Sudan Journal of Science and Technology
Journal homepage:
http://jst.sustech.edu/

Using Data Mining Techniques to Establish Standard Sizing System for Sudanese Army Officers Poshirt

Elawad F. Elfaki Elnour

Department of Textile Engineering, College of Engineering Technology of Industries, Sudan University of Science and Technology (SUST)
E-mail: abuzaid207@gmail.com and alawad_alfaki@sust.edu

ARTICLE INFO

ARTICLE HISTORY

Received: 1/4/2018
Accepted:1 /6/2018
Available online: June 2018

KEYWORDS:

Data mining, anthropometric variables,
garments, industry,
national, sizing,
chart,
clustering

ABSTRACT

The purpose of this research was to use data mining techniques in establishing standard sizing system for Sudanese army officers garment especially (Poshirt) based on an anthropometric body measurements variables SUD and to compare it with Sur Military Clothing Factory in Sudan (SUR),USA, and EUR national sizing charts. The data set collected for 841 army officers and for each individual 13 anthropometric variables was involved. In this work data mining methods (WEKA and SPSS) were used for clustering and establishing sizing system. The K-means algorithm was implemented to determine the final cluster classification. Cluster analysis using chest and waist as a control anthropometric variables revealed a new established sizing system SUD which consists of eight distinct clusters namely; XS, S, M, L, XL, XXL, XXXL and XXXXL. Size codes and upper and lower size limits are generated. The new proposed sizing system SUD profile was compared with profile of SUR, USA, and EUR national standards sizing charts. The results revealed that the proposed sizing system SUD followed approximately the same profile as the three others sizing systems but there was slight differences between the new established systems SUD and the USA and EUR national standard size charts

INTRODUCTION:

To understand the term 'data mining' it is useful to look at the literal translation of the word: to mine means to extract. The verb usually refers to mining operations that extract from the earth its hidden, precious recourses. The link of this word with data suggests an – in depth search to find additional information which

previously were un-noticed in the mass of data available. From the viewpoint of scientific research, data mining is relatively new discipline that has developed mainly from studies carried out in other disciplines such as computing, marketing and statistics. In order to automatically generate useful knowledge from a variety of data, and presented it in human oriented forms, a powerful tools is strongly needed. Researchers have been exploring ideas and methods in different areas as efforts to satisfy this need. Such areas include; data mining, text mining, machine learning, statistical data analysis, data visualization, and pattern recognition. Data mining techniques have been used in many fields, but little research has been done on its application to sizing systems for the manufacturing of garments. (Norsaadah et al, 2008) established a sizing system for the manufacture of garment using decision tree based data mining to determine the pants size of army soldier's uniform. Hsu (2008) conducted an empirical study in apparel industry in order to support manufacturing decision for production management as well as marketing with various customers' needs. More recently (Bagherzadeh et al, 2010) introduced a study for developing sizing systems by data mining techniques using anthropometric data. A three – stage data mining procedure were employed to develop sizing system for lower body figure type of Iranian male. (Hai et al, 2008) studied the application of data mining techniques for developing a sizing system for army soldiers' uniform in Taiwan. In Malaysia; (Norsaadah et al, 2008) used data mining techniques to explore anthropometric data for the development of sizing system using the whole data was analyzed using descriptive analysis of average, mean and standard deviation. The data obtained was further explored using the factor analysis method.

Materials and Methods:

In this work an anthropometric data of 841 army officers with age ranging from 18 to 60 years were collected from Sur Military Clothing Factory in Sudan. These records have existing measurements authorized by an anthropometric expert using a measuring tape. The measurements for the anthropometric variables were in inches with decimals. Therefore, they were converted into integer centimeters in order to ease the comparison with the commonly used national garment sizing standard units. Further, they were preprocessed i.e. they were examined and purified to omit the outliers in order to increase the efficiency and ensure the accuracy of its analysis during the processing and transformation. A total of 813 numbers of instances were processed, with 13 different attributes. The attributes for the jacket were, cuff, sleeve, shoulder, length, waist, chest and collar. For the trouser the attributes were length1, foot, knee, thigh, hip, and waist1. For the trouser, the term length1 and waist1 were used in order to distinct them from those of the Jacket. The measurements of anthropometric data followed the ISO 85591/1989 body measurement standard.

The chosen methods for this work were WEKA 3.6.9 and SPSS version 18.0.Because WEKA does not work with numbers, these (813) instances for (jacket + trouser) were categorized in numerical groups. The distribution of the proposed groups was based on waist and length attributes. The database was converted to (size.csv) format to be processed in WEKA. The text file

describes a list of instances that sharing a set of attributes. After processing the (size.csv) file in WEKA a list of all attributes, statistics, and other parameters can be utilized as shown in figure: 1.

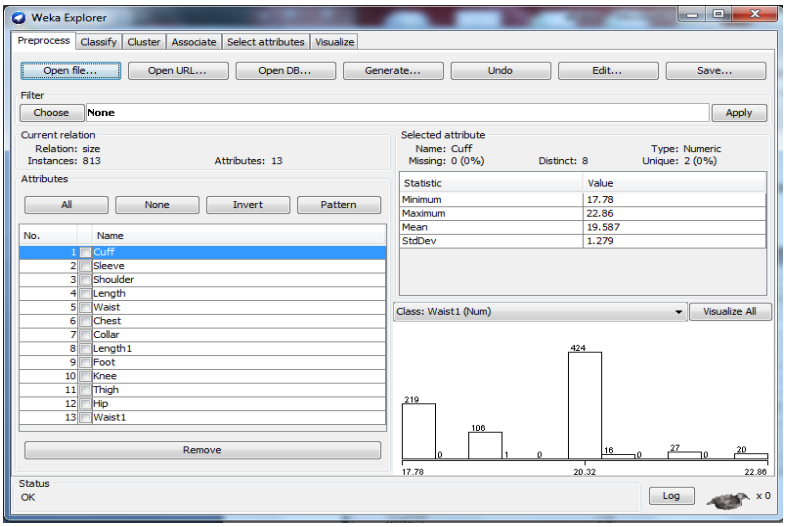


Figure: 1 Processed (size.csv) file in WEKA

Results and Discussion:

Clusters 7, 8 and 9 include the attributes from 0-6, 0-7, and 0-8 respectively.

In cluster 7 the percentage covered by the proposed sizing systems was 97%. However, the size figures 4XL and 5XL were not selected by the WEKA. This may be due to the fewer number of persons in these size figures. On the other hand, in cluster 8 the percentage covered by the proposed sizing system was 96.7% and there was only one size figure (5XL) with no classes that was not represented. This may be due to the fewer number of persons in this size figure. In cluster 9 the percentage covered by the proposed sizing system was 76.7%, but there were four sizing figures XL, 3XL, 4XL and 5XL were not represented. This may be due to the fewer number of persons in these size figures.

As can be seen from the results, in the three clusters (SPSS, 1999, ISO 8559, 1989 and Beazley, 1997) the size 5XL was omitted in all clusters. This may be attributed to the nature of the Sudanese male body shape. Therefore cluster 8 seems to be the best sizing system that represents the data collected from Sur factory. This is because it covered nearly 96.7% of the data and it includes 8 figure types.

After the eleven figure sizes were classified by the WEKA, the new established size system of the eight figures size were determined following step by step SPSS (1999) and the results are given in Table1.

Table 1: The Proposed New Established Size System

	XS M-2STD	S M-1STD	M Mean	L Mean+1 STD	XL M+2STD	XXL M+3STD	XXXL M+4STD	XXXXL M+5STD
Cuff	17	18.3	19.6	20.9	22.2	23.5	24.8	26.1
Sleeve	56.6	60.1	63.6	67.1	70.6	74.1	77.6	81.1
Shoulder	40.6	43.6	46.6	49.6	52.6	55.6	58.6	61.6
Length	68.6	73.4	78.2	83	87.8	92.6	97.4	102.2
Waist	73.5	85.1	96.7	108.3	119.9	131.5	143.1	154.7
Chest	92	100.5	109	117.5	126	134.5	143	151.5
Collar	36.1	39	41.9	44.8	47.7	50.6	53.5	64.9
Length 1	97.1	101.7	106.3	110.9	115.5	120.1	125	129.3
Foot	40	42.9	45.8	48.7	51.6	54.5	57.4	60.3
Knee	46.1	50	53.9	57.8	61.7	65.6	69.5	73.4
Thigh	63.8	69.5	75.2	80.9	86.6	92.2	97.8	103.7
Hip	93.2	99.2	105.2	111.2	117.2	123.2	129.2	135.2
Waist 1	72.6	83.3	94	104.7	115.4	126.1	136.8	147.5

n = 813 all values are in centimeters

Figure (2) shows the relevant scatter plots of chest on the X-axis verse the waist on the Y-axis and the interval was 4 cm to demonstrate the distribution of all figures type. It has been reported that, Cooklin, (1992) the chest is the most important anthropometric variable in establishing sizing systems in the field of garment making. The waist is also an important variable for sizing male garments in many countries. Figure (3) illustrates the differences between the eight types for the new established sizing systems. The figure was plotted as a line graph to yield a better insight into the differences between the new established sizing systems.

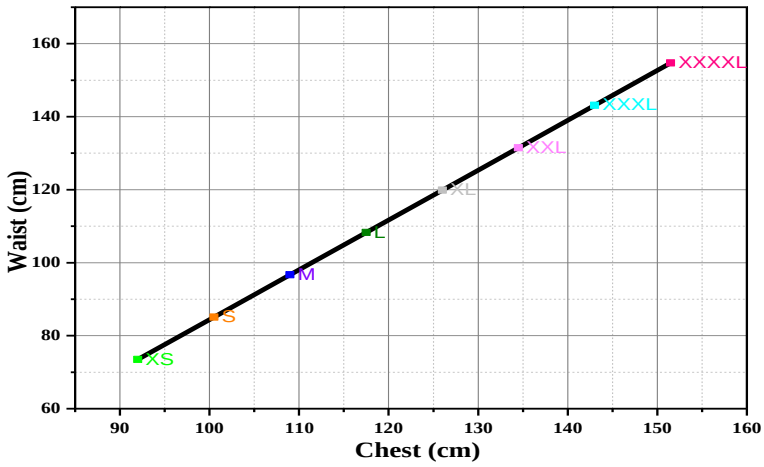


Figure 2: Scatter plot of chest verse waist for the proposed new established size system

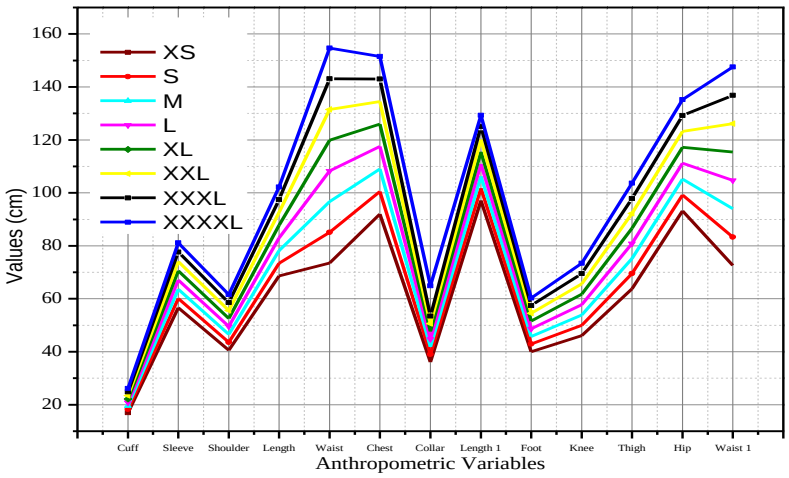


Figure 3: The distinct anthropometric variables between the proposed new established size systems

The new size code was termed (SUD) abbreviated for Sudan, and then the code was added to establish the 8 size figures for Sudanese army officers (poshirt uniform). Therefore, the new size codes are, SUD XS, SUD S, SUD M, SUD L, SUD XL, SUD XXL, SUD XXXL and SUD XXXXL.

In order to know what percentages of the populations are covered by each size we needed to establish lower and upper limits. The value obtained for each size code was used as a midpoint and the lower and upper limit were determined from it. The half value of the standard deviation of each body dimension added or subtracted to the midpoint value. A value of 0.01 was subtracted from the figure obtained below the midpoint to create limits between the lower value of the next size and the upper value of the previous one. To avoid overlapping of figures with the next size value of 0.01 was subtracted from the upper limit making it less than the next value. (Beazley, 1997) and other researchers used this procedure.

A Comparison between the New Established Sizing Systems SUD and SUR, USA, EUR Sizing Systems Specifications:

In order to compare the new established sizing systems with others sizing systems, five figure types; SX, S, M, L and XL, and five variables; chest, waist, hip, collar and sleeves were chosen from the eight figure types, see table (2). As can be seen from tables (3),(4),and(5) and figures (4), the new size system (SUD) followed the size chart for the three sizing systems (SUR, USA and EUR).

Table (2): Five Steps Size Ranges Obtained from the New Established Sizing Systems for the new established sizing systems

Size Body Dimensions (cm)	XS Mean -2STD	S Mean -1STD	M Mean	L Mean +1STD	XL Mean +2STD
Chest	92	100.5	109	117.5	126
Waist	73.5	85.1	96.7	108.3	119.9
Hip	93.2	99.2	105.2	111.2	117.2
Collar	36.1	39	41.9	44.8	47.7
Sleeve	56.6	60.1	63.6	67.1	70.6

□ n = 813 all values are in centimeters

Table (3): SUR Military Clothing Factory Poshirt (U4) Size Specification

Size	XS (cm)	S (cm)	M (cm)	L (cm)	XL (cm)
Chest	104.1	106.7	109.2	111.8	114.3
Waist	84	89	94	99	104
Hip	99.8	104.8	109.8	114.8	119.8
Collar	36.8	38.1	39.4	40.6	41.9
Sleeve	61	62.2	63.5	64.8	66

All values are in centimeters

Table (4): USA size charts specification for (men)

Size	XS (cm)	S (cm)	M (cm)	L (cm)	XL (cm)
Chest	96.5	101.6	106.7	111.8	116.8
Waist	81.3	86.4	91.4	94	106.7
Hip	88.9	104.1	109.4	114.3	119.4
Neckband (collar)	38.1	39.4	40.6	41.9	43.2
Sleeve	83.8	83.8	86.4	86.4	88.9

Table (5): EUR size charts specification for (men)

Size	XS (cm)	S (cm)	M (cm)	L (cm)	XL (cm)
Chest	92	97	107	112	117
Waist	81	87	92	99	107
Hip	99	104	109	114	119
Neckband Collar	38	39	40	42	43
Sleeve	84	84	87	87	89

A line graph was plotted and it shows significant differences in chest and waist among the eight figure types of the new proposed established sizing systems. This was appear very clearly when New SUD compared with national SUR, USA, and EUR size systems.

Therefore, the scatter plot of chest verse waist for the new size system followed the same trend that was plotted in figure (5). For the new established size figure (SUD) figure (6) plots a distribution the line for the (M and L) figure types. Also there was a deviation in the line for graph of chest on the X-axis verse waist on the Y-axis to demonstrate the distribution of all five figure types. Figure (3) compared with figures (6) for (SUR, USA and EUR) respectively. It can be seen that in figure (4) for both the new established sizing systems (SUD and SUR) the plot of chest verse waist fall in straight line. However, when compared (SUD,SUR,USA and EUR) as shown in figure (6) in the USA size systems there was a deviation in the (S and M) figure types in the case of the EUR size systems.

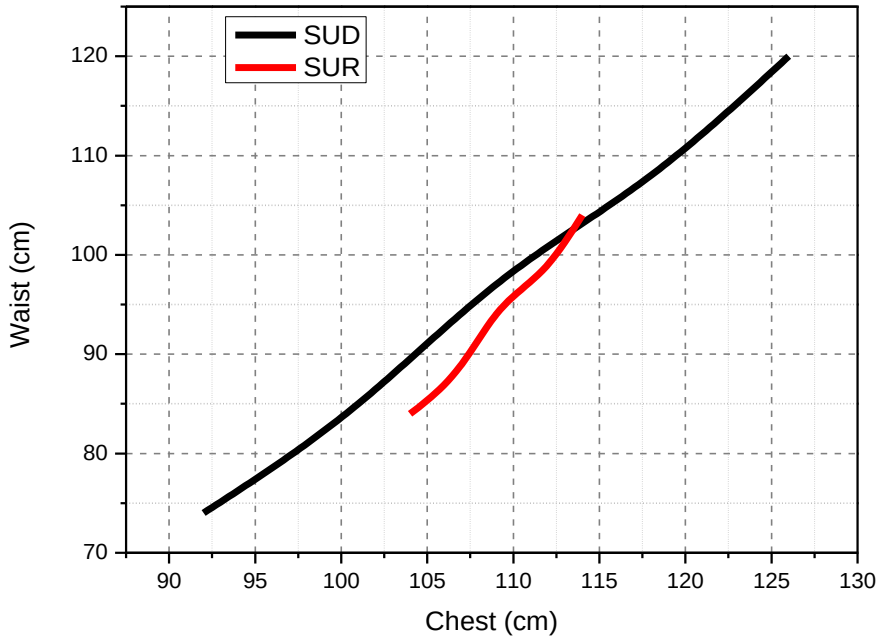


Figure 4: The New SUD and SUR scatter plot of chest verse waist for 5 figures Types

Figure 6 shows the differences between the five figure types for the new established size systems (SUD, SUR, USA and EUR) respectively. All the sizing systems followed the same profile, but there was slight differences between the new established systems SUD and the others sizing systems as shown in figure (6).

This may be due to:

a) The national standards (SUR + USA + EUR) deal essentially with the size designation of clothing are not directly concerned with the sizing systems as such. In other words, the establishment of a size designation system that indicates the body size of a person that a garment is intended to fit. Therefore, the takes the measurements. Also this difference may be due to the nature of size designation system is based on body and not garment measurements. Choice of garments measurements is normally left to the garment designer and manufacture, who are concerned with style, cut and other fashion elements, and he must make due allowance for garments normally worn beneath a specific garment.

b) The new established sizing systems deals with clothing designer and manufactures where the measurements are taken manually and for every individual. Therefore, the measurements may not be accurate. The accuracy of measurements depends on the skill of the person who

the Sudanese body size compared with other nationalities.

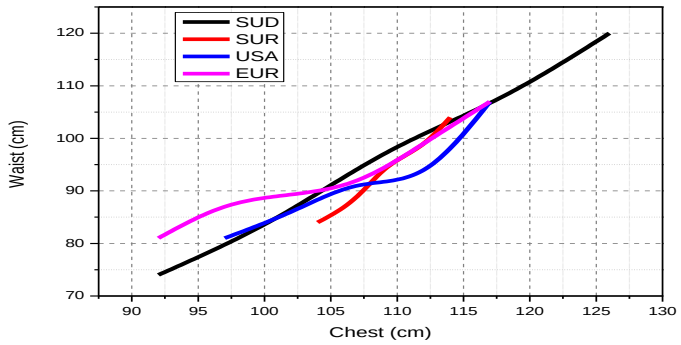
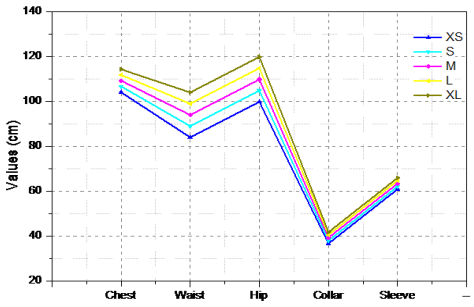
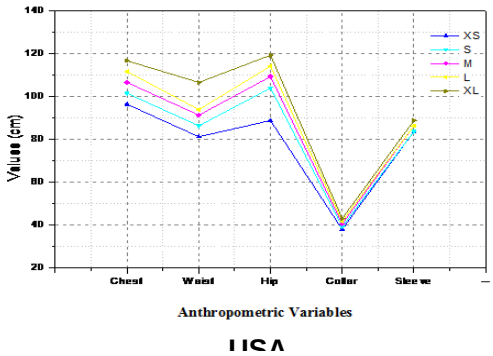


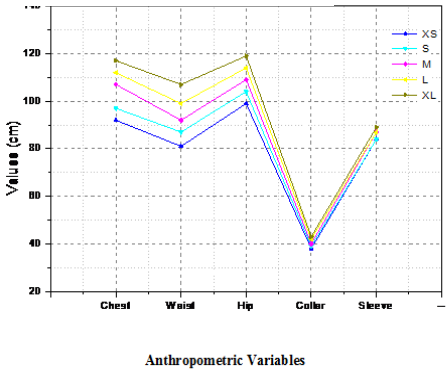
Figure (5): The sum of New SUD, SUR, EUR and USA scatter plot of chest verse waist for 5 figure types



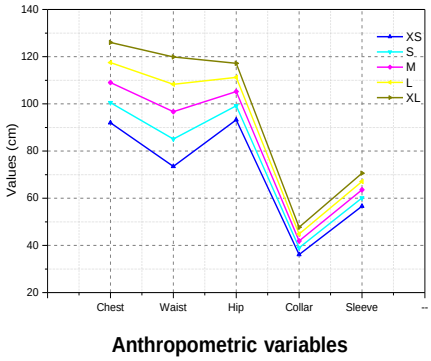
Anthropometric Variables
SUR 5 figure types and corresponding anthropometric variables



Anthropometric Variables
USA 5 figure types and corresponding



Anthropometric Variables
EUR 5 figure types and corresponding anthropometric variables



Anthropometric Variables The new SUD established sizing systems 5 figure types and corresponding anthropometric variables

Figure (6): The sum of new SUD, SUR, EUR and USA 5 figure types and corresponding anthropometric variables

Conclusion:

- This work contributes largely to knowledge of size chart by providing a detailed procedure involved in establishing standard sizing system based on anthropometric variables and will serve as the basis for other future research in garments industry in Sudan.
- Data mining methods (WEKA and SPSS) were applied and helped to establish sizing systems for Sudanese army officers' uniform (Poshirt).
- The new established sizing systems SUD as compared with SUR, USA and EUR standard national sizing charts were followed approximately the same profile of the three others sizing charts.
- The percentage of army officers who fall in a certain figure type and sizes can serve as a good reference to indicate the quantity of garments to be produced for specific market.
- Realistic plan for producing male army officer's uniforms can be established
- K-means algorithm is useful to determine the final cluster classification.

References

Beazley, A. (1997) 'Size and Fit: Procedures in undertaking a survey of body measurements'. *Journal for Fashion Marketing and Measurement.* 2(1)(1) 55 – 85.

Cooklin, G., (1992), Pattern Grading for Men's Clothes, Blackwell, Oxford.

Fayyad, U. M., Piatetsky – Shapiro, G., and Smyth, P., R. (Eds.),(1996), Advances in knowledge Discovery and Data Mining. Cambridge, M. A: AAA1 Press/MIT Press.

Hai- F., Lin., Chin – H. H., Mao – J. J. W., and Yu-C. Lin,(2008), An Application of Data Mining Technique in Developing System for Army Soldiers in Taiwan, *Issue 4, Volume 7, April 2008,*

Hsu, C. H, Lin. H F and Wang, M.I.J (2008) Developing Female Size Charts for facilitating garments Production by using data mining. *Journal of the Chinese institute of industrial engineers.Vol 24, Issue3, 245-251.*

ISO 8559, 1989. Garment Construction and Anthropometric Surveys – Body Dimensions, International Organization for Standardization.

Norsaadah Z., Jamil S.M., Nasir T., Yong Y.T. and Yap B. W. (2008). Using Data Mining Technique to Explore Anthropometric Data towards the Development of Sizing System.

R. Bagherzadeh, M. Latis and A.R Faramarzi, (2010), employing a Three- Stage Data Mining Procedure to Develop Sizing System, *Worled Applied Sciences Journal 8(8):923-929.*

SPSS, (1999), CRISP-DM 1.0 Step- By- Step Data Mining Guide.

www.simplicity.com,men's for men of average builds about 5 feet 10 inches without shoes: size U.S, sizes EUR.