Processes*, 15(6): 1033-1045.  
DOI: 10.1002/hyp.259