Statistical evaluation of groundwater geochemistry: a case study between Chinnakuppam and Kulathur, South Chennai, Tamil Nadu, India

Amadeius Abin\textsuperscript{a}, Sibin Raj\textsuperscript{a}, T.B. Aghil\textsuperscript{b}, N.S. Magesh\textsuperscript{b}, S.G.D. Sridhar\textsuperscript{a}

\textsuperscript{a}Department of Applied Geology, University of Madras, Chennai, Tamil Nadu 600 025, India
\textsuperscript{b}Centre for Geotechnology, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu 627 012, India

ABSTRACT

The groundwater quality in the southeast coast of India is severely affected by sea water intrusion and it is a growing concern to the masses. Lack of sufficient recharge with only one prevailing monsoon season, the quality and quantity of the groundwater aquifers in this region are deteriorating. A geochemical study was conducted on the southern part of Chennai metropolitan city in the state of Tamil Nadu, India along the coastal reaches on developing sub-urban population areas. Groundwater samples from different locations along the coast were collected using random sampling and analyzed for various physico-chemical parameters. The results were tabulated and further statistical work was carried out. The statistical approach provides a means to estimate the distribution characteristics of the data thus provide a means for the development of hypothetical models. The primary objective of the analysis was to verify certain observation found in the earlier analysis about the distribution properties of geochemical data. Tests were conducted for the goodness of fit on log-normal distribution with 5% confidence interval. Eight parameters were tested on the probability distribution plot to estimate the best fit. It has been confirmed that the geochemical values does follow the log-normal distribution. As such, after obtaining the required parameters of suitable sample size, hypothetical models can be developed using this observation. The analysis support the earlier assumptions, suggesting an effective method of evaluation and estimation of errors.

1. Introduction

The numerical analysis of geochemical data is still an ill ventured horizon in the geological society and it is often discarded for the surreptitious mathematical complex equations that often accompany these analyses. Apart from stifling mathematical equations, the scope for statistical examination is valid without indulging in the extreme complexities.

The proximity to urban landscapes exposes the in-situ groundwater resources to rapid encroachment, depletion and contamination. The natural influx of the saline water to the groundwater aquifers is a gruesome factor that further deepens the growing concern. Improper planning and rapid industrialization of the suburbs have paid heavy lapse on the groundwater potentials instilling fresh flow of saline water.

Hydro-geochemical processes such as dissolution, precipitation, ion-exchange, and the residence time along the flow path control the chemical composition of the groundwater in the shallow alluvial aquifers (Apodaca et al. 2002). Groundwater condition along the south-east coast is a great concern due to the declining water level and increasing risk of salinity. The occurrence of only one prevailing monsoon underlies these concerns. The ever increasing quench of the limited water resources crush the strategic aquifers and pose serious threat for future generation. Geochemical qualitative analysis is a direct estimate of the contamination providing vital clues regarding the quality and saline intrusion in the coastal region. The critical information regarding the complexity and vitality of the groundwater can be derived from these values. The sample collection and laboratory analysis often display heterogeneity in error that etches the final observation. However, for quantitative measurement of this error and to validate the final observations there is a need to exist a mathematical model. The observed data figures often follow a standard distribution. It has been mentioned by several authors that geochemical data are approximately log-normally distributed (Finney 1941; Ahrens 1965; Allegre and Lewin 1995).

The aim of the present work is to study the distribution of the data and test the goodness of fit of the cumulative frequency distribution to the standard distribution curves, with the help of present day computer applications.

2. Study area

The study area is situated along the northeastern coast of Tamil Nadu that includes Chinnakuppam, Vilambur, Tembakkam, Marakanam, and Kolathur from North to South. These sites fall in Kanchipuram district along the coast that stretches for about 30 km. Geographically, it is located between 12°20’ and 12°52’30” N latitude...
and from 79°20’ to 80°17’30” E longitudes. The area is characterized by sand dunes of quaternary and recent period, Cuddalore sandstone of mio-pliocene age, shales and sandstone of upper gondwannas and charnockites of archean era (Chidambaram et al. 2011). The greater part of the area mainly consists of metamorphic crystalline rocks of archean age that belongs to the charnockite and khondalite groups (CGWB 2001). The upper most formation is coastal alluvium which is few meters thick and is underlined by thick sequences of crystalline rock. The quaternary/recent sediments weathered and fractured crystalline charnockite function as an unconfined aquifer system. The groundwater head follows the topography, i.e., it flows towards east and west from the central part of the area. The groundwater requirement of this area is mostly met by wells penetrating up to upper quaternary/recent sediments. During the study it is noticed that most of the dug well are tapping ground water from recent alluvium at a depth ranging from 0.5 m to 9.8 m. The study area map is shown in the figure1.

The coastal area of Kanchipuram district is characterized by several strand lines, lagoons, mangroves, salt marsh, estuaries, creeks, barrier dunes, spits, beach terraces. The major water bearing formation being the coastal sands extending 30-40 feet, and the sand thickness is increasing towards the coast. In many places, the coastal water is brackish in nature due to the presence of Buckingham canal running parallel to the coast. The major zone of groundwater occurrence is the Palar mouth where the Palar river bed acts as a potential formation. The important mineral found in this stretch is silica sands along the Mamallapuram coast.

3. Methodology

A total of 30 water samples were collected from Chinnakuppam in the north to Kolathur in the south of the study area during pre-monsoon (June, 2011) and post-monsoon (January, 2012) from bore wells and dug wells to assess the quality of groundwater. The water samples were numbered starting from 1 to 30 that are collected from Chinnakuppam to Kolathur along East Coast Road (ECR). The EC and pH were measured on the field. According to APHA (1995) the groundwater samples were analyzed in the geochemical laboratory to find out anions and cations. Pre-cleaned polyethylene bottles were used for collecting water samples from open wells and bore wells. Groundwater samples were collected using simple random sampling method. Electrical Conductivity (EC) and pH were measured using conductivity and pH meters. TDS was computed from EC multiplied by 0.64. Na+ and K+ ions were determined by flame photometer. Total hardness (TH) as CaCO3, total alkalinity (TA) as CaCO3, carbonates, bicarbonate and chloride were analyzed by volumetric methods. Mg was calculated from TH and Ca contents. SO4 were estimated by Nephelometer techniques. Fluoride (F) was determined using the SPADNS method. All concentrations are expressed in milligrams per liter (mg/l). The results were tabled and statistical parameters were calculated. Means and standard deviations were estimated using the robust log-probability (MR) method, percentile concentrations and inter-quartile ranges (IQR) were estimated by the maximum likelihood estimation (MLE) method, as described by Helsel (1990) and Helsel and Hirsch (1992). Computer application Minitab was used to test the goodness of fit with 5% confidence intervals.

4. Results and Discussions

Data represented in table 1 demonstrate each statistical parameters used in the present study. Analysis was primarily concentrated on the parameters mean, SD, variance and distribution parameters: skewness and kurtosis. Highest mean values are calculated for the EC values with median at 490. TDS similarly has a higher value; however in the case of elemental concentration sodium (Na) and bicarbonate (HCO3) have major peaks. TDS also display a high variation with SD of 276.95 while EC has highest standard deviation among variable analyzed. The increase in TDS in some well may be attributed by the nature of overlying rock material. In area having more clay content the TDS generally increases with recharge due to rainfall. It is also possible that rising level in the backwater contributes to the increase
of ionic concentration. The high values of sodium and bicarbonate are observed in the distribution analysis with sodium having a mean of 71 and bicarbonate with 147. The SD is also having a higher value in the range of 59 and 110 which is in lieu with observed mean values. The median values are 54 and 126 for these ions.

Table 1. Statistical variation of different physico-chemical parameters in the study area

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>StDev</th>
<th>Variance</th>
<th>CVariance</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Sum1</th>
<th>SSQ1</th>
<th>Skewness1</th>
<th>Kurtosis1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDS</td>
<td>438.40</td>
<td>276.95</td>
<td>7670.90</td>
<td>68.71</td>
<td>338.00</td>
<td>366.00</td>
<td>400.00</td>
<td>490.00</td>
<td>13152.00</td>
<td>1.05</td>
<td>1.48</td>
</tr>
<tr>
<td>EC</td>
<td>631.60</td>
<td>400.42</td>
<td>160338.80</td>
<td>52.49</td>
<td>490.00</td>
<td>519.50</td>
<td>599.00</td>
<td>771.00</td>
<td>16617382.00</td>
<td>1.08</td>
<td>1.58</td>
</tr>
<tr>
<td>pH</td>
<td>6.81</td>
<td>0.53</td>
<td>0.28</td>
<td>7.74</td>
<td>6.54</td>
<td>6.95</td>
<td>7.20</td>
<td>204.22</td>
<td>1398.25</td>
<td>-1.06</td>
<td>0.99</td>
</tr>
<tr>
<td>HCO3</td>
<td>147.30</td>
<td>110.51</td>
<td>12212.08</td>
<td>57.10</td>
<td>126.00</td>
<td>190.50</td>
<td>4419.00</td>
<td>1005069.00</td>
<td>1.46</td>
<td>2.43</td>
<td></td>
</tr>
<tr>
<td>TH</td>
<td>137.00</td>
<td>75.29</td>
<td>5668.90</td>
<td>54.96</td>
<td>118.50</td>
<td>196.00</td>
<td>4110.00</td>
<td>727468.00</td>
<td>0.48</td>
<td>-0.49</td>
<td></td>
</tr>
<tr>
<td>Ca</td>
<td>33.30</td>
<td>18.64</td>
<td>347.32</td>
<td>55.97</td>
<td>30.00</td>
<td>52.00</td>
<td>999.00</td>
<td>43339.00</td>
<td>0.41</td>
<td>-0.59</td>
<td></td>
</tr>
<tr>
<td>Mg</td>
<td>13.00</td>
<td>7.28</td>
<td>53.03</td>
<td>56.02</td>
<td>11.00</td>
<td>19.00</td>
<td>390.00</td>
<td>66080.00</td>
<td>0.60</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td>Na</td>
<td>71.53</td>
<td>59.21</td>
<td>3506.05</td>
<td>82.78</td>
<td>54.00</td>
<td>116.00</td>
<td>2146.00</td>
<td>255186.00</td>
<td>1.14</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>SO4</td>
<td>14.20</td>
<td>21.73</td>
<td>472.23</td>
<td>153.03</td>
<td>4.00</td>
<td>20.50</td>
<td>426.00</td>
<td>19744.00</td>
<td>2.12</td>
<td>4.18</td>
<td></td>
</tr>
</tbody>
</table>

Relatively lower concentrations are observed for the other elements namely magnesium, sulphate and calcium. The mean values for these elements are 33 for calcium, 13 for magnesium, and 14.2 for sulphate. Estimated mean concentrations are often seen to exceed the median values and this reflects the right-skewed cumulative distribution (Gilbert 1987; Helsel 1990). Probability graphs for TDS, pH, TH, Ca, Mg, Cl, Na, HCO3, SO4 are plotted with 95% confidence interval (Fig. 2). The general observation is all the data values tend to follow the log-normal distribution. The probability plot for EC, display sufficient variation with lower p<0.05, suggesting the log normal distribution may not be a best fit for the values. Step like frequency percentiles are observed for the values of sulfates and sodium suggesting a recurrence in data values.
However, a drawback of using probabilistic methods may include the time and expense required to gather a sufficient number of data to accurately represent the modeled system.

Fig. 2 Probability plot of various physico-chemical parameters of groundwater in the study area.

5. Conclusion

Geochemical data represent the overall quality of the groundwater resources and suggest the degree of contamination the aquifers are subjected into. However, the absence of well-developed statistical techniques in the interpretation often leads to anomalous conclusions and errors in the results. Very small sample populations and the absence of supporting data often limit statistical analysis and hinder the final outcomes. The result of this test confirms to the earlier assumptions of the geochemical data following the log-normal distribution. The increasing concentrations of sodium and bicarbonate suggest zones of saline intrusion and contaminated groundwater aquifers. However, development of hypothetical model need to be supported with a large number of samples but will provide an excellent tool in tackling the issues of sustainable development along the coastal regions of the southeastern India.

References


