Journal of ISOSS
2017 Vol. 3(2), 229-242

ANALYSIS OF THE REGIONAL POTENTIALS OF DEVELOPMENT IN THE WESTERN, SOUTHERN AND EASTERN STATES OF THE SUDAN THROUGH USING FACTOR AND CLUSTER ANALYSIS, 2011

Mozamel Elnair Somi Kakitla[§] and **Adam Hassan Takola Omer**
Department of Economics and Rural Development
Faculty of Community Development, Dalanj University, Sudan
Email: [§]mozamel-20@hotmail.com

ABSTRACT

The study aims at throwing lights on two methods of analysis that have multiple variables: Factor analysis through using principal components and cluster analysis through using hierarchical method. The study uses factor analysis to Study the available variables which affect the potentials of regional development in some States of the Sudan namely Western, Southern and Eastern States, attempting to reach the best classification of the factors that affect the potentials of the regional development. While, the study uses cluster analysis to group the States in development categories based on the variables of the study. The data collected in 2011 from the government institutions. The internal variables of the study were (34), such as: Demographic, natural, educational, health, agricultural, animal resources, services and general utilities. The findings of the first analysis are: Multi-potentials factor, which includes (human, animal resources, demographic, educational, general utilities factor), factor of livestock, Rainfall and general services, factor of health and police services, factor of general utilities, factor of civil defense and poultry, geographical factor and climate factor. While findings of the second analysis is that the states are grouped into four regions consequently: South Darfur stands first, North Kordofan comes the next and followed by South Kordofan and Gadarif, and lastly Kassala, West Darfur, Blue Nile, North Darfur and Red Sea. The study recommends that classification of regions using method of potentials development necessary to be applied in the different levels of the government institutions, i.e. Localities, Administrative Units and Towns.

KEYWORDS

Factor analysis, principal components, cluster analysis, hierarchical method, Sudan States.

1. INTRODUCTION

In this study the focus is on nine States of the Sudan namely; South Darfur State, State of North Kordofan, South Kordofan, West Darfur, Kassala, Gadarif, Blue Nile, North Darfur, and Red Sea State. The Subject of the analysis is the nine States of the Sudan, which characterized by large differences. This study intends to point to the use of factor and cluster analysis in exploring the characteristics of individual States and in grouping these States in to homogenous groups (Ivana, 2005). The main hypothesis of the

study is that, using the methods of factor and cluster analysis singles out those States that represent a more suitable subject matter of regional potentials of development (Ivan, 2005). Results lead to the identification of the main dimensions of variables of the states (factor analysis), and the grouping of the States of similar dimensions into clusters (cluster analysis) (Ivana, 2005). Disparities in regional development levels leads to many problems including, for example, urban economic, social and environmental problems,… etc. So, the highlighting of regional variation of the distribution of various phenomena is the most important proper planning requirements and has a direct impact in depicting the comprehensive development and regional development policies (Mozamel, 2015).

2. LITERATURE REVIEW

2.1 Factor Analysis

Factor analysis is that branch of multivariate analysis which deals with internal structure of matrices of covariance and correlation (Maxwell, 1963). Factor analysis starts with a set of observations obtained from a given sample by means of such a priori measures. It is a method of analyzing this set of observations from their inter-correlations to determine whether the variations represented can be accounted for adequately by a number of basic categories smaller than that with which the investigation was started. Thus, data obtained with a large number of a priori measures may be explained in terms of a smaller number of reference variables (Benjamin, 1954).

2.1.1 Factor Analysis Models

A set of linear random variables can be represented in form of matrix equations. If the variables in the x matrix are linear function of variables in the z x matrix, it can be represented by the following equation:

$$X = \mu + AZ, \tag{2.1}$$

where A is factor of Z. The above equation can be expressed as the following matrix:

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

Using matrix multiplication laws the equation for each variable in the x matrix can be given as follows:

$$\begin{aligned} X_1 &= \mu_1 + a_{11}z_1 + a_{12}z_2 + \ \ldots\ + a_{1n}z_n \\ X_2 &= \mu_2 + a_{21}z_1 + a_{22}z_2 + \ \ldots\ + a_{2n}z_n \\ &\ \ \vdots \\ X_p &= \mu_p + a_{p1}z_1 + a_{p2}z_2 + \ \ldots\ a_{pn}z_n \end{aligned} \tag{2.2}$$

The above equations commonly used to represent the linear statistical models such as factor analysis. The random vector $\underline{x}$ with p components has μ mean and Σ

covariance. Factor analysis assumes that $\underline{x}$ depends linearly on a number of random variables: $F_1, F_2, \ldots, F_m$ which are called common factors, in addition to p variance sources: $\varepsilon_1, \varepsilon_2, ..., \varepsilon_p$ which are called errors or specific factors, therefore the factors analysis model is:

$$\begin{aligned} X_1 - \mu_1 &= \lambda_{11}F_1 + \lambda_{12}F_2 + \ldots + \lambda_{1m}F_m + \varepsilon_1 \\ X_2 - \mu_2 &= \lambda_{21}F_1 + \lambda_{22}F_2 + \ldots + \lambda_{2m}F_m + \varepsilon_2 \\ &\vdots \\ X_p - \mu_p &= \lambda_{p1}F_1 + \lambda_{p2}F_2 + \ldots + \lambda_{pm}F_m + \varepsilon_p \end{aligned} \tag{2.3}$$

where

$\mu_i =$ The mean of the variable i .

$\varepsilon_i =$ The i specific factor.

$F_i =$ The i common factor.

The coefficients λ_{ij} are the loading of the variable i on the factor j , so the Λ matrix is the matrix of factor loading and can be written as follows:

$$\left(X - \mu\right) p_{X_1} = \Lambda_{p_X m} + F_{mXm} + Ep_{X_1} \tag{2.4}$$

In case of standardization of data measurement units and that is accustomed in factor analysis, consequently the general mean will be zero be zero and the equation number (2.4) will be as follows:

$$X = \Lambda F + E \tag{2.5}$$

2.1.2 Principal Component Method (PCM)

The Principal components (PC) method is the most widely used in the extraction of factors. The p of inter-related variables are transformed into unrelated small groups called principle components. The PCM is used to reveal and explain the dependability between the variables and to examine the relationships that may exist between them. Analysis of the principal components of m of the original variables produces m of new variables which called principal components:

$$PC_1,\ PC_2, \ldots,\ PC_m$$

Each principal component is a linear fitting of scores on the original variables, as follows:

$$\begin{aligned} PC_1 &= b_{11}x_1 + b_{12}x_2 + \ldots + b_{1m}x_m = xb_1 \\ PC_2 &= b_{21}x_1 + b_{22}x_2 + \ldots + b_{2m}x_m = xb_2 \\ &\vdots \\ PC_m &= b_{m1}x_1 + b_{m2}x_2 + \ldots + b_{mm}x_m = xb_m \end{aligned} \tag{2.6}$$

or it can be a matrix as follows:

$$PC = XB \tag{2.7}$$

Each column of B contains coefficients related to PC. The PC_1 coefficients are selected to have variance as greater as possible, also PC_2 coefficients are chosen to have a great variance on condition that the scores on PC_1 and PC_2 are unrelated. In general, the PC_i coefficients are selected with greater variance as possible not to be related to PC_1 through PC_{i-1}.

2.2.1 Cluster analysis (CA)

Cluster analysis is the art of finding groups in data. Such groups are called clusters, and to discover them is the aim of cluster analysis. Basically, one wants to form groups in such way that objects in the same group are similar to each other, whereas objects in different groups are as dissimilar as possible (Kaufma. Rousseeuw, 1967). Objects that differ in insignificant details are given the same name, can be treated the same, and can be expected to act the same. The word clustering is almost synonymous with classification (Hartigan, 1974). The aim of cluster analysis is to cluster a given set of data or objects into cluster (subsets, groups, classes). This clustering should have the following properties: Homogeneity within the cluster, i.e. data that belong to the same cluster should be as similar as possible and heterogeneity between clusters i.e. data that belong to different clusters should be as different as possible (Hoppne, Klawon, Krus, Runkler, 1999).

2.2.2 The Data Matrix

Suppose there are n objects to be clustered, which may be persons, flowers, words, countries, or whatever. Clustering algorithms typically operate on either of two input structures; The first represents the objects by means of p measurements or attributes, such as height, weight, sex, color,... etc. These measurements can be arranged in an n-by-p matrix, where the rows correspond to the objects and the columns to the attributes (objects-by-variables matrix), since the row and column entities are different. These measurements can be organized in an n-by-p matrix, where the rows correspond to the objects (or cases) and the columns correspond to the variables. When the fth measurements of the ith objects is denoted by x_{if} (where $i = 1,...,n$ and $f = 1,...,p$) this matrix looks like:

$$\begin{array}{c} \\ n \text{ objects} \end{array} \begin{array}{c} p \text{ variable} \\ \begin{pmatrix} x_{11} & \cdots & x_{1f} & \cdots & x_{1p} \\ \vdots & & \vdots & & \vdots \\ x_{i1} & \cdots & x_{if} & \cdots & x_{ip} \\ \vdots & & \vdots & & \vdots \\ x_{i1} & \cdots & x_{if} & \cdots & x_{ip} \end{pmatrix} \end{array}$$

2.2.3 Hierarchical Clustering Technique

It is one of the traditional methods of cluster analysis. In (HCT) the data is not divided into cluster in a single step, we need instead to the successive stages and it is divided into Agglomerative Methods (AM) and Divisive Methods (DM). The (AM) begins with a series of successive mergers of the n units, which turn into groups, while (DM) divides n n units sequentially to precise divisions. The most (AM) commonly used are Single Linkage Method, Complete Linkage, Group Average, Centroid Cluster, Median Cluster, Ward's and Lance and Williams Flexible Method. In the Group Average Method, which will be used in this paper, the distance between each two clusters is defined as unaverage distance between each pair of items.

2.2.4 Distance Measures

The distance Measures used when portioning the items into clusters. The idea is that the clusters contain items which separated by small distances compared to the distances between the clusters. These distances depend on the single dimension or the multiple dimensions. The data is usually transformed into distance matrix (DM) or similarity matrix (SM) for clustering n items by determination of the distance between clusters centers and observations a aggregations. The most important distance measures are Euclidean distance, Manhattan (City-Block), Mahalanobis, and Minkowski. Euclidean distance are the most commonly used, and calculated as follows:

$$dE\left(y_i, y_j\right) = \sqrt{\sum_{K=1}^{p} \left(y_{ik} - y_{jk}\right)^2}$$

for each $i, j \in n$

y_i and y_j represent the ith and jth item in p dimension.

3. METHOD AND MATERIALS

Factor analysis through component principle and cluster analysis through hierarchical method and agglomerative method through the average linkage (between groups) were used to study the variables. In this Study, the data were collected from government institutions in Western, Southern and Eastern States of the Sudan in 2011, 33 variables were studied, namely Demographic, natural, educational, health, agricultural, animal resources, services and general utilities.

3.1 States and Variables

Table 1
List of States and Variables with Respective Codes

Codes	States	Codes	Variables
X1	Kassala	Y1	Population estimates
X2	Red sea	Y2	Maximum Temperatures
X3	Gadarif	Y3	Minimum Temperatures
X4	North Kordofan	Y4	Rainfall
X5	South Kordofan	Y5	Number of cows
X6	South Darfor	Y6	Number of sheep
X7	West Darfor	Y7	Number of goats
X8	North Darfor	Y8	Number of camels
X9	Blue Nile	Y9	Number of horses
		Y10	Number of Poultry
		Y11	Number of donkeys
		Y12	Number of hospitals
		Y13	Number of hospitals beds
		Y14	Number of health centers
		Y15	Number of doctors
		Y16	Number of primary schools
		Y17	Goys in primary schools
		Y18	Girls in primary schools
		Y19	Teachers in primary schools
		Y20	Secondary schools
		Y21	Boys in secondary schools
		Y22	Girls in secondary schools
		Y23	Teachers in secondary schools
		Y24	Number of illiterates
		Y25	Grain production
		Y26	Agricultural banks
		Y27	Localities
		Y28	Administrative units
		Y29	Police stations
		Y30	Police points
		Y31	Universities
		Y32	Civil defense units
		Y33	Units of traffic
		Y34	Geographical area

Source: The Researcher

3.2 Factor Analysis Results

Factor analysis begins by calculating total variance explained which gives the number of components extracted as shown in Table 3.

Table 3
Total Variance Explained

Component	Initial Eigen Values			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	12.886	37.900	37.900	12.886	37.900	37.900	10.250	30.148	30.148
2	6.751	19.857	57.757	6.751	19.857	57.757	5.477	16.109	46.257
3	5.183	15.245	73.002	5.183	15.245	73.002	5.268	15.493	61.751
4	3.168	9.319	82.320	3.168	9.319	82.320	4.193	12.331	74.082
5	2.395	7.044	89.364	2.395	7.044	89.364	3.685	10.838	84.920
6	1.815	5.339	94.704	1.815	5.339	94.704	2.256	6.635	91.555
7	1.065	3.131	97.835	1.065	3.131	97.835	2.135	6.280	97.835

Source: The Researcher from SPSS

The Initial Eigen Values represent factor variation and determine the factors that will remain in the analysis. The first factor explains most of the variation and run cross with the greater Eigen Value. There are seven components each has Eigen value greater than one and 97.8% of the total variance is due to it. The Extraction Sums of Squared Loadings contains: Eigen Value for each factor, participation rate of each factor in the total variation and cumulative rate of each factor in the total variation. The components with Eigen Value less than one was deleted.

Table 4
Rotated Component Matrix

	Component						
	1	2	3	4	5	6	7
y1	.961						
y2							.879
y3		-.896					
y4		.525					.603
y5				.887			
y6		.732			.523		
y7		.830					
y8			.789				
y9	.552	.526	-.523				
y10					.971		
y11		.878					
y12					.515	.585	
y13			.561		.598		
y14			.941				
y15			.707				
y16	.770						

	Component						
	1	2	3	4	5	6	7
y17	.868						
y18	.852						
y19	.869						
y20	.966						
y21	.970						
y22	.948						
y23	.903						
y24		.504	.491		.590		
y25						.553	.698
y26	.512			.544			
y27	.850			.511			
y28		.583					
y29		.732		.522			
y30	.495		.859				
y31				.954			
y32			.692		.662		
y33				.820			
Y34						.938	

Source: The Researcher from SPSS

Table 4 shows Component Matrix with seven components, each component is linear structure of the scores on the original variables. The following are the seven components from the 34 variables:

$$\mathbf{PC_1} = 0.961y_1 + 0.552y_9 + 0.770y_{16} + 0.868y_{17} + 0.852y_{18} + 0.869y_{19} + 0.966y_{20} + 0.970y_{21} + 0.948y_{22} + 0.903y_{23} + 0.512y_{26} + 0.850y_{27} + 0.495y_{30}$$

$$\mathbf{PC2} = -0.896y_3 + 0.525y_4 + 0.732y_6 + 0.830y_7 + 0.526y_9 + 0.878y_{11} + 0.504y_{24} + 0.583y_{28} + 0732y_{29}$$

$$\mathbf{PC_3} = 0.789y8 - 0.523y9 + 0.561y13 + 0.941y14 + 0.707y15 + 0.491y24 + 0.859y30 + 0.692y32$$

$$\mathbf{PC_4} = 0.887y5 + 0.544y26 + 0.511y27 + 0.522y29 + 0.954y31 - 0.820y33$$

$$\mathbf{PC_5} = 0.523y_6 + 0.971y10 + 0.515y12 + 0.598y13 + 0.590y24 + 0.662y32$$

$$\mathbf{PC_6} = 0.585y_{12} + 0.553y_{25} + 0.938y34$$

$$\mathbf{PC_7} = 0.879y2 + 0.603y4 + 0.698y25$$

3.3 Factors Extraction

Table 5
The Loading of the First Factor

N	Variables	Loadings
1	Population Estimates	.961
2	Number of Horses	.552
3	Number of Primary Schools	.770
4	Boys in Primary Schools	.868
5	Girls in Primary Schools	.852
6	Teachers in Primary Schools	.869
7	Secondary Schools	.966
8	Boys in Secondary Schools	.970
9	Girls in Secondary Schools	.948
10	Teachers in Secondary Schools	.903
11	Agricultural Banks	.512
12	Localities	.850
13	Police Points	.495

Source: The Researcher from SPSS

This factor can be named as factor of multi- potentials, because it reflects potentials of human, animal resources, demography, educational and general utilities.

Table 6
The Loading of the Second Factor

N	Variables	Loadings
1	Minimum Temperatures	-.896
2	Rainfall	.525
3	Number of Sheep	.732
4	Number of Goats	.830
5	Number of Horses	.526
6	Number of Donkeys	.878
7	Number of Illiterates	.504
8	Administrative Units	.583
9	Police Stations	.732

Source: The Researcher from SPSS

This factor can be named factor of livestock, rainfall and general services.

Table 7
The Loading of the Third Factor

N	Variables	Loadings
1	Number of Camels	.789
2	Number of Horses	-.523
3	Number of Hospitals Beds	.561
4	Number of Health Centers	.941
5	Number of Doctors	.707
6	Number of Illiterates	.491
7	Police Points	.859
8	Civil Defense Units	.692

Source: The Researcher from SPS

This factor can be called factor of health and police services.

Table 4
The Loading of the Fourth Factor

N	Variables	Loadings
1	Number of Cows	.887
2	Agricultural Banks	.544
3	Localities	.511
4	Police Stations	.522
5	Universities	.954
6	Units of Traffic	.820

Source: The Researcher from SPSS

This factor can be called general utilities.

Table 9
The Loading of the Fifth Factor

N	Variables	Loadings
1	Number of Sheep	0.523
2	Number of Poultry	.971
3	Number of Hospitals	.515
4	Number of Hospitals Beds	.598
5	Number of Illiterates	.590
6	Civil defense Units	.662

Source: The Researcher from SPSS

The fifth factor is factor of civil defense and poultry.

Table 10
The Loading of the Six Factor

N	Variables	Loadings
1	Number of hospitals	.585
2	Grain production	.553
3	Geographical area	.938

Source: The Researcher from SPS

This factor can be named geographical factor.

Table 11
The Loading of the Seventh Factor

N	Variables	Loadings
1	Maximum Temperatures	.879
2	rainfall	.603
3	Grain production	.698

Source: The Researcher from SPSS

This factor can be called climate factor.

3.4 Cluster Analysis Results

Table 12
Average Linkage (Between Groups) Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	2	28.133	0	0	2
2	1	9	37.739	1	0	3
3	1	8	42.480	2	0	5
4	3	5	49.203	0	0	7
5	1	7	52.379	3	0	7
6	4	6	54.725	0	0	8
7	1	3	57.769	5	4	8
8	1	4	89.558	7	6	0

Source: The Researcher from SPSS

Table 12 determines the states that has been linked together in each step of the analysis. In the first step Kassala and Red Sea are linked because the distance between them is the smallest (28.133), the small value of the coefficient shows homogeneity of the groups linked.

Table 13
Cluster Membership

Case	4 Clusters	3 Clusters	2 Clusters
1 : 1	1	1	1
2 : 2	1	1	1
3 : 3	2	2	1
4 : 4	3	3	2
5 : 5	2	2	1
6 : 6	4	3	2
7 : 7	1	1	1
8 : 8	1	1	1
9 : 9	1	1	1

Source: The Researcher from SPSS

The rang of solution in SPSS is determined from 2 to 4, so, the results distribution is shown in three columns in table 13. In case of clustering to two groups, the first group contains kassala, Red sea, Gadarif, South Kordofan, West Darfur, North Darfur and Blue Nile, the second group comprises North Kordofan and South Darfur. If clustering to three groups, the first group includes Kassala, Red sea, West Darfur, North Darfur and Blue Nile, the second group covers Gadarif and South kordofan, the third group consists of North Kordofan and South Darfur. In case of clustering to four groups, the first group involves Kassals, Red sea, West Darfur, North Darfur and Blue Nile, the second group contains Gadarif, and South Kordofan, the third group comprises only North Kordofan and the fourth group involves only South Darfur.

4. DISCUSSION OF THE RESULTS

The factor analysis method produces seven factors: Firstly, the factor of multi-potentials, which consists of thirteen variables, it is of great significance because it explains 37.900% of the total variance; therefore, it is the basic factor that explains the correlation Matrix and the variables with greatest impact on this factor are: Boys in secondary schools, number of secondary schools, population estimates, girls in secondary schools and teachers in secondary schools with loadings 0.970, 0.966, 0.961, 0.948, 0.903 respectively, and the variables that come second on impacting the first factor are: Teachers in primary schools, boys in primary schools, girls in primary schools and Localities with loadings 0.869, 0.868, 0.852, 0.852 respectively, thirdly, come number of primary schools, number of horses, agricultural banks and Police points with loadings 0.770, 0.552, 0.512, 495 respectively. The structure of this factor reflects the impact of education in different fields, precisely, the positive high loadings of education variables means direct relation between it and animal resources, demographic and general utilities variables, therefore, this factor is named multi-potentials.

The second factor is factor of livestock, rainfall and general services, and it explains 19.857% of the total variance, and the variables with high loadings in this factor were: Number of donkeys, number of goats, number of sheep, police stations and administrative units with loadings 0.878, 0.830, 0.732, 0.732, 0.583 respectively.

The relative important of the third factor comes from that it account for 15.245% of the total variance, it comprises eight variables, only seven contribute in the structure of this factor, these variables are: Number of health centers, police points, number of camels, number of doctors, civil defense units, number of hospitals beds and number of illiterates with loadings 0.941, 0.859, 0.789, 0.707, 0.692, 0.561, 0.491 respectively. This factor shoes the impact of the health and police services in increasing animal food production, so it can be called health and police services factor.

The fourth factor contains six variables, all with positive loadings and with great impact, these variables are: Universities, number of cows, units of traffic, agricultural banks, police stations and o Localities with loadings 0.954, 0.887, 0.820, 0.544, 0.522, 0.511 respectively. This factor can be named general utilities factor.

The fifth factor contains six variables, only two with great impact, they are number of Poultry and Civil defense units with loading 0.917 and 0.662 respectively, so, it can be named defense and poultry factor.

The six factor includes three variables, only one with great impact, it is Geographical area variable, with loadings 0.938, so, it can be named Geographical factor.

The last factor shoes the impact of climate in grain production, the maximum temperatures has the higher loading, it is 0.879, so, it can be called climate factor.

The cluster analysis method gives three solutions, the first solution groups the states into two regions, the first region includes kassala state, Red sea, Gadarif, South Kordofan, West Darfur, North Darfur And Blue Nile as similar in potentials development, the second region involves North Kordofan and South Darfur, they are similar.

The second solution gives three regions, the first region comprise Kassala state, Red sea, West Darfur, North Darfur and Blue Nile, these states are homogeneous due to potentials development, the second region contains Gadarif and South Kordofan as similar states, the third region covers North Kordofan and South Darfur.

The third solution gives four regions, in which Kassals state, Red sea, West Darfur, North Darfur and Blue Nile constitute the first region, Gadarif, and South Kordofan state are in the second region, North Kordofan and South Darfur construct the third and fourth region respectively.

5. RECOMMENDATIONS

1. Ubtating of ministries and institutions Database to allow a greater number of variables for the researchers.
2. Construction of a strong statistical database in each state of the Sudan.
3. The dimensions of variables that determine the regional developments were: demographic, service, economic and natural resources and recommended to be included in Regional disparity studies.
4. Classification of regions using method of potentials development ought to be applied in the different levels of the government (Localities, Administrative Units and Towns).

6. CONCLUSION

In This Study , 34 variables were reduced to seven factors using factor analysis, these factors are: Multi-potentials factor, which includes human, demographic, educational, health and general utilities factors, factor of livestock, Rainfall and general services, factor of health and police services, factor of general utilities, factor of civil defense and poultry, geographical factor and the last is climate factor. By implementing cluster analysis the nine states were clustered into four developmental regions, these regions are: South Darfur, then, North Kordofan, followed by South Kordofan and Gadarif and lastly Kassala, West Darfur, Blue Nile, North Darfur and Red Sea.

REFERENCES

1. Fruchter, B. (1954). *Introduction to Factor Analysis*. D. Van Nostrand Company Inc., New York.
2. Lawley, D.N. and Maxwell, A.E. (1963). *Factor analysis as A statistical Method*, Butterworths, London.
3. Hoppner, F., Klawonn, F., Kruse, R. and Rankler, T. (1999). *Fuzzy Cluster Analysis: Methods for Classification Data Analysis and Image Recognition*. John Wiley & Sons Inc. New York.
4. Hartigan, J.A. (1974). *Clustering Algorithms*, John Wiley and Son Ltd, New York.
5. Bakaric, I.R. (2005). *Uncovering Regional Disparities - The use of Factor and Cluster Analysis*, Institute of Economics, Zegreb.
6. Kaufman, L. and Rousseeuw, P.J. (1989). Finding Groups in Data: An Introduction to Cluster Analysis, John Wiley and Son Ltd, New York.
7. Somi, M.E. (2015). *The Analysis of the Regional Potentials of Development in the Sudan through Factor and Cluster Analysis*. Sudan University of Science and Technology, Khartoum.