Multivariate and Geostatistical Analysis of Groundwater Quality in Palar River Basin

P.J. Sajil Kumar, P. Jegathambal, E.J.James
Water Institute, Karunya University, Coimbatore
Email: esther.jegatha@gmail.com

Abstract— The knowledge of the occurrence of groundwater, its replenishment, physical and chemical characteristics have special significance in arid and semi-arid zones where groundwater is the main source of water. Assessing the quality of groundwater is important in determining its suitability for different purposes. In recent years, multivariate analysis is widely applied to identify the underlying structure of the groundwater quality data. Also the geostatistical tool is mostly used to get the spatial distribution map of a particular pollutant in the specified region. The results obtained through above mentioned tools will be helpful for the decision makers to adopt suitable remedial measures to protect the groundwater sources. In this study, the effect of discharge of tannery effluents in the Palar river basin was studied using factor analysis and geostatistics. Based on the results, it is concluded that the groundwater is not suitable for drinking in the northeast and southwest areas of the Palar river basin.

Keywords. Water quality analysis, factor analysis, geostatistics, kriging, spatial distribution.

I. Introduction

Provision of quality water to the various sectors such as domestic, irrigation, sanitation, environment and industry is the primary goal of any water resources project. The ratio of world population and water withdrawals during the twentieth century is found to be 3:7. So the protection of groundwater is necessary from the its socio-economic point of view. In recent years, groundwater quality analysis gained importance to understand the processes contributing to pollution. The factors behind this contamination may be natural or anthropogenic. Important natural processes contributing to pollution in groundwater are rock-water interactions, dissolution, precipitation, sorption and geochemical reactions. Anthropogenic activities such as waste disposals, leaching of salts, fertilizers,pesticide from the agricultural fields and salt intrusion due to over exploitation contribute to groundwater pollution. Groundwater investigation consists of both quality and quantity determination. Numerous analytical tools are available to facilitate the interpretation and presentation of geochemical analysis. The analyses shows water-rock interactions; reflects the differences in mineral composition of the aquifer, existence of fissures, faults and cracks which affect the groundwater movement in the subsurface medium. Analysis is also useful for estimating residence time and recharge zone [1]. Many graphical methods such as scatter diagram, ionic concentration diagram, piper diagram, percentage of composition diagram, US Salinity Diagram, Dendogram are useful in interpretation and presentation of groundwater quality in a specified location. In all the methods, samples with similar chemical characteristics are grouped together that can be correlated with location. These types of plots demonstrate how the chemical quality of groundwater varies spatially over time. Scatter diagrams are used to understand the trend in a particular species of ions with different water quality parameters. Using Stiff diagram, the quality of water can be compared based on the distinctive graphical shape of the diagram. In the piper diagram, the ion concentrations are plotted as percentages on a single graph. It is mostly used for groundwater quality analysis since it can show the mixing of two waters from different sources. In recent years, multivariate statistical methods are widely applied to extract the underlying information about the hydrochemical data sets.

Factor analysis is one of the tools used to identify the underlying structure of the data, which cannot be identified by the usual graphical methods such as piper diagrams, stiff diagrams etc .There are 3 stages in the factor analysis, a) generation of a correlation matrix for all the variables, b) extraction of factors from the correlation matrix based on the correlation coefficients of the variables, c) rotation of these factors to maximize the relation between them and other variables [2]. Multivariate analysis itself is a wide branch of statistical analysis. Based on the nature of the data, problem and objectives, an appropriate tool is selected. After identifying the major factors responsible for the groundwater contamination, geostatistics tool can be applied to obtain a continuous surface through the interpolation using a set of sample points. It is also capable of producing error or uncertainty surface.

For past few decades, many works have been carried out on the interpretation of water quality parameters using Factor analysis, principle component analysis and multiple linear regressions. Factor analysis technique is used for explaining the local and regional variations in the hydrogeochemical processes and distinguishing the geogenic and anthropogenic sources by selecting the index wells for the long term monitoring [3]. The changes in the surface water quality data during low flow conditions and high flow conditions were explained using factor analysis [4]. A study on interrelations and sources of pollution of groundwater of North Chennai (India) by factor analysis and grouping of the trace metal species using multiple linear regression was carried out [5]. The industrial effluent quality parameters such as salt stress, salt type, heavy metals and potassium effects were identified as critical factors using principle component analysis [6]. The author [7] attempted to carry out the water quality analysis at the Agackoy monitoring station on the Porsuk tributary in the Sakarya river basin using principle component analysis. The hydrochemical data were analysed in order to explore the composition of the phreatic aquifer groundwater and origin of water mineralization using mathematical modeling [8]. Factor analysis was applied to identify the critical pollution indicators for prospecting and delineating the boundaries due to saltwater intrusion and arsenic pollution [10]. Geostatistical tool of ArcGIS was used to predict the nutrient pollution in groundwater [12]. Geostatistical methods were used to study the spatial changes in the nitrate concentration which is the main source of industrial waste water [13]. Nowadays geostatistics is increasingly used for mapping of hydrological parameters. In this study, the watershed of Upper Palar basin was created using spatial analyst in ArcGIS and 62 well locations with 13 set of water quality parameters were demarcated. Since the data had multivariate nature and several of the water quality variables were correlated, factor analysis was used for identifying the underlying major factors that are responsible for groundwater pollution. Spatial distribution mapping was done using geostatistics analyst in ArcGIS 9.3.

II. STUDY AREA

Palar river basin which is flowing through the North Arcot district of Tamil Nadu state lies between 12° 28' 0” N latitude and 80° 10' 0” E longitude. River Palar is the major drainage system in the district, which rises near Nandhidurg in Karnataka and enters the district about 7 km west of Vaniyambadi. During pre-monsoon, this area experiences high temperature of around 44°C and minimum of around 20°C during post-monsoon. The major source of rainfall is northeast monsoon (October to December). The annual recharge of ground water is 127822 ha-m and the present extraction works out to be 94076 ha-m [15]. The total annual rainfall in this area is 800 mm/year. It is an industrial area which exports around 70% of the leather goods from India. Tanneries are mainly small scale industries located in Vaniyambadi, Ambur and Ranipet which are discharging huge amount of salts and heavy metals like chromium that affect the quality of water and soil. Ranipet in Palar basin has been identified as one of the top ten dirtiest and polluted cities in the world according to the New York-based Blacksmith Institute. Tanneries discharging their effluents directly or indirectly into the Palar river. Water quality analysis of the untreated effluent shows that the Total Dissolved Solids (TDS) is in the range of 20000-30000mg/l [14]. Other major constituents of pollution are sodium, chloride, magnesium, sulphate and chromium. Due to high permeable sand along the river course, there exist many tanneries along the river course with good groundwater potential [18]. Bore hole data show that there is only one layer which is made up of alluvium. High hydraulic conductivity in alluvium causes high contaminant transport in the subsurface media. The study area of Palar Basin is given in Fig 1.

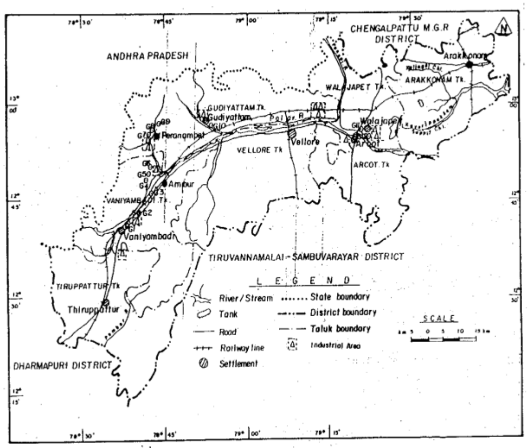

Fig 1. Study area map

A. *Geology and Hydrogeology of the Study Area*

Geologically, the sub-basin is covered by Peninsular Gneisses and embraces a variety of granitic rocks such as massive and porphyritic biotite granite, biotite gneiss of foliated varieties. They are massive, coarse grained granites containing quartz, feldspars in varying proportions with biotite and hornblende as the common ferromagnesium minerals [16]. They have a general trend of N.N.E. – S.S.W. and dip at fairly steep angles to the E.S.E. The alluvium and fractured crystalline rocks acts as aquifer system. Groundwater is occuring in unconfined condition in both alluvium and underlying weathered and fractured rocks. Alluvium occurring in the study area consists of sand, gravel and sandy clay with a thickness of 1 to 6m. The aquifer parameters of the area is shown in the Table.1

Table 1. Aquifer parameters of the study area

Parameters	**Value**
Hydraulic conductivity	20-69 m/day
Transmissivity	80 m^2 / h (along river course) Range 1- 30 m^2 / h
Specific yield value	0.037 – 0.32
Porosity	0.2
Longitudinal dispersivity	30
Lateral dispersivitiy	10

III. MATERIALS AND METHODS

A. *Factor Analysis*

Factor analysis is a statistical approach that can be used to analyze interrelationships among a large number of variables and to explain these variables in terms of their common underlying factors. This approach reduces data into a small set without losing the key information from the original data. In order to carry out the factor analysis, the raw data should be standardized. This step helps in increasing the influence of variables having small variance and reducing the influence of the variables having large variance. Then the correlation coefficients are evaluated which will be helpful in explaining the structure of the underlying system which produced the data [4]. The degree of mutually shared variability between individual pairs of water quality parameters can be represented by correlation coefficient. The correlation coefficient is given as,

$$r_{x,y} = \sum (x-x_m)(y-y_m) / \sqrt{[\sum(x-x_m)]^2[\sum(y-y_m)]^2} \qquad (1)$$

In this expression the correlation coefficients ($r_{x,y}$) is simply the sum (over all samples) of the products of the deviations of the *x*-measurements and the *y*-measurements on each sample, from the mean values of *x* and *y*, respectively, for the complete set of samples [10]. The eigenvalues and factor loadings for the correlation matrix are determined and scree - plot is drawn. One of the most important steps in factor analysis is the extraction of factors based on the variances and co-variances of the variables. The eigenvalues and eigenvectors are evaluated which represent the amount of variance explained by each factor. The factors with eigenvalues greater than one depict more variation in data than individual variable. Finally, by the process of rotation, the loading of each variable on one of the extracted factors is maximized and the loadings on all the other factors are minimized. These factor loadings are useful in grouping the water quality parameters and providing information for interpreting the data.

B.Geostatistics

Using geostatistics tool, a continuous surface can be created when the sample points at different locations are given. Deterministic techniques use mathematical functions for interpolation that are directly based on the surrounding measured values. Geostatistics methods depend on both statistical and mathematical functions that include autocorrelation. Kriging in geostatistics is similar to inverse distance weighting except that the weights are based not only on the distance between the measured sampling points but also on the overall spatial arrangement among the sampling points. The basic assumption in kringing is that the sampling points that are close to each other are similar than those are away. As the first step in geostatistics modeling, the above mentioned assumption is verified using empirical semivariogram. Then the best fit is provided by the line that represents points in the empirical semivariogram cloud graph which quantifies the spatial autocorrelation. Based on the spatial autocorrelation among the measured and predicted locations, kriging weights that are assigned to each measured parameter are determined. Once the kriging weights are determined, the value of the parameter for any unknown location can be calculated. The predicted value for any location can be given by the formula

$$\hat{Z}\ (S_0) = \sum_{i=1}^{N} \lambda i\ Z\ (Si) \qquad (2)$$

Where S_0 = prediction location, λi = unknown weight for the measured sample location, Si = measured value at the *i*th location, N = number of sample locations.

IV. Results and discussion

The statistical summary and descriptive statistics of ground water represented by 13 water quality parameters for 62 wells in the study area during post-monsoon and pre-monsoon seasons of the year 2007 are given in Tables 2 and 3. Piper diagram has been extensively used to understand the similarities and differences in the composition of waters and to classify them into certain chemical types. The concentrations are plotted as percentages with each point representing a chemical analysis. The cations and anions are shown by separate ternary plots. The apexes of the cations plot are calcium, magnesium and sodium plus potassium ions. The apexes of the anion plot are sulphate, chloride and carbonate plus bicarbonate ions. The two ternary plots are then projected onto a diamond. The diamond is a matrix transformation of a graph of the anions (sulfate + chloride/ total anions) and cations (sodium + potassium/total cations). The water quality data is used to plot the piper trilinear diagram which demonstrates that Na^+- K^+ and HCO_3^-- Cl^- facies are most predominant parameters in the groundwater of Palar region. The depicted result indicates that the source of Na+ and Cl- ions is mainly the effluents of tanneries, since lot of sodium chloride is being used for processing raw hides. The origin of K+ may be from the geologic formations, such as potassium bearing feldspars. But the concentration of K+ in the samples is not high. Piper diagrams for Post-monsoon and Pre-monsoon (2007) are depicted in the Fig 2 and 3.

Factor analysis was applied to the data set of the study area to examine the relationship between water parameters and to identify the factors that influence the concentration of each parameter. The univariate statistics corresponding to all water quality parameters, eigenvalues for different factors, % variance and cumulative % variance, correlation matrix, scree plot, factor scores, component loadings were obtained through factor analysis. The components having an eigenvalue greater than 1 (Kaiser Criterion) are considered while analyzing the highly correlated factors. Factors having lower eigenvalues contribute little to the explanation of the variables. Varimax rotation was done to create a new set of variables that replaces original set of data for further analysis. Each component loading greater than 0.6 is taken into consideration during interpretation. Principle components that are identified through component scores are the direct indicators of the hydrogeological processes occurring in the subsurface. Zero or near zero score indicates that the areas are affected at an average degree by the hydrochemical processes. Extreme negative scores indicated that the areas are unaffected by the processes. The regression analysis was done for the data sets obtained for both post and pre-monsoon and the correlation matrix is presented in Tables 4 and 5. From the values given in the table, it is observed that the groundwater quality is mainly controlled by TDS, Hardness, EC, Cl , Na ,Ca ,Mg and SO_4. Again the high positive correlation of hardness with Mg^{2+}(r=0.96), Ca^{2+} (r=0.82) and Cl^- (r=0.9) indicates the dominance of these ions, and the negative correlation shows that K^+ is not a major contributor of pollution. All values are in mg/l except pH and EC (μS/cm).

The variables or factors having eigenvalues greater than one percent of variance and cumulative percent of total variance of groundwater data are presented in the Table 6. From the above table, it is inferred that the first four components together account for 77.40 and 77.04% of the total variance in post and pre-monsoon respectively. The first factor has a high loading of Na, Cl, Mg, Ca, hardness, EC and TDS and accounts for 57.07% (post-monsoon) and 51.6 (pre-monsoon) of the total variance. The association of TDS with higher concentrations of Ca^{2+}, Na^+, Cl^- is due to tanneries which use calcium carbonate, sodium chloride, sodium sulphide, sodium dichromate and sulphuric acid to process raw hides and discharge their effluents into the river course. Ca and Mg have moderate factor loading during pre-monsoon and are slightly less than that of post-monsoon. This is probably due to the precipitation of calcium carbonate in the pre-monsoon and more dissolution of calcium in the post-monsoon. Here calcium and magnesium are responsible for the temporary hardness and both of them are important constituents of the most of the igneous and metamorphic rocks. This is also due to carbonate rock dissolution and rock water interaction that occur in the aquifer with granite, peninsular gneiss and carbonate rich rocks.

The second factor explains 11.78% (post-monsoon) and13.8% (pre-monsoon). Potassium and bicarbonate ions have a factor loading of 0.93 and 0.66 respectively during post-monsoon. The existence of potassium is mainly due to rock water interactions on Potassium bearing feldspars, clay minerals such as illlite and biotite-rich minerals. Bicarbonate mainly originates from the dissolution of carbonate rocks present in the aquifer. It also results from the reaction of CO_2 with silicate minerals. The third factor accounts for 8.55% (post-monson) and (pre-monsoon) of total variance. During pre-monsoon, the concentration of fluoride has high loading factor of 0.88. The important source of fluoride in groundwater is fluoride bearing minerals such as fluorite, apatite, amphiboles and micas. In the Palar region, the source of fluoride may be due to ion exchange of F, leaching of F containing minerals, higher evapotranspiration and longer residence time of water in the aquifer.

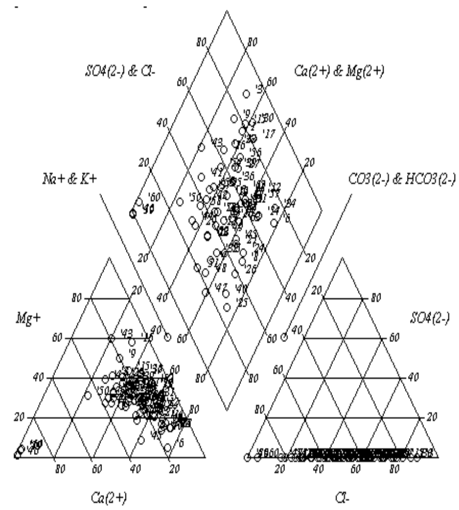

Fig. 2. Piper trilinear plot of Palar river basin (post- monsson)

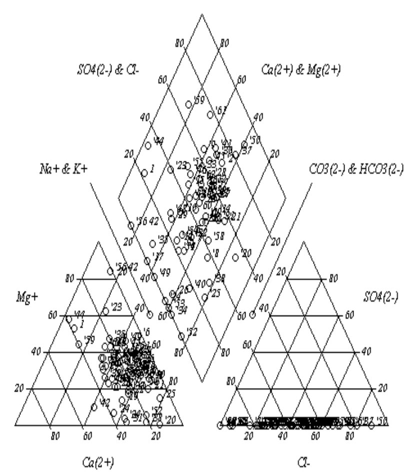

Fig. 3 Piper trilinear plot of Palar river basin (pre-monsson)

The groundwater quality mapping for six water quality parameters that were identified by factor analysis was done using ArcGIS 9.3. Before performing the spatial modeling, it is necessary to check for the normality of the experimental distribution of the available data [11]. If any non-normality is identified, using logarithmic or Box-Cox transformations, the data are driven to normal distribution. In this study, all the data related to six water quality parameters were transformed to be close to the normal distributed data using log transformation as given in Fig 4. The statistical summary of the groundwater quality parameters is represented in Table 7. The semivariogram cloud graph was created then to identify the local and global outliers and check the abnormalities in the groundwater quality data. Quantifying the spatial structure of the data and making the prediction are done by kriging. Then the process of fitting a semivariogram model is done and the performance of the different models (circular, spherical, tetraspherical, exponential, guassian) have been compared based on the nugget-still rate which is used in the classification of spatial dependency. Based on the values of the ratio, the strength of the spatial dependency is determined. If the value is less than 0.25, then there can be strong spatial dependency and rate between 0.25 and 0.75 shows moderate dependency. A weak spatial dependency exists in the classification when the ratio is above 0.75 [13], [17]. The suitable model for fitting the experimental variogram was selected based on less RSS value. The model and the parameters of the groundwater quality variograms are given in Table 8. In this study, ordinary kriging method is used for the prediction and estimation of groundwater quality parameters such as hardness, EC, Cl^-, $Ca^{2+,}$ Mg^{2+}, Na^+. The spatial distribution map of the concentrations of the above mentioned parameters are mentioned in Fig 5 and 6. From the spatial distribution maps, it is observed that the highest concentrations of Na, Cl^-, and conductivity occur in the southwest and northeast regions of the study area where many tanneries are located. The concentration of Mg^{2+} is high in southwestern region during post-monsoon season. Concentration of potassium is high in southeast region when compared to south west region of the study area. The groundwater quality is found to be hard in south west region during the post-monsoon than during pre-monsoon because of high dissolution of ions. The concentration of Ca^{2+} is high in both southwest and northeast regions during the post-monsoon due to the discharge of pollutants and rock-water interaction and is high in the central region during the pre-monsoon due to groundwater flow.

V. CONCLUSION

The groundwater quality analysis was carried out on the data obtained for 62 wells in the region of Palar river basin using multivariate analysis and geostatistics tool. The underlying structure of the groundwater system was explained by three major factors in which Na^+, Ca^{2+},K^+, TDS, EC, Cl^- and hardness were classified under the first major factor having high factor loading. Then using ordinary kriging, a geostatistic model and concentration distribution map for the above mentioned parameters were prepared to describe the spatial and temporal behavior of the hydrochemical parameters. Initially, the groundwater quality data was lognormally distributed. The spherical model was identified to be the best model to represent the spatial variability of Ca^{2+}, TDS, hardness and EC, whereas exponential model was found to be best for Mg^{2+}, $K^{+.}$. From the concentration distribution maps, it was observed that the southwest and northeast regions of the study area are affected by all the water quality parameters estimated.

REFERENCES

[1] A. A. Cronin, J. A. C. Barth, T Elliot, R M Kalin , "Recharge velocity and geochemical evolution for the Permo-Triassic Sherwood sandstone", Northern Ireland, J. Hydrology., 2005, 315 (1-4), 308 -324.

[2] A. K. Gupta, S. K. Gupta and Patil. R.S, "Statistical analyses of coastal water quality for a port and harbour region in India". *Environ. Monit. Asses.* 102, 2005,pp.179-200.

[3] N .Subba Rao, D. John Devadas and K..V.Srinivasa Rao, "Interpretation of groundwater quality using principal component analysis from Anantapur district, Andhra Pradesh, India",Environmental Geosciences, v. 13, no.4, 2006, pp. 239–259.

[4] Hülya Boyacioglu "Surface water quality assessment using factor analysis", ISSN 0378-4738 = Water SA Vol. 32 No. 3, 2006,pp.389-394.

[5] M. Kumaresan and P. Riyazuddin "Factor analysis and linear regression model (LRM) of metal speciation and physico-chemical characters of groundwater samples", Environ Monit Assess 138, 2008,pp. 65–79.

[6] Madhumitha Das "Identifintificcation of effluent quality indicators for use in irrigation", journal of scientific and industrial resrech,Vol.68, 2009,pp 634-39.

[7] Nazire Mazlum, Adem ¨Ozer, and Suleyman Mazlum, "Interpretation of Water Quality Data by Principal Components Analysis", Tr .J. of Engineering and Environmental Science. 23, 1999, 19 -26.

[8] I. Chenini and S. Khemiri, "Evaluation of groundwater quality using multiple linear regression and structural equation modeling", Int. J. Environ. Sci. Tech., 6 (3), 2009,pp.509-519.

[9] J. K. Pathak*, Mohd Alam and Shikha Sharma, "Interpretation of Groundwater Quality Using Multivariate Statistical Technique in Moradabad City, Western Uttar Pradesh State, India", E-Journal of Chemistry Vol. 5, No.3, 2008,pp. 607-619.

[10] Chen-Wuing Liu, Kao-Hung Lin, Yi-Ming Kuo. "Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan", The Science of the Total Environment 313, 2003, pp. 77–89

[11] Salah, Hamad , "Geostatistical analysis of groundwater levels in the south Al Jabal Al Akhdar area using GIS Ostrava, 25,2009.

[12] A.H. Gehan..Sallam and Mohamed Embaby, "Geostatistical analyst as a tool to predict nutrients pollution in water",Thirteenth International Water Technology Conference, Hurghada, Egypt, IWTC 13, 2009,pp.1201-1211

[13] Mevlut Uyan, and Tayfun Cay, "Geostatistical methods for mapping groundwater nitrate concentrations", 3rd International Conference on Cartography And GIS, Nessebar, Bulgaria, 2010.

[14] C. P Gupta, M. Thangarajan and V. V. S. G. Rao, "Groundwater Pollution in the Upper Palar Basin Tamilnadu", National Geophysical Research Institute, 1994.

[15] S. Mohan and M. Muthukumaran.. "Modelling of pollutant transport in groundwater". I E. *(I)* Journal-EN , 2004 , pp. 22-32.

[16] N. Rajmohan and L. Elango"Identification and evolution of hydrogeochemical processes in the groundwater environment in an area of the Palar and Cheyyar River Basins", Southern India, Environmental Geology (2004) 46:47-61

[17] R. Taghizadeh Mehrjardi M. Zareian Jahromi, Sh. Mahmodi and A. Heidari, Spatial Distribution of Groundwater Quality with Geostatistics (Case Study: Yazd-Ardakan Plain), World Applied Sciences Journal 4(1):09-17,2008

[18] M Thangrajan, Quantification of pollutant migration in the groundwater regime through mathematical modeling, Research Communications, Current Science, Vol 76, No, 1, 1999.

Table 2. Statistical summary of hydrochemical parameters of groundwater

	Postmonsoon			Premonsoon		
	Min	Max	Average	Min	Max	Average
Ca^{++}	28	300	79	18	240	75
Mg^{++}	12	316	72	24	199	70
Na^{+}	41	644	217	5	621	199
K^{+}	1	246	16	2	196	15
HCO_3^-	134	781	375	238	946	468
SO_4^-	29	648	147	7	394	126
Cl^-	46	1383	313	25	1170	266
NO_3^-	1	129	20	0	46	11
F	0	2	1	0	3	1
TDS	368	3518	1125	388	2686	1033
Hardness	155	1900	493	145	1080	477
EC	550	6290	1887	670	4210	1628
pH	7	9	8	7	9	8

Table 3. Descriptive statistics

Variable	N	Postmonsoon			Premonsoon		
	N	Mean	Std Dev.	Std Err	Mean	Std Dev.	Std Err
Ca^{++}	62	79.302	45.964	5.791	75.365	47.140	5.939
Mg^{++}	62	71.587	58.602	7.383	70.254	36.010	4.537
Na^{+}	62	216.714	142.267	17.924	199095	115.039	14.494
K^{+}	62	15.841	37.622	4.740	14.746	31.161	3.926
HCO_3^-	62	375.048	149.469	18.831	467825	143.778	18.114
SO_4^-	62	147.349	115.564	14.560	125556	77.687	9.788
Cl^-	62	313.444	263.040	33.140	265873	211.091	26.595
NO_3^-	62	20.206	24.040	3.029	10.810	10.331	1.302
F	62	0.957	0.444	0.056	0.803	0.732	0.092
TDS	62	1125.365	669.865	84.395	1033.095	485.261	61.137
Hardness	62	492.508	325.914	41.061	477302	217.088	27.351
EC	62	1886.984	1124.009	141.612	1628254	743.094	93.621
Ph	62	8.049	0.314	0.040	7.763	0.278	0.035

Table 4. Correlation matrix (postmonsoon)

Chemical variables	Ca^{++}	Mg^{++}	Na^{+}	K^{+}	HCO_3^-	SO_4^-	Cl^-	NO_3^-	F	TDS	Hardness	EC	pH
Ca^{++}	1.000												
Mg^{++}	0.631	1.000											
Na^{+}	0.545	0.561	1.000										
K^{+}	0.029	0.036	0.089	1.000									
HCO_3^-	0.319	0.381	0.611	0.448	1.000								
SO_4^-	0.484	0.548	0.751	0.099	0.361	1.000							
Cl^-	0.773	0.862	0.813	0.005	0.398	0.636	1.000						
NO_3^-	0.637	0.622	0.513	0.019	0.27	0.299	0.617	1.000					
F	0.066	0.284	0.353	0.072	0.425	0.199	0.262	0.093	1.000				
TDS	0.763	0.821	0.903	0.146	0.584	0.757	0.944	0.69	0.3	1.000			
Hardness	0.819	0.962	0.607	0.018	0.394	0.576	0.91	0.684	0.233	0.876	1.000		
EC	0.756	0.842	0.889	0.127	0.564	0.746	0.957	0.674	0.3	0.996	0.889	1.000	
pH	0.494	0.183	0.172	0.049	0.056	0.176	0.308	0.219	-0.03	0.268	-0.31	0.278	1.000

Table 5. Correlation matrix (premonsoon)

Chemical Variables	Ca^{++}	Mg^{++}	Na^{+}	K^{+}	HCO_3^-	SO_4^-	Cl^-	NO_3^-	F	TDS	Hardness	EC	pH
Ca^{++}	1.000												
Mg^{++}	0.321	1.000											
Na^{+}	0.369	0.540	1.000										
K^{+}	0.036	0.002	0.189	1.000									
HCO_3^-	0.114	0.362	0.598	0.439	1.000								
SO_4^-	0.486	0.654	0.698	0.086	0.397								
Cl^-	0.666	0.719	0.799	0.080	0.235	0.646	1.000						
NO_3^-	0.259	0.292	0.439	0.234	0.364	0.354	0.281	1.000					
F	0.276	0.286	0.050	0.289	0.255	0.276	0.149	0.062	1.000				
TDS	0.618	0.731	0.914	0.248	0.556	0.799	0.912	0.493	0.011	1.000			
Hardness	0.763	0.857	0.570	0.022	0.309	0.711	0.853	0.340	0.045	0.836	1.000		
EC	0.651	0.750	0.883	0.217	0.490	0.778	0.938	0.470	0.033	0.991	0.867	1.000	
pH	0.424	0.106	0.150	0.002	0.013	0.110	0.284	0.246	0.374	0.249	-0.302	0.279	1.000

Table 6. Varimax rotated factor loadings

Variable	Postmonoon				Premonsoon			
	Factor 1	Factor 2	Factor 3	C	Factor 1	Factor 2	Factor 3	C
Ca^{++}	0.868	0.052	-0.240	0.814	0.652	-0.551	-0.043	0.730
Mg^{++}	0.869	-0.048	0.177	0.789	0.850	0.192	0.037	0.762
Na^{+}	0.757	0.281	0.388	0.804	0.750	0.011	0.470	0.784
K^{+}	-0.046	0.932	-0.135	0.889	-0.104	-0.205	0.840	0.758
HCO_3^-	0.386	0.662	0.431	0.772	0.342	0.275	0.747	0.751
SO_4^-	0.667	0.215	0.253	0.555	0.829	0.146	0.185	0.742
Cl^-	0.949	0.037	0.138	0.922	0.894	-0.283	0.098	0.889
NO_3^-	0.747	0.022	-0.117	0.572	0.358	-0.067	0.502	0.384
F	0.199	-0.046	0.808	0.695	0.196	0.882	-0.108	0.828
TDS	0.941	0.244	0.204	0.987	0.903	-0.114	0.394	0.984
Hardness	0.948	-0.018	0.046	0.902	0.935	-0.168	0.003	0.903
EC	0.946	0.216	0.203	0.983	0.917	-0.168	0.331	0.979
pH	-0.437	0.072	0.429	0.380	-0.235	0.682	0.010	0.521
Eigen value	7.419	1.532	1.112		6.713	1.803	1.499	
% of Var.	57.072	11.782	8.552		51.641	13.867	11.532	
Cum. %	57.072	68.854	77.406		51.641	65.508	77.040	

Table 7. Statistical summary

Variable	Postmonsoon						Premonsoon					
	Min	Max	Mean	Std.dev	Kurtosis	Skewness	Min	Max	Mean	Std.dev	Kurtosis	Skewness
Ca^{++}	3.332	5.7032	4.2418	0.48957	3.1686	0.45817	2.89	5.480	4.142	0.58388	2.6011	0.1581
Mg^{++}	2.48	5.5755	4.0269	0.6922	2.9186	0.26852	3.1781	5.293	4.149	0.45582	2.7792	0.40714
K^{+}	0	5.5053	1.8806	1.0593	5.4903	1.5573	0.6931	5.278	1.966	0.9499	6.0441	1.8048
TDS	5.908	8.165	6.8825	0.5385	2.4276	0.28864	5.951	7.895	6.844	0.4355	2.2272	0.4069
Hardness	5.043	7.5496	6.0396	0.54773	3.0915	0.49137	4.9767	6.984	6.075	.043491	2.5282	0.18417
EC	6.309	8.7467	7.4002	0.5366	2.5785	0.25203	6.5073	8.345	7.303	0.42525	2.2469	0.37702

Table 8. Semivariance parameters for the logarithmically transformed groundwater quality data

S No	Groundwater Quality Parameter	Model	Nugget	Still	Nugget/Still	R^2
1	Ca^{++}	Spherical	0.09117	0.6151	0.145	0.9181
2	Mg^{++}	Exponential	0.06973	0.9895	0.0704	0.9144
3	K^{+}	Exponential	0.28103	1.476	0.190	0.9179
4	TDS	Spherical	0.144899	0.4693	0.308	0.9697
5	Hardness	Spherical	0.14412	0.55018	0.261	0.9485
6	EC	Spherical	0.1471	0.5129	0.2868	0.9278

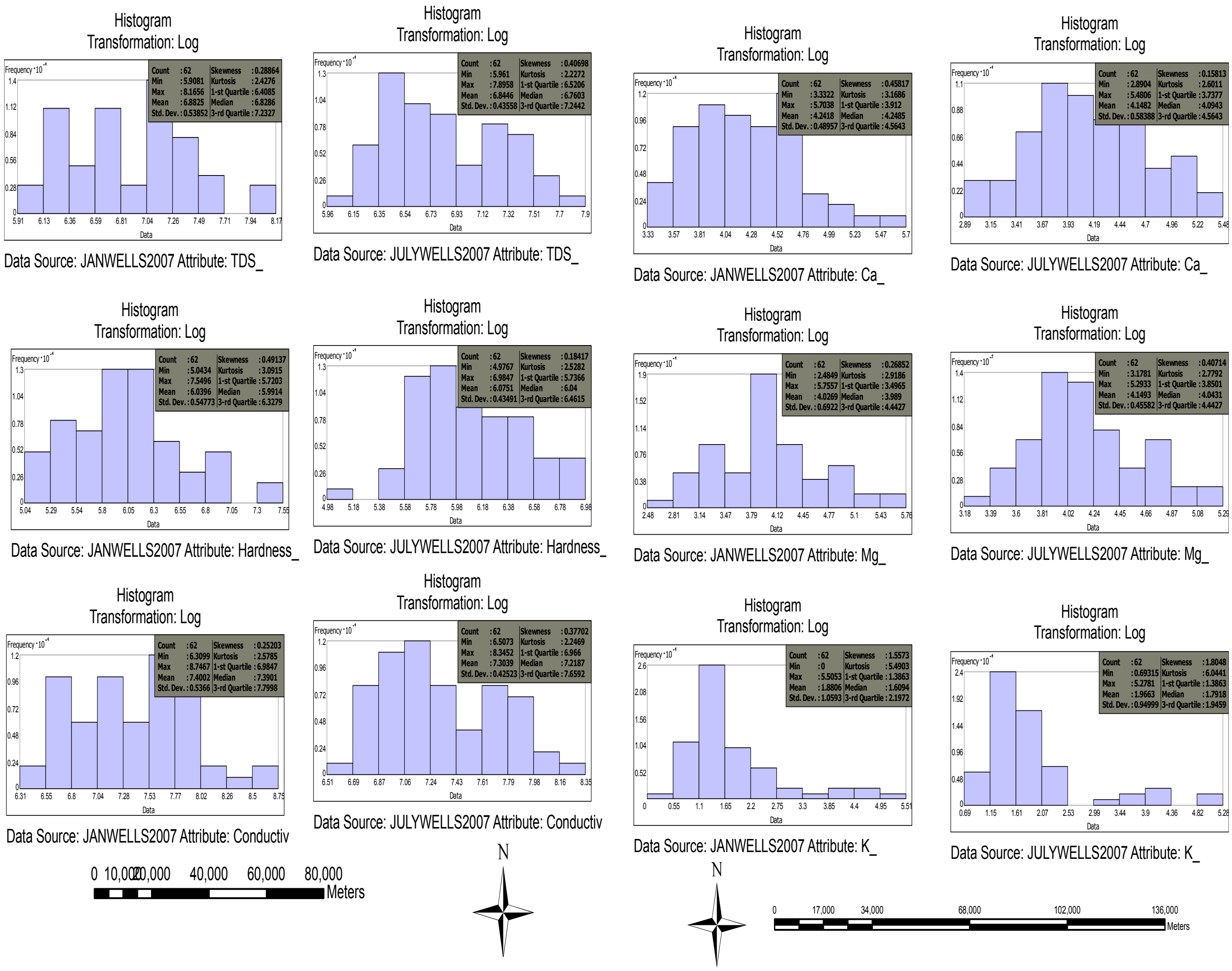

Fig 4. Histograms of log-transformed groundwater quality data

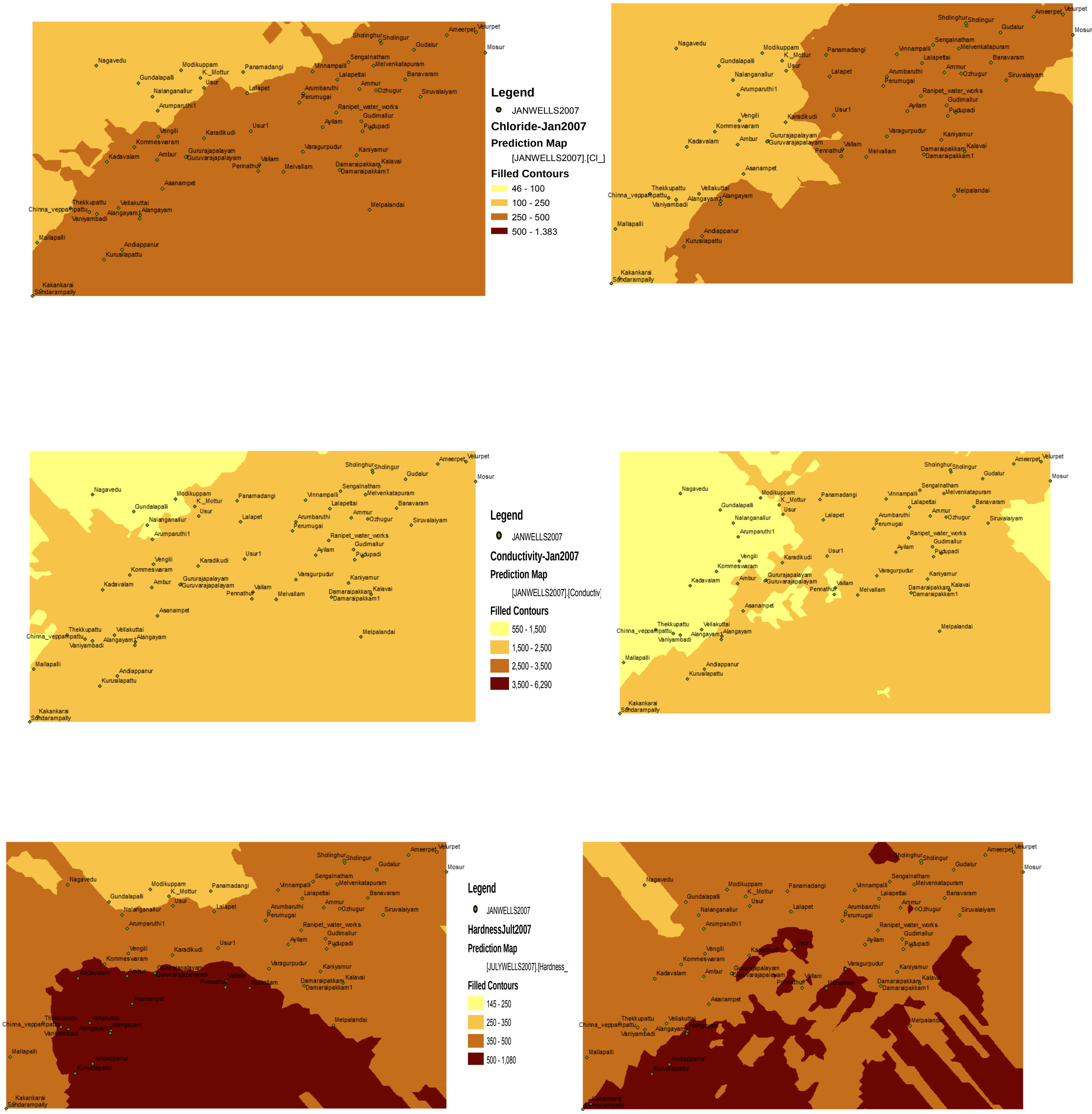

Fig 5. Spatial distribution map of groundwater quality parameters (hardness, EC, Cl^-)

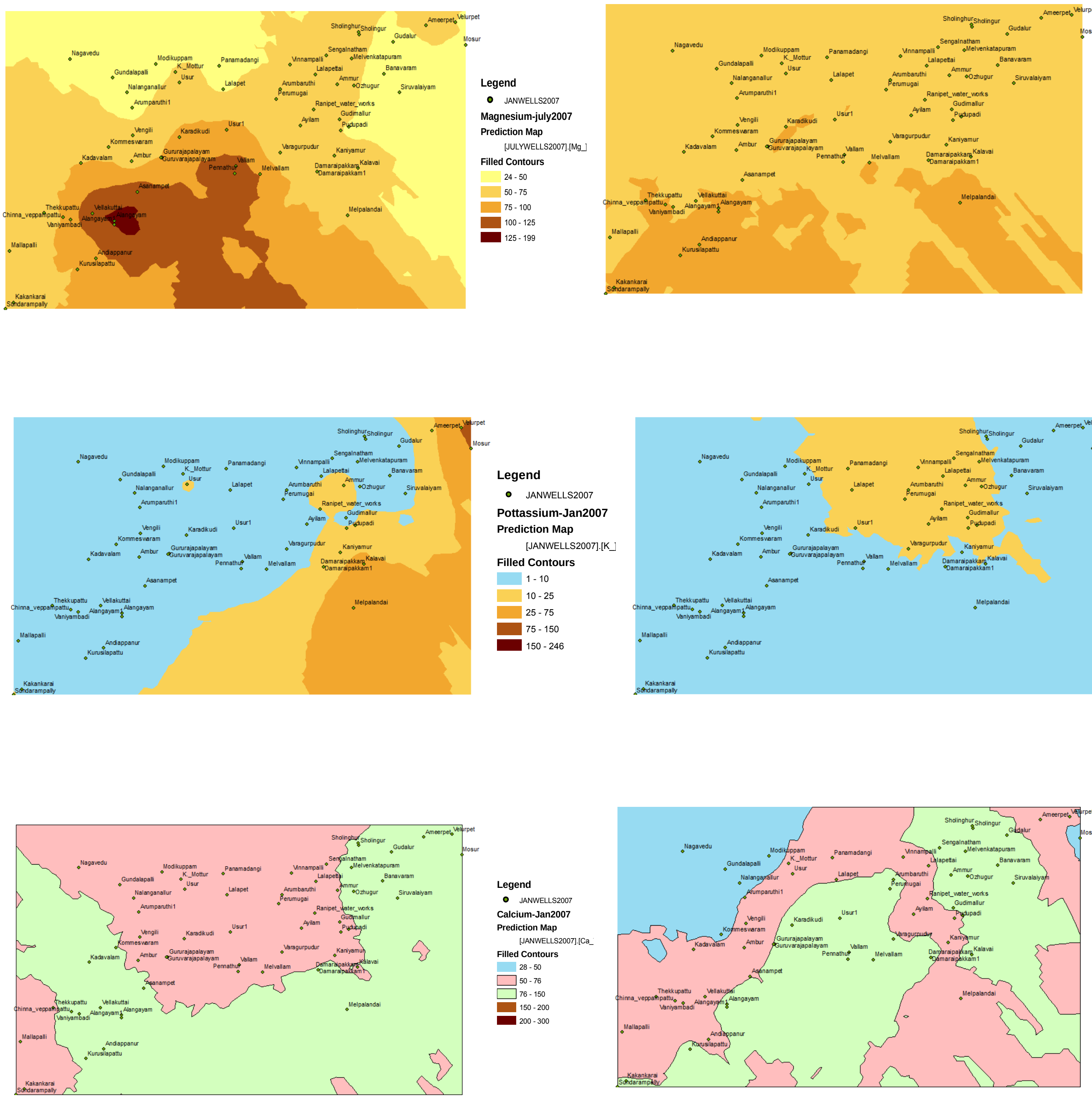

Fig 6: Spatial distribution map of groundwater quality parameters (hardness, EC, Cl^{-})