

## ARTICLE OPEN



# Natural recharge transcends anthropogenic forcing that influences arsenic vulnerability of the quaternary alluviums of the Mid-Gangetic Plain

Ashwin Singh <sup>1</sup>, Arbind Kumar Patel <sup>2</sup>, Jyoti Prakash Deka<sup>3</sup> and Manish Kumar <sup>2</sup>✉

We evaluated the anthropogenic forcing and exceedance probability of arsenic vulnerability in the Quaternary Alluviums of the Mid-Gangetic Plain through the coupled application of hydrogeochemical analyses, inverse modelling, probability analyses, kriging, bivariate and multivariate statistical analysis (MVA). Fifty-seven groundwater samples (total 171) each were collected during the pre-monsoon (April 2015), monsoon (July 2015) and post-monsoon (January 2016). Seasonal cyclicality of ion exchange and reverse ion exchange was observed, with the former being common during pre-monsoon while the latter was dominant during post-monsoon. PHREEQC modelling showed the involvement of the agrochemicals such as calcium nitrate and calcium phosphate and other isolated incidences of chemical usage like bleaching powder as well as the probability of wet deposition of the anions like  $\text{SO}_4^{2-}$  and  $\text{NO}_3^-$ . Kriging-based arsenic mapping revealed that rainfall recharge-led dilution plays the most dominant role in its mobilization. Owing to natural recharge in the alluvium plain, the exceedance probability of arsenic concentration above  $5 \mu\text{g L}^{-1}$  falls drastically from more than 0.8 in the pre-monsoon to 0.5 during the post-monsoon. Study implies that pre-monsoon pumping of groundwater must be regulated in combination with proper disposal of pH and ORP affecting chemicals so that the high natural recharge should not significantly induce arsenic mobilization. Further, we recommend that vulnerability estimation should not be based solely on the present scenarios but must include the features of probable exceedance and saturation status of groundwater in this belt. We put forward a comprehensive model to explain the operative processes governing arsenic mobilization.

npj Clean Water (2020)3:27; <https://doi.org/10.1038/s41545-020-0075-5>

## INTRODUCTION

The Holocene quaternary alluviums of Mid-Gangetic Plain (MGP) supports around 60 million population of India. The region while apparently fertile for agriculture is enriched with high arsenic deposits, which gets into the groundwater system through over-pumping and dissolution of iron-oxy-hydroxides (Fe-OOH) via fluctuating redox conditions<sup>1–3</sup>. Arsenic release is suspected to be microbial degradation of sedimentary organic deposits under the presence of  $\text{O}_2$  and  $\text{NO}_3^-$ <sup>4–6</sup>. The process further accelerates with suitable mineralogical and geological conditions<sup>7–9</sup>. On the other hand, human-induced anthropogenic activities like overutilization of groundwater disrupt the equilibrium between the oxic and anoxic zones causing the increased formation of metal oxy-hydroxides. Owing to the non-desiring ubiquitous nature of arsenic, the uptake mechanism by the animals and plants is purely adventitious in the biotic environment.

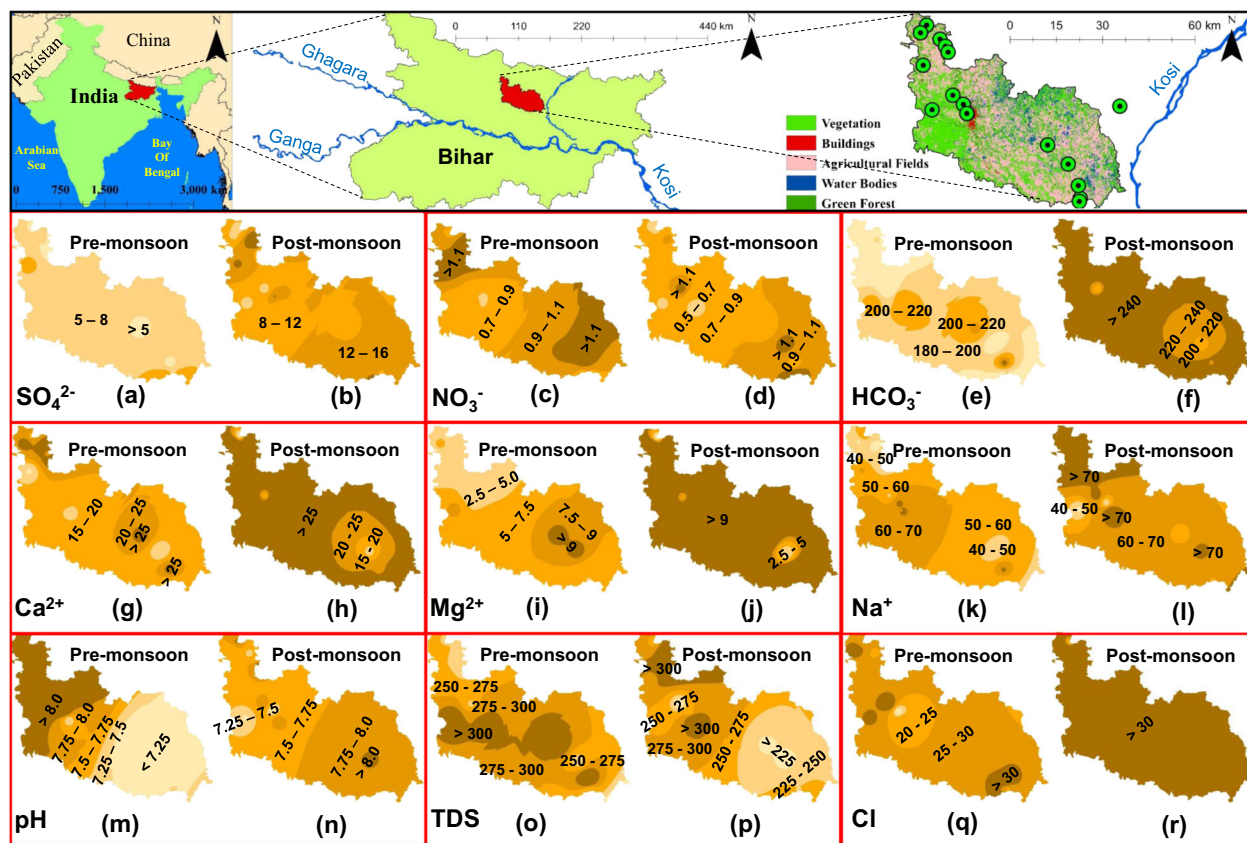
Both the toxic pentavalent As (V) and trivalent As (III) forms intrude into the cell cytosol by mimicking the properties of nutrients and essential substrates like phosphates, glucose and glycerol<sup>10</sup>. While trivalent aromatic arsenic forms, carry the highest toxicity but the stability deteriorates as a function of temperature and sample matrix when precursors like methylated forms take the dominance. Significant risk exists when anoxic conditions vis-à-vis in the groundwater interaction interfaces persist<sup>11</sup>. In neutral solutions, arsenic exists as arsenite ( $\text{As}(\text{OH})_3$ ), more chemically similar to glycerol<sup>12</sup>, causing the un-intentional conduction through the AQP's bi-directional channels<sup>13</sup>. The exposed population of MGP has exhibited significant amount of carcinogenicity,

ranging from skin ailments like keratosis and hyperpigmentation to neurological disorders like brain diseases<sup>14,15</sup>. Further evidence shows that prolonged and short-term exposure quantified through arsenic aggregation in hair, nail and urine carries the risk of developing cancer shortly in the near future<sup>16,17</sup>. The situation is grimmer for the MGP and the other fluvio-deltaic regions of the subcontinent as there is an average background exposure of 15.9 and 10.3  $\text{mg kg}^{-1}$  in the onshore and offshore water samples respectively, with the finer sediments having higher concentration at increasing depth<sup>18</sup>. Studies since 2000 are continuously indicating elevated arsenic concentration exceeding WHO limits of  $10 \mu\text{g L}^{-1}$ <sup>19,20</sup>.

While the hydrogeochemical investigation of groundwater can reveal much about the underlying mechanism of arsenic release, an alternate method involves using soil and its secondary characteristics as predictors for evaluating the arsenic vulnerability<sup>21</sup>. Calcisols have been believed to affect the concentration of geogenic anions in the groundwater. The release is usually triggered under the increased presence of sodium bicarbonate through reduction under alkaline conditions. Low concentration of  $\text{Ca}^{2+}$ , acting as a sink during silicate weathering, generally complements the release process<sup>21</sup>. Physical processes like precipitation and land surface temperature also quantitatively impact in changing the arsenic vulnerability through changes in the evapotranspiration progressing through the aquifer<sup>22,23</sup>.

The high presence of arsenic in the groundwater of MGP is further complemented due to (i) spatial geomorphological variation, (ii) climatic differentiation, and (iii) land use and landcover patterns. A landcover of agricultural crop field

<sup>1</sup>Discipline of Civil Engineering, Indian Institute of Technology, Gandhinagar, Gujarat 382355, India. <sup>2</sup>Discipline of Earth Sciences, Indian Institute of Technology, Gandhinagar, Gujarat 382355, India. <sup>3</sup>Discipline of Environmental Sciences, Gauhati Commerce College, Guwahati, Assam 781021, India. ✉email: manish.env@gmail.com



**Fig. 1** Map illustrating the study area of Darbhanga and the spatial distribution of various water quality parameters in  $\text{mg L}^{-1}$ . The concentration of anions such as  $\text{SO}_4^{2-}$  (a, b),  $\text{NO}_3^-$  (c, d) and  $\text{HCO}_3^-$  (e, f) has increased significantly from pre to post monsoon. Alkali earth metals such as  $\text{Ca}^{2+}$  (g, h),  $\text{Mg}^{2+}$  (i, j) and alkali metals like  $\text{Na}^+$  (k, l) have similar enrichment trends after monsoon. Change in parameters like pH (m, n), TDS (o, p) and Cl (q, r) signify the activeness of various anthropogenic activities.

accompanied with changes in the irrigation pattern impels the spatial differentiation in the natural recharge. An excess recharge enriched with anthropogenic nitrate added via chemical fertilizers can change the weathering patterns causing contamination of arsenic<sup>2,24</sup> and fluoride<sup>20</sup>. Excessive abstraction of groundwater under these circumstances not only degrades the quantity but can also lead to aggravation of geogenic pollution as can be seen from the Bangladeshi arsenic epidemic of the late twentieth century<sup>1,18,25–27</sup>. Co-occurrences of arsenic and fluoride along with contaminants of emerging concern have potentially hazardous consequences for different life forms. Thus, a proper understanding of the groundwater hydro-geochemistry is very imperative with implications on influencing the decision and policy-making process affecting groundwater conservation<sup>28</sup>.

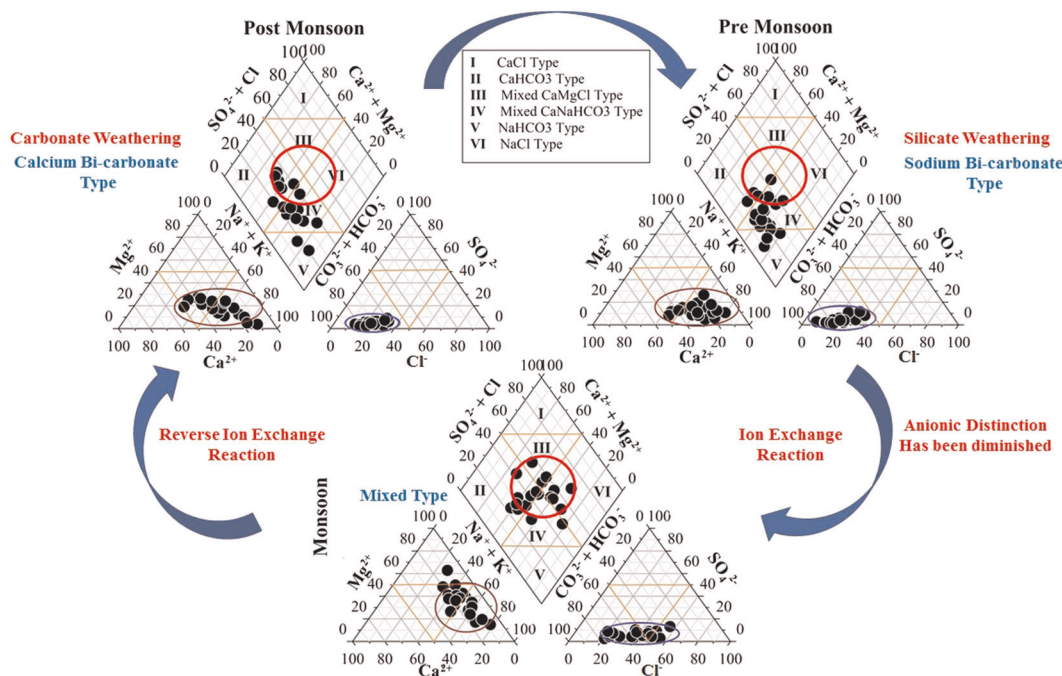
In this regard, deriving a meaningful interpretation from the multi-parameter dataset is a challenge. Multivariate analysis (MVA), like factor analysis (FA), principal component analysis (PCA) and hierarchical cluster analysis (HCA), have been commonly used in such hydrogeochemical investigations<sup>29–32</sup>. While reducing the dimensions of datasets, the discreet knowledge about the data may get lost, but the essence of the data or the broad meaning conveyed by the data gets preserved in the form of data clusters. In light of the stated facts, the current study was undertaken (i) to understand the impact of seasonal change i.e. monsoon on the hydro-geochemistry of Darbhanga district through integrated application of hydrogeochemical, bivariate and multivariate statistical analysis (MVA), (ii) to propose a probabilistic model-based seasonal mapping of arsenic occurrence and exceedance, and (iii) to put forward the conceptual model to explain the observed arsenic spatial and seasonal variation. Overall, this study puts an effort to evaluate whether

natural recharge or anthropogenic influence is more critical for arsenic mobilization in the Alluvium plains of MGP.

## RESULTS AND DISCUSSION

### Prevalence of ions, water facies and hydro-geochemistry

Major ions chemistry of the area indicates a groundwater state that is mostly within the acceptable drinking water range when compared to the World Health Organisation<sup>33</sup> and Bureau of Indian Standards (BIS)<sup>34</sup> standards (Supplementary Table 1). However, the dominating influence of monsoon precipitation on groundwater quality, visible through drastic changes in the ORP and pH, is likely to enhance arsenic pollution vulnerability in the groundwater. There was a noticeable increase in EC and TDS from pre-monsoon to post-monsoon, implying active weathering and dissolution through the infiltrating rainwater (Supplementary Table 1). This is supported by the enhanced levels of  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{SO}_4^{2-}$ ,  $\text{HCO}_3^-$  and  $\text{Cl}^-$  in the post-monsoon (Fig. 1). The involvement of weathering and dissolution is further revealed through the Gibbs plots<sup>16,19</sup> (Supplementary Fig. 1). All the groundwater samples were found to fall within the rock–water interaction or the weathering-dominated region of the plot. On closer observation, we noticed a decrease in  $(\text{Na}^+ + \text{K}^+ + \text{Ca}^{2+})$  and  $(\text{Cl}^-)/(\text{Cl}^- + \text{HCO}_3^-)$  ratios towards the post-monsoon season, which was indicative of elevation in carbonate weathering due to precipitation. Piper diagrams (Fig. 2) further substantiate the fact by indicating that the dominant water types during pre-monsoon season were dominantly of mixed Ca-Na- $\text{HCO}_3$  type which changes towards Ca- $\text{HCO}_3$  type in the post-monsoon owing to significant recharge.



**Fig. 2** Piper plot showing the evolution of groundwater. The dominance of carbonate weathering during post monsoon succeeds the activeness of silicate weathering during pre monsoon. Seasonality has propelled a major shift in the groundwater evolution of the region with reverse ion exchange reaction becoming prominent during post monsoon along with water type changing to calcium bi-carbonate.

#### Spatial hydrogeochemical profiling

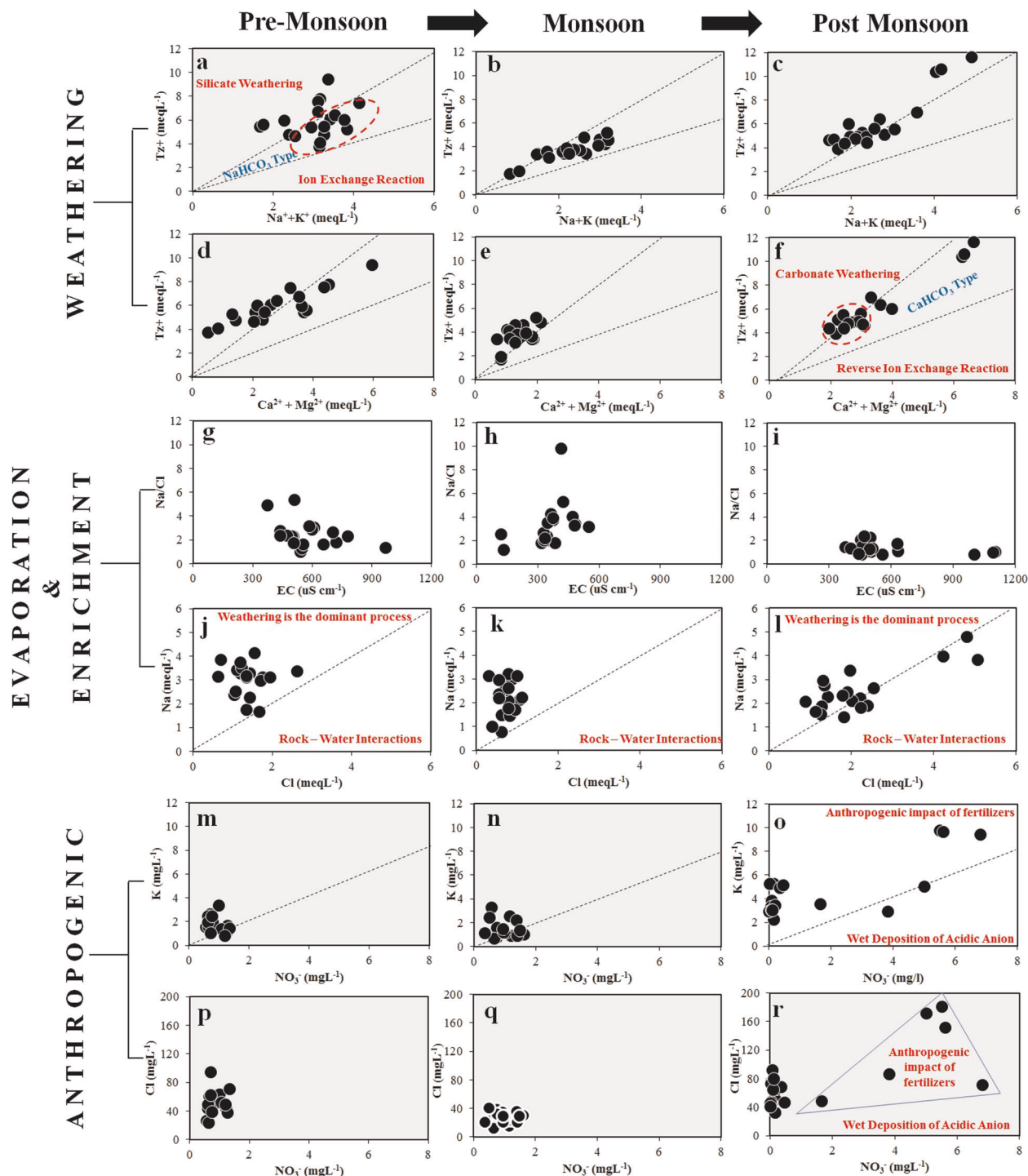
The spatial distribution of  $\text{SO}_4^{2-}$  shows that the concentration in the entire study area is way below the permissible drinking standards of BIS 10500<sup>34</sup> ( $<200 \text{ mg L}^{-1}$ ) and WHO 2011<sup>33</sup> ( $<250 \text{ mg L}^{-1}$ ) (Fig. 1). However, the concentration of  $\text{SO}_4^{2-}$  has increased during the post-monsoon as compared to pre-monsoon. A significant change takes place in the north of Darbhanga district where there is an almost 100% increase in the concentration values. The spatial distribution of  $\text{NO}_3^-$  shows that the part of the problem of introduction of  $\text{NO}_3^-$  in Darbhanga could be the increased usage of fertilizers during the pre-monsoon Rabi crops as compared to less fertilizer application during the post-monsoon Kharif crops. Consistent concentration of  $\text{NO}_3^-$  in the north-western region of the district signifies that Darbhanga city has little contribution in changing the dynamics of  $\text{NO}_3^-$  across different seasons, and hence agricultural fields play an essential role in its cyclicity. Within the study area, it is evident that during the post-monsoon, the northern region of the district experiences a more substantial carbonate weathering as compared to the southern part of the district. The change in the spatial distribution of TDS during the pre- and post-monsoon shows that the recharge flux does have the ability to change the overall drinking suitability of groundwater (Fig. 1). Decrease in TDS in the south-eastern region shows that agricultural fields are more susceptible to infiltration of rainwater and hence more dilution will occur leading to decreasing levels of TDS in the groundwater. The northern region has more presence of built-up regions leading to less dilution during the post-monsoon period.

#### Seasonality in the hydro-geochemistry

The seasonal aspect of the hydro-geochemistry was evaluated by utilizing bivariate methods like scatter plots and correlation analysis. The plots of total cations ( $\text{Tz}^+$ ) versus ( $\text{Na}^+ + \text{K}^+$ ) (Fig. 3a–c) and ( $\text{Ca}^{2+} + \text{Mg}^{2+}$ ) (Fig. 3d–f) confirms the higher involvement of silicate weathering and ion exchange during the pre-monsoon season and an elevation in carbonate weathering towards the post-monsoon. During the pre-monsoon, majority of

the samples plot closer to the 2:1 line (Fig. 3a), while towards the post-monsoon the spread decreases along the 2:1 line (Fig. 3b). This was also confirmed through the strong correlation of  $\text{Na}^+$  with both EC and TDS (Supplementary Table 2). Negative correlation of  $\text{K}^+$  with both TDS and EC suggests that the silicate species involved in this process appears to be sodium feldspar or albite<sup>7,8</sup>. What we also observed was the strong correlation of  $\text{Na}^+$  with  $\text{HCO}_3^-$ , suggestive of ion exchange with Ca-Mg- $\text{HCO}_3$  type water during the drier pre-monsoon i.e. lower water table leading to the gradual shift to mixed Ca-Na- $\text{HCO}_3$  water type<sup>7</sup>. During the post-monsoon, the elevation in carbonate weathering is also supported by the strong correlations of  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  with EC and TDS, indicating the involvement of dolomite weathering<sup>22</sup>. An increasing  $\text{Na}^+:\text{Cl}^-$  ratio with EC also points towards the excess of  $\text{Na}^+$  due to silicate weathering, which appears to be dominating during the pre-monsoon season (Fig. 3g–i). An immediate decrease in the  $\text{Na}^+:\text{Cl}^-$  ratio towards the post-monsoon season (Fig. 3j) was a case of halite dissolution promoted by precipitation.  $\text{Na}^+$  vs.  $\text{Cl}^-$  (Fig. 3j–l) exhibit an excess of  $\text{Na}^+$  compared to  $\text{Cl}^-$  implying silicate weathering while samples close to the 1:1 line are indicative of halite dissolution.

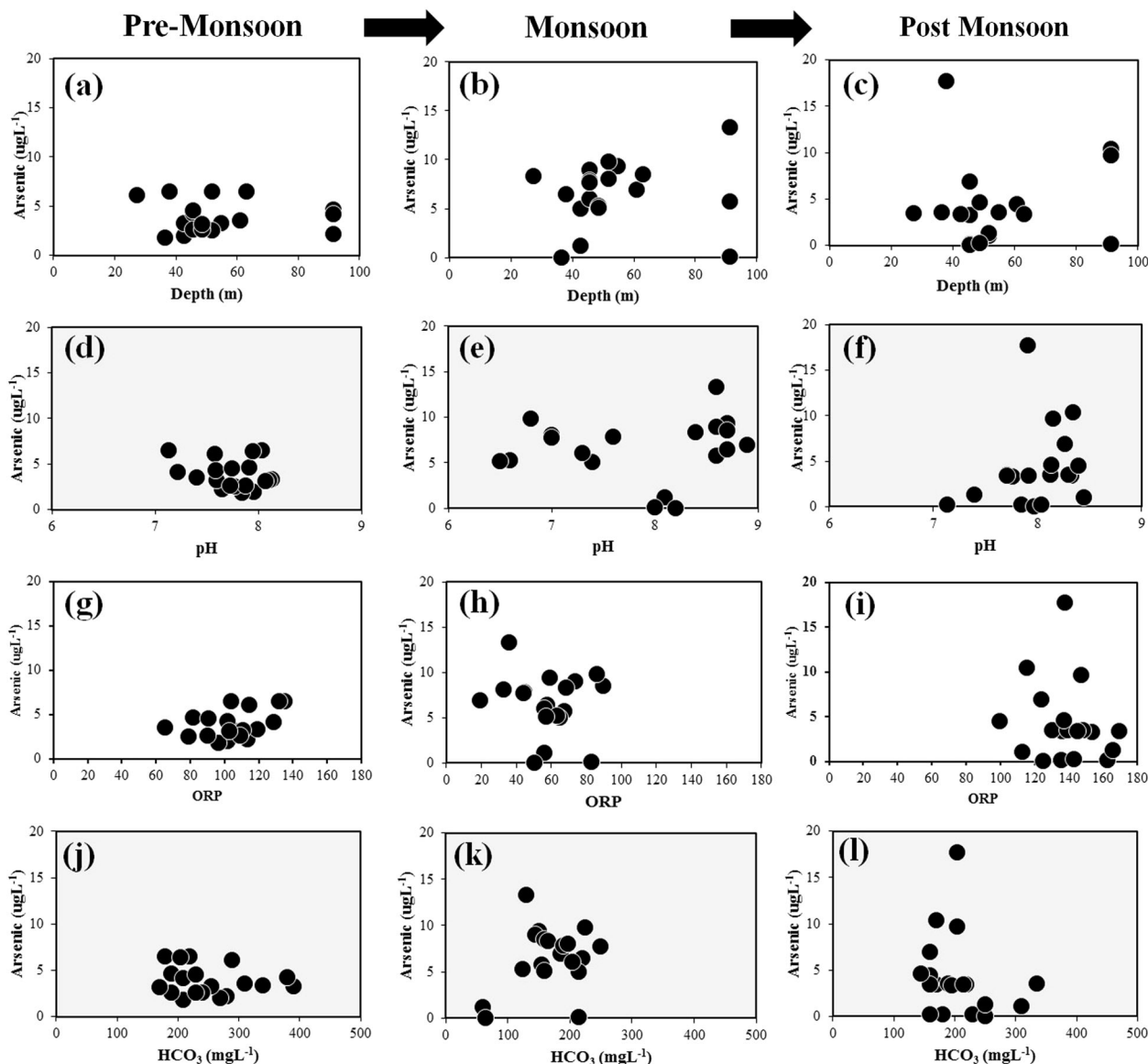
On the other hand, anthropogenic influences were evaluated by plotting  $\text{K}^+$  vs.  $\text{NO}_3^-$  (Fig. 3m–o) and  $\text{Cl}^-$  vs.  $\text{NO}_3^-$  (Fig. 3p–r). Levels of both  $\text{K}^+$  and  $\text{NO}_3^-$  are quite low in the groundwater; moreover, for most samples, there is no visible relationship between the two. We have already stated that the background levels of  $\text{NO}_3^-$  is very low; however, the reasons behind the low concentrations of  $\text{K}^+$  are still not clear. A few of the samples that fall along the uniline could be released due to agricultural usage of potassium and  $\text{NO}_3^-$ -based fertilizers<sup>8</sup>. During the pre-monsoon season,  $\text{Cl}^-$  levels in the groundwater remain lower than the post-monsoon. Also, in the post-monsoon, there was not only a spike in  $\text{Cl}^-$  level but also a weak relationship with  $\text{NO}_3^-$  being observed. Given the fact that  $\text{Cl}^-$  also displayed positive correlations with  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$  and  $\text{SO}_4^{2-}$  during the post-monsoon, it is likely that it could be an anthropogenic occurrence along with incidences of reverse ion exchange. Use of chemicals like bleaching powder<sup>35</sup>, wet deposition from the various industrial



**Fig. 3** Scatter diagrams tracing the hydro-chemistry of the groundwater. Possible change in weathering from silicate to carbonate has been detected (a–f). Mineral mediated enrichment of alkali metals has weakened with monsoon (g–i). The role of background anthropogenic interferences has been found to be relatively less (m–r).

estates could be some of the plausible sources of  $\text{Cl}^-$  in groundwater during this period. The molar ratio of  $\text{Ca}^{2+}:\text{Mg}^{2+}$  which remains closer to 2 for several samples during the pre-monsoon (Supplementary Fig. 2) again reaffirms the dominance of silicate weathering during the pre-monsoon season<sup>17</sup>. During monsoon, there was a significant shift in this ratio. However, for a wide range of samples, this ratio remained between 1 and 2, denoting a mix of calcite and dolomite weathering. Based on scientific literature we can attribute this to the dominance of

carbonate weathering<sup>8</sup>, which we have already discussed earlier. A ratio close to 1 indicates dolomite dissolution, while a ratio close to 2 shows calcite weathering. Higher rates of carbonate weathering are also supported by the  $(\text{Ca}^{2+} + \text{Mg}^{2+})$  vs.  $(\text{SO}_4^{2-} + \text{HCO}_3^-)$  plot (Supplementary Fig. 2) in which most of the groundwater samples plot below 1:1 line<sup>5,23</sup>. The relationship of  $\text{Cl}^-$  with the alkaline earth elements  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  is not significant during the pre-monsoon season; however, as observed earlier in the correlation matrix, this relationship becomes much



**Fig. 4** Dominant parameters impacting the occurrence of arsenic. Scatter diagrams between arsenic and depth (a–c), pH (d–f), ORP (g–i) and  $\text{HCO}_3^-$  (j–l).

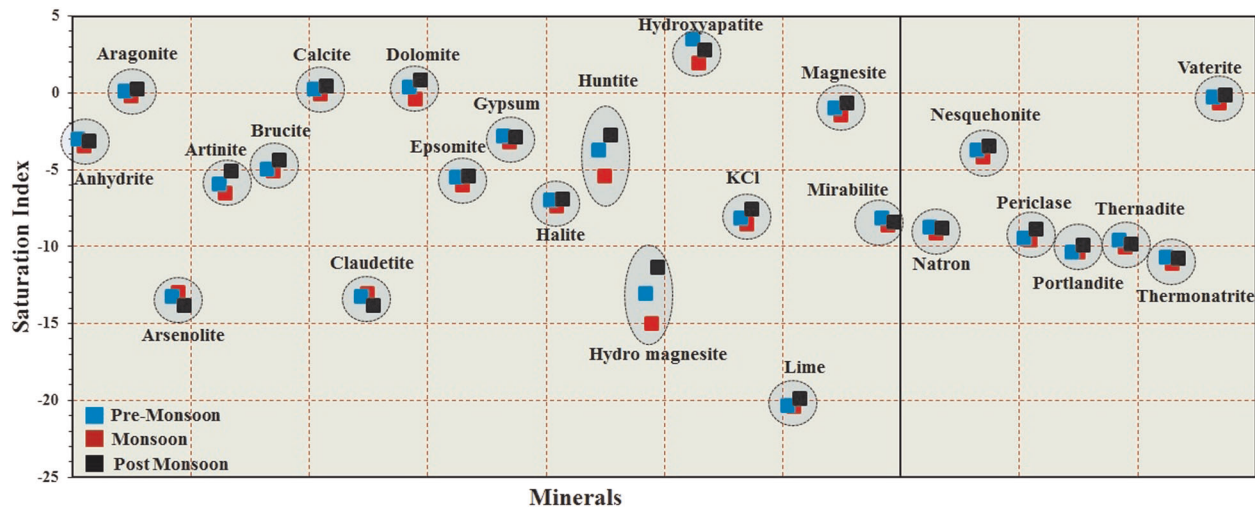
more prominent during the post-monsoon. We suspect the role of various anthropogenic factors behind enhancing  $\text{Cl}^-$  in groundwater. However, a more critical analysis points towards reverse ion exchange reaction<sup>1</sup>. The related discussion has been further elaborated in Supplementary Fig. 2 [( $\text{Ca}^{2+} + \text{Mg}^{2+}$ )/ $\text{HCO}_3^-$  vs.  $\text{Cl}^-$ ]<sup>35</sup>.

The subdued nature of carbonate weathering during the pre-monsoon season can also be observed based on the correlations of  $\text{HCO}_3^-$  with  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  which remained low during the pre-monsoon ( $r = 0.37$  and  $0.47$  respectively, Supplementary Table 2) and becomes stronger during post-monsoon ( $r = 0.71$  and  $0.66$  respectively). The occurring exchange reactions could be traced from the relationship of  $\text{Cl}^-$  with  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  which has been analysed in Supplementary Fig. 2 [( $\text{Ca}^{2+} + \text{Mg}^{2+}$ )/ $\text{HCO}_3^-$  vs.  $\text{Cl}^-$ ]. The increase in ( $\text{Ca}^{2+} + \text{Mg}^{2+}$ ): $\text{HCO}_3^-$  ratio with salinity denotes an increase in the alkaline earth metals due to the process of reverse ion exchange. Speculations concerning  $\text{HCO}_3^-$  depletion must be ruled out, as the conditions in the groundwater were alkaline during both pre-monsoon and post-monsoon seasons. The occurrence of arsenic in the oxic groundwater conditions

could be one of the causes behind its low exposure in the region (Fig. 4). Drawing from the empirical relationship between pH and ORP, it could be seen that during pre-monsoon the As (III) and As (V) have a high probability of occurring as  $\text{H}_3\text{AsO}_3^0$  and  $\text{HAsO}_4^{2-}$  respectively. The occurrence of  $\text{H}_3\text{AsO}_3^0$  usually mediates in acidic water with a very high sensitivity to  $\text{HCO}_3^-$  concentration. High monsoon recharge leads to severe redox changes in the groundwater, causing species to be distributed among various oxidation forms. As (V)-dominated Alkali-Oxic (AO) water could be seen triggered due to sodic conditions under desorption mobilization mechanism.

Scenario analyses through PHREEQC-based saturation index estimation

The stability of the groundwater is dictated by the thermodynamic equilibrium of the various minerals present in it. The estimation of the saturation index (SI) becomes important in order to quantify the dissolution/precipitation potential of the minerals besides helping in understanding the reaction pathways<sup>22,27</sup>. SI estimation



**Fig. 5 Saturation index of various minerals obtained through PHREEQC modelling and their variations during the different seasons. Undersaturation of arsenolite ( $SI < 0$ ) shows high vulnerability for the future.**

is a sequential process requiring the estimation of molality ( $m$ ), valency ( $Z$ ), activity coefficient, and ionic strength ( $I$ ). SI is defined as the ratio of  $K_{iap}$  (ion activity product) and  $K_{sp}$  (solubility product).  $SI > 1$  ( $K_{iap} > K_{sp}$ ) denotes supersaturation,  $< 1$  ( $K_{iap} < K_{sp}$ ) denotes undersaturation, and  $K_{iap} = K_{sp}$  denotes equilibrium condition.

For our study, the seasonal SI mean of most minerals with exception of aragonite, calcite, dolomite, hydroxyapatite remains below zero denoting a state of undersaturation causing ion release (Fig. 5). In case of aragonite, it could be seen that as post-monsoon comes ( $SI > 0$ ), there is precipitation of  $Ca^{2+}$  and  $CO_3^{2-}$  to form the mineral, possibly providing a sinking effect to the  $Ca^{2+}$  concentration. However, near equilibrium conditions ( $SI \sim 0$ ) dictate that this effect could be marginal. Further, weak undersaturation of aragonite, calcite and dolomite during only monsoon as opposed to the other two seasons provides a clear indication that the rainfall recharge is the critical factor controlling the ion release. Over-saturation of hydroxyapatite ( $Ca_{10}(PO_4)_6(OH)_2$ ) explains that despite significant availability of  $Ca^{2+}$  ions, its sequestration through precipitation occurs through entire monsoon. The strongest correlation between  $Ca^{2+}$  and  $HCO_3^-$  ( $r = 0.86$ ; Supplementary Table 2) and between  $Ca^{2+}$  and  $SO_4^{2-}$  ( $r = 0.80$ ) is observed during the monsoon season, which provides more basis to our understanding that availability of  $Ca^{2+}$  mostly occurs through four minerals namely aragonite, anhydrite, calcite and dolomite. Similar undersaturation trends are observed in arsenolite ( $As_4O_6$ ) and claudetite ( $As_2O_3$ ) signifying that oxidation of arsenic sulfides is still active. Seasonal consistency in the correlation between arsenic and  $SO_4^{2-}$  ( $r = -0.32$ ,  $-0.26$  and  $-0.32$  during pre-monsoon, monsoon and post-monsoon respectively) and relatively stronger correlation between arsenic and  $Ca^{2+}$  during pre-monsoon ( $r = 0.53$ ) point towards minerals like gypsum and anhydrite which dissociate under the present alkalioxic (AO) conditions. Although it is important to emphasize that anions like  $SO_4^{2-}$  are naturally available in the atmosphere but an exceptionally high correlation between TDS and  $SO_4^{2-}$  during only monsoon season shows that the contribution of rainfall (rich in sulfate) in driving this mechanism is highest.

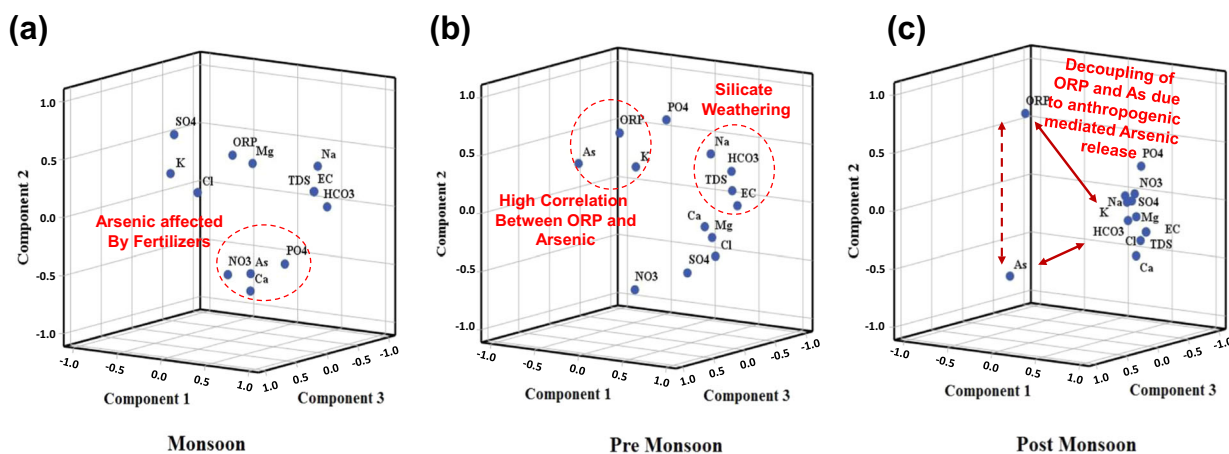
A critical analysis of the various sodium minerals shows that during pre-monsoon and monsoon, natron ( $Na_2CO_3 \cdot 10H_2O$ ) could be one of the sources for providing  $Na^+$  and  $HCO_3^-$  in the groundwater ( $SI \sim -10$ ;  $r = 0.82$  during pre-monsoon;  $r = 0.83$  during monsoon). This is also supported by strong correlation between  $HCO_3^-$ ,  $Na^+$  and TDS signifying that the ions in the groundwater must be a direct manifestation of natron dissolution.

The cumulative weakening of this relationship during the post-monsoon and increase in correlation between  $PO_4^{3-}$  and  $Na^+$  ( $r = 0.46$ ) shows that agrochemical leaching has disturbed the existing geogenic balance. It is worth noting that ORP, which is a dominant arsenic controlling factor, shows highest correlation with  $NO_3^-$  ( $r = -0.49$ ) and  $PO_4^{3-}$  ( $r = 0.38$ ) during post-monsoon.

#### Principal component analyses

PCA was done on the seasonal data to decompose it into five principal components (PCs) with a total of 86.1% variance explaining capability out of which the first three PCs can explain 74.3% variance (Supplementary Table 3). Parallel comparison revealed several differences in the governing hydro-chemistry of the pre-monsoon and post-monsoon, which were also analysed earlier in the bivariate and correlation analyses. PC1 of the pre-monsoon season has the highest variance of 37% with strong positive loadings imparted by EC, TDS,  $Na^+$  and  $HCO_3^-$  (Fig. 6). This PC is governed by silicate weathering, showing that ion exchange and sodium enrichment are predominantly mediated via silicate weathering during pre-monsoon<sup>36</sup>. This appears to be a cyclic phenomenon altering between the pre-monsoon and the post-monsoon as reverse ion exchange becomes the norm during the post-monsoon season. Two previous studies have also revealed that the soil of the Indo-Gangetic Plains is sodic<sup>6,9,14,26</sup>. Presence of sodium carbonate was found to immobilize  $Ca^{2+}$  resulting in a high sodium adsorption ratio (SAR) and ultimately, a high exchangeable sodium percentage (ESP). PC2 is represented by positive loading from  $Mg^{2+}$  and negative loading from pH with 20% variance. The exact hydro-chemistry of this PC was not clear; therefore, it was studied further by the application of hierarchical cluster analysis.

Although the level of  $NO_3^-$  in the groundwater was very low, however, it imparted a strong loading on PC3 during the pre-monsoon season. The lone representation of  $NO_3^-$  was considered to be anomalous and was studied further with the application of HCA. PC4 with 9.3% variance is represented by high positive loading from  $PO_4^{3-}$  and negative loading due to ORP representing anthropogenic factors primarily of the agricultural origin, as seen from its positive correlation with  $Ca^{2+}$  ( $r = 0.5$ , Supplementary Table 2) along with a low positive loading of 0.49 in PC4. The causative agrochemical appears to be calcium phosphate besides geogenic contributions from minerals such as hydroxyapatite ( $Ca_{10}(PO_4)_6(OH)_2$ ). PC5 contributes to 7.9% variance and has significant positive loading only due to  $Cl^-$ . Since the contributing mechanism is not clear, therefore, it has been



**Fig. 6** Three-dimensional visualization of PCA loadings. There exists a characteristic difference in the processes active during **a** monsoon, **b** pre-monsoon and **c** post-monsoon season.

studied further with HCA (Supplementary Fig. 3). PC1 with 41.4% variance for post-monsoon has a much more dominant representation of carbonate weathering as evident from the positive loadings of  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ , EC and TDS. Both calcite and dolomite weathering were found to be equally prevalent as revealed through the scatter plots and correlation.  $\text{Cl}^-$  was also found to impart positive loading on PC1 during the post-monsoon season which again reaffirms the occurrence of reverse ion exchange reaction with changing seasons. An influx of fresh rainwater with higher DO could lead to an elevation in ORP, while at the same time the low levels of  $\text{NO}_3^-$  may be further diminished due to dilution in the post-monsoon. This explains the reason behind the positive loading of ORP and negative loading of  $\text{NO}_3^-$  in PC2 during the post-monsoon season. The last PC of post-monsoon with 12.8% variance is represented by  $\text{Na}^+$  and pH reflecting the sodic nature of the soils in the Gangetic plains.

#### Hierarchical cluster analyses

HCA was applied to gain a better understanding and validate the findings of the bivariate and PCA analyses (Supplementary Fig. 3). The Dendrogram generated for the pre-monsoon season has three distinct clusters. The first cluster is represented by EC, TDS,  $\text{HCO}_3^-$  and  $\text{Na}^+$ , which shows silicate weathering and ion exchange resulting in  $\text{Na-HCO}_3$  type groundwater. The second cluster is represented by  $\text{Ca}^{2+}$ ,  $\text{PO}_4^{3-}$ ,  $\text{NO}_3^-$  and pH. The strong loading of  $\text{NO}_3^-$  in PC3 during pre-monsoon together with PC3 and PC4 represent agrochemicals<sup>16,19</sup>. A single giant cluster is represented by  $\text{K}^+$ ,  $\text{SO}_4^{2-}$ ,  $\text{Mg}^{2+}$ ,  $\text{Cl}^-$  and ORP, which also points towards a different class of agrochemicals for the correction of potassium ( $\text{K}_2\text{SO}_4$ ) and magnesium ( $\text{MgCl}_2$ ) deficiencies in the soil. In the post-monsoon season three clusters were identified; the first group consists of two sub-clusters; the first being EC, TDS and  $\text{HCO}_3^-$ , and the second being  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$  and  $\text{Cl}^-$ .

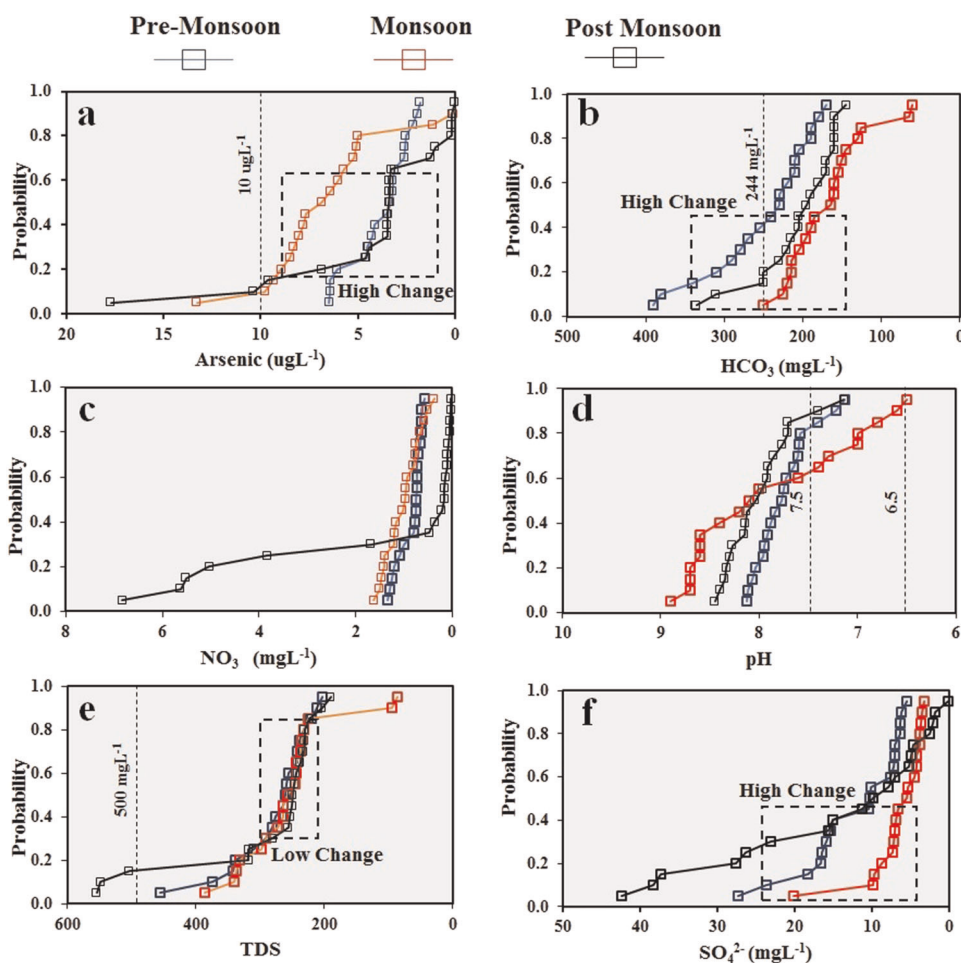
As explained in the PCA, this group represents the dominance of carbonate weathering during the post-monsoon with the secondary involvement of reverse ion exchange. The second cluster consists of  $\text{SO}_4^{2-}$  and  $\text{NO}_3^-$ . Although the correlation matrix does not reveal any trend between the two but the HCA offers a more transparent look at the possible occurrence of wet deposition of these anions during the rainy season. Darbhanga district is home to several industrial estates which could be linked to this problem. The last cluster consists of two sub-clusters; one consisting of pH and  $\text{Na}^+$ , and the other consisting of ORP,  $\text{PO}_4^{3-}$  and  $\text{K}^+$ . This implies high sodic nature of the soil along with agrochemical usage during the post-monsoon season.

#### Probabilistic exceedance and seasonal occurrences

The primary objective behind mapping the probability exceedance of critical parameters is to validate their response with a physical/chemical process change. Figure 7 correctly identifies the arsenic,  $\text{HCO}_3^-$ , pH,  $\text{NO}_3^-$  and  $\text{SO}_4^{2-}$  anomalies existing during different seasons. The monsoon probability trends are steeper and more distant as compared to other season trends (Fig. 7a). This validates the role of rainfall recharge when we see the complementing trends of  $\text{SO}_4^{2-}$ ,  $\text{HCO}_3^-$  and pH. It could also be stated that arsenic in the concentration range of  $8\text{--}5\ \mu\text{g L}^{-1}$  is much more susceptible to a physical process such as precipitation-led recharge mediated primarily through  $\text{SO}_4^{2-}$  (Fig. 7). pH in the range of 6.5 and 7.5 can also be validated as a contributing factor when there is a shift in the  $\text{HCO}_3^-$  concentration between 150 and  $350\ \text{mg L}^{-1}$ .

The presence of arsenic above  $5\ \mu\text{g L}^{-1}$  has been mapped in terms of a probability for all the three seasons (Fig. 8). The Akaike information criteria (AIC) for monsoon is 98.88, which signifies that the model has retained relatively higher information of the original data while fitting (Supplementary Table 4). In terms of model fitting, a comparison between the three seasons show that pre-monsoon spherical model is the best fit model (AIC = 66.79) (Supplementary Figs. 4–6). However, apart from model fitting, the capability of prediction of the pre-monsoon spherical model is relatively weak compared to monsoon Gaussian model (RMSS = 0.976; RMSE = 0.087). The post-monsoon arsenic model performs relatively weak as compared to pre-monsoon and monsoon season models, as the variance in the datasets diminishes greatly. A comparative analysis of standard deviation shows that pre-monsoon models have a peak distribution and less spread ( $\sigma = 1.51$ ) as compared to other season models further validating its fitting capabilities. The increase in the spread of region with high arsenic concentration during the monsoon season after the pre-monsoon shows that rainfall-led dilution is having its impact in the transportation of arsenic. The drainage and the general lineament of the region play a role in transporting arsenic in the northward's direction.

The very fact that the probability of occurrence of arsenic above  $5\ \mu\text{g L}^{-1}$  shrinks to 0.5 from above 0.8 during post-monsoon shows that arsenic is being stabilized in the bed and the presence of iron-oxy-hydroxides (Fe-OOH) has become marginalized as the concentration of Fe reaches to almost zero<sup>37</sup>. Using a GIS-based overlay technique, the study region has been mapped for the Water Quality Index (WQI) (Supplementary Table 5). A WQI below 50 is considered *excellent* with regards to drinking standards and a WQI between 50 and 100 is regarded as *very good*. The entire region though falling in *excellent* to the *very good* category has geogenic arsenic contamination as the worrying factor. The



**Fig. 7 Exceedance probability of various parameters impacting arsenic concentration.** Corresponding change in exceedance values of (a) arsenic and (b)  $\text{HCO}_3^-$  validates the role of rainfall recharge in arsenic mobilization process. Factors like (c)  $\text{NO}_3^-$ , (d) pH, (e) TDS and (f)  $\text{SO}_4^{2-}$  show isolated yet confirmatory signatures of anthropogenic intrusion.

impact of consuming water contaminated with arsenic is much more severe as compared to consuming water with excessive TDS. Therefore, through spatial distribution mapping, it has been tried to understand the areas where arsenic remains a potent challenge in changing the water quality dynamics. The results present a concerning reality where it has been found that arsenic is a dominant factor which could change the WQI from 25 to 35% in almost the entirety of the study area during pre-monsoon, even if its ability to cause mutagenicity and carcinogenicity is ignored. At small places in the extreme north and the central of Darbhanga district the percentage rise to almost 35%.

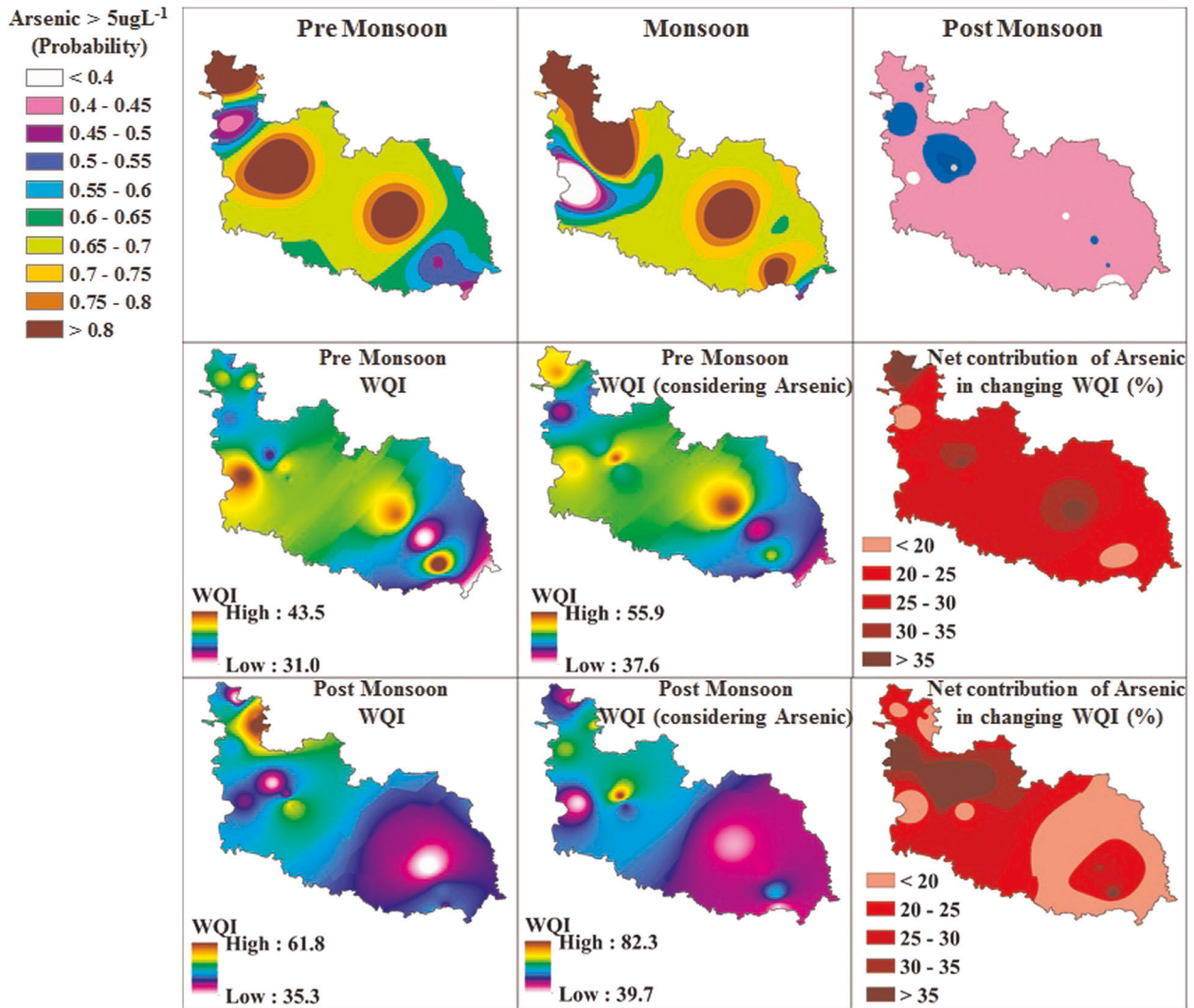
#### Conceptual modelling and its implications

By following an integrated approach involving all the analytical and statistical techniques, we have developed a conceptual understanding of the region vis-à-vis the operative processes active for arsenic mobilization (Fig. 9). It should be noted that our study was conducted under three economic settings, i.e. (1) agriculture, (2) urban and (3) rural, which differ significantly in terms of arsenic susceptibility and groundwater exploitation. As opposed to the popular belief, the agriculture region showed less arsenic ( $<10 \mu\text{g L}^{-1}$ ) which shows that the arsenical pesticides are not used as a possible mitigation strategy against crop weeds. Also, it shows that monsoon-led irrigation is still prevalent and the pressure on groundwater exploitation has not reached the levels of West Bengal and Lower Gangetic Plains. The high fertility of the alluvium sediments has also caused less exposure of fertilizers as

the concentration of nitrate and phosphates are mostly within the permissible levels. However, the presence of arsenic above  $10 \mu\text{g L}^{-1}$  in the urban setting shows that the threat of arsenic contamination can loom quickly if over-exploitation of groundwater takes place continuously. A significant problem is also associated with lowering of pH due to high concentration of anion such as sulfate, which may make the conditions optimum for arsenic adsorption over iron-oxy-hydroxides. This is a big challenge for regions where fewer checks and balances are put in place in terms of curbing industrial effluxes into the open system. The rural region which is characterized by a mix of agricultural and commercial activities may prove hotbeds of arsenic toxicity as could be seen through the trends obtained in different seasons (Fig. 8). Also, the high undersaturation of arsenolite and claudetite ( $\sim 15$ ) (Fig. 5) is a worrying cause as activation of arsenic breakdown can occur through micro-organism if triggers like nitrate and sulfate are not restricted for their intrusion in the groundwater.

As far as the hydro-chemistry of the groundwater is concerned, silicate weathering was found to be prominent during the pre-monsoon; however, it would be important to point out here that the  $\text{K}^+$  in the groundwater could not be traced to silicate weathering as was seen from the low values as well as the lack of correlation between  $\text{K}^+$  and  $\text{Na}^+$ . Based on the previous studies, it could be inferred that soda feldspar or albite are the dominant minerals in the aquifers of the study area as the soils of the Ganga plains are sodic in nature. This has further been supported by both





**Fig. 8** Probability distribution of arsenic above  $5 \mu\text{g L}^{-1}$ . There is a marked difference in the suitability of groundwater when arsenic is considered as one of the parameters in WQI.

the PCA and the HCA. MVA also revealed the cyclic change from ion exchange during pre-monsoon to reverse ion exchange during the post-monsoon. As post-monsoon advanced, carbonate weathering was found to take precedence over silicate weathering leading to higher concentrations of  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ . Statistically significant correlations of  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  with  $\text{Cl}^-$  also highlights the dominance of reverse ion exchange during post-monsoon. Overall, anthropogenic forcing on the groundwater quality was found to be marginal and mostly consisted of agrochemical (dominant being calcium nitrate and calcium phosphate fertilizers) usage and isolated incidences of chemicals like bleaching powder. However, one of the key findings of the HCA analysis was the possibility of wet deposition of the anions ( $\text{SO}_4^{2-}$  and  $\text{NO}_3^-$ ) during the post-monsoon season. Eventually, the present study clearly displayed the crucial role played by MVA; not only were these techniques supportive of the bivariate analyses but in several incidences where the observed trends were unclear, the MVA offered a much clearer explanation (Supplementary Fig. 7).

Due to the high population density and agriculture intensiveness of the region, excessive abstraction of groundwater and its contamination due to nutrient leaching from agricultural inputs go hand in hand. The detection of arsenic above  $10 \mu\text{g L}^{-1}$  at only

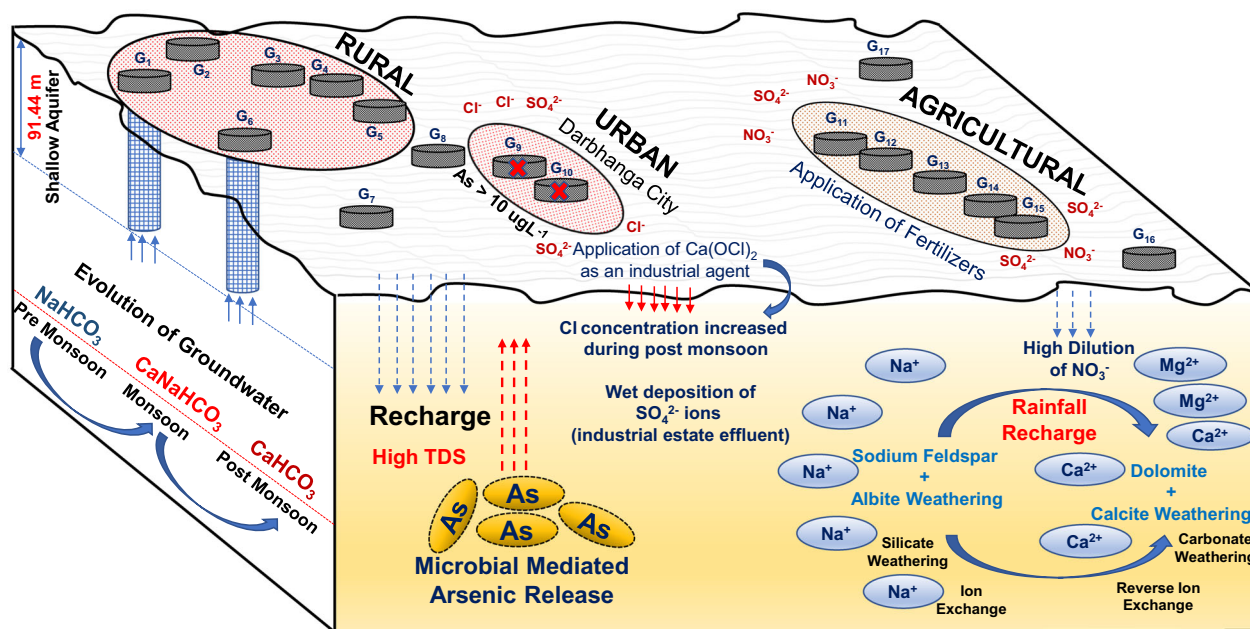
50 m depth signifies that the situation can be difficult to manage if a policy limiting the groundwater abstraction is not put in place. Further, changes in pH and ORP over seasons signifies that the industrial effluents during monsoon are also finding their pathway in the aquifer system.

Regions with high risk of arsenic contamination are also regions with high population density; therefore, there will be risk for the vast population to develop cancer through both dermal and oral exposure. At a policy level we, therefore, suggest that there is a dire need to create hazard buffers where the activities of groundwater extraction could be limited as well as the industrial extrusion be quantified and checked.

## METHODS

### Study area

Darbhanga district in Bihar extends between  $85^{\circ}40'$  E and  $86^{\circ}10'$  E latitudes and between  $26^{\circ}30'$  N and  $26^{\circ}0'$  N longitudes (Fig. 1) under the Ganga floodplains of the Bhagmati sub-basin. The lineament of the subsurface of the study region is straight with the geomorphological origin being of fluvial nature (<https://bhuvan.nrsc.gov.in>). The region has very less variation in the elevation from the mean sea level, with the terrain being flat and consisting of vast expanses of alluvial plains. The soil in the study



**Fig. 9 Conceptual model of the operative processes active in the study region.** The governing hydro-chemical processes exhibit significant deviations due to leaching of  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$  and other pH affecting chemicals.

region consists of a mixture of sand, silt and clays (calcareous) of different proportions<sup>1</sup>. The general drainage of the study region slopes slightly towards the south with a slight depression in the elevation at the centre of the study region (ASTER DEM 30 m spatial resolution <https://earthexplorer.usgs.gov/>). Despite having a tropical three-season climate (summer, winter and monsoon), the summer season predominates. May is the hottest month with temperature recorded above 40 °C. The region receives almost the same rainfall (1142.3 mm) as the average rainfall received by India (1200 mm) with 90% of the rainfall demand being met during the monsoon season (The National Oceanic and Atmospheric Administration (NOAA), <https://www.noaa.gov/>). The aquifers that consist of water-bearing sub-layers are of quaternary sediment origin. The total width of these formations at places reaches to 2 kilometers.

#### Sampling strategy and laboratory analyses

A total of 171 groundwater samples were collected during the pre-monsoon (57—during April 2011), monsoon (57—during June 2011) and post-monsoon (57—during January 2013). Collection was done in pre-washed polyethylene bottles after pumping the tube wells for 5–10 min. Raw samples were used for anion analyses, and filtered through Millipore (0.45  $\mu\text{m}$ ) and acidified to a pH  $\leq 2$  with  $\text{HNO}_3$  (2 M ultra-pure grade, Merck, India) used for cation analyses. Fourteen water quality parameters were analysed, out of which  $\text{HCO}_3^-$ , pH, EC, TDS and ORP were analysed in situ using Hanna-Multi-parameter Water Quality portable metre (H19828).  $\text{Na}^+$  and  $\text{K}^+$  were analysed using flame photometer (Systronics-128), and  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  using ICP-OES (Perkin Elmer Optima 2100 DV). Chloride was analysed using argentometric methods<sup>38</sup>, while  $\text{SO}_4^{2-}$ ,  $\text{PO}_4^{3-}$  and  $\text{NO}_3^-$  were analysed using UV Spectrophotometry (Shimadzu UV-1700)<sup>38</sup>. Arsenic concentration was determined using AAS (Thermo-Scientific Spectrometer 3000). Data validation and accuracy of the analytical methods was checked using normalized ionic charge balance. Bivariate plots were prepared using Microsoft Excel 2016<sup>39</sup> (Microsoft, Redmond, Washington, USA). SPSS 20 (IBM Armonk, New York, USA) was utilized for generating the correlation matrix. Arc-GIS 10.3 was used for spatial mapping and generating probabilities of spatial arsenic occurrence.

#### Saturation index

The saturation index (SI) is defined as the ratio of  $K_{\text{iap}}$  (ion activity product) and  $K_{\text{sp}}$  (solubility product). In the present study, the SI of anhydrite ( $\text{CaSO}_4$ ), gypsum ( $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ ), halite ( $\text{NaCl}$ ), epsomite, calcite, and dolomite was studied for the three seasons. For which, first of all, the ionic strength ( $I$ ) was estimated using the molality ( $m$ ) and

valency ( $Z$ ), as per Eq. 1<sup>20</sup>.

$$I = 0.5 \times \sum_{i=1}^n m \times Z^2. \quad (1)$$

Using the value of ionic strength, the activities of the individual solute species were computed by Eq. 2.

$$\text{Activity} = m \times a, \quad (2)$$

where  $a$  is the activity coefficient estimated using the Debye–Huckel equation

$$\log a = \frac{-A \times Z^2 \times \sqrt{I}}{1 + a_0 \times B \times \sqrt{I}}, \quad (3)$$

where  $a_0$  is the ionic radius and  $A$  and  $B$  are constants dependent on temperature and pressure.

#### Probability distribution of arsenic

The concentration of arsenic was predicted for different locations using kriging<sup>40</sup>. Measured As concentration of each location  $Z(x_0, x_1, x_2, x_3, \dots)$  along with their spatial correlation was used as predictors. Further by associating a weight  $\lambda$  based on spatial arrangement and the distance between the sampling points (Eq. 4), an empirical fitting was done<sup>41</sup>.

$$Z(S_0) = \sum_{i=1}^N \lambda_i \times Z(S_i), \quad (4)$$

where  $Z(S_0)$  corresponds to the location at which prediction must be done.  $i$  is the number of sampled locations varying from 1 to 19 for each season.  $\lambda_i$  is the associated weight, the value of which has been determined by fitting a suitable variogram model and  $Z(S_i)$  is the value of concentration at a known location<sup>42</sup>. In order to generate a prediction surface, an empirical semi-variogram was estimated based on the different relationships between the various pairs of sampling points (Eq. 5)<sup>43</sup>.

$$\text{Empirical semi-variogram}(\text{distance}) = 0.5 \times \text{average}((v_i - v_j))^2. \quad (5)$$

In the next step we fitted a modelled semi-variogram to the points which form the empirical semi-variogram, which was used for computing three parameters namely sill, nugget and range. For estimating the probability of As concentration, two variables namely (i) indicator code ( $I$ ) and uniform value ( $U$  is also known as standardized rank) are defined as the main and the auxiliary variable respectively<sup>44</sup>.

$$U(x) = \frac{r}{n}, \quad (6)$$

where  $r$  denotes the rank and  $n$  is the number of observations (19 points representing 57 samples for pre-monsoon, monsoon and post-monsoon each). In terms of uniform value  $U(x)$ , Eq. 4 can be further modified as given by Eq. 7.

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Ux - Ux + h]^2. \quad (7)$$

For different pairs of observation points, the cross semi-variogram can be estimated as follows.

$$\gamma_C(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \{[lx - lx + h] * [Ux - Ux + h]\}^2. \quad (8)$$

The final estimation of probability is based on Eq. 9 (modifying Eq. 8)

$$P_n = \sum_{i=1}^n l(x) \times \lambda_i + \sum_{i=1}^n U(x) \times \lambda_u, \quad (9)$$

where  $\lambda_i$  and  $\lambda_u$  are the respective weights assigned to the indicator code ( $l(x)$ ) and uniform value ( $U(x)$ ) function respectively. In order to accurately predict the concentration values, the variance is kept minimum and the weight factors  $\lambda_i$  and  $\lambda_u$  are solved by Lagrange constraint method under un-biased conditions assuming  $\sum \lambda_i = 1$  and  $\sum \lambda_u = 0$  by Eqs. 10 and 11.

$$\sum_{j=1}^n \lambda_{ij} \times \gamma(x_j - x_i) + \sum_{p=1}^n \lambda_{up} \times \gamma_C(x_p - x_i) = \gamma(x_0 - x_i) - \mu_i, \quad (10)$$

$$\sum_{j=1}^n \lambda_{ij} \times \gamma_C(x_j - x_i) + \sum_{p=1}^n \lambda_{up} \times \gamma_U(x_p - x_i) = \gamma_C(x_0 - x_i) - \mu_u, \quad (11)$$

where  $\mu_i$  and  $\mu_u$  are the Lagrange multipliers,  $\gamma$  and  $\gamma_C$  are the semi-variance and cross-variance of  $l(x)$  and  $\gamma_U$  is the semi-variance of  $U(x)$ <sup>39</sup>.

A threshold concentration value of  $5 \mu\text{g L}^{-1}$  is selected for As-mapping distribution zones based on probability ranging from 0 to 1. The probability distribution maps are demarcated into various zones and corresponding change in the zones is studied for the three seasons, i.e. pre-monsoon, monsoon and post-monsoon. Usually the datasets are not normally distributed; therefore, normal score transformation is applied where the data are structured like standard normal distribution<sup>45</sup>. It is important to note that the present study has been designed specifically to understand the potential of monsoon recharge in alleviating the arsenic concentration in the groundwater. The role of anthropogenically induced heavy metal toxicity has not been accounted for based on the understanding of the region that dominantly depends on agriculture for its livelihood. Therefore, the role of metals like Pb, Hg, Cd, etc., which may have a role in arsenic cyclicity, does not come under the scope of this study.

## DATA AVAILABILITY

All other relevant data are available from the corresponding author on request.

Received: 19 January 2020; Accepted: 7 May 2020;

Published online: 04 June 2020

## REFERENCES

- Patel, A. K., Das, N. & Kumar, M. Multilayer arsenic mobilization and multimetal co-enrichment in the alluvium (Brahmaputra) plains of India: a tale of redox domination along the depth. *Chemosphere* **224**, 140–150 (2019).
- Kumar, M. et al. Arsenic enrichment in groundwater in the middle Gangetic Plain of Ghazipur District in Uttar Pradesh, India. *J. Geochem. Explor.* **105**, 83–94 (2010).
- Das, N., Das, A., Sarma, K. P. & Kumar, M. Provenance, prevalence and health perspective of co-occurrences of arsenic, fluoride and uranium in the aquifers of the Brahmaputra River floodplain. *Chemosphere* **194**, 755–772 (2018).
- Nickson, R. T., McArthur, J. M., Ravenscroft, P., Burgess, W. G. & Ahmed, K. M. Mechanism of arsenic release to groundwater, Bangladesh and West Bengal. *Appl. Geochem.* **15**, 403–413 (2000).
- Kumar, M., Goswami, R., Patel, A. K., Srivastava, M. & Das, N. Scenario, perspectives and mechanism of arsenic and fluoride co-occurrence in the groundwater: a review. *Chemosphere* **249**, 126126 (2020).
- Borah, R., Taki, K., Gogoi, A., Das, P. & Kumar, M. Contemporary distribution and impending mobility of arsenic, copper and zinc in a tropical (Brahmaputra) river bed sediments, Assam, India. *Ecotoxicol. Environ. Saf.* **161**, 769–776 (2018).
- Kumar, M., Ramanathan, A. L., Rao, M. S. & Kumar, B. Identification and evaluation of hydrogeochemical processes in the groundwater environment of Delhi, India. *Environ. Geol.* **50**, 1025–1039 (2006).
- Kumar, M., Kumari, K., Ramanathan, A. L. & Saxena, R. A comparative evaluation of groundwater suitability for irrigation and drinking purposes in two intensively cultivated districts of Punjab, India. *Environ. Geol.* **53**, 553–574 (2007).
- Mazumder, P., Sharma, S. K., Taki, K., Kalamdhad, A. S. & Kumar, M. Microbes involved in arsenic mobilization and respiration: a review on isolation, identification, isolates and implications. *Environ. Geochem. Health* **1–27**, <https://doi.org/10.1007/s10653-020-00549-8> (2020).
- Garbinski, L. D., Rosen, B. P. & Chen, J. Pathways of arsenic uptake and efflux. *Environ. Int.* **126**, 585–597 (2019).
- Yang, H.-C. & Rosen, B. P. New mechanisms of bacterial arsenic resistance. *Biomed. J.* **39**, 5–13 (2016).
- Dopp, E., Von Recklinghausen, U., Diaz-Bone, R., Hirner, A. V. & Rettenmeier, A. W. Cellular uptake, subcellular distribution and toxicity of arsenic compounds in methylating and non-methylating cells. *Environ. Res.* **110**, 435–442 (2010).
- Jensen, M. Ø., Tajkhorshid, E. & Schulten, K. The mechanism of glycerol conduction in aquaglyceroporins. *Structure* **9**, 1083–1093 (2001).
- Saha, D. Arsenic groundwater contamination in parts of middle Ganga plain, Bihar. *Curr. Sci.* **97**, 753–755 (2009).
- Kumar, M. Understanding the remobilization of copper, zinc, cadmium and lead due to ageing through sequential extraction and isotopic exchangeability. *Environ. Monit. Assess.* **188**, 381 (2016).
- Goswami, R., Kumar, M., Biyani, N. & Shea, P. J. Arsenic exposure and perception of health risk due to groundwater contamination in Majuli (river island), Assam, India. *Environ. Geochem. Health* **42**, 443–460 (2019).
- Kumar, M. et al. Understanding the seasonal dynamics of the groundwater hydrogeochemistry in National Capital Territory (NCT) of India through geochemical modelling. *Aquat. Geochem.* **22**, 211–224 (2016).
- Ravenscroft, P., Burgess, W. G., Ahmed, K. M., Burren, M. & Perrin, J. Arsenic in groundwater of the Bengal Basin, Bangladesh: distribution, field relations, and hydrogeological setting. *Hydrogeol. J.* **13**, 727–751 (2005).
- Das, A. & Kumar, M. Arsenic enrichment in the groundwater of Diphu, North-east India: coupled application of major ion chemistry, speciation modeling, and multivariate statistical techniques. *CLEAN—Soil Air Water* **43**, 1501–1513 (2015).
- Kumar, M., Rao, M. S., Kumar, B. & Ramanathan, A. Identification of aquifer-recharge zones and sources in an urban development area (Delhi, India), by correlating isotopic tracers with hydrological features. *Hydrogeol. J.* **19**, 463–474 (2011).
- Singh, A., Patel, A. K., Ramanathan, A. & Kumar, M. Climatic influences on arsenic health risk in the metamorphic precambrian deposits of Sri Lanka: a re-analysis-based critical review. *J. Clim. Change* **6**, 15–24 (2020).
- Singh, A. et al. Prediction of arsenic vulnerable zones in the groundwater environment of a rapidly urbanizing setup, Guwahati, India. *Geochemistry* 125590, <https://doi.org/10.1016/j.chemer.2019.125590> (2019).
- Singh, A., Patel, A. K. & Kumar, M. in *Emerging Issues in the Water Environment during Anthropocene* 285–300 (Springer, 2020).
- Chatterjee, D. et al. Assessment of arsenic exposure from groundwater and rice in Bengal Delta Region, West Bengal, India. *Water Res.* **44**, 5803–5812 (2010).
- Anawar, H. M., Akai, J., Mostofa, K. M. G., Safiullah, S. & Tareq, S. M. Arsenic poisoning in groundwater: health risk and geochemical sources in Bangladesh. *Environ. Int.* **27**, 597–604 (2002).
- Harvey, C. F. et al. Groundwater arsenic contamination on the Ganges Delta: biogeochemistry, hydrology, human perturbations, and human suffering on a large scale. *Comptes Rendus Geosci.* **337**, 285–296 (2005).
- Patel, A. K., Das, N., Goswami, R. & Kumar, M. Arsenic mobility and potential co-leaching of fluoride from the sediments of three tributaries of the Upper Brahmaputra floodplain, Lakhimpur, Assam, India. *J. Geochem. Explor.* **203**, 45–58 (2019).
- Nieto, P., Custodio, E. & Manzano, M. Baseline groundwater quality: a European approach. *Environ. Sci. Policy* **8**, 399–409 (2005).
- Chen, K., Jiao, J. J., Huang, J. & Huang, R. Multivariate statistical evaluation of trace elements in groundwater in a coastal area in Shenzhen, China. *Environ. Pollut.* **147**, 771–780 (2007).
- Cloutier, V., Lefebvre, R., Therrien, R. & Savard, M. M. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J. Hydrol.* **353**, 294–313 (2008).

31. Menció, A. & Mas-Pla, J. Assessment by multivariate analysis of groundwater–surface water interactions in urbanized Mediterranean streams. *J. Hydrol.* **352**, 355–366 (2008).
32. Baig, J. A. et al. Speciation and evaluation of Arsenic in surface water and groundwater samples: a multivariate case study. *Ecotoxicol. Environ. Saf.* **73**, 914–923 (2010).
33. World Health Organization. World report on disability 2011. [https://www.who.int/disabilities/world\\_report/2011/report/en/](https://www.who.int/disabilities/world_report/2011/report/en/) (WHO, Geneva, 2011).
34. Bureau of Indian Standards. IS 10500: 2012 Drinking Water — Specification (second revision) (BIS, New Delhi, 2012).
35. Das, N. et al. Effect of river proximity on the arsenic and fluoride distribution in the aquifers of the Brahmaputra floodplains, Assam, northeast India. *Groundw. Sustain. Dev.* **2**, 130–142 (2016).
36. Chae, G.-T., Yun, S.-T., Kim, K. & Mayer, B. Hydrogeochemistry of sodium-bicarbonate type bedrock groundwater in the Pocheon spa area, South Korea: water–rock interaction and hydrologic mixing. *J. Hydrol.* **321**, 326–343 (2006).
37. Solaiman, A. R. M., Meharg, A. A., Gault, A. G. & Charnock, J. M. Arsenic mobilization from iron oxyhydroxides is regulated by organic matter carbon to nitrogen (C: N) ratio. *Environ. Int.* **35**, 480–484 (2009).
38. American Public Health Association. Standard methods for examination of water and wastewater, method 9221 B (APHA, Washington DC, 2012).
39. Juang, K.-W. & Lee, D.-Y. Simple indicator kriging for estimating the probability of incorrectly delineating hazardous areas in a contaminated site. *Environ. Sci. Technol.* **32**, 2487–2493 (1998).
40. Yao, X. et al. Comparison of four spatial interpolation methods for estimating soil moisture in a complex terrain catchment. *PLoS ONE* **8**, e54660 (2013).
41. Oliver, M. A. & Webster, R. Kriging: a method of interpolation for geographical information systems. *Int. J. Geogr. Inf. Syst.* **4**, 313–332 (1990).
42. Matheron, G. *Les variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature* (Masson et CIE, 1965).
43. Cressie, N. Statistics for spatial data. *Terra Nova* **4**, 613–617 (1992).
44. Eldeiry, A. A. & Garcia, L. A. Using indicator kriging technique for soil salinity and yield management. *J. Irrig. Drain. Eng.* **137**, 82–93 (2011).
45. Perttunen, C. D. & Stuckman, B. E. The normal score transformation applied to a multi-univariate method of global optimization. *J. Glob. Optim.* **2**, 167–176 (1992).

## ACKNOWLEDGEMENTS

This study is supported by the DST-SERB Young Scientist Grant to M.K.

## AUTHOR CONTRIBUTIONS

A.S.: Analyses, Methodology, Original draft, Data curation and Revision; A.K.P.: Sampling, Analyses, Data curation and Editing. J.P.D.: Visualization, Sampling, Analyses and Data curation. M.K.: Conceptualization, Visualization, Reviewing, Fund management, Project administration, Editing and Supervision.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41545-020-0075-5>.

**Correspondence** and requests for materials should be addressed to M.K.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020