Study of groundwater contamination and drinking suitability in basaltic terrain of Maharashtra, India through PIG and multivariate statistical techniques

Vasant Wagh, Shrikant Mukate, Aniket Muley, Ajaykumar Kadam, Dipak Panaskar and Abhay Varade

ABSTRACT

The integration of pollution index of groundwater (PIG), multivariate statistical techniques including correlation matrix (CM), principal component analysis (PCA), cluster analysis (CA) and various ionic plots was applied to elucidate the influence of natural and anthropogenic inputs on groundwater chemistry and quality of the Kadava river basin. A total of 80 groundwater samples were collected and analysed for major ions during pre- and post-monsoon seasons of 2012. Analytical results inferred that Ca, Mg, Cl, SO4 and NO3 surpass the desirable limit (DL) and permissible limit (PL) of Bureau of Indian Standards (BIS) and the World Health Organization (WHO) in both the seasons. The elevated content of total dissolved solids (TDS), Cl, SO4, Mg, Na and NO3 is influenced by precipitation and agricultural dominance. PIG results inferred that 52.5 and 35%, 30 and 37.5%, 12.5 and 20%, 2.5 and 5% groundwater samples fall in insignificant, low, moderate and high pollution category (PC) in pre- and post-monsoon seasons, respectively. PC 1 confirms salinity controlled process due to high inputs of TDS, Ca, Mg, Na, Cl and SO4. Also, PC 2 suggests alkalinity influence by pH, CO3, HCO3 and F content. PIG and statistical techniques help to interpret the water quality data in an easier way.

Key words | groundwater, hydrogeochemistry, India, Kadava river, Maharashtra

INTRODUCTION

Nowadays, groundwater has become a vital and reliable source for drinking and agricultural uses in many countries due to inadequate freshwater (Adimalla 2019). Generally, groundwater composition is influenced by geological formations and anthropogenic inputs (Yuan et al. 2018; Mukate et al. 2019a, 2019b). Concerning India, urbanization, industrialization, agricultural chemicals, population stress, etc. pose severe threats to the groundwater quality as extracted from aquifer systems (Pawar et al. 2008; Rao et al. 2014; Panaskar et al. 2016; Wagh et al. 2016a; Varade et al. 2018; Barakat et al. 2019). In recent decades, the drinking of polluted water has raised many health problems such as the toxicity of fluoride, arsenic, boron, nitrate, etc. (Hossain et al. 2013; Pandith et al. 2017; Adimalla et al. 2018a, 2018b; Kadam et al. 2019). Hence, it is essential to evaluate and monitor the groundwater quality to diminish health-related problems and to protect groundwater quality for a sustainable approach. In general, several techniques
have been widely used to identify the geogenic and human-induced actions altering groundwater chemistry (Adimalla 2019; Rao et al. 2019). In this study, the categorization of groundwater has been investigated through multivariate statistical tools (CM, PCA, CA) to identify the dominance of ion exchange, weathering, evaporation, etc., processes controlling groundwater composition and hydrochemical interpretation. This method has been widely employed to resolve hydrochemical, geochemical, geo-environmental problems, etc. at diverse scales (Aryafar & Ardejani 2013; Ebrahimi et al. 2018). Generally, cluster and factor analyses are widely used for representing a complex association among several variables (Davis 1986). PCA is a technique in which the utility of significant parameters are judged by the performance; also, it contains linear transformation of original variables to new variables and their association with each other (Davis & Schwartz 1987). Factor analysis is used in hydrochemistry to identify the considerable affiliation of water quality variables along with other variables and their combined influence on groundwater chemistry (Love et al. 2004). These techniques have been used by many researchers to understand the chemical composition, sources of water contaminants and association between different parameters which influence the water quality (Ko et al. 2010; Rao et al. 2014; Pazand & Javanshir 2016; Kant et al. 2018).

The present research work has been carried out in Kadava river basin which is a tributary of the Godavari River. The total geographical area is 1,053 km², with latitude 19°55’N to 20°25’N and longitude 73°55’E to 74°15’E located in the Nashik district of Maharashtra (Figure 1). The average annual precipitation is 700 mm, mainly received from the south-west monsoonal wind. The area experiences a semi-arid climate and temperature varies from 5 °C to 42 °C in winter and summer seasons with humidity around 45–62%. The study area has an undulating topography with moderate slopes tending towards the south-east direction (CGWB 2014). Agriculture is the prime occupation with a variety of crops, namely, grapes, sugarcane, onions, vegetables, etc., growing throughout the year. The alluvium thickness is around 20–25 m at the south and central parts which contain mostly sand, gravel and kankar deposition. The rock formation is Upper Cretaceous to Lower Eocene age and comprises pahoehoe and a’a lava flows situated in the south-east part of Deccan volcanic province (GSI 2001). Groundwater occurs in unconfined to confined conditions at 20–25 m depth comprising weathered and fractured parts (CGWB 2014). In the river basin, groundwater is largely used for drinking, domestic and irrigation uses; thus, the quality of groundwater is closely allied with human health. However, in the study area, very limited information is available on the groundwater quality and factors responsible for altering hydrogeochemistry; nonetheless, studies by a few central government agencies and researchers reported the elevated content of nitrate and heavy metals in groundwater (CGWB 2014; Wagh et al. 2017a, 2017b, 2018a, 2018b). Generally, shallow alluvium aquifers are highly vulnerable to contamination compared to deep aquifers in basaltic terrain due to the high permeability and porosity behaviour (Pawar et al. 2008). Hence, constant monitoring and categorization of groundwater quality is constantly needed to protect from further groundwater contamination problems which may potentially affect human health (Gaikwad et al. 2019). However, in the study area, the influence of natural and anthropogenic factors on groundwater quality has not been widely available to date. Thus, the present study aims: (1) to categorize the influence of natural and anthropogenic processes on groundwater quality through multivariate statistical techniques; (2) to ascertain the physicochemical characteristics of groundwater for drinking use; (3) to use the pollution index of groundwater (PIG) to characterize the variation in groundwater quality. The outcomes of the study may help in hydrochemical interpretation, source of water contaminants and development of integrated management of water resources in the river basin.

MATERIAL AND METHODS

To understand the hydrogeochemical nature of groundwater, 80 groundwater samples were collected and analysed during pre-monsoon (PRM) and post-monsoon (POM) seasons of 2012. Pre-treated plastic containers were used for sample collection and each well was pumped for 2–3 minutes to diminish the influence of stored pipe water and then transported to the laboratory for further chemical analysis. pH and electrical conductivity (EC) were recorded...
in situ by a handheld multi-parameter tester. Further, major ions, namely, calcium (Ca), magnesium (Mg), sodium (Na) and potassium (K), carbonate (CO₃), bicarbonate (HCO₃), sulphate (SO₄), nitrate (NO₃) and chloride (Cl) were determined by using standard procedures of APHA (APHA 2005). Total dissolved solids (TDS) were computed from EC by multiplying a factor of 0.65 (Hem 1985). Total hardness was calculated by Ca (meq) + Mg (meq) × 50. For analytical precision, ion balance errors (IBE) were computed with the range of error within ±10% (Berner & Berner 1987). The Survey of India (SOI) topographic sheets numbers 46 L/3, L/4, 46 H/15 and H/16 at 1:50,000 scales were used for preparation of the study area base map with ArcGIS 9.3v software. Multivariate statistical analysis was performed by MS-Excel, SPSS 22.0 and R software.

RESULTS AND DISCUSSION

The summary of physicochemical parameters of PRM and POM seasons and their comparison with the Bureau of Indian Standards (BIS) drinking standards is summarized in Table 1.

Figure 2 explores the statistical summary of the cations and anions of PRE and POM seasons of 2012 through box plots. It simply explains the temporal variation of ions in PRM and POM seasons. These plots were prepared using R software and combined the ions’ data of both the seasons in one platform. The line across the box demonstrates the median content of ions in both the seasons. The median value, first and third quartile, range values were determined for each parameter. The box plot of pH, CO₃, HCO₃, Cl, K
Physicochemical characteristics of groundwater

Groundwater shows pH from 7.8 to 8.9 and 7.7 to 8.5 with an average of 8.5 and 8.1 in PRM and POM seasons of 2012, indicative of the alkaline nature owing to loss of CO₂, precipitation and dissolution processes (Pawar & Kale 2006). In general, water containing a high pH would not affect human health; however, it may change the taste of the groundwater and it is closely associated with other water elements (Wagh et al. 2016b). As per the BIS standards, 40 and 2.5% of samples surpass the PL in PRM and POM seasons. EC values vary from 810 to 6,180 μS/cm with an average of 2,173.53 in PRM; conversely, it shows a wide variation of 824–8,120 μS/cm (average 2,741.38 μS/cm) in the POM season. Elevated EC is attributed to dissolution of salts and load of anthropogenic inputs such as agricultural and domestic activities (Morrison et al. 2001). The BIS has not specified any threshold limit for EC; but, according to the World Health Organization (WHO 2011) standards, 23 (57.5%) of the PRM and 30 (75%) of the POM groundwater samples surpass the PL in both the seasons. The mean content of TDS is 739.7 and 861.17 mg/L in PRM and POM seasons. According to BIS standards, 75 and 97.5% samples exceed the DL; moreover, only 2.5% of samples surpass the PL limit in PRM and POM seasons, respectively, due to the influence of agricultural and anthropogenic inputs. As per WHO standards, only 5% of samples exceed the PL in both the seasons. Drinking water containing high TDS may lead to some potential gastrointestinal problems (Mukate et al. 2019a). The calcium content ranges from 13.63 to 98.07 mg/L (average 39.78 mg/L) and 21.04 to 118.34 mg/L (average 50.94 mg/L) in PRM and POM seasons. According to BIS and WHO standards, only 7.5 and 22.5% samples exceed the DL; moreover, only 2.5% of samples surpass the PL limit in PRM and POM seasons, reflecting the precipitation of calcite (Pawar et al. 2008). Magnesium content ranges widely from 20.22 to 231.03 mg/L and 17.35 to 255.86 mg/L with an average value of 71.48 and 93.06 mg/L; however, 15 and 37.5% of samples surpass the PL limit of BIS in PRM and POM seasons. The enrichment of magnesium occurs in the downstream part owing to geological control and Thakurwadi formation.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>6.5–8.5</td>
<td>6.5–8.5</td>
<td>7.8–9</td>
<td>8.5</td>
</tr>
<tr>
<td>EC</td>
<td>500–1,500</td>
<td>–</td>
<td>810–6,180</td>
<td>2,173.53</td>
</tr>
<tr>
<td>TDS</td>
<td>500–1,500</td>
<td>500–2,000</td>
<td>356.82–2,272.96</td>
<td>739.7</td>
</tr>
<tr>
<td>Ca</td>
<td>75–200</td>
<td>75–200</td>
<td>13.65–98.07</td>
<td>39.78</td>
</tr>
<tr>
<td>Mg</td>
<td>50–100</td>
<td>30–100</td>
<td>20.92–231.03</td>
<td>71.48</td>
</tr>
<tr>
<td>Na</td>
<td>200–600</td>
<td>200</td>
<td>36.2–501.4</td>
<td>111.2</td>
</tr>
<tr>
<td>K</td>
<td>10</td>
<td>12</td>
<td>0.3–8.3</td>
<td>2.49</td>
</tr>
<tr>
<td>Cl</td>
<td>250–500</td>
<td>250–1,000</td>
<td>31.24–1,085.8</td>
<td>210.4</td>
</tr>
<tr>
<td>SO₄</td>
<td>200–250</td>
<td>200–400</td>
<td>31.54–261.41</td>
<td>111.9</td>
</tr>
<tr>
<td>NO₃</td>
<td>45</td>
<td>45</td>
<td>12.94–69.76</td>
<td>49.09</td>
</tr>
<tr>
<td>F</td>
<td>1–1.5</td>
<td>1–1.5</td>
<td>0.1–0.7</td>
<td>0.35</td>
</tr>
<tr>
<td>TH</td>
<td>100–500</td>
<td>300–600</td>
<td>131.95–1,028.92</td>
<td>392.43</td>
</tr>
</tbody>
</table>

All values in mg/L except pH; EC in (μS/cm); DL, desirable limit; PL, permissible limit.
containing picritic horizons and olivine and pyroxene mineral released from basalt weathering (Beane et al. 1986; Rabemanana et al. 2005). The consumption of high calcium and magnesium-rich water leads to colon cancer (Yang et al. 1997). Total hardness (TH) content in 15 and 25% of groundwater samples exceeded the PL of drinking water in PRM and POM seasons and is thus restricted for drinking in a few locations. High TH in drinking water leads to scale formation, encrustation in water supply pipes and urolithiasis problems in humans (Wagh et al. 2016b). Sodium content ranges from 36.2 to 501.4 mg/L (average 11.2 mg/L) and 31 to 449.4 mg/L (average 113.88 mg/L) in PRM and POM seasons, respectively. It is revealed that only 15 and 12.5% of samples in the downstream catchment are beyond the allowable limit of the BIS. High content of Na is due to basalt-derived salts along with Salher formation enfolding plagioclase feldspar and anorthite (Stallard & Edmond 1983). The consumption of sodium content in water may cause high blood pressure, vomiting, headache, hyperosmolarity and muscle stiffness in humans (Prasanth et al. 2012).

The K content in groundwater is within the PL of BIS and WHO except for sample number 12; hence, it is fit for drinking. An excessive amount of K in water results in a bitter taste to drinking water (Wagh et al. 2016b). This sample is mainly influenced by K-rich fertilizers from agricultural inputs. Chloride content shows 31.24–1,085.8 mg/L (average 210.4 mg/L) and 63.3–829.94 mg/L (average 229.66 mg/L) in PRM and POM seasons,
respectively. As per the BIS standards, 27.5 and 32.5% of samples exceeded the DL; however, only sample number 38 surpassed the PL in the PRM season. High content of chloride is due to weathering of halite mineral, domestic waste, fertilizers, leachate pollution, etc. (Loizidou & Kapetanos 1993). The ground water quality beyond threshold level may lead to alterations in taste, corrosion and palatability (Mukate et al. 2019b). The values of sulphate content are within the PL; however, only 10 and 30% of samples exceed the DL (200 mg/L) in PRM and POM seasons. The high content of sulphate encountered in the agriculture area is indicative of sulphate contained in fertilizers. The drinking of sulphate-rich water may be the cause of catharsis, laxative effect, gastrointestinal and dehydration problems (WHO 2011). Nitrate content demonstrated a variation of 12.94–69.76 mg/L and 31.65–65.22 mg/L with an average of 49.09 and 52.65 mg/L in PRM and POM seasons. Analytical results confirmed that 67.5 and 75% of samples exceed the PL (45 mg/L) in PRM and POM seasons. Elevated nitrate content corresponds to the use of nitrogen complex fertilizers, domestic waste and raw sewage (Wagh et al. 2019a). An elevated amount of nitrate in drinking water may lead to ‘blue baby syndrome’ or methemoglobinemia, which is a typical state where a baby’s skin becomes blue due to a reduced amount of haemoglobin in the blood (WHO 2011). The values of fluoride content are within the PL of the BIS and WHO (1.5 mg/L) in PRM and POM seasons. Only sample number 24 is more than the PL in the POM season and is attributed to fertilizer inputs. A high content of fluoride in drinking water can lead to dental and skeletal fluorosis (BIS 2012).

Pollution index of groundwater (PIG)

The study applied the PIG proposed by Subba Rao (2012) to ascertain the degree of pollution of groundwater at each sample location. This index measures the status of relative impact on individual water quality parameters. This index is represented in numerical scale and helps to quantify the degree of water contamination. This index suggests a composite pressure of a single water quality variable on overall groundwater quality of a location. The index has been computed by considering the water quality variables, namely, pH, EC, TDS, TH, Ca, Mg, Na, K, Cl, SO₄, NO₃, F and TH. Initially, the relative weight (Rw) of one to five scales was assigned to each parameter depending on their importance in drinking water and public health. Therefore, the minimum weight of 1 was assigned to K; 2 was given to EC and Ca; 3 was considered for HCO₃; 4 was assigned to Na and Cl; and the highest weight of 5 to pH, TDS, TH, Mg, F, SO₄ and NO₃. Further, the weight parameter (Wp) was calculated for each water quality variable to ascertain its relative share on overall quality of groundwater (Equation (1)). Moreover, the status of concentration (Sc) was determined by dividing the content of each water quality parameter by its respective drinking water standards (Ds) (Equation (2)). Also, overall water quality (Ow) was calculated by multiplying the values of Wp with Sc (Equation (3)). Finally, to ascertain the influence of contaminants on the groundwater quality, PIG was computed by taking the sum of all Ow values (Equation (4)). Table 2 summarizes the considered parameters, values of Rw, Wp and Ds. In evaluation of the PIG, the relative input of water quality variables of every water sample was considered. Given the overall quality of water (Ow) is >0.1; then, it accounts for 10% of the value of 1.0 of the PIG which signifies the influence of contamination on the quality of groundwater. Rao et al. (2018) suggested a PIG classification into five categories, i.e.: insignificant pollution (<1.0); low pollution

<table>
<thead>
<tr>
<th>Chemical parameter</th>
<th>Relative weight (Rw)</th>
<th>Weight parameters (Wp)</th>
<th>Drinking water standards (Ds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>5</td>
<td>0.094</td>
<td>7.5</td>
</tr>
<tr>
<td>EC</td>
<td>2</td>
<td>0.038</td>
<td>500</td>
</tr>
<tr>
<td>TDS</td>
<td>5</td>
<td>0.094</td>
<td>500</td>
</tr>
<tr>
<td>TH</td>
<td>5</td>
<td>0.094</td>
<td>300</td>
</tr>
<tr>
<td>Ca</td>
<td>2</td>
<td>0.038</td>
<td>75</td>
</tr>
<tr>
<td>Mg</td>
<td>5</td>
<td>0.094</td>
<td>30</td>
</tr>
<tr>
<td>Na</td>
<td>4</td>
<td>0.075</td>
<td>200</td>
</tr>
<tr>
<td>K</td>
<td>1</td>
<td>0.019</td>
<td>10</td>
</tr>
<tr>
<td>Cl</td>
<td>4</td>
<td>0.075</td>
<td>250</td>
</tr>
<tr>
<td>F</td>
<td>5</td>
<td>0.094</td>
<td>1.5</td>
</tr>
<tr>
<td>HCO₃</td>
<td>3</td>
<td>0.057</td>
<td>300</td>
</tr>
<tr>
<td>SO₄</td>
<td>5</td>
<td>0.094</td>
<td>150</td>
</tr>
<tr>
<td>NO₃</td>
<td>5</td>
<td>0.094</td>
<td>45</td>
</tr>
<tr>
<td>Sum (Σ)</td>
<td>53</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
This classification has been used to classify the groundwater quality into insignificant pollution to very high pollution in a numerical way.

\[
W_p = \frac{R_w}{\sum R_w}
\]

\[
S_c = \frac{C}{D_s}
\]

\[
O_w = W_p \times S_c
\]

\[
PIG = \sum O_w
\]

In the PRM season, the PIG values vary from 0.60 to 2.90 with an average value of 1.10; however, in the POM season, they deviate from 0.70 to 2.96 with an average value of 1.32. According to the classification of PIG (Table 3), 52.5 and 35% of groundwater samples from PRM and POM seasons fall into the insignificant category of pollution. Also, 30, 37.5% and 12.5, 20% of samples come under the low and the moderate pollution categories in PRM and POM seasons, respectively. Conversely, 2.5 and 5% samples fall in the high pollution category in PRM and POM seasons. It is observed that only 2.5% of samples (sample number 38) come under the very high category of pollution in both the seasons. It is confirmed that most of the samples in the vicinity of an agricultural area are highly influenced by anthropogenic activities.

Multivariate statistical analysis

This study uses multivariate statistical analysis to identify the significant association among water quality variables, their positive or negative influence with each other and most affecting parameters that are responsible for alteration of groundwater quality. In view of this, a total of 14 physicochemical parameters were used for the design of the CM, PCA and CA.

Correlation matrix (CM)

In hydrochemical study, CM is commonly used to identify the positive or negative affiliation among different groundwater quality parameters and their collective influence on hydrochemistry of the area (Islam et al. 2017). Generally, those variables show correlation coefficients >0.7; 0.5–0.7; <0.3, denoted as strong, moderate and weak association between them. Figure 3 explores the distribution of each parameter and is shown on the diagonal of the matrix plot. At the bottom, bivariate scatter plots with a fitted line are displayed and on the top the value of the correlation plus the significance level shown as star symbols indicating p-values (0, 0.001, 0.01, 0.05, 0.1, 1) and symbols (***, xx, x, , ”), respectively.

In the PRM season (Figure 3(a)), EC has strong correlation with TDS (0.72), Cl (0.77), Na (0.72) and moderate relation with TH (0.65) and Mg (0.68), suggesting a dissolution process which alters the salt content (Wagh et al. 2016b). TDS is strongly correlated with TH (0.92), Mg (0.91), Na (0.95), Cl (0.94) and SO₄ (0.74). Also, TH has shown strong correlation with Mg (0.97), SO₄ (0.71), Na (0.77) and Cl (0.89), which indicates a permanent type of

\(\text{Table 3 | Classification of PIG values for PRM and POM seasons of 2012}\)

<table>
<thead>
<tr>
<th>Range of PIG</th>
<th>Pollution classification</th>
<th>Sample numbers</th>
<th>% of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1.0</td>
<td>Insignificant</td>
<td>1,3–8,10–12,14–16,18,24,25,28–30,34,36</td>
<td>PRM 52.5 POM 35</td>
</tr>
<tr>
<td>1.0–1.5</td>
<td>Low</td>
<td>2,9,13,17,19,21–23,26,31,35,40</td>
<td>PRM 30 POM 37.5</td>
</tr>
<tr>
<td>1.5–2.0</td>
<td>Moderate</td>
<td>20,27,32,33,37</td>
<td>PRM 12.5 POM 20</td>
</tr>
<tr>
<td>2.0–2.5</td>
<td>High</td>
<td>39</td>
<td>PRM 2.5 POM 5</td>
</tr>
<tr>
<td>&gt;2.5</td>
<td>Very high</td>
<td>38</td>
<td>PRM 2.5 POM 2.5</td>
</tr>
</tbody>
</table>
hardness (Wagh et al. 2019b). Mg demonstrates strong correlation with Cl (0.92) and Na (0.78), but moderate association with sulphate (0.66) which indicates that these ions are responsible for increased hardness of water. Na and Cl present a strong affiliation correlation (0.92) that indicates evaporation dominance which controls salinity of water due to the semi-arid environment. Cl shows moderate association with SO$_4$ (0.58) because their sources are independent. It is observed that pH, Ca, K, CO$_3$, HCO$_3$ and SO$_4$ have negative correlation with other physicochemical parameters. It is confirmed that the inputs of TDS, Cl, SO$_4$, Mg, Na and NO$_3$ are due to precipitation and agricultural inputs. In the POM season (Figure 3(b)), EC and TDS show significant association (0.76) with Mg.
The EC and TDS depict a positive relationship with Ca, TH, Na, SO₄, Cl and Mg, which demonstrates an increase in salinity and salt dissolution increases EC of groundwater. TDS depicts a positive association with Cl (0.95), TH (0.84), Na (0.86), SO₄ (0.81), Mg (0.84) and HCO₃ (0.55) indicative of dissolution of ions which contribute to TDS. TH of groundwater is influenced by dissolved salts of Ca, Mg, Cl and SO₄ ions, supporting strong correlation with Mg (0.93), Cl (0.86), SO₄ (0.72) and HCO₃ (0.53). In general, interpretation confirms that these ions probably result from processes like ion exchange and weathering of silicate minerals. SO₄ exhibits a positive relation towards EC, TDS, TH, Mg, Na and Cl, indicating human-induced activities are altering the hydrochemistry within the area. Also, Cl shows good correlation with EC (0.72), TDS (0.95), TH (0.86), Mg (0.89) and Na (0.83), which is suggestive of evaporation dominance. The values of Na, Cl and TDS are increased due to inputs from intensive agriculture, domestic waste, weathering and soil erosion. Also, it was found that nitrate is negatively correlated with other parameters due to an independent source, i.e., agricultural inputs. It is demonstrated that Na and Cl exhibit positive association in both the seasons due to evaporation dominance.

Cluster analysis (CA)

CA is widely used to determine the groups or clusters of groundwater samples based on their similarities and dissimilarities according to hydrochemical characteristics (Danielsson et al. 1999; Lattin et al. 2005). In CA, similar objects represent the same class based on its similarity level and the dendrogram constituted and analysed. In the PRM season (Figure 4(a)), K, F, pH, Ca, NO₃ and CO₃ are jointly associated with Mg, SO₄, HCO₃ and Na; further, HCO₃ is related with another group which is linked with TDS and TH. Also, NO₃ and Cl are associated with EC. In the POM season, K, F, pH, Ca, NO₃ and CO₃ are closely associated with Mg, SO₄, HCO₃ and Na; further, HCO₃ is related with another group and this group shows association with TDS and TH jointly with associated Cl. Further, they

![Figure 4](http://iwaponline.com/aqua/article-pdf/69/4/398/724808/jws0690398.pdf)
are associated with EC (Figure 4(b)). K-means cluster analysis is the most commonly used method and is implemented with R programming. Here, k-means cluster analysis has been used for partitioning the standardized data into a number of groups and clusters distinguished for their interpretation of similarities and dissimilarities in groundwater data. Figure 5 represents k-means cluster analysis and three clusters were formed in PRM and POM seasons of 2012. Cluster 1 represents sample nos. 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 17, 21, 23 in PRM and 1, 4, 17, 18, 19, 20, 23, 24, 26, 27, 29, 35, 36, 39, 40 in POM season. Cluster 2 exhibits samples 20, 27, 31, 32, 33, 37, 38, 39 and 8, 10, 11, 21, 31, 32, 33, 37, 38 in PRM and POM seasons. Cluster 3 represents 19 samples in PRM, i.e., 1, 2, 3, 12, 15, 16, 18, 19, 22, 24, 25, 26, 28, 29, 30, 34, 35, 36, 40 and 2, 3, 5, 6, 7, 9, 12, 13, 14, 15, 16, 22, 25, 28, 30, 34 in POM seasons. In both seasons groundwater sample numbers 20 and 37 fall in the same cluster; however, these samples are adversely affected due to their occurrence in the same flow direction and agricultural runoff may be responsible.

**Principal component analysis (PCA)**

PCA was carried out by SPSS 22.0 v to identify the most affecting water quality parameters to differentiate the hydrochemical data based on natural characteristics of groundwater. Here, R-mode factor analysis was used to elucidate the association among variables in terms of simple relations (Saager & Sinclair 1974). Figure 6 explores the plots of physicochemical parameters in the first two principal components in both the seasons. Figure 7 demonstrates the loading of ions for both the seasons. It simply elaborates the association between each parameter. The direction of loading corroborates the nature of positive or negative relations of the first two components. Table 4 summarizes the results of rotated component matrix of various parameters in PRM and POM seasons. The first component (PC 1) accounts for the total variance of 40.24% with 40.24% cumulative and eigen value of 5.63. It is positively loaded with EC, TDS, TH, Mg, Na, Cl, SO4 with values of 0.78, 0.95, 0.92, 0.94, 0.90, 0.98 0.69, respectively, in the PRM season. It is confirmed that the elevated content of TDS is influenced by ions like Na, Mg, SO4 and Cl. The groundwater in the study area has a permanent type of hardness due to the influence of Mg, Cl and SO4 suggesting a high loading of TDS. High loading of Na in excess of Ca is indicative of the ion exchange process between calcium and sodium (Drever 1988). Moreover, Na and Cl reveal high loadings due to solubility behaviour (Rao 2014). There is also high loading of Mg, Na, Cl and SO4 ions due to them being derived from natural processes.
and anthropogenic inputs. The second component (PC 2) exhibits a moderate load of Ca (0.72), CO₃ (0.58), HCO₃ (0.60), F (0.59) and SO₄ (0.48) with 14.29% of total variance indicating groundwater chemistry being controlled by fresh water infiltrated with soil CO₂ to form H₂CO₃, and long-term irrigation practices also increasing the alkalinity (Rao 2014). PC 2 loading suggests that alkalinity is the main causative factor in influencing groundwater quality. PC 3 shows an increased load of fluoride (0.44), HCO₃ (0.43) and NO₃ (0.63) with 12.72% of the total variance, cumulative 67.26% and eigen value of 1.78. The load of F indicates anthropogenic inputs from fluoride-rich fertilizers. The high load of nitrate demonstrates application of nitrogen complex fertilizers and agrochemicals, thus reflecting anthropogenic origins. In the POM season, PC 1 exhibits a high association between EC, TDS, TH, Mg, Na, Cl, CO₃, HCO₃ and SO₄ (0.75, 0.95, 0.86, 0.89, 0.86, 0.93, 0.48, 0.71, 0.80) with 44.35% of total variance, cumulative 44.35% and eigen values of 6.2. It is revealed that high loading of EC and TDS is due to contaminant percolation in groundwater along with surface runoff from agricultural fields as a result of which alkalinity increased. Moreover, salinity and alkalinity is controlled by inputs of Na, Cl, SO₄ and HCO₃ derived from natural and agricultural activities. PC 2 shows loading of pH (0.77) and CO₃ (0.48) having 15.80% of total variance which is indicative of rain water percolating into aquifers which controls the alkalinity. The high loadings of F (0.79) suggest the fluoride is sourced from chemical fertilizers used in agricultural fields. PC 3 demonstrates a high loading of K (0.91) indicating K-rich fertilizer applications. PC 4 demonstrates moderate loading of Ca (0.44) due to a diminutive dissolution of calcite minerals.
and NO₃ (0.92) because of excessive use of nitrogen-rich fertilizers.

**Hydrogeochemostral categorization**

To understand the hydrochemical nature in the study area, to identify geology and mineralogy of rocks, agricultural inputs, dissolution, precipitation and climate condition influence the groundwater chemistry. Generally, during groundwater recharge within a river basin, reactions with soil, rock–water interactions, various types of weathering and significant anthropogenic inputs may change groundwater chemistry (Todd 1980). The scatter plot of EC vs TDS (Figure 8(a)) represents moderate to good correlation. Electrical conductivity of water is attributed to the dissolution of salts and inorganic pollution load in the water, since it is the sum of all the ions or total dissolved solid content (Wagh et al. 2018c). Cl vs SO₄ (Figure 8(b)) indicates moderate correlation in PRM and POM seasons. The high content of chloride and low sulphate content possibly signify sulphate diminution in groundwater (Datta & Tyagi 1996). The linear relationship between Mg vs Na (Figure 8(c)) confirms that magnesium concentration is increasing due to silicate weathering. The proportion of Na vs Cl (Figure 8(d)) recognizes the influence of the salinity mechanism in an arid environment (Singh et al. 2005). This ratio suggests that weathering of silicate minerals is one of the possible sources of Na, and if the ratio is <1 it confirms the ion exchanges of Na with Ca and Mg ion in soil particles (Tiwari & Singh 2017). Cl vs Ca/Mg plot (Figure 8(d)) indicates that salinity is increased due to chloride being related with amplifying the contents of Ca + Mg (Raju et al. 2016). Cl vs Na/Cl (Figure 8(e)) shows moderate to good correlation which suggests that salinity increased due to chloride being directly proportional to Ca + Mg and inversely proportional to Na/Cl due to the ion exchange process (Raju et al. 2016). Ca + Mg vs HCO₃ + SO₄ plot (Figure 8(f)) is applied to characterize the influence of processes like ion exchange and weathering; however, if the equiline ratio is 1:1 there is dissolution of minerals, namely, calcite, dolomite.

| Table 4 | Rotated component matrix of PRM and POM seasons of 2012 |
| Parameters | PRM season | POM season |
| | Components | 1 | 2 | 3 | Components | 1 | 2 | 3 | 4 |
| pH | 0.03 | 0.00 | −0.82 | −0.34 | 0.77 | 0.04 | −0.06 |
| EC | 0.78 | −0.12 | 0.36 | 0.75 | −0.31 | −0.00 | 0.19 |
| TDS | 0.95 | 0.24 | 0.08 | 0.95 | −0.15 | 0.04 | 0.02 |
| TH | 0.92 | 0.23 | −0.01 | 0.86 | −0.16 | −0.20 | 0.22 |
| Ca | 0.10 | 0.72 | 0.01 | 0.08 | −0.65 | −0.15 | 0.44 |
| Mg | 0.94 | 0.05 | −0.02 | 0.89 | 0.08 | −0.15 | 0.06 |
| Na | 0.90 | 0.09 | 0.11 | 0.86 | 0.11 | 0.15 | −0.13 |
| K | −0.16 | 0.11 | −0.27 | −0.12 | −0.08 | 0.91 | 0.06 |
| Cl | 0.98 | −0.06 | 0.04 | 0.93 | −0.04 | −0.08 | 0.10 |
| F | 0.12 | 0.59 | 0.44 | 0.12 | 0.79 | −0.31 | 0.03 |
| CO₃ | −0.00 | 0.58 | −0.28 | 0.48 | 0.46 | 0.38 | 0.02 |
| HCO₃ | 0.13 | 0.60 | 0.45 | 0.71 | 0.22 | −0.09 | −0.05 |
| SO₄ | 0.69 | 0.48 | −0.07 | 0.80 | −0.31 | 0.08 | −0.17 |
| NO₃ | −0.04 | 0.14 | 0.63 | 0.02 | −0.09 | 0.09 | 0.92 |
| Eigen values | 5.63 | 2.00 | 1.78 | 6.20 | 2.21 | 1.22 | 1.22 |
| % of variance | 40.24 | 14.29 | 12.72 | 44.30 | 15.80 | 8.73 | 8.71 |
| Cumulative % | 40.24 | 54.33 | 67.26 | 44.30 | 60.15 | 68.89 | 77.60 |
and gypsum (Cerling et al. 1989). The plot represents linear correlations in PRM and POM seasons which indicate the ion exchange process is dominant. The ionic content in meq/L falls above the equiline signifying carbonate and silicate weathering. Also, it shows that HCO$_3$ is in excess of Ca$^+$Mg, possibly indicating the process of ion exchange reaction (Rajmohan & Elango 2004). Cl vs HCO$_3$ + Cl (Figure 8(g)) ratio exhibits strong positive correlation ($r^2 = 0.97$ and 0.95) in PRM and POM seasons which indicates the influence of salinization due to agricultural inputs. All the groundwater samples fall in equiline 1:1 in PRM and POM seasons.

The logarithmic bivariate plot of HCO$_3$/Na vs Ca/Na (Figure 8(h)) and Mg/Na vs Ca/Na (Figure 8(i)) are used
to find the weathering phenomenon. The higher solubility of Na than Ca amplifies the sodium content; thus, Ca vs Na ratio is assumed to be low, which infers silicate weathering as the existing geochemical process controls the groundwater quality (Raju et al. 2016). However, a few samples from both the seasons suggested carbonate weathering is due to the presence of lime kankar in the area. The plot of (Na + K) – Cl/(Ca + Mg) – (HCO₃ + SO₄) represents a dominant cation exchange process; if the samples fall close to zero then the ion exchange process is absent, whereas samples falling on slope line 1 represent an exchange of ions (Kortatsi 2007). It is observed that samples fall on the equiline with slopes of −1.026 and −1.057 in both the seasons confirming the exchange of Na, Ca and Mg (Figure 8(j)). The plot of Na + K vs (TZ⁺) (Figure 8(k)) represents good correlation ($r^2 = 0.88$ and 0.75) in PRM and POM seasons. Sodium and potassium confirm positive correlation with total cations, particularly at higher concentrations. It is observed that silicate weathering is the leading process controlling the hydrochemistry in the area. The scatter plot of Ca + Mg vs TZ⁺ (Figure 8(k)) illustrates good correlation ($r^2 = 0.88$ and 0.77) in PRM and POM seasons. It is confirmed that many of the samples fall under the equiline and reflect weathering of carbonate minerals which is the foremost origin of calcium and magnesium ions (Kant et al. 2018). The overall explanation confirms that groundwater is mainly influenced by silicate weathering followed by carbonate weathering and ion exchange process in the study area.

CONCLUSIONS

In this study, PIG and multivariate statistical techniques like CM, PCA and CA have been utilized to determine the influence and classify the groundwater quality in the Kadava river basin. The hydrochemical results inferred that groundwater is alkaline in nature and possesses permanent hardness. According to BIS drinking standards, pH, TDS, TH, SO₄, Mg, Na, Cl, Ca, NO₃ and F ions exceed the threshold limit of drinking in a few groundwater samples. The high nitrate contents are found in 62.5 and 75% of samples from PRM and POM seasons due to excess use of fertilizers in agriculture. The elevated EC is attributed to dissolution of salts and the inorganic pollution load in the water. PIG results inferred that 52.5 and 35%, 30 and 37.5%, 12.5 and 20%, 2.5 and 5% groundwater samples came under insignificant, low, moderate, high pollution categories in PRM and POM seasons, respectively. However, only one sample (number 38) is in the very high pollution category of pollution in both the seasons. The interpretation of PIG confirms that groundwater samples located in the vicinity of intense agricultural areas were found to be problematic; hence, particular remedial measures are required. The multivariate statistical techniques, namely, CM, PCA and CA are appropriate here and useful tools for identifying and interpreting the behaviour of groundwater and its possible sources of contamination. The CM depicts that TDS content in groundwater is increased by ions like Mg, Cl, Na and SO₄. TH is closely related with Mg, Cl and SO₄ content and, as a result, water became permanently hard. PCA results corroborated that high loadings of Na over Ca indicate the ion exchange process is dominant which influences the water chemistry. In the POM season, high loadings of magnesium, sodium chloride and sulphate confirm that groundwater is influenced by human-induced activities taking place in the area. Moreover, a positive loading of CO₃, HCO₃ and pH advocates an alkalinity controlled process due to precipitation influence. CA infers that sample numbers 20 and 37 fall in the same cluster in both seasons due to their occurrence in surface flow direction and agricultural runoff may be responsible for diminishing the groundwater quality. Dendrogram analysis depicts that K, F, pH, Ca, NO₃ and CO₃ are jointly associated with Mg, SO₄, HCO₃ and Na; further, HCO₃ is coupled with another group which is linked with TDS and TH. Also, NO₃ and Cl are associated with EC in both the seasons. The linear correlation of Ca + Mg/Na + K is weak, signifying accumulation of calcium and magnesium ions in the groundwater originating from the weathering of silicate-rich basalts. Ca/Mg ratio suggests that salinity increased due to chloride and is linked with an increase in Ca + Mg content and decline in Na/Cl due to the control of the ion exchange process. The Cl vs HCO₃ + Cl ratio exhibits salinization dominance due to agricultural inputs. In a nutshell, the results indicate that geo-genic processes including ion exchange, silicate weathering, evaporation and agricultural inputs are the main controlling factors which alter
groundwater chemistry within the area. However, those aquifers are problematic for drinking and agriculture, and require particular remedial measures like least use of chemical fertilizers, rain water harvesting, water purification set-up implemented in order to sustain water quality for future use. The outcomes of the study may help local authorities to make integrated basin water management plans to sustain groundwater quality in semi-arid areas of the river basin.

REFERENCES


Davis, J. C. 1986 Statistics and Data Analysis in Geology, 5th edn. Wiley, New York, USA.


Rao, N. S., Sunitha, B., Adimalla, N. & Chaudhary, M. 2019 Quality criteria for groundwater use from a rural part of Wanaparthy District, Telangana State, India, through ionic spatial distribution (ISD), entropy water quality index (EWQI) and principal component analysis (PCA). Environmental Geochemistry and Health 42, 1–21.


