Available online at www.elixirpublishers.com (Elixir International Journal)

Environment and Forestry

Elixir Environment and Forestry 40A (2011) 5637-5642

ISSN: 2229-712X

Application of Factor analysis & GIS for Groundwater Water Quality Assessment in Integrated Dharmapuri District, India

[1]Manjula Devi.K.K and [2]S.Balasubramanian

[1]Research Scholar, Department of Environmental Sciences, Bharathiar University, Coimbatore-641046.

[2]Professor, Department of Environmental Sciences, Bharathiar University, Coimbatore-641046

ARTICLE INFO

Article history:

Received: 8 September 2011;

Received in revised form:

10 November 2011;

Accepted: 12 November 2011;

Keywords

Ground Water,

Factor analysis,

TWAD,

Water Quality Parameters,

GIS.

ABSTRACT

The presence of elements to a certain level support the physiological processes in all living organisms including human beings, but when exceeds the limit they pose health threats. This study was initiated to identify the groundwater quality parameters like pH, Ec, Tds, Na, Mg, Ca, SO4, Fl ,TH ,Cl present in Integrated Dharmapuri district. Two seasonal data of January(winter) and May(Summer) month was obtained from Tamilnadu Water Supply and Drainage Board (TWAD) from 2003 to 2005 for three years .The obtained data was used for factor analysis and compared the loaded factors season wise and mapped using Geographical Information System (GIS).

1.0 Introduction

The collection and analysis of groundwater quality data to achieve a management strategy is a laborious task. To overcome such problems data reduction methodology (PCA and Factor Analysis) is normally adopted by the researchers. Factor analysis and PCA multivariate techniques provide direct insight into the inter relationship among variables for addressing conceptual issues relating to the underlying structure of the data, as well as a process for data reduction. These multivariate techniques involve the computation and analysis of the variance and co – variance structure of a set of variables Shameeri *et al.* (2000); Johnson and Wichern (2002); and Maragatham (2003) have applied the techniques to highlight the information that is not despaired using univariate techniques. Johnson and Wichern (2002) have reported that these techniques may be performed for the following purposes:

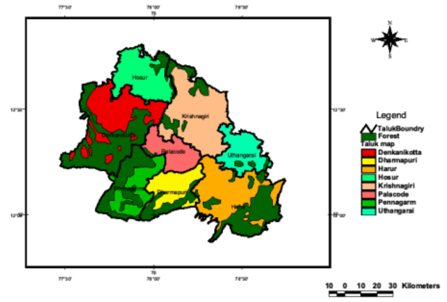
- ⇒ Data reduction and structural simplification to make interpretation easier.
- ⇒ Sorting and grouping whereby groups of similar variables are created
- ⇒ Investigation of the dependence among variables.
- ⇒ To predict relationships among variables.
- ⇒ For hypothesis construction and testing.

Few attempts were made in previous studies to investigate groundwater problems using Factor/PCA multivariate analysis and for evaluation of the impact of hydro chemical process involved in groundwater quality. Dawdy *et al.* (1967); Klovan (1975); Ashley *et al.* (1978); Dalton *et al.* (1978); Usnoff *et al.* (1989); and Guler *et al.* (2002); have applied this technique to make proper interpretations for understanding the need of strength and weaknesses.

2.0. Study Area

The study area lies at a geographical extent from 11° 45’ to 12° 53’ N and 77° 13’ to 78 °45’ E. The total area of the district is 9581.26sq.km. The district is surrounded by Vellore, Tiruvannamalai and Villupuram districts in the East, Salem district in the South, the states of Karnataka and Andhra Pradesh in the North (Map 3.1) The Dharmapuri district has been recently bifurcated into two in the year 2004 to form an additional Krishnagiri district*. The altitude of the district ranges from 300-1200m above mean sea level. the integrated Dharmapuri district has been divided into 8 taluks namely Dharmapuri, Krishnagiri, Pennagaram, Harur, Hosur, Palacode, Denkanikotta, and Uthangarai. The two taluks were bifurcated from Harur (Papiredipatti) and Uthangarai (Pochampalli taluk) .The population of the district as per 2001 census was 28,56,300 and the total area of the district is 9581.26sq.km.

Map.3.1 .Talukwise map of Dharmapuri district



* The area for the present work is Integrated Dharmapuri district (With Krishnagiri district).

3.0. Data Used and Methodology

For the present study, the groundwater quality data along with fluoride were obtained from Tamil Nadu Water Supply and Drainage Board (TWAD), Chennai, for 71 locations (January and May,2003-2005). The average of three years fluoride mean value for January and May (n=71) were taken.

3.1. The Model

Multi co-linearity occurs when two independent variables in a matrix are perfectly correlated and show a similar pattern of correlations with the other variables. When subjecting the variables to principle component are used as variables in subsequent analysis.

Multivariate analysis refers to all statistical methods that simultaneously analyze multiple measurements on each individual or object under investigation. The method is used to measure, explain and/to predict the degree of relationship among variables (weighted combinations of parameters). Multivariate methods have been applied to problems arising in the social, medical, physical, soil and environmental sciences.

The principle component analysis is one of the multivariate techniques concerned with explaining the variance structure through a few linear combinations of the original variables. Its general objectives are data reduction and interpretation. Those scores are used for factor analysis to describe the covariance relationships among the variables in terms of few underlying, but unobservable, random quantities called factors.

Correlation and Covariance matrices can also be diagnosed. When a matrix is diagnolised, it is transformed into a matrix with numbers in the positive diagonal, which represent variance for the original matrix that has been repackaged by manipulating it as follows (Johnson and Deen, 1992).

$$L=V'RV \tag{1}$$

where,

V = eigenvector

R = correlation matrix

V' = Transfer of V

The calculation of eigen values and eigenvectors is done by reorganized equation.

$$R=VLV' \tag{2}$$

Then $V\sqrt{L}$ is called A and $\sqrt{L}V$ is A'

Now (2) is rewritten as

$R = AA'$

Once the Eigen values and its vectors are known, the unrotated factor loading matrix can be derived by straight forward matrix multiplication as,

$$A=V\sqrt{L} \tag{3}$$

By using the transformation matrix

$$A = \frac{\text{Cos } \Psi\text{- Sin } \Psi}{\text{Sin } \Psi \text{ Cos } \Psi}$$

R = Reproduced correlation matrix

R res = R-1 – R

The regression co – efficient for producing factor scores is a product of the inverse of the correlation matrix and factor – loading matrix.

B = R -1 – A and

F = ZB

Z = FA'

where,

F = Factor scores

A' = Factor loading and

Z = scores on variables.

Scree Test Criterion

The Scree test Criterion was used to identify the optimum number of factors that can be extracted before the amount of unique variance begins to dominate the common variance structure. The Scree test was derived by plotting the latent roots against the number of factors in their order of extraction and the shape of the resulting curve was used to evaluate the cutoff points.

Deriving factors and assessing overall fit

The following methods/ criteria are used for designing, assessing and interpreting the factors.

1. ***Correlation analysis***: The correlation analysis is used to explain inter – correlations among the variables (parameters).
2. ***Measures of sampling adequacy (MSA)*****:** Measure calculated both for the entire correlation matrix and also for individual variables to evaluate the appropriateness of applying factor analysis.
3. ***Rotation***: Orthogonal – Varimax rotation is used to transform the coordinates to an angle of 90° at right angles. The correlation between the factors is determined to be 0.
4. ***Community***: Total amount of variance on the selected original variable is shown with the other entire variables included in the analysis.

4.0. Result and Discussion

When dealing with the groundwater samples, bore wells in Dharmapuri district were subjected to factor analysis, Four factors are responsible for January and May (each n=71), respectively.

The scree plot presented in figure.1, shows the near flatness, which renewed that the first four factors have maximum total variance. This confirms that the extracted four factors are mainly responsible to explain the underlying mechanism of groundwater in January data of Dharmapuri district.

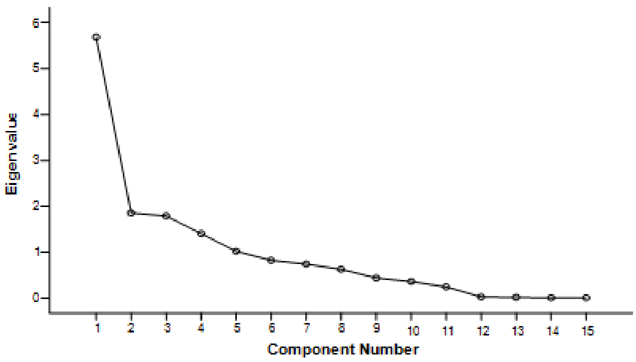


Figure.1.Scree plot for water quality parameters loaded in four factors in the month of January.

The parameters along with its component loading values (> 0.5) were considered for interpretations after the Varimax rotation. In factor 1, EC, TDS, and Cl were loaded heavily (> 0.8) and the other parameters loaded in factor 1 are Na, Mg, TH and Ca. Four parameters were loaded in factor 2 viz, Mg, TH, Ca and turbidity. In factor 3, fluoride was heavily loaded, followed by No3 and Potassium (K). In factor 4, Fe and SO_4 were loaded highly. Among these parameters fluoride is loaded negatively, which implies that the source of contamination may be unique and the interaction with other parameters are very less (Table 1).

The loaded components (factor 1 to 4) for January (winter) data were spatially presented as maps 5.3.9, 5.3.10, 5.3.11 and 5.3.12, respectively.

The physico-chemical parameters are highly concentrated in small patches of Denkanikotta, Krishnagiri and Harur taluks and moderately concentrated in Palacode, Dharmapuri and Hosur taluks in the district (Map 5.3.9).

The second factor which is loaded by turbidity, TH, Calcium and Magnesium were highly concentrated in Uthangarai, Denkanikotta, Krishnagiri, Hosur, Dharmapuri and Harur taluks which is represented in the form of small patches in Map 5.3.10. The moderate concentrations of the parameters are found in all the taluks except in few places of Pennagaram, Dharmapuri and Harur taluks.

The third factor spatial distribution map reveals that the parameters (fluoride, NO_3 and potassium) were loaded and highly concentrated in patches of Harur, Krishnagiri and Hosur taluks. The moderate concentrations of the above parameters are seen in all the taluks except the western part of the Denkanikotta and Hosur taluks which is presented in Map 5.3.11.

The fourth factor spatial distribution map reveals that the concentration of iron and sulphate is highly concentrated in Pennagaram, followed by a portion of Dharmapuri and Palacode taluks. The moderate concentrations are seen in the central part of the district, starting narrowly from the Krishnagiri taluk and ending broadly in Pennagaram and Dharmapuri taluks, which is presented in Map 5.3.12.

The methodology of factor analysis and the criteria used for overall fit of the model, validation and also interpredicting the factor (winter) January are presented in the table 1 and 2 . It is observed that the total variance explained by the four factors are 72% out of which 32% is occupied by the first factor whereas second, third and fourth factors occupy 15.8%, 12.7% and 10.8%, respectively.

Similarly, the scree plot for May is presented in fig.2, which shows the near flatness that renewed that the first four factors have maximum total variance. This confirms that the four extracted factors are mainly responsible to explain the underlying mechanism of groundwater in May of integrated Dharmapuri district.

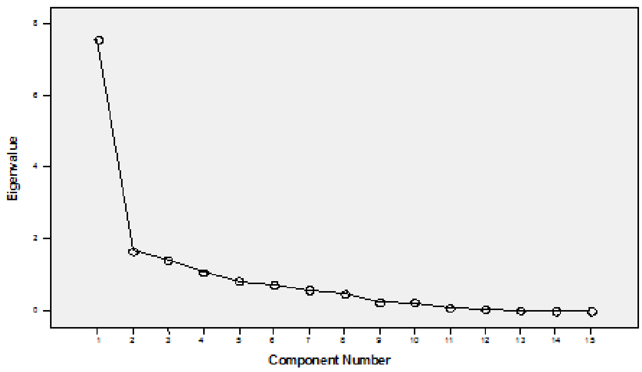


Figure 2. Scree plot for water quality parameters loaded in four factors in the month of May

In factor 1, out of the fifteen parameters, ten parameters are loaded heavily (EC, TDS, TH, Ca, Mg, CL, Talk, Na and Potassium) K, NO3 were loaded in factor 2, whereas in factor 3, turbidity and fluoride were loaded. In factor 4, pH and Iron were loaded. In summer season, the spatial distributions of Factor 1 to 4 were presented as maps (5.3.14, 5.3.15, 5.3.16, 5.3.17).

From the map 5.3.14 (factor 1), it is observed that physico-chemical parameters are highly concentrated in a patch of Harur taluk. The moderate concentrations of the parameters are seen in the major parts of the Harur, followed by Krishnagiri and Dharmapuri, while it is seen as very little patches in Denkanikotta, Hosur, Uthangarai and Palacode taluks.

The spatial distribution of Factor 2 in map 5.3.15 shows that, the parameters (NO_3 and K) loaded are highly concentrated in Harur, and moderately in the southern part of the district that covers the taluks, Denkanikotta, Palacode, Pennagaram, Dharmapuri and Harur, respectively.

The fluoride and turbidity were loaded in factor 3, spatial distribution in map 5.3.16 reveals that the parameters are present throughout the district and randomly distributed in small pitches in the district and the maximum in eastern part of the district of Krishnagiri and Uthangarai taluks. The moderate concentrations of parameters are seen randomly in all the taluks except in few patches of the district.

In factor 4, the iron and pH of groundwater is spatially distributed highly in Denkanikotta, Hosur, and Krishnagiri taluks and small patches are seen in Harur taluks, Pennagaram and Uthangarai taluks. The moderate concentrations of parameters are seen in all the eight taluks except a middle portion of Palacode taluks Map5.3.17. The spatial distribution of groundwater confirms the results observed earlier.

Similar studies were also adopted for the groundwater quality of summer season (n = 71). The Varimax rotated components for four factors, and the percentage of variance explained by each factor components and the respective communalities were presented in Table 4.

The total variance explained by the four factors is nearly 78% out of this, factor 1, explained 47.29%, which is half of the total variance. The second, third and fourth factors explained nearly equal proportions of the total variance. (Table 4).

Based on the result out of the 15 parameters, including the fluoride under consideration, five parameters i.e., TDS, TH, Calcium, Magnesium and fluoride were found to exceed the Indian standard limit in winter as well as in summer seasons. It is observed that the parameters exceeding the BIS drinking water standards are to be considered, for removal methods.

5.0 Conclusion

From the results of table 1 to 4 and the maps 5.3.9 to 5.3.17 for winter (Januay) and summer (May)seasons, it is concluded that, Four factors were controlling the groundwater chemistry in both seasons, of integrated Dharmapuri district. However,the maximum parameters were loaded in factor 1 i.e., eight parameters in winter and ten parameters in summer. All the parameters were positively correlated with the overall mineralization of groundwater. Fluoride is negatively loaded in winter and positively loaded in summer. From the obtained results to conclude that the fluoride in groundwater of integrated Dharmapuri district is through the weathering of fluoride bearing rocks.

Table 1. Varimax Rotated Factor loading, communalities for Groundwater quality Parameters in Dharmapuri District for the month of January (n=71)

S.No	Parameters	Fact 1	Fact 2	Fact 3	Fact 4	h^2
1.	Turbidity	0.57	0.604	-0.166	-0.332	0.506
2.	EC	0.930	0.168	0.072	0.113	0.910
3.	TDS	0.931	0.168	0.053	0.111	0.909
4.	pH	0.197	-0.469	0.027	0.125	0.275
5.	Talk	0.845	0.028	-0.139	-0.162	0.761
6.	TH	0.535	0.739	0.197	0.213	0.917
7.	Ca	0.512	0.707	0.226	0.302	0.906
8.	Mg	0.552	0.750	0.167	0.113	0.907
9.	Na	0.655	-0.161	0.379	-0.343	0.716
10.	K	0.493	-0.167	0.569	0.047	0.597
11.	Fe	-0.220	-0.089	0.354	0.642	0.593
12.	No3	0.234	0.017	0.765	0.056	0.643
13.	Cl	0.806	0.192	0.090	-0.039	0.696
14.	Fl	0.195	-0.161	-0.723	0.002	0.587
15.	So4	0.122	-0.032	-0.152	0.878	0.809

KMO test: 0.600 df: 105 Chi-square: 1573.414 Sig: 0.01

Table 2. Components with the % of variance and cumulative % of groundwater quality parameters for the month of January

Component	Extraction Sum of Square Loadings			Rotation Sum of square Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.690	37.932	37.932	4.817	32.114	32.114
2	1.853	12.351	50.283	2.378	15.850	47.965
3	1.788	11.918	62.201	1.905	12.701	60.665
4	1.402	9.350	71.551	1.633	10.886	71.551

Table 3. Varimax Rotated Factor loading, communalities for Groundwater quality Parameters in Dharmapuri District for the month of May (n=71)

S.No	Parameters	Fact 1	Fact 2	Fact 3	Fact 4	h^2
1	Turbidity	0.120	0.209	0.787	0.049	0.680
2	EC	0.968	0.185	0.093	-0.037	0.981
3	TDS	0.968	0186	0.092	-0.037	0.981
4	pH	-0.039	-0.369	-0.027	0.607	0.507
5	Talk	0.721	-0.016	0.378	-0.249	0.725
6	TH	0.949	0.138	0.118	-0.074	0.938
7	Ca	0.945	0.167	0.101	-0.066	0.934
8	Mg	0.914	0.096	0.137	-0.076	0.869
9	Na	0.513	0.480	-0.131	0.274	0.586
10	K	0.623	0.646	-0.012	0.142	0.825
11	Fe	-0.052	0.154	0.287	0.730	0.641
12	No3	0.068	0.862	0.094	-0.180	0.788
13	cl	0.893	0.216	-0.116	0.154	0.881
14	Fl	0.103	-0.157	0.747	0.163	0.620
15	So4	0.769	-0.369	0.108	0.110	0.752

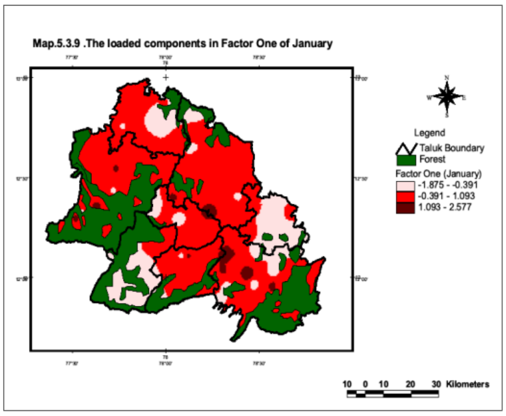
KMO test: 0.715 df: 105 Chi-square: 2141.963 Sig: 0.01

Table 4. Components with the % of variance and cumulative % of groundwater quality parameters for the month of May

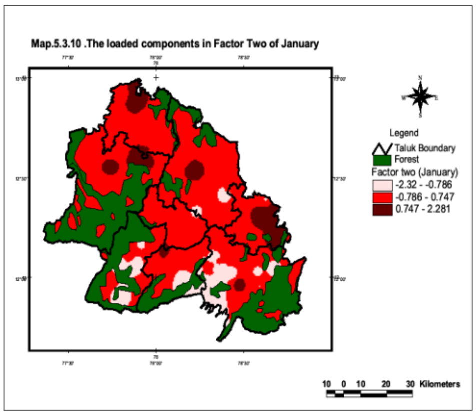
Component	Extraction Sum of Square Loadings			Rotation Sum of square Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.549	50.329	50.329	7.094	47.291	47.291
2	1.670	11.135	61.464	1.927	12.844	60.135
3	1.415	9.436	70.900	1.515	10.099	70.234
4	1.073	7.154	78.054	1.173	7.820	78.054

Table 5. Drinking water quality standards and its remarks for the month of January and May

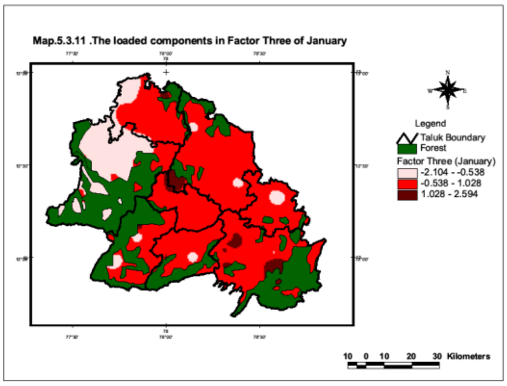
S.No	Parameters	Drinking water quality Standards (BIS)	January	Remarks	May	Remarks
1	Turbidity	10 NTU	1.9	< BIS limit	1.6	< BIS limit
2	EC	1400	1357.23	< BIS limit	1154.86	< BIS limit
3	TDS	500mg/l	952.23	> BIS limit	808.42	> BIS limit
4	pH	6.5-8.5	7.6	< BIS limit	7.4	< BIS limit
5	TALK	200mg/l	354.11	< BIS limit	338.01	< BIS limit
6	TH	200mg/l	463.74	> BIS limit	370.59	> BIS limit
7	Ca	75mg/l	106.49	> BIS limit	84.94	> BIS limit
8	Mg	30mg/l	47.66	> BIS limit	37.95	> BIS limit
9	Na	-	99.61	-	82.49	-
10	K	-	6.53	-	5.47	-
11	Fe	0.30mg/l	0.12	< BIS limit	0.03	< BIS limit
12	NO_3	45mg/l	40.94	< BIS limit	37.14	< BIS limit
13	Cl	250mg/l	163.88	< BIS limit	122.46	< BIS limit
14	F	1.0mg/l	1.12	> BIS limit	1.15	> BIS limit
15	SO_4	150mg/l	54.05	< BIS limit	30.50	< BIS limit



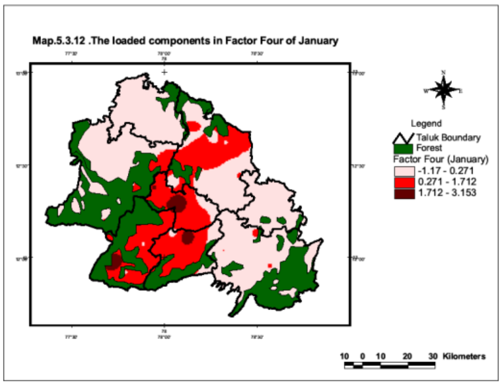
Map.5.3.9 .The loaded components in Factor One of January



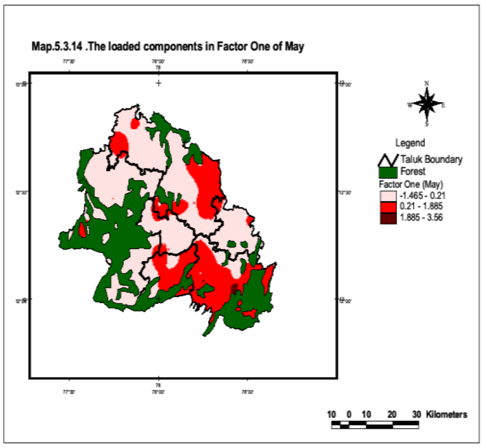
Map.5.3.10 .The loaded components in Factor Two of January

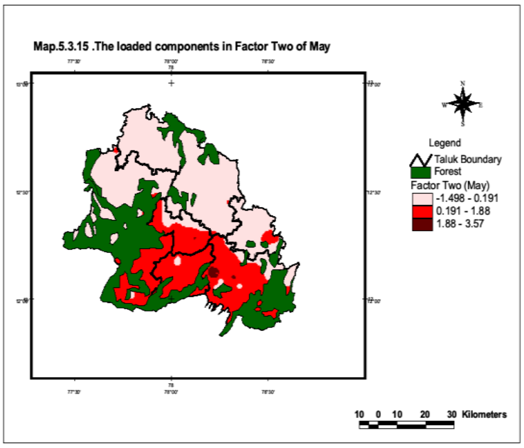


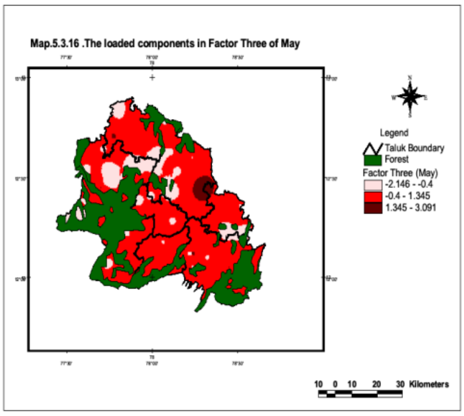
Map.5.3.11 .The loaded components in Factor Three of January

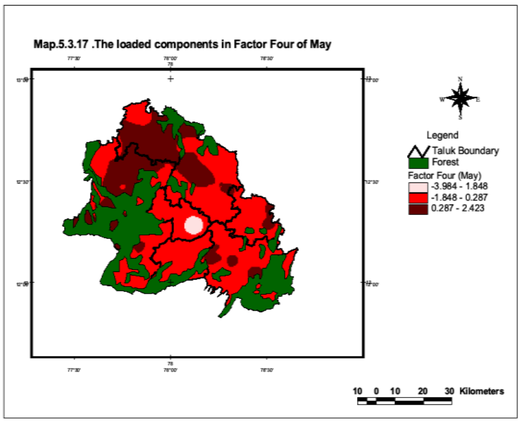


Map.5.3.12 .The loaded components in Factor Four of January









6.0. References

1. Shameeri F, Geetha P, and Balasubramanian S (2000) "Application of factor analysis to interpret the effect of sewage water chemistry on plant/soil". *Res.J.Chem.Environment* 4(1):17-21.
2. Johnson R and Wichern D (2002) "Applied Multivariate Statistical Analysis". Prentice Hall, New Jersey, 426-476.
3. Maragatham, (2003) "Studies on the correlation of selected bio chemical socio-Economic and Anthropometric Factors on Pregnancy Outcome in South Indian Women" Ph.D., Thesis submitted to Bharathiar University, Coimbatore.
4. Dawdy D R and Feth J H (1967) "Application of factor analysis in the study of chemistry of Groundwater quality", Mojave River valley, California. *Water resource* 3: 505 – 510.
5. Klovan J E (1975) "R – Mode and Q – Mode factor analysis". In McCammon RB (Ed) concepts in geostatistics, Springes, Berlin, Heidelberg, New York, 168.
6. Ashley R P and Lloyd J W (1978) "An example of the use of Factor analysis and Cluster analysis in Groundwater Chemistry Interpretation". *J Hydrol* 39: 355 – 364.
7. Dalton M G and Upchurch S B (1978) "Interpretation of Hydro chemical faces by factor analysis", *Groundwater* 16(4):228-233.
8. Unsoff J, Guzman – Guzman A (1989) "Multivariate analysis in hydrochemistry, an example of the use of factor and correspondence analysis". *Groundwater 27*(1): 27:34.
9. Guler C, Thyne G D, McCray J E and Turner A K (2002) " Evaluation of Graphical and Multivariate Statistical method for classification of water chemistry data" H*ydrological Journal :* 10(4): 455 – 474.