Sindh Univ. Res. Jour. (Sci. Ser.) Vol.44 (3) 411-416 (2012)

SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)

Rural Population Distribution in Sindh, Pakistan-A Quantitative Analysis

S. AKHTAR, M.R. DHANANI*

Department of Geography, University of Karachi, Karachi, Pakistan

Received 21st January 2012 and Revised 19th July 2012

Abstract: Rural Population distribution has remained an important field of study of researchers in less developed countries like Pakistan where is about 68 per cent of population live in rural areas. It is an important measure of determining man- land relationship. Therefore rural population distribution is generally linked with the distribution of cultivated land and cropped area. The quantitative analysis like population density, location quotient, population potential, Lorenz curve and Gini coefficient are generally used by population researchers to analyse rural population distribution. In this paper an attempt is made to apply the above quantitative techniques in the study of rural population distribution in Sindh which is one of the province of Pakistan. The principal component analysis has been applied to extract variables and to determine components which explain rural population distribution.

Keywords: rural population, population potential, component analysis, location quotient

1. INTRODUCTION

Like other less developed countries of the World, majority of the people of Pakistan live in rural areas. Although Sindh is the second most populated province of Pakistan its rural population distribution is different from other provinces of the country. On one hand its share of rural population in the total population of the province is just above 50 per cent but on the other hand out of 17 districts, 13 districts have share of rural population is more than 70 per cent **(Table 1).** The other interesting picture of the highly uneven distribution of rural population is that the two districts show extreme opposite picture of rural population. In Tharparkar about 95 per cent of populations of the district lives in rural areas while in contrary in Karachi only 5 per cent of population of the district belong to rural areas. Man-land relationship is an important measure of analyzing distribution of rural population (Clarke, 1973). Agriculture is the dominant land use of Sindh. Therefore cultivated land and crops production are important affecting factors rural population distribution. Quantitative analysis of population distribution has been important methods of population studies (Medvedkov, 1970). In Pakistan there is dearth of studies on the application of mathematics in rural population distribution. Therefore it is worthwhile to apply different quantitative tech. to analyse rural population distribution in Sindh.

Table 1: Percentage of rural population by districts in Sindh

Districts	Rural population as per cent of total population in district (1998)
Jacobabad	75.63
Shikarpur	75.92
Larkana	71.10
Sukkur	49.10
Ghotki	83.60
Khairpur	76.30
Hyderabad	49.20
Dadu	78.60
Naushehro Feroze	82.30
Nawabshah	73.60
Sanghar	77.20
Badin	83.50
Tharparkar	95.60
Thatta	88.80
Mirpurkhas	66.80
Umerkot	84.60
Karachi	5.24
Sindh(Total)	51.28

2. MATERIALS AND METHODS

The present study is primarily based upon the data of District Population Census, 1998 of Sindh province. Data about cultivated land and cropped area of various districts of Pakistan were obtained from Agriculture Census of Sindh. The cartographic technique and quantitative methods were applied to analyse rural population distribution in Sindh. These mathematical methods are as follows.

1. To examine man-land relationship, rural population densities of districts of Sindh were computed. Correlation and regression analysis were used to determine rural population density and cultivated land relationship.
2. Location quotient technique has proved a good alternative measure of population density was applied in Sindh.
3. Population potential model, Lorenz curve and Gini coefficient were used to explain concentration and dispersal of rural population in Sindh.
4. For extraction of different variables and their relationship Principal component analysis was used.

++Corresponding author. M. R. Dhanani, E-mail: mrdhanani@gmail.com Cell,No.+92-453541162

*Department of Geography, University of Sindh, Jamshoro, Pakistan

5. The computed results were mapped through cartographic technique.

3. RESULTS AND DISCUSSION

In the last population Census which was conducted in 1998, the province of Sindh accounted for 15.6 million of rural population which was unevenly distributed in 17 administrative districts of the province. The three districts Hyderabad, Larkana and Dadu accounted for 26 per cent of the total rural population of Sindh **(Table 2).** All are located in the fertile plain of River Indus.

Rural population distribution has been explained by population researchers in different quantitative parameters like population density, location quotient, Gini coefficient, population potential models etc. (Robinson, 1957 and Clarke, 1973). Let us examine these quantitative measures on Sindh.

Table 2: Components of population distribution by districts in Sindh, 1998

Districts	Rural population (000) (1)	Rural population as percent of total rural population (2)	Rural population Density (person per sq.km) (3)	Rural Population Potential (4)	Rural Location Quotient (5)	Cultivated area as percent of total area (6)	Cropped area (Acres) (7)	Total Area (Sq.km) (8)
Jacobabad	1078	6.91	204.2	432.6	1.86	46.9	946	5278
Shikarpur	668	4.28	266.1	301.5	2.40	55.4	465	2512
Larkana	1370	8.77	184.5	435.5	1.66	34.8	982	7423
Sukkur	446	2.86	86.4	182.2	0.78	17.5	389	5165
Ghotki	812	5.20	133.4	273.0	1.20	43.3	713	6083
Khairpur	1181	7.57	74.25	541.2	0.67	14.9	850	15910
Hyderabad	1422	9.11	257.7	712.3	2.32	47.4	832	5519
Dadu	1328	8.50	69.6	520.9	0.62	9.0	495	19070
Naushehro Feroze	895	5.13	303.9	505.0	2.75	51.7	590	2945
Nawabshah	789	5.05	175.3	380.3	1.58	35.3	434	4502
Sanghar	1121	7.19	104.5	593.8	0.94	25.8	1041	10728
Badin	597	6.08	88.8	233.6	0.79	49	972	6726
Tharparkar	874	5.60	44.5	227.9	0.40	13.7	242	19638
Thatta	988	6.30	56.9	358.8	0.51	10	515	17355
Mirpurkhas	605	3.88	207.1	288.5	1.87	65.8	574	2925
Umerkot	551	3.53	98.3	189.0	0.88	29.5	335	5608
Karachi	517	3.33	146.6	141.7	1.32	3.1	40	3527
Sindh Total	15600	100.0	110.7	-----	-----			

Source: Population Census of Pakistan, 1998

Measures of Population Density and Location quotient

The most commonly and simplest method of explaining rural population distribution is population density. The concept of population density refers to numbers of people to the space occupied by them. Two main criticisms have generally been made on these measurements are: i) densities are merely averages ii) Construction of density map is dependent upon the criteria used in the selection of class intervals. Despite these criticisms this method is popularly used by Population geographers because of its simplicity and common understanding. About 50 per of the area of Sindh rural population density lies in between 50 to 100 persons per sq. km which contains about 33 per cent of the total rural population of Sindh (**Table 3)**. The presence of Thar Desert in the eastern districts like Tharparkar, Umerkot, Sanghar, Khairpur and Ghotki, the dry rugged hilly topography in Dadu, Thatta and Karachi and marshy lands in Thatta and Badin districts are the main reasons of generally low population density.

Table 3: Share of population and area in percentage in different population density categories,1998

Rural Population Density (persons per sq. km)	Percentage of Rural Population	Percentage of total Area
Less than 50	5.73	13.07
50—100	33.37	49.61
100—150	16.09	14.37
150—200	14.18	8.48
200---250	11.01	5.77
Over 250	19.59	7.76

Source: Population Census of Pakistan, 1998

The population density map clearly indicates that population densities of districts located in the eastern desert region, western rugged hilly region and southern marshy region have relatively low population densities that is less than 50 persons per sq. km where as districts located in the Indus plain like Hyderabad, Shikarpur, Naushehro Feroze have very high population density i.e. more than 250 persons per sq.km **(Fig 1).**

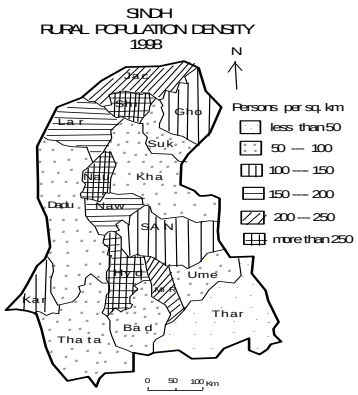


Fig 1: Map of Rural Population density in Sindh

It comes out from the study that rural population of Sindh largely depends upon agriculture. The canals irrigated districts of Shikarpur, Jacobabad, Naushehro Feroze, Hyderabad and Ghotki have high percentage of cultivated land i.e. above 40 per cent of total land of districts. These districts also show high population density i.e. above 200 persons per sq. km. On the other hand Dadu, Tharparkar and Thatta have very low percentage of cultivated land i.e. about 10 per cent of the land of district **(Table 2)**. Pearson correlation coefficient and regression analysis confirm significant relationship between rural population density and percentage of cultivated land in districts of Sindh.

Correlations (Pearson)
Correlation of X and Y = 0.735
Where X is cultivated area as percent of total area

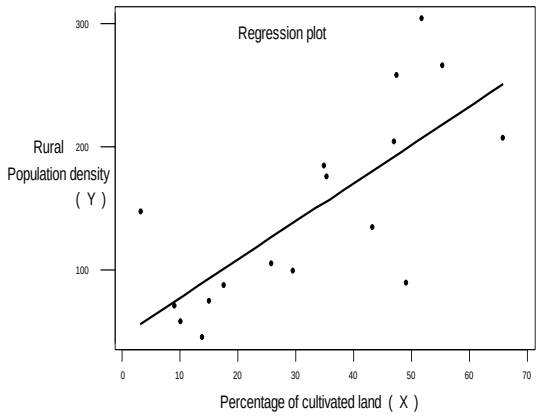
Y is Rural Population density Regression Analysis

The regression equation is

```
Y = 46.2 + 3.10 X
```

Predictor	Coef	St.Dev	T	P
Constant	46.25	27.58	1.68	0.114
X	3.1023	0.7387	4.20	0.001

S = 55.76 R-Sq. = 54.0%
R-Sq. (adj) =51.0%



Location quotient is another quantitative measure of population distribution used as alternative of population density map. It is simply calculated by dividing percentage of total rural population in district to percentage of total area expressed as follows.

$$Location\ Quotient = \frac{\text{Percentage of total rural population in district}}{\text{Percentage of total area in district}}$$

The values of location quotients of districts of Sindh were mapped. Location coefficient values ranged above 2 that are Hyderabad, Shikarpur and Naushehro Feroze have highest concentration of rural population. Similarly location quotient values less than 0.5 indicate least concentration of rural population. Districts of Tharparkar, Dadu, Thatta, Khairpur and Badin lie in this category **(Fig 2).** The desert and rugged topography and saline marshy land are the main controlling factors of least population concentration in these districts.

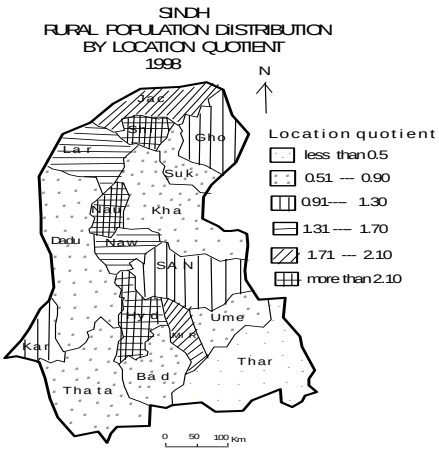


Fig 2: Rural population distribution in Sindh by location quotients of districts,1998

Population potential model

Population potential model is an important method of measuring dispersal of rural population. It was formulated by Stewart in 1958 on the basis that laws of Newtonian physics may be applied to the interaction of people. Following the formula for gravitational force, Stewart defines demographic force. He used a term population potential which refers to a point as a measure of the nearness of people to that point. At a given point, the population potential is the sum of the reciprocals of the distances of all individuals from the point. It is expressed mathematically as follows:-

$$Population\ Potential = \sum \frac{P}{r}$$

Where p is rural population at i point and r is distance from all population points to i point.

shown on map with the help of isolines. The equipotential lines represent distribution of rural population in thousands of persons per kilometer. Hyderabad has highest population potential values i.e. 712 which show centre of gravity of rural population in Sindh (Table 2). The maps of equipotential lines clearly indicate two gravitational cells of rural population concentration in Sindh. The primary centre is located in Hyderabad and secondary in Naushehro Feroze. The concentration of rural population in these two districts can be explained that these two districts account for highest cropped area and percentage of cultivated land in the province. Rural population disperses from these two population gravitational cells outward as distance increases **(Table 2 and Fig 3).**

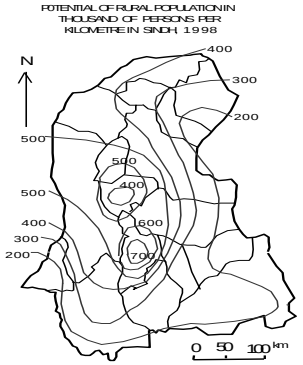


Fig 3: Equipotential lines in thousands of persons per kilometer in Sindh,1998.

Lorenz curve and Gini coefficient

The unevenness and concentration of rural population distribution can be mathematically and graphically explained by Lorenz curve and Gini coefficient. Administrative districts of Sindh were arranged in order of decreasing rural population density and then population and areas of districts were totaled for each density class. Cumulative percentages of population were plotted on Y-axis and cumulative percentages of area were taken on X-axis. The curve clearly indicates that rural population distribution is highly uneven **(Fig 4).** Highest concentration of rural population lies in the category of population density 50-100 persons per sq. km. About 33 percent of rural population of Sindh is distributed in the 50 per cent area of Sindh.

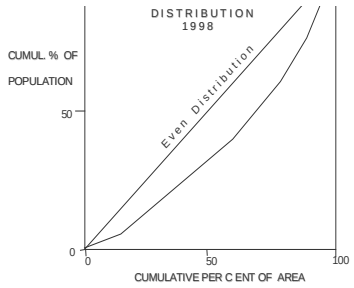


Fig 4: Lorenz curve of population density in Sindh, 1998.

The Gini coefficient is an important measure to compute unevenness and concentration in the distribution. It varies from 0 to 1. One represents maximum concentration. The Gini coefficient is mathematically expressed as follows;

$$G = \frac{1}{100 \times 100}\left[\sum_{i=0}^{i=n} x_i y_{i+1} - \sum_{i=0}^{i=n} x_{i+1} y_i\right]$$

Where X is cumulative percentage of area
Y is cumulative percentage of population

Result: Gini Coefficient = 0.301 = 30.1 per cent

The Lorenz curve and value of Gini coefficient clearly explain the uneven distribution of rural population in Sindh.

Principal Component Analysis

There are various quantitative factors that affect the distribution of rural population in Sindh. It was worked out that eight interrelated variables may affect rural population distribution (Table 2). To extract the relevant variables and components on the basis of high correlation coefficients and rotated factors principal component analysis was applied. The computer software SPSS was used for computation and analysis of principal component analysis. There are three stages in the component analysis:

1. Aa correlation matrix is generated for all variables. A correlation matrix is a rectangular array of the correlation coefficients of the variables with each other.

2. Factors are extracted from the correlation matrix based on the correlation coefficients of the variables.

3. The factors are rotated in order to maximize the relationship between the variables.

There are eight variables, taken from 17 districts of Sindh which are: rural population, percentage of rural population in the total population of district, rural population density, rural location

quotient, percentage of cultivated land in district, cropped area and total area of district.

Results of SPSS

Correlation matrix was computed. Values of correlation coefficients ranging between -1 to +1. The correlation coefficient between a variable and itself is always 1 **(Table 4).** The negative values show negative relationship. Coefficient value 0 indicates no relationship while 1 shows perfect strong relationship between them.

Table 4: Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	.00	0.95	0.07	0.86	0.07	--0.09	0.57	0.45
2	.95	1.00	--0.00	0.81	-0.01	--0.05	0.67	0.46
3	.07	--0.00	1.00	0.30	1.00	0.73	0.14	--0.72
4	.86	0.81	0.30	1.00	0.30	0.12	0.61	--0.21
5	.07	--0.01	1.00	0.30	1.00	0.73	0.14	--0.72
6	-.09	--0.05	0.73	0.12	0.73	1.00	0.41	--0.66
7	.57	0.67	0.14	0.61	0.14	0.41	1.00	--0.01
8	.45	0.46	--0.72	0.20	-0.72	--0.66	-0.01	1.00

Source: Computed by Authors on SPSS

The Kaiser–Meyer-Olkin (KMO) test measures the sampling adequacy which should be greater than 0.5 for satisfactory component analysis. In this case it is 0.572 **(Table 5).**

Table 5: Result of Kaiser-Meyer-Olkin(KMO) test

Kaiser- Meyer – Olkin Measure of Sampling Adequacy	0.572
Bartlett's Test of Sphericity Approx. Chi –Square Degree of freedom: df Significance	 239.341 280.000

Source: Computed by Authors on SPSS

In the next step all the factors extractable from the analysis along with their Eigen values, the percent of variance attributable to each factor, and the cumulative variance of the factor and previous factors were computed. Eigen value greater than 1 is significant. The first factor accounts for 43.51 % of the variance and the second 41.950%. All other factors are not significant **(Table 6).**

Table 6: Total Variance Explained

Compo-nents	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.481	43.512	43.512	3.481	43.512	43.512
2	3.356	41.950	85.462	3.356	41.950	85.462
3	0.728	9.100	94.562			
4	0.226	2.825	97.387			
5	0.141	1.759	99.146			
6	0.048	0.604	99.750			
7	0.020	0.250	100.000			
8	0.00001	0.0002	100.000			

Extraction Method: Principal Component Analysis; Source: Computed

The screen plot is a graph of the Eigen values against all the factors. The graph is useful for determining how many factors to retain. The point of interest is where the curve starts to fall sharply. It can be seen that the curve begins to fall sharply between 2 and 3 **(Fig 5)** The rotated component matrix extracted two components. Component 1 comprises percentage of rural population, rural population, rural population potential and cropped areawhile component 2 comprises rural population density, rural location quotient, percentage of cultivated area and total area **(Table 7)**.

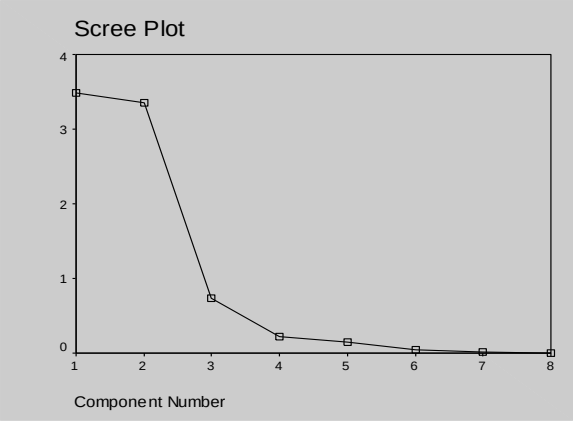


Fig 5: Screen graph

Table 7: Rotated Component Matrix

	Component 1	Component 2
1 Rural population as percent of total rural population	0.964 0.960	
2 Rural population	0.920	
3 Rural Population Potential	0.749	
4 Cropped area		0.948
5 Rural population density		0.948
6 Rural location Quotient		
7. Cultivated area as percent of total area		0.869 -0.836
8 Total Area		

Extraction Method: Principal component analysis
Rotation Method: Varimax with Kaiser normalization
Source: Computed by Authors on SPSS

The component score coefficient matrix was computed **(Table 8).** The component scores were computed on district basis and results were plotted on maps **(Figs 6 and 7).**

Table 8: Component Score Coefficient Matrix

	Component 1	Component 2
1 Rural population	0.280	--0.039
2. Rural population as percent of total rural population	0.282	--0.050
3. Rural population density	0.035	0.278
4. Rural Population Potential	0.267	0.040
5. Rural location Quotient	0.035	0.278
6. Cultivated area as percent of total area	0.019	0.255
7 Cropped area	0.217	0.600
8. Total Area	0.110	--0.256

Extraction Method: Principal component analysis
Rotation Method: Varimax with Kaiser normalization

the data variance. The map of scores on component 1 comprises rural population, percentage of rural population, rural population potential and cropped areas indicate highest scores i.e. above 2.26 in Hyderabad and Sanghar. Umerkot and Sukkur have least scores. The second high score region

This component accounts for 42 percent of the data variance comprises Larkana, Jacobabad, Dadu and Khairpur where scores ranged 1.91 to 2.25 (**Fig 6).**

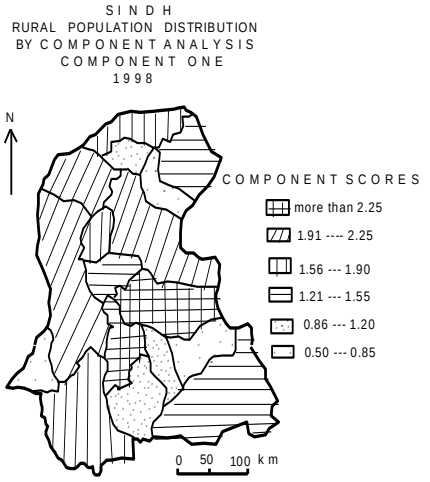


Fig 6: Map of scores on component 1 based on component matrix scores (Table 8).

The map of scores on component 2 comprises rural population density, rural location quotient, percentage of cultivated area and total area indicate high scores above 0.76 in Naushahro Firoze, Hyderabad and Shikarpur and least in Tharparkar, Dadu and Thatta **(Fig7).**

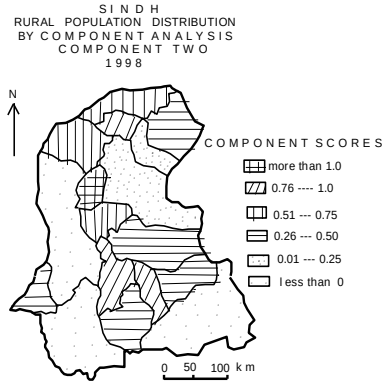


Fig 7: Map of scores on component 2 based on component matrix scores (Table 8).

4. CONCLUSION

The quantitative analysis of rural population distribution in Sindh reveals that location quotient is a good measure of analyzing rural population distribution. Population potential, Lorenz curve and Gini coefficient are good measures of explaining population dispersal and population concentration. Principal component analysis has proved a good technique for extracting variables affecting rural population distribution in Sindh. It can be concluded from the results that districts located in the heart of the Indus plain like Hyderabad, Naushehro Feroze, Shikarpur, Jacobabad are core of rural population concentration having high population densities. These regions have also high percentage of cropped area and cultivated land. The peripheral region of rural population concentration is located in the desert district of Tharparkar, marshy and saline land of Thatta district and rugged hilly district of Dadu.

REFERENCES:

Anderson T.R. (1956) "Potential models and the spatial distribution of population", American Geographers, Vol. **(47):** 379-391.

Bogue D.J. (1959) "Population distribution", University of Chicago Press, Chicago, 156-170.

Clarke J. I.(1973) " Population Geography", pp 20-24, Pergamon press, Oxford. Journal of Regional Science Vol. **(34):** 175-182.

Government of Pakistan (1998) District Population Census Report", Islamabad, 30-45.

Government of Pakistan (2001) "Agriculture Census Report of Sindh province", 56-67.

Khan F.K. (2005) "Pakistan; Economy, People and Environment", Oxford University Press, Karachi, 54-67.

Medvedkov V. (1970) "Application of mathematics to population geography", Population geography: A reader, 144-153.

Stewart J. Q. (1958) "Physics of population distribution", J. of Regional Sci. Vol. **(34):** 99-123.