IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 30, 2019, accepted September 10, 2019, date of publication September 30, 2019, date of current version October 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2944763

Malaysian Women Shoe Sizing System Using Multivariate Normal Probability Distribution

MOHD FAIZAL MOHD HAMZAH[1], OMAR MOHD RIJAL[2], KOZO KIMURA[3], AND NORLIZA MOHD NOOR[1]

[1]Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur 54100, Malaysia
[2]Independent Researcher, Kuala Lumpur 51100, Malaysia
[3]I-Ware Laboratory Company Ltd., Osaka 562-0035, Japan

Corresponding author: Norliza Mohd Noor (norliza@utm.my)

This work was supported in part by the Ministry of Education Malaysia, in part by the Universiti Teknologi Malaysia, and in part by the University of Malaya.

ABSTRACT The Malaysian women population frequently face the problem of finding the best fitting shoes. This problem is created by the absence of a Malaysian women shoe sizing system. Standard statistical methods involving the Multivariate Normal distribution are used in a novel process of addressing issues related to the creation of a shoe sizing system, in particular, the problem of defining categories of shoe sizes. This study focused on the use of five-foot measurement namely, foot length (FL), foot breadth (FB), foot's ball girth (BG), instep length (IL), and fibulare instep length (FIL). Univariate hypothesis testing was performed taking advantage of the existence of normal probability distribution. For brevity, details for FL, FB, and BG are shown in this paper, followed by a comparison of performance results between (FL, FB, BG) and (FL, FB, BG, IL, FIL). Our results were compared to a similar study showing almost the same aggregate loss and coverage percentage. The result shows that a modest sample size of 160 was sufficient to define categories of shoe sizes to help develop a prototype shoe sizing system using the proposed novel approach. The proposed prototype shoe sizing system provides information for the planning, design, and manufacturing of Malaysian women's footwear with implications for better fitness and comfort.

INDEX TERMS Shoe size variation, clustering analysis, mixture distribution, hypothesis testing, shoe sizing system.

I. INTRODUCTION

The search for the best fitting shoe has been motivated by the need for comfort and foot health [1], avoiding injury that led to a better quality of life [2], excellence in sport [3], higher workplace performance such as in a factory [4] or hospital [5]. Many methods [6] and approaches [7], [8], [9] have been considered to develop a shoe sizing system. Improved technology has resulted in the development of devices such as a two dimensional (2D) foot scanner [10], three dimensional (3D) foot scanner [11], [12] and 3D non-contact human body laser scanner [13] to easily and readily obtain foot measurements, for example foot length, foot breadth and ball girth.

Whilst it is possible to create the best fitting shoe for a given individual, which usually involve considerable expense, the problem of the best-fitting shoe is still prominent for the general masses. The latter remark is strongly supported by the fact that different countries have different shoe sizing systems [14].

Malaysian women have a smaller foot shape compared to other regions [15]. The current practice in Malaysia is to use the United Kingdom (UK) standard, the United States of America (US) standard or the Japan Standard for selecting shoe size. Most of the time, the adjusted shoe size does not fit well to foot shape and size [16], [17]. Kong *et al.* [18] suggested that the shoe size standard should be country-specific.

One survey done by Shariff *et al.* [19] found that 60% of Malaysian women did not have enough choices in choosing shoe size while 66.3% were dissatisfied and faced difficulties in choosing the right shoe size. The survey also shows that more than two-thirds of the sample thought that Malaysia should have a newly-developed standard for shoe sizing.

Few statistical approaches have been used in determining different types of shoe sizes. An approach of differentiating the shoe sizes use a single foot measurement such as foot

The associate editor coordinating the review of this manuscript and approving it for publication was Nikhil Padhi.

length as used in ISO 9407:1991, [20]. Previous work using simple regression on two types of foot measurements investigated the relationship between foot length and joint girth to study the shoe size variation [21]. The logistic regression was used to predict footwear satisfaction with respect to different shoe sizes [22]. A hierarchical clustering method with Ward's inter-group distance measure was used to investigate the shoe size variation for 2867 German children [23]. The combination of principal component analysis (PCA) and K-means cluster analysis was applied to classify Taiwanese male's shoe sizes by using 3 types of foot measurements; foot breadth, foot length and navicular height [24].

One important issue concerning different shoe-sizing systems is the problem of measuring and quantifying shoe size variability which is addressed in this paper. This paper presents a novel process to investigate shoe size variation conditional on adopting the Japanese standard (JIS S-5037). Initial work on validation of clusters of shoe size measurement was done in [25]. A statistical approach was used where clustering and graphical methods will provide an initial indication of data clusters, and the Multivariate Normal probability distribution provides a model for the data where statistical tests are carried out. An initial shoe sizing system is created from the clustering method which is then improved using the Multivariate Normal mixture distributions. For the proposed shoe sizing system, univariate hypothesis testing was also performed to investigate the significant separation of subdivisions of shoe size, taking advantage of the existence of the normal probability distribution. To the best of our knowledge, the use of statistical probability distribution and statistical hypothesis testing to create shoe size categories has not been done before.

II. MATERIALS AND METHODS

A. DATA DESCRIPTION

A total of 160 randomly selected Malaysian women volunteers from different backgrounds were recruited and all participants formally consented to be in this study. A sample size of 160 is commonly regarded as sufficient to satisfy the statistical requirement for model fitting and hypothesis testing done in this study. Their feet were scanned using the I-Ware USB High Type Foot Scanner. The age of the women volunteers ranged between 19-71 years old. The descriptive statistics of the individuals are given in Table 1 where the BMI range from 15.2 to 46.3.

Only left foot measurements will be used in this study since there is no significant difference between the left foot and the right foot [16], [26]. Figure 1 illustrates all five-foot measurements being measured in centimeter (cm) for the left foot for each woman volunteer. In this study, five-foot measurements which is foot length, ball girth, foot breadth, instep length and fibulare instep length (labelled as 1, 2, 3, 4 and 5 in Figure 1) are used to represent shoe size as according to the Japanese standard (JIS S-5037) and further supported by [27], [9], [12].

TABLE 1. Descriptive statistics of individuals (N=160).

Characteristics	Mean	Minimum	Maximum	Standard Deviation
Age (years)	36.09	19	71	11.15
BMI	24.9	15.2	46.3	6.05

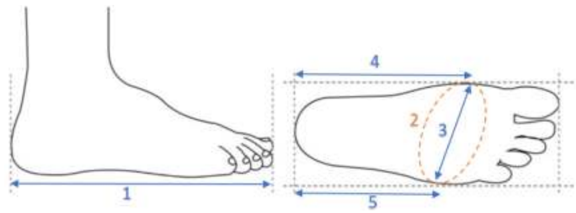


FIGURE 1. Illustration of Foot Measurements. 1: Foot length, 2: Ball girth, 3: Foot breadth, 4: Instep length, 5: Fibulare instep length.

In this study, agglomerative hierarchical clustering analysis was used to obtain the initial grouping of the data. A normal probability mixture model was utilized to confirm data grouping. Statistical tests were carried out on the derived mixture model to define the categories of shoe sizes. For brevity, in this paper, the details are illustrated for the case of three-foot measurements namely, FL, FB, and BG. The comparison between the use of three-foot measurements (FL, FB, BG) and the use of five-foot measurements (FL, FB, BG, IL, FIL) were illustrated with the performance measure namely, percentage of coverage and aggregate loss.

B. AGGLOMERATIVE HIERARCHICAL CLUSTERING ANALYSIS

An agglomerative hierarchical clustering method is one of the common tools used in shoe size modeling [28], [29], [23]. Several distance measures such as Euclidean distance, Manhattan distance, Maximum distance, Canberra distance, and linkage methods like single linkage, complete linkage, group average linkage, centroid linkage, median linkage and Ward linkage method were applied to all possible combination of the selected foot measurements.

The availability of many types of distance measures and linkage methods may result in conflicting results [30]. This study will exploit the Cophenetic Correlation Coefficient (CPCC) index that is one of the internal validation criteria with range [−1,1] in choosing the optimal distance measure and linkage method. The CPCC index compares the original distance measure matrix, D with the generated cophenetic distance matrix, C. The cophenetic distance between any two observations, x_i, x_j is defined to be the intergroup dissimilarity when these two observations are combined into a single cluster for the first time. The large value of the CPCC index indicates that the selected linkage method and distance measure yield optimal clusters [31], [32]. The results in a previous study [25] showed that Euclidean distance and Average linkage provide the best clustering method for this dataset using the CPCC index.

$$CPCC = \frac{\frac{1}{M}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d\left(x_i, x_j\right) c\left(x_i, x_j\right) - \mu_D\mu_C}{\sqrt{(\frac{1}{M}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d^2\left(x_i, x_j\right) - \mu_D^2)(\frac{1}{M}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} c^2\left(x_i, x_j\right) - \mu_C^2)}} \tag{1}$$

For any two observations, x_i and x_j, the CPCC index is defined as (1), as shown at the top of this page, where $d\left(x_i, x_j\right)$ is a distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, $c(x_i, x_j)$ is the dendrogrammatic distance between the model points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$,. This distance is the height of the node at which these two points are first joined together. $M = \frac{n(n-1)}{2}$ and $\mu_D = \frac{1}{M}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d\left(x_i, x_j\right)$, $\mu_C = \frac{1}{M}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} c\left(x_i, x_j\right)$.

C. MIXTURE MODEL

The result from agglomerative hierarchical clustering seldom gives a concrete and conclusive outcome [33]. The s-component mixture model is then fitted to the s-clusters obtained. Let $\boldsymbol{x}^{\boldsymbol{T}} = (x_1, x_2, x_3)$ denote foot measurements such that $\boldsymbol{x}$ are considered independent and identically distributed random vectors having probability density function (pdf)

$$f\left(\boldsymbol{x}; \boldsymbol{\pi}, \boldsymbol{\theta}\right) = \sum_{r=1}^{s} \pi_r g_r(\boldsymbol{x}; \boldsymbol{\theta_r}) \tag{2}$$

where $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \ldots, \pi_s)$, $\boldsymbol{\theta}^T = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_s)$, $\sum_{r=1}^{s} \pi_r = 1, 0 \le \pi_r \le 1$ and $g_r(.) \ge 0$, $\int g_r(.)\, dx = 1$ for $r = 1, 2, \ldots, s$.

π_r is called the mixing proportion, $\boldsymbol{\theta_r}$ is the estimated parameter, $g_r(.)$ is the component density of the distribution, and s is the number of clusters [34], [35]. Parameter estimation is carried out using the Expectation Maximization (EM) algorithm. The EM algorithm was introduced to find the maximum likelihood estimates (MLE) for incomplete data [36]. It gains popularity vastly due to its simplicity and faster computation compared to the Newton Raphson method, despite its slow convergence [37]. To show that data fits the mixture normal model, the Likelihood Ratio Test (LRT), Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used.

The likelihood ratio λ [38] computed from the EM algorithm is used to calculate $-2log\lambda$ which then will be used to justify the number of clusters, s, gained from the EM algorithm, after carrying out the following hypothesis test, $H_0 : s = s_0 vs H_1 : s = s_0 + 1$, given s_0 is number of cluster and the test statistic is $-2log\lambda = -2(logL(\hat{\theta}|x)_{s_0+1} - logL(\hat{\theta}|x)_{s_0}) \sim \chi^2_{m(s_0+1)-m(s_0)}$ where m is the number of parameter and $m(s_0 + 1) - m(s_0)$ is the difference of number of parameters under alternative and null hypothesis. Finally, $\chi^2_{m(s_0+1)-m(s_0)}$ is the Chi-square distribution with one degree of freedom. The p-value for this test is also calculated [38].

Two other information criteria, Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC), were also used to justify the number of components. Suppose $s = s_0$, then AIC and BIC are as follows,

$$AIC = -2logL\left(\hat{\theta}\,|\,x\right) + 2m\left(s_0\right) \tag{3}$$

$$BIC = -2logL\left(\hat{\theta}\,|\,x\right) + 2m\left(s_0\right) log\left(n\right) \tag{4}$$

such that $m(s_0)$ is the number of parameters in model with $s = s_0$ and n is the sample size. Both AIC and BIC penalized the model for increasing use of estimated parameters which in turn avoid over-fitting for the data. However, there are studies showing that AIC tends to overestimate the number of components and it should be compared to BIC, as BIC put more weight on punishment for more parameters being used (log n) [38]. The smallest value of AIC or BIC (for a fixed number of shoe size measurement) suggests mixture model fit data, which in turn is a way of verifying the parameter, in particular, s which is the number of clusters [38].

D. STUDYING SHOE SIZE WITH MIXTURE MODEL

The value of s in (2), section II(C) is the first definition of shoe size variation. If s equals two, the original sample is divided into two groups of individuals with significant different shoe sizes. Within each group, a univariate hypothesis test of the difference between sub-samples can be carried out, as we can take advantage that shoe size measurements are normally distributed.

To test the hypothesis concerning the equality of two group variance; $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$, the test statistic used is $F = \frac{s_1^2}{s_2^2}$ where s_1^2 is the larger sample variance. The two-sample t-test [39] is used to determine if two population means are equal. For the case of equal variance assumption, the common standard deviation can be estimated by the pooled standard deviation: $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$. The test statistic is $t^* = \frac{\bar{x}_1-\bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ with the degree of freedom equal to $df = n_1 + n_2 - 2$.

For the case of unequal variance assumption, the test statistic is $t^* = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$ with the degree of freedom equal to $df = \frac{(n_1-1)(n_2-1)}{(n_2-1)C^2+(1-C)^2(n_1-1)}$ (round down to nearest integer), where $C = \frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$

An alternate, conservative option to using the exact degrees of freedom calculation can be made by choosing the smaller of $(n_1 - 1)$ or $(n_2 - 1)$. We may also test the hypothesis that the groups mean are equal, H_0: $\mu_1 = \mu_2$ vs H_a: $\mu_1 \neq \mu_2$.

E. A BLIND SHOE SIZING SYSTEM

Given that we used no information from the data, a blind shoe sizing system can be created. In this section, all shoe sizing system assume that the 3D foot measurements are treated independently. For each shoe size measurement, twelve intervals of that measurement values can be constructed as follows. Without loss of generality consider foot lengths; $\mu_k + j\sigma_k, k = 1, 2$ and $j = -3, -2, -1, 0, 1, 2, 3$, where μ_k and σ_k is the mean and the standard deviation of foot length for group k (k = 1, 2). If the actual value of foot measurement (for example FL) of an individual falls in any interval shown above, the individual will belong to that interval. However, given three-foot measurements, the total number of combinations will be $12 \times 12 \times 12$ which is 1728 combinations as illustrated in Figure 2. If the blind shoe sizing system is used, the data size should be larger than 1728. This is not practical and therefore we proposed a novel procedure involving clustering and mixture distribution as explained in section II(B) and section II(C).

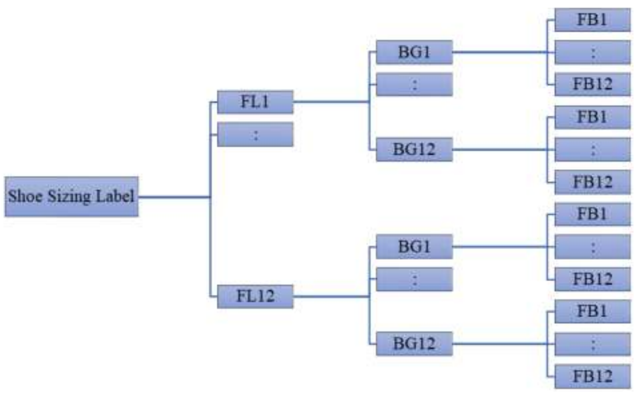


FIGURE 2. Diagram on how the shoe sizing system is to be labeled.

F. A NOVEL SHOE SIZING SYSTEM

The blind shoe sizing system has certain disadvantages in particular when the shoe size measurements are considered independently. The number of subdivisions of shoe sizes is large. An improvement may be achieved if these subdivisions are determined from a 3D individual instead of considering their measurements separately. A novel procedure is proposed where the 3D data is clustered as explained in section II(B), followed by fitting of normal mixture distribution as explained in section II(C). The CPCC value is used to indicate the optimal cluster whilst the AIC, BIC, and p-value indicate optimal mixture distributions.

G. VALIDATION OF SHOE SIZING

The final phase of this paper will validate the proposed shoe size system based on three criteria which are the percentage of coverage, number of shoe size and aggregate loss as suggested by Zakaria [40]. The good shoe sizing system should have a higher percentage of coverage, a smaller number of shoe sizes and aggregate loss less than benchmark value. There is a relationship between the percentage of coverage and the number of shoe sizes, where the percentage of coverage is high when many shoe sizes are used. To accommodate this issue, the shoe size with more than 2% of coverage is selected as an acceptable shoe size [41].

The aggregate loss was employed as the third factor in accessing the goodness of fit of the proposed shoe size as used in [24]. The aggregate loss was calculated using the following formula (5), as shown at the bottom of this page.

The benchmark value used in this paper is $\sqrt{n}$ inch where n is the number of foot measurements used for developing a size system [42]. In this study n is equal to 3. Explicitly, the benchmark value used is 43.99 mm.

III. RESULT

A. A SHOE SIZING SYSTEM FROM AGGLOMERATIVE CLUSTERING METHOD

The Agglomerative clustering method using Average linkage and Euclidean distance is illustrated in Fig. 3. The stopping point for the clustering is determined by the CPCC index less than 0.63 and the p-value greater than 0.2. Five initial clusters (groups) labeled as A, B, C, D, and E were obtained. Cluster C and cluster E will be regarded as two separate subdivisions of shoe size, as we cannot perform any statistical test of differences simply because of small sample size.

For cluster A, cluster B, and cluster D, further subdivisions of shoe size may be created by studying the pair-wise correlations of FL, FB, and BG. The foot length may be regarded as uncorrelated with the other two shoe size measurements (Table 2). Henceforth within each of cluster A, cluster B and cluster D, the division of FL based on the Mondopoint sizing system is done first, where the interval of foot length is 5mm shown in Table 3. For each FL interval, four subdivisions of (FB and BG) are created (Table 4). Figure 4 illustrates a possible shoe sizing system developed from the Agglomerative Clustering method.

B. A NOVEL SHOE SIZING SYSTEM FROM MIXTURE MODEL

The AIC and BIC show the goodness of fit of the data to the mixture distribution as illustrated in Table 5. If the hypothesis

$$Aggregate\ Los = \frac{\sum\left(\sqrt{(Assigned\ FL - Actual\ FL)^2 + (Assigned\ FB - Actual\ FB)^2 + (Assigned\ BG - Actual\ BG)^2}\right)}{Number\ of\ individuals\ in\ the\ category} \quad (5)$$

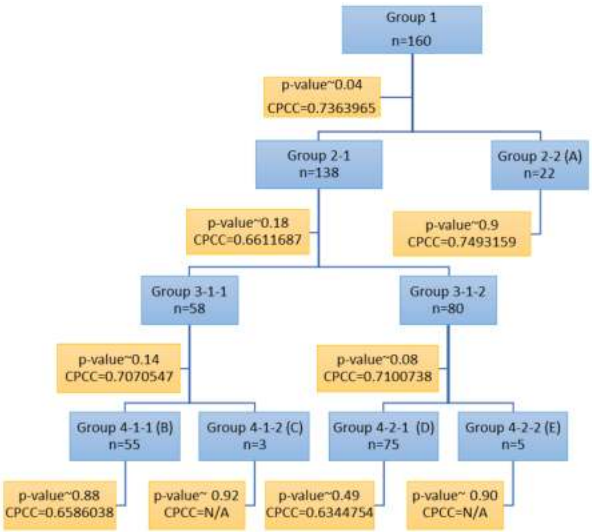


FIGURE 3. The CPCC index and p-value for clustering analysis showing clusters (A), (B), (C), (D) and (E).

TABLE 2. The pairwise correlation between foot length, foot breadth and ball girth for cluster A, cluster B and cluster D.

Cluster A	**FL**	**FB**	**BG**
FL	1	0.162534	0.216439
FB	0.162534	1	0.865752
BG	0.216439	0.865752	1
Cluster B	**FL**	**FB**	**BG**
FL	1	0.11117	0.076637
FB	0.11117	1	0.758554
BG	0.076637	0.758555	1
Cluster D	**FL**	**FB**	**BG**
FL	1	-0.00733	0.041638
FB	-0.00733	1	0.851549
BG	0.041638	0.851549	1

TABLE 3. The FL division based on the Mondopoint system.

FL label	Length (mm)	FL label	Length (mm)
FL01	205≤FL<210	FL08	240≤FL<245
FL02	210≤FL<215	FL09	245≤FL<250
FL03	215≤FL<220	FL10	250≤FL<255
FL04	220≤FL<225	FL11	255≤FL<260
FL05	225≤FL<230	FL12	260≤FL<265
FL06	230≤FL<235	FL13	265≤FL<270
FL07	235≤FL<240	FL14	270≤FL<275

test $H_0 : s = 1\ vs H_1 : s = 2$ is considered (second row of Table 5), the null hypothesis is rejected due to a small p-value (by using a significance level of 0.10). This result is supported

TABLE 4. The subdivisions of FB and BG.

Subdivisions	**FB**	**BG**
Subdivisions 1	FB> mean FB	BG> mean BG
Subdivisions 2	FB> mean FB	BG< mean BG
Subdivisions 3	FB< mean FB	BG> mean BG
Subdivisions 4	FB< mean FB	BG< mean BG

TABLE 5. Likelihood ratio test (LRT) result and information criteria for three shoe size measurements.

Shoe size Measurements	No of group	Log-likelihood	$-2\log\lambda$	LRT p-value	AIC	BIC
FL, BG, FB	1	1660.3	-	-	3326.7	3335.9
	2	-1606.7	107.2	0.05	3227.4	3249
	3	-1605.3	1.4	0.36	3232.6	3266.4

by the BIC value that shows a smaller value for two groups (3249 less than 3335.9). However, if the hypothesis test $H_0 : s = 2\ vs H_1 : s = 3$ is considered, the p-value is 0.36, which suggests accepting the null hypothesis. These results imply that the optimal number of groups is two if three shoe size measurements are used.

To improve this shoe sizing system results from [25] will be used. The mixture distribution, $f(\underline{x}) = p_1 N(\mu_1, \sum_1) + p_2 N(\mu_2, \sum_2)$ with $\underline{x}^T = (x_1, x_2, x_3)$ where x_1 is foot length, x_2 is foot breadth and x_3 is ball girth is given as follows;

$$f(\underline{x}) = 0.865\, N\left[\begin{bmatrix} 232.34 \\ 94.21 \\ 225.98 \end{bmatrix}, \begin{bmatrix} 114.630 & 29.906 & 71.959 \\ 29.906 & 21.830 & 49.887 \\ 71.959 & 49.887 & 125.984 \end{bmatrix}\right] + 0.135N\left[\begin{bmatrix} 246.53 \\ 107.24 \\ 257.02 \end{bmatrix}, \begin{bmatrix} 127.210 & 9.473 & 25.448 \\ 9.473 & 10.527 & 21.469 \\ 25.448 & 21.469 & 53.440 \end{bmatrix}\right] \tag{6}$$

Geometrically, the creation of a shoe sizing system is the sub-division of two ellipsoids, $\{N(\mu_1, \sum_1)$ and $N(\mu_2, \sum_2)\}$. If we regard the probability distribution from (6) as representing the population, the first ellipsoid represents 86.5% of the population whilst the other represents 13.5% of the population.

The t-test and F-test (section II(D)) is used to investigate the separation between ellipsoids. From Table 6, there is a clear separation between the two ellipsoids. All foot measurements, foot length, foot breadth, and ball girth have an unequal mean. The measurement FL shows equal variances whilst BG and FB shows unequal variances. These remarks show that both ellipsoids are significantly different. Each ellipsoid then can be sub-divided to form a new sizing system. Table 7 shows how eight sub-divisions are created and Figure 5 shows the membership of these sub-divisions.

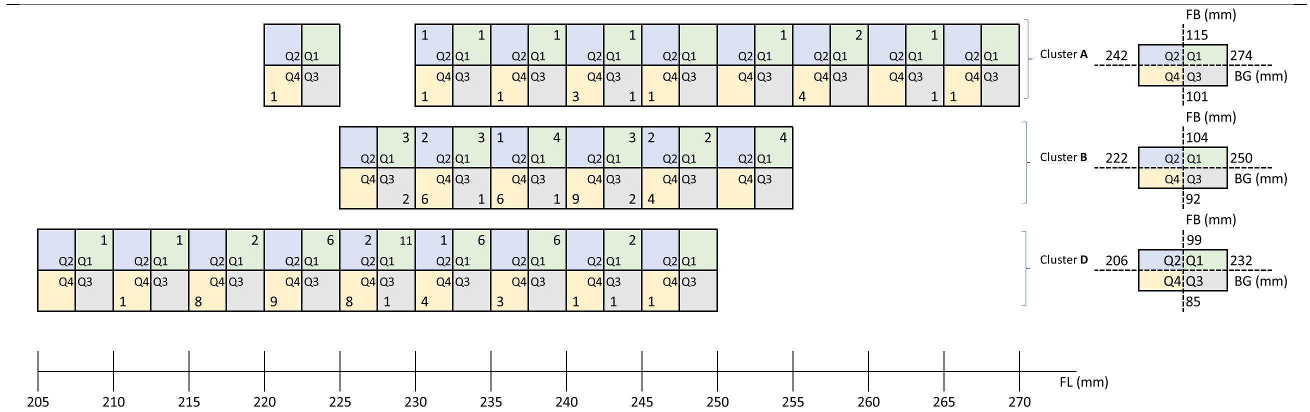


FIGURE 4. A shoe sizing system developed from the Agglomerative Clustering method. The sub-division of FL based on the Mondopoint sizing system is based on Table 3. For each FL interval, the FB and BG are represented by the quadrants Q1, Q2, Q3 and Q4, as defined is Table 4.

TABLE 6. Testing the equality of variance and mean for ellipsoids.

Shoe size measurements	Test equality of variance	Test equality of mean when variance equal	Test equality of mean when variance not equal
FL	$F = \frac{\sqrt{127.21}^2}{\sqrt{114.63}^2} =$ 1.11 $F <$ $F_{0.05}$~1.84, accept H_0 Both variance are equal	Pooled std dev: $s_p = 32.392$ $T = -5.507$ $\|T\| > t_{1-0.05/2,v}$~1.960 Reject H_0 Both mean are unequal	Not applicable
BG	$F = \frac{\sqrt{125.984}^2}{\sqrt{53.440}^2} =$ 2.36 $F >$ $F_{0.05}$~1.84, reject H_0 Both variance are unequal	Not applicable	Pooled std dev: $s_p = 10.925$ $T = -16.426$ $\|T\| > t_{1-0.05/2,v}$~1.960 Reject H_0 Both mean are unequal
FB	$F = \frac{\sqrt{21.830}^2}{\sqrt{10.527}^2} =$ 2.07 $F >$ $F_{0.05}$~1.84, reject H_0 Both variance are unequal	Not applicable	Pooled std dev: $s_p = 4.6357$ $T = -15.794$ $\|T\| > t_{1-0.05/2,v}$~1.960 Reject H_0 Both mean are unequal

The sub-division can be further refined by creating subdivisions of foot length following the Mondopoint sizing system. Then, four quadrants are created following Table 4 based on the mean obtained from two ellipsoids. The refined sizing systems developed are shown in Figure 6.

TABLE 7. The subdivisions based on FL, FB and BG for given Ellipsoid.

Subdivisions	FL	FB	BG
Subdivisions 1	FL< mean FL	FB> mean FB	BG> mean BG
Subdivisions 2	FL< mean FL	FB> mean FB	BG< mean BG
Subdivisions 3	FL< mean FL	FB< mean FB	BG> mean BG
Subdivisions 4	FL< mean FL	FB< mean FB	BG< mean BG
Subdivisions 5	FL> mean FL	FB> mean FB	BG> mean BG
Subdivisions 6	FL> mean FL	FB> mean FB	BG< mean BG
Subdivisions 7	FL> mean FL	FB< mean FB	BG> mean BG
Subdivisions 8	FL> mean FL	FB< mean FB	BG< mean BG

C. TEST OF SEPARATION OF SUB-DIVISION IN ELLIPSOID

Without loss of generality, the sub-divisions G1 (with 43 members), G2 (19 members), G3 (18 members) and G4 (46 members) in ellipsoid 1 (Figure 5) are investigated. The test $H_0 : \mu_1 = \mu_2$ versus the alternative $H_a : \mu_1 \neq \mu_2$ (see section II (D)) will be considered. Table 8 shows the result for the T-test carried out from JMP software [43]. All four sub-divisions show significant separation. However, these tests are not carried out for the sub-divisions shown in Figure 6 since there are sub-divisions with no members or with very few members. This latter remark suggests that other test of separation is needed, but not done in this study.

IV. DISCUSSION

Variation of shoe size in the shoe-sizing problem is frequently defined as the existence of different subdivisions of shoe size measurements. A major effort towards the development of a shoe sizing system is to allocate individuals into different size subdivisions. Table 9 and Table 10 show plots that illustrate fundamental issues often ignored when searching for distinct size subdivisions. When 2-D plots are studied, the definition of subdivisions (groups of a point) depends on the selected pair of shoe size measurements. When BG and FB are used,

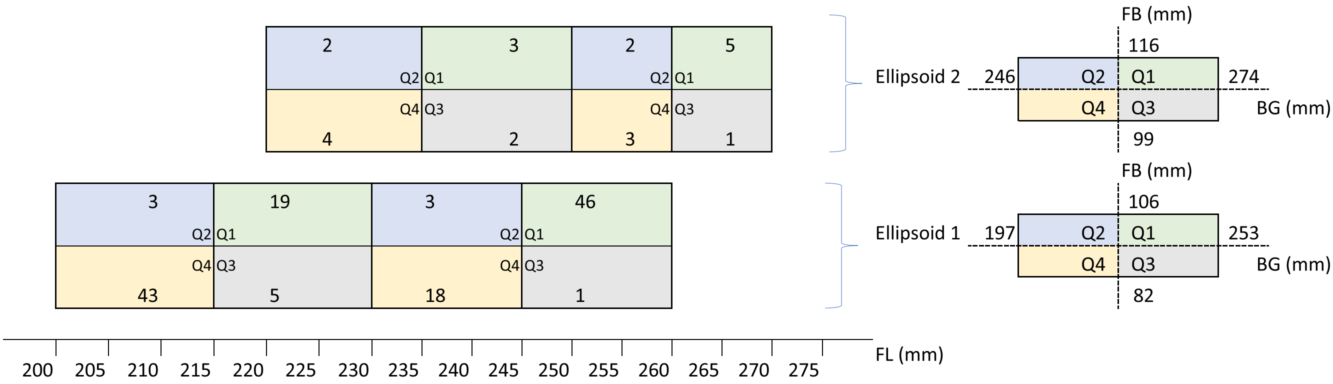


FIGURE 5. The sub-division of the two probability ellipsoids.

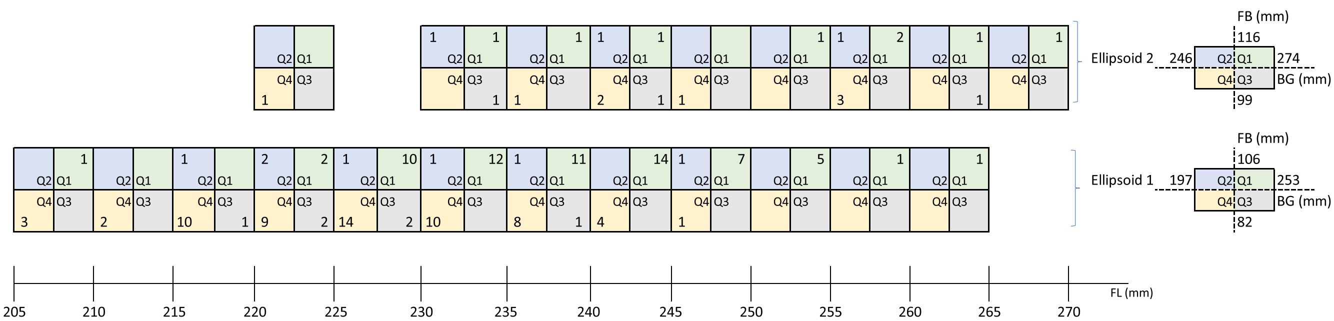


FIGURE 6. A novel shoe sizing system developed from the mixture probability model.

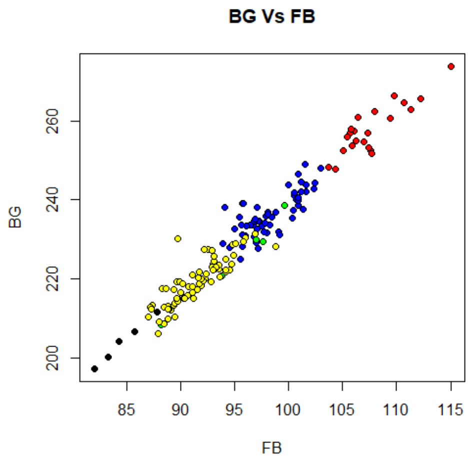
TABLE 8. Hypothesis testing for separation of selected sub-divisions in ellipsoid 1.

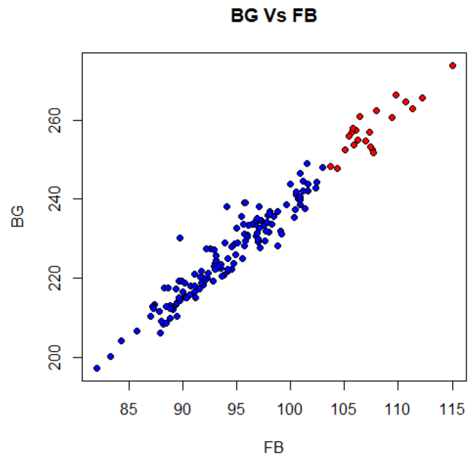
Sub-divisions compared	Foot Measurements: FL			Foot Measurements: FB			Foot Measurements: BG		
	T-test statistic		Result	T-test statistic		Result	T-test statistic		Result
	For equal variance	For unequal variance		For equal variance	For unequal variance		For equal variance	For unequal variance	
(G1,G2)	0.0249	0.0207	Not equal mean	<0.0001	<0.0001	Not equal mean	<0.0001	<0.0001	Not equal mean
(G1,G3)	<0.0001	<0.0001	Not equal mean	0.0163	0.0104	Not equal mean	0.0042	0.0009	Not equal mean
(G1,G4)	<0.0001	<0.0001	Not equal mean	<0.0001	<0.0001	Not equal mean	<0.0001	<0.0001	Not equal mean
(G2,G3)	<0.0001	<0.0001	Not equal mean	<0.0001	<0.0001	Not equal mean	<0.0001	<0.0001	Not equal mean
(G2,G4)	<0.0001	<0.0001	Not equal mean	0.0368	0.0342	Not equal mean	0.0272	0.0324	Not equal mean
(G3,G4)	0.0147	0.0025	Not equal mean	<0.0001	<0.0001	Not equal mean	<0.0001	<0.0001	Not equal mean

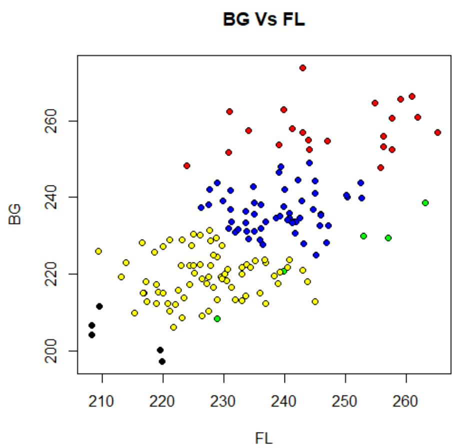
subdivisions tend to be linear whilst the other pairwise combination yields non-linear shoe size subdivisions. It is also clear the clustering result show different subdivisions compare to those obtained from the mixture model. Similar remarks may be made from the plots in Table 10, in particular, the three-dimensional subdivisions are more distinct than their two-dimensional counterparts. Once these issues are attended to, then the search for shoe size subdivisions are carried out.

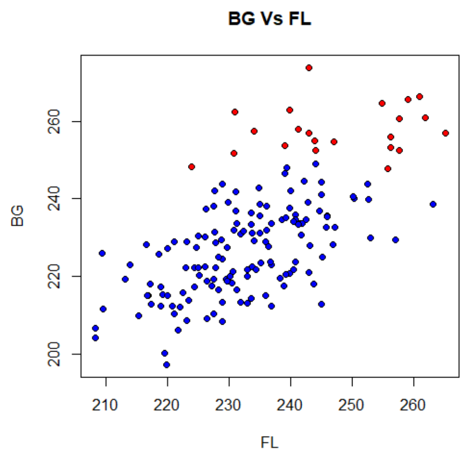
In the search for shoe size subdivisions, clustering methods are often applied. Two clusters represent two groups of individuals with different shoe-sizes. Further in each cluster, subdivisions are carried out, for example, foot length may be divided into six intervals, which in turn is an indicator of foot length variation [20]. Most of the methods do not allow a formal procedure in testing a significant difference between sub-divisions. In particular, are two

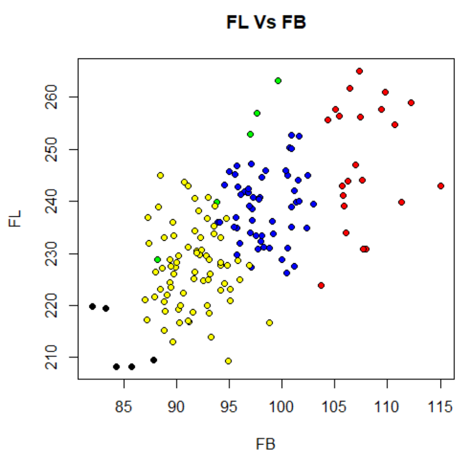
TABLE 9. **Two dimensional plot of 160 individuals from clustering method and mixture model.**

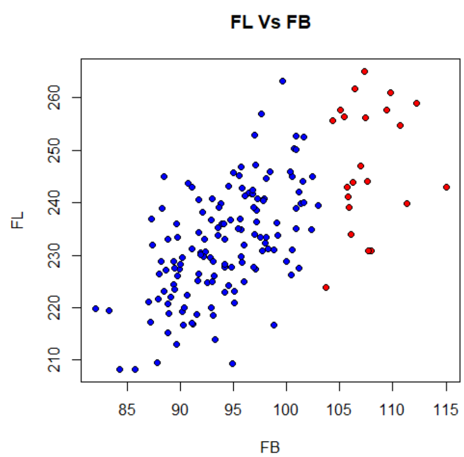
<table>
<tr><td>2-D plot type</td><td>Grouping from Clustering method</td><td>Grouping from Mixture model</td></tr>
<tr><td>2-D plot BG Vs FB</td><td>


</td><td>


</td></tr>
<tr><td>2-D plot BG Vs FL</td><td>


</td><td>


</td></tr>
<tr><td>2-D plot FL Vs FB</td><td>


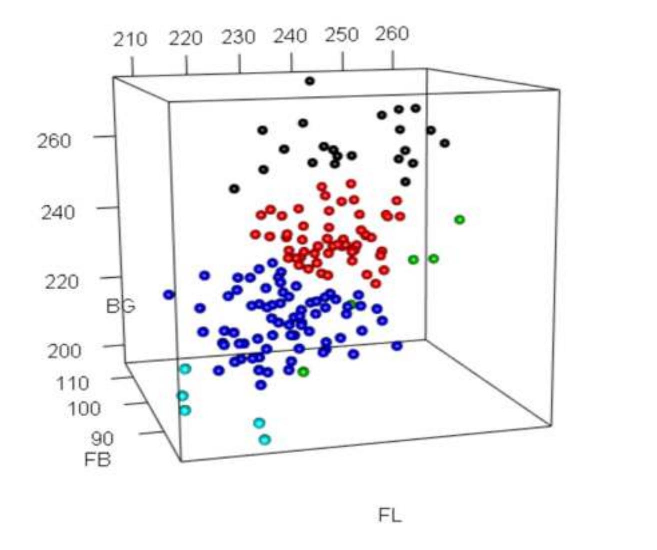
</td><td>


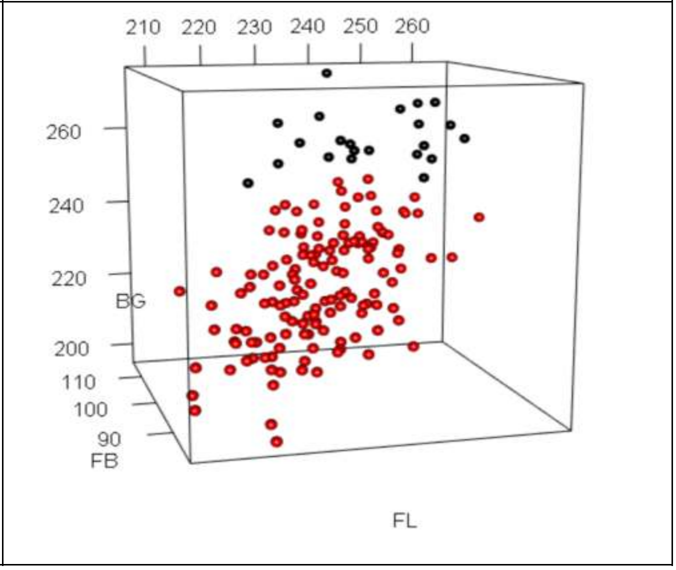
</td></tr>
</table>

* Red represent individual from cluster A, blue for cluster B, green for cluster C, yellow for cluster D and black for cluster E for shoe sizing system from clustering method. For mixture model, red represent individual from first ellipsoid and blue for second ellipsoid.

TABLE 10. Three dimensional plot of 160 individuals from clustering method and mixture model.

Grouping from Clustering method	Grouping from Mixture model





* Red represent individual from cluster A, blue for cluster B, green for cluster C, yellow for cluster D and black for cluster E for shoe sizing system from clustering method. For mixture model, red represent individual from first ellipsoid and blue for second ellipsoid.

TABLE 11. Percentage of coverage and aggregate loss for shoe sizing developed from the Agglomerative clustering method. Total coverage is 72.4%.

Cluster	A	A	B	B	B	B	B	B	B	B	B
Size Label	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
Assigned FL, mm	245	260	230	235	235	240	240	245	245	250	255
Assigned FB, mm	108	108	104	98	104	98	104	98	104	98	104
Assigned BG, mm	258	258	250	236	250	236	250	236	250	236	250
No. of subject	3	4	3	6	3	6	4	9	3	4	4
Coverage,%	2.0	2.6	2.0	3.9	2.0	3.9	2.6	5.9	2.0	2.6	2.6
Aggregate loss, mm	3.95	7.13	10.01	5.78	10.67	5.72	8.15	5.30	6.48	8.29	10.17
Benchmark value, mm	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99

Cluster	D	D	D	D	D	D	D	D	D
Size Label	A12	A13	A14	A15	A16	A17	A18	A19	A20
Assigned FL, mm	220	225	225	230	230	235	235	240	240
Assigned FB, mm	92	92	99	92	99	92	99	92	99
Assigned BG, mm	219	219	232	219	232	219	232	219	232
No of subject	8	9	6	8	11	4	6	3	6
Coverage, %	5.3	5.9	3.9	5.3	7.2	2.6	3.9	2.0	3.9
Aggregate loss, mm	5.92	8.10	8.51	6.45	8.75	6.31	12.86	6.23	12.01
Benchmark value, mm	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99

adjacent intervals (sub-divisions) really different? Some form of formal statistical hypothesis testing should be carried out.

This problem is magnified when there is no obvious way of selecting the best clustering method, for example, average linkage or single linkage. In this study, the CPCC index

TABLE 12. Percentage of coverage and aggregate loss for a novel shoe sizing developed from mixture probability model. Total coverage is 71.25%.

Ellipsoid	1	1	1	1	1	1	1	1	1	1	1	1
Size Label	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Assigned FL, mm	220	225	230	230	235	235	240	240	245	245	250	255
Assigned FB, mm	94	94	94	106	94	106	94	106	94	106	106	106
Assigned BG, mm	225	225	225	253	225	253	225	253	225	253	253	253
No of subject	10	9	14	10	10	12	8	11	4	14	7	5
Coverage, %	6.25	5.63	8.75	6.25	6.25	7.50	5.00	6.88	2.50	8.75	4.38	3.13
Aggregate loss, mm	15.10	14.00	9.34	20.72	8.28	20.55	7.47	18.10	5.96	18.88	19.45	15.80
Benchmark value, mm	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99	43.99

TABLE 13. Comparison of validation result between shoe sizing from mixture probablity model and [20].

	The proposed mixture model with 3-foot measurements	(Lee et al., 2012)	The proposed mixture model with 5-foot measurements
Number of subjects	160	1835	160
Types of foot measurement	Foot length, foot breadth and ball girth	Foot length, foot breadth and navicular height	Foot length, foot breadth, ball girth, Instep length, Fibulare Instep length
Method used	Mixture probability	Hierarchical and nonhierarchical clustering method	Mixture probability
Percentage of coverage	71.25%	72.96%	69.00%
Number of shoe size proposed	12	6	12
Aggregate loss	All less than benchmark value	All less than benchmark value	All less than benchmark value

gives a rough indication of obtaining appropriate clusters. Once clusters are obtained it is common practice to develop a shoe sizing system. Information of correlation between shoe size measurements is combined with the clusters obtained to yield an approximate sizing system as illustrated in Figure 4 where for example (in cluster A), there is one individual with length FL [220mm,225mm), FB [101mm,108mm) and BG [242mm,258mm). In particular, our approach seeks (BG, FB, FL) intervals by looking for (BG, FB) values for fixed FL values. A total of 96 shoe size subdivisions were obtained with several subdivisions having no individual. When using five-foot measurements, we obtained a total of 224 shoe size subdivisions with a lot of subdivisions having no individual.

Unfortunately, it is not easy to formally test or verify distinct shoe size subdivisions. As such, the information obtained from clusters, for example, cluster mean and cluster variance, become the input for estimating the parameters of the mixture normal distribution which in turn is used to find distinct shoe size subdivisions. The AIC index, BIC index, and p-value show that a two-component normal mixture probability distribution is the best to model the data. This mixture model in itself provides information about the variation of shoe size ((6) from section III(B)) where 86.5% have a shorter length (average 232.34 mm) compare to 13.5% with longer feet (average 246.53 mm). However, their individual variances suggest considerable overlap. Table 6 further illustrates the different situation where inequality of mean and (or) variance complicate the description of shoe size variation. This remark shows that using length alone may not be sufficient to indicate shoe size variability. In other words, the number of shoe size measurements selected is crucial. Henceforth, development of a shoe sizing system which directly makes use of the two-component multivariate normal mixture probability distribution (in three measurements) should take account of the different mean-variance combination. This study divides each component of the multivariate normal mixture probability distribution as defined in Table 7 with the result illustrated in Figure 5 where for example 43 individuals have length [200-215 mm) with FB [82-94mm) and BG [197-225 mm). Further Table 10 illustrates that we can investigate whether the subdivision is distinct (depending on the number of subdivisions membership).

A prototype sizing system is proposed and illustrated in Figure 6. Table 11 and Table 12 summarize the validation for shoe sizing systems developed from the Agglomerative cluster method and (or) multivariate normal mixture probability model from the data used. The aggregate loss and coverage percentage for all sizes using the Agglomerative cluster method is comparable to the method using multivariate normal mixture probability. However, the latter method uses much fewer subdivisions of FL, FB ad BG. In addition, this proposed system is considered superior to other systems, the blind shoe sizing system (Figure 2) and the system developed solely from the clustering method in Figure 4; from two aspects. Firstly, the test of group separation has been carried out, and secondly, fewer ''empty'' or zero membership subdivisions obtained. The term ''prototype'' is emphasized in the sense that a more ''powerful'' test of group separation can be developed but not applied here. A sample size of 160 was considered sufficient to develop a prototype shoe sizing system since the statistical test of group separation used needed only moderate sample sizes.

Table 13 compares the performance of the proposed mixture probability model and the method in [20]. From the result, the coverage percentage and aggregate loss are comparable but a very much smaller sample size is needed for our proposed method. The comparison between using three-foot measurements (FL, FB, BG) and five-foot measurements (FL, FB, BG, IL, FIL) showed very similar performances.

V. CONCLUSION

The outcome of this work is a novel method in addressing the issue of Malaysian women's shoe size variation. The methods proposed are a combination of an improved clustering method and hypothesis testing of mixture models. A prototype shoe sizing system was then proposed (illustrated by Figure 6) which has a percentage of coverage and aggregate loss similar to a method using three-foot measurements with a much larger sample size. A comparison of the use of five-foot measurements and three-foot measurements showed that the latter was sufficient for this study. The proposed prototype shoe sizing system provides information for the planning, design, and manufacturing of Malaysian women's footwear with implications for better fitness and comfort.

ACKNOWLEDGMENT

The authors would like to thank guidance and constructive suggestions from Makiko Kouchi, Digital Human Research Group, National Institute of Advanced Industrial Science and Technology, Japan.

REFERENCES

[1] E. Y. L. Au and R. S. Goonetilleke, ''A qualitative study on the comfort and fit of ladies' dress shoes,'' *Appl. Ergonom.*, vol. 38, no. 6, pp. 687–696, 2007.
[2] J. C. Menant, J. R. Steele, H. B. Menz, B. J. Munro, and S. R. Lord, ''Effects of footwear features on balance and stepping in older people,'' *Gerontology*, vol. 54, no. 1, pp. 18–23, 2008.
[3] D. Bourgit, G. Y. Millet, and J. Fuchslocher, ''Influence of shoes increasing dorsiflexion and decreasing metatarsus flexion on lower limb muscular activity during fitness exercises, walking, and running,'' *J. Strength Conditioning Res.*, vol. 22, no. 3, pp. 966–973, 2008.
[4] C.-L. Lin, M.-J. J. Wang, and C. G. Drury, ''Biomechanical, physiological and psychophysical evaluations of clean room boots,'' *Ergonomics*, vol. 50, no. 4, pp. 481–496, 2007.
[5] M.-C. Chiu and M.-J. J. Wang, ''Professional footwear evaluation for clinical nurses,'' *Appl. Ergonom.*, vol. 38, no. 2, pp. 133–141, 2007.
[6] M. Pantazi, A. M. Vasilescu, A. Mihai, and D. Gurau, ''Statistical-mathematical processing of anthropometric foot parameters and establishing simple and multiple correlations. Part 1: Statistical analysis of foot size parameters,'' *Revista de Piel rie Încălţ minte*, vol. 17, no. 4, pp. 199–208, 2017. doi: 10.24264/lfj.17.4.3.
[7] M. Hill, R. Naemi, H. Branthwaite, and N. Chockalingam, ''The relationship between arch height and foot length: Implications for size grading,'' *Appl. Ergonom.*, vol. 59, pp. 243–250, Mar. 2017.
[8] A. Luximon, Y. Zhang, Y. Luximon, and M. Xiao, ''Sizing and grading for wearable products,'' *Comput.-Aided Des.*, vol. 44, no. 1, pp. 77–84, 2012.
[9] A. I. Shugaba, F. Shinku, I. M. Gambo, M. B. Mohammed, C. B. Uzokwe, E. R. Damilola, A. M. Rabiu, and Y. M. Usman, ''Relationship between foot length, foot breadth, ball girth, height and weight of school children aged 3-5 years old,'' *J. Biol. Chem. Res.*, vol. 30, no. 1, pp. 107–114, 2013.
[10] A. S. Rodrigo, R. S. Goonetilleke, and C. P. Witana, ''Model based foot shape classification using 2D foot outlines,'' *Comput.-Aided Des.*, vol. 44, no. 1, pp. 48–55, 2012.
[11] Y.-C. Lee, G. Lin, and M.-J. J. Wang, ''Comparing 3D foot scanning with conventional measurement methods,'' *J. Foot Ankle Res.*, vol. 7, no. 1, p. 44, 2014.
[12] C. P. Witana, S. Xiong, J. Zhao, and R. S. Goonetilleke, ''Foot measurements from three-dimensional scans: A comparison and evaluation of different methods,'' *Int. J. Ind. Ergonom.*, vol. 36, no. 9, pp. 789–807, 2006.
[13] A. S. J. Roodbandi, H. Naderi, N. Hashenmi-Nejad, A. Choobineh, M. R. Baneshi, and V. Feyzi, ''Technical report on the modification of 3-dimensional non-contact human body laser scanner for the measurement of anthropometric dimensions: Verification of its accuracy and precision,'' *J. Lasers Med. Sci.*, vol. 8, no. 1, p. 22, 2017.
[14] M. Mauch, K. J. Mickle, B. J. Munro, A. M. Dowling, S. Grau, and J. R. Steele, ''Do the feet of German and Australian children differ in structure? Implications for children's shoe design,'' *Ergonomics*, vol. 51, no. 4, pp. 527–539, 2008.
[15] M. Kouchi, ''Foot dimensions and foot shape: Differences due to growth, generation and ethnic origin,'' *Anthropol. Sci.*, vol. 106, pp. 161–188, Jan. 1998. doi: 10.1537/ase.106.supplement_161.
[16] S. B. Bari, M. Othman, and S. N. Mohd, ''Foot anthropometry for shoe design among preschool children in Malaysia,'' *Pertanika J. Social Sci. Humanities*, vol. 18, no. 1, pp. 69–79, 2010.
[17] D. Chaiwanichsiri, N. Tantisiriwat, and S. Janchai, ''Proper shoe sizes for Thai elderly,'' *Foot*, vol. 18, no. 4, pp. 186–191, 2008.
[18] P. W. Kong, C. Y. Lim, R. Ding, and T. Sterzing, ''Subjective evaluation of running footwear depends on country and assessment method: A bi-national study,'' *Ergonomics*, vol. 58, no. 9, pp. 1589–1604, 2015.
[19] S. M. Shariff, Y. B. Bong, A. A. Shariff, M. Kouchi, K. Kimura, A. F. Merican, O. B. M. Rijal, and N. M. Noo, ''Development of a foot sizing system for Malaysian women,'' in *Proc. 5th Int. Conf. 3D Body Scanning Technol.*, 2014, pp. 93–98.
[20] *Shoe Sizes—Mondopoint System of Sizing and Marking*, Standard ISO 9407:1991(en), 1991.
[21] F.-T. Cheng and D.-B. Perng, ''A systematic approach for developing a foot size information system for shoe last design,'' *Int. J. Ind. Ergonom.*, vol. 25, no. 2, pp. 171–185, 2000.
[22] B. Nácher, S. Alemany, J. C. González, E. Alcántara, J. García-Hernández, S. Heras, and A. Juan, ''A footwear fit classification model based on anthropometric data,'' presented at the 8th Annu. Digit. Hum. Modelling Design Eng. Symp., Lyon, France, Jul. 2006.
[23] M. Mauch, S. Grau, I. M. C. Krauss, and T. Horstmann, ''A new approach to children's footwear based on foot type classification,'' *Ergonomics*, vol. 52, no. 8, pp. 999–1008, 2009.
[24] Y.-C. Lee, W.-Y. Chao, and M.-J. Wang, ''Foot shape classification using 3D scanning data,'' in *Proc. Southeast Asian, Netw. Ergonom. Soc. Conf. (SEANES)*, Jul. 2012, pp. 1–6.
[25] M. F. M. Hamzah, O. M. Rijal, and N. M. Noor, ''Validation of clusters for M-foot shape measurement: Malaysian women foot shape,'' presented at the Region Conf., IEEE TENCON, 2017, pp. 906–911.
[26] I. P. D. Manna, S. Ghosh, K. S. Kumar, and P. Dhara, ''A comparative study of foot dimension between adult male and female and evaluation of foot hazards due to using of footwear,'' *J. Physiol. Anthropol. Appl. Hum. Sci.*, vol. 20, no. 4, pp. 241–246, 2001.

[27] K. Baba, "Foot measurement for shoe construction with reference to the relationship be tween foot length, foot breadth, and ball girth," *J. Hum. Ergol.*, vol. 3, no. 2, pp. 149–156, 1974.

[28] A. Bataller, E. Alcántara, J. González, A. Garcia, and S. Alemany, "Morphological grouping of Spanish feet using clustering techniques," presented at the 5th Symp. Footwear Biomech., 2001.

[29] Y.-C. Lee and M.-J. Wang, "Taiwanese adult foot shape classification using 3D scanning data," *Ergonomics*, vol. 58, no. 3, pp. 513–523, 2015.

[30] B. S. Everitt, *Finite Mixture Distributions*. Hoboken, NJ, USA: Wiley, 1981.

[31] B. S. Everitt, *Cluster Analysis*, 3rd ed. New York, NY, USA: Edward Arnold and Halsted Press, 1993.

[32] S. M. Vahidipour, A. Mirzaei, and M. Rahmati, "Comparing weighted combination of hierarchical clustering based on Cophenetic measure," *Intell. Data Anal.*, vol. 18, no. 4, pp. 547–559, 2014.

[33] C. Fraley and A. E. Raftery, "Model-based clustering, discriminant analysis, and density estimation," *J. Amer. Statist. Assoc.*, vol. 97, no. 458, pp. 611–631, 2002.

[34] D. Reynolds, "Gaussian mixture models," in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds. Boston, MA, USA: Springer, 2015.

[35] D. M. Titterington, A. F. M. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distributions*. Hoboken, NJ, USA: Wiley, 1985.

[36] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. B, Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.

[37] M. Watanabe and K. Yamaguchi, *The EM Algorithm and Related Statistical Models*. Boca Raton, FL, USA: CRC Press, 2003.

[38] G. McLachlan and D. Peel, *Finite Mixture Models*. Hoboken, NJ, USA: Wiley, 2004.

[39] G. W. Snedecor and W. G. Cochran, *Statistical Methods*. Ames: Iowa State College Press, 1980.

[40] N. Zakaria, "Sizing system for functional clothing—Uniforms for school children," *Indian J. Fibre Textile Res.*, vol. 36, no. 4, pp. 348–357, 2011.

[41] O. Kwon, K. Jung, H. You, and H.-E. Kim, "Determination of key dimensions for a glove sizing system by analyzing the relationships between hand dimensions," *Appl. Ergonom.*, vol. 40, no. 4, pp. 762–766, 2009.

[42] D. Gupta and B. Gangadhar, "A statistical model for developing body size charts for garments," *Int. J. Clothing Sci. Technol.*, vol. 16, no. 5, pp. 458–469, 2004.

[43] J. Sall, A. Lehman, M. L. Stephens, and L. Creighton, *JMP Start Statistics: A Guide to Statistics and Data Analysis Using JMP*. SAS Institute, 2012.

MOHD FAIZAL MOHD HAMZAH received the B.Eng. degree (Hons.) in electrical engineering and the M.Sc. degree in statistics from the University of Malaya, Malaysia, in 2007 and 2015, respectively. He is currently pursuing the Ph.D. degree with Universiti Teknologi Malaysia.

OMAR MOHD RIJAL received the B.Sc. degree in mathematics and operational research from Ulster New University, Ireland, and the Ph.D. degree in applied statistics from the University of Glasgow, U.K. He was an Associate Professor with the Institute of Mathematical Sciences, University of Malaya, Malaysia. He currently freelances as an Independent Researcher. His research interests include applied statistics, and signal and image analysis for medical and industrial applications.

KOZO KIMURA is currently the President of I-Ware Laboratory Company Ltd. He holds several patents on foot shape and footprint information distribution systems. His research interest includes digital foot measurement systems.

NORLIZA MOHD NOOR received the B.Sc. degree from Texas Tech University, Lubbock, TX, USA, and the master's degree (by research) and the Ph.D. degree from Universiti Teknologi Malaysia (UTM), all in electrical engineering, where she is currently an Associate Professor with the Razak Faculty of Technology and Informatics. Her research interests include machine learning, and signal and image analysis for medical and industrial applications.