Water Research 185 (2020) 116257

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Rising arsenic concentrations from dewatering a geothermally influenced aquifer in central Mexico

Peter S.K. Knappett[a,*], Yanmei Li[b], Isidro Loza[b], Horacio Hernandez[c], Manuel Avilés[k], David Haaf[d,e], Santanu Majumder[a], Yibin Huang[a], Brian Lynch[a], Viridiana Piña[l], Jianjun Wang[f], Lenny Winkel[d,e], Jürgen Mahlknecht[g], Saugata Datta[h], William Thurston[i], Dylan Terrell[i], D. Kirk Nordstrom[j]

[a] Dept. Geology & Geophysics, Texas A&M University, College Station, TX 77843, United States
[b] Dept. Mines, Metallurgy and Geology Engineering, University of Guanajuato, Guanajuato 36000, México
[c] Dept. Geomatic and Hydraulic Engineering, University of Guanajuato, Guanajuato 36000, México
[d] Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland
[e] Dept. Environmental Systems Science, ETH, 8092 Zürich, Switzerland
[f] School of Environmental Studies, China University of Geosciences, Wuhan 430074, China
[g] Centro del Agua para América Latina y el Caribe, Escuela de Ingenieria y Ciencias, Tecnologico de Monterrey, Monterrey, 64849, México
[h] Dept. Geological Sciences, University of Texas at San Antonio, San Antonio, TX, 78249, United States
[i] Caminos del Agua, San Miguel de Allende, Guanajuato 37712, Mexico
[j] United States Geological Survey, Boulder, CO 80303, United States
[k] Master Program of Water Science, University of Guanajuato, Guanajuato, 36000, Mexico
[l] Doctoral Program of Water Science and Technology, University of Guanajuato, Guanajuato, 36000, Mexico

ARTICLE INFO

Article history:
Received 23 January 2020
Revised 29 July 2020
Accepted 30 July 2020
Available online 31 July 2020

Keywords:
Arsenic
Groundwater
Fluoride
Geogenic
Nitrate
Geothermal
Wellbore mixing

ABSTRACT

This study identifies causes of rising arsenic (As) concentrations over 17 years in an inter-montane aquifer system located just north of the Trans-Mexican-Volcanic-Belt in the Mesa central physiographic region that is extensively developed by long-screened production wells. Arsenic concentrations increased by more than 10 μg/L in 14% (3/22) of re-sampled wells. Similarly, in a larger scale analysis wherein As concentrations measured in 137 wells in 2016 were compared to interpolated, baseline concentrations from 246 wells in 1999, As concentrations rose more than 10 μg/L in 30% of wells. Between 1999 and 2016, the percentage of all wells sampled in each basin-wide sampling campaign exceeding the World Health Organization's 10 μg/L drinking water limit increased from 38 to 64%. Principal Components Analysis (PCA), step-wise multiple regression, and Random Forest modeling (RF) revealed that high As concentrations are closely associated with high pH and temperature, and high concentrations of fluoride (F), molybdenum (Mo), lithium (Li), sodium (Na) and silica (Si), but low calcium (Ca) and nitrate (NO_3) concentrations. Pumping-induced mixing with hot, geothermally impacted groundwater generates alkaline water through hydrolysis of silicate minerals. The rising pH converts oxyanion sorption sites from positive to negative releasing As (and Mo) to pore waters. The negative correlation between nitrate and As concentrations can be explained by conservative mixing of shallow, young groundwater with geothermally influenced groundwater. Therefore water carrying an anthropogenic contaminant dilutes water carrying geogenic contaminants. This process is enabled by long well screens. Over-exploitation of aquifers in geothermal regions for agriculture can drive As concentrations in water from production wells to toxic levels even as the total dissolved solids remain low.

1. Introduction

The aim of Target 6.1 within the United Nation's Sustainable Development Goal 6 is to achieve universal, convenient access to safe, and affordable drinking water. Global Indicator 6.1.1 tracks progress towards achieving universal access to safe drinking water and states that priority contaminants like arsenic (As) and fluo-

* Corresponding author.
E-mail addresses: knappett@tamu.edu (P.S.K. Knappett), jurgen@tec.mx (J. Mahlknecht), saugata.datta@utsa.edu (S. Datta).

https://doi.org/10.1016/j.watres.2020.116257
0043-1354/

ride (F) should be progressively monitored in household drinking water (United Nations-Water, 2017). The testing and subsequent mitigation will require massive investment by governments, non-governmental organizations (NGOs), private households and businesses. The success of mitigation strategies which are designed based on the knowledge that is generated through these testing campaigns will be susceptible to the stability of As and F concentrations in groundwater. Although groundwater chemistry is commonly stable over time, As concentrations have been observed to change over time (Mihajlov et al., 2020). Seasonal concentration fluctuations may occur in shallow aquifers influenced by seasonal surface water table fluctuations (Schaefer et al., 2016) and long-term increases in As concentrations occur in production wells that are pumped at high rates (Ayotte et al., 2015). Long-term and large-scale studies suggest that changes in As concentrations can be substantial across a region within a wide variety of aquifer settings under the influence of intensive pumping (Appleyard et al., 2006; Ayotte et al., 2015, 2011; Erban et al., 2013; Harvey et al., 2005; Smith et al., 2018; Zhang et al., 2018). As rampant groundwater depletion for agriculture in arid regions continues around the world, rising concentrations of geogenic contaminants and resulting chronic diseases may result (Hellegers et al., 2011; Ling et al., 2011). The impacts of intensive pumping on concentrations of geogenic As in over-exploited, inter-montane aquifer systems under the influence of geothermal waters, however, has not to our knowledge been systematically studied at a regional scale.

Three types of aquifers commonly contain naturally concentrations of dissolved As that exceed the World Health Organization (WHO) drinking water limit of 10 μg/L (Smedley and Kinniburgh, 2002): 1) iron-reducing, neutral pH ground waters with abundant dissolved organic carbon (DOC), for example within the Ganges-Brahmaputra delta (Fendorf et al., 2010; Nickson et al., 1998); 2) high pH waters in semi-arid or arid regions, for example in Pakistan (Naseem and McArthur, 2018); and 3) geothermally influenced ground waters (Grassi et al., 2014; Guo et al., 2008) which are typical of western Latin America (Bundschuh et al., 2012; Munoz et al., 2015; Nordstrom, 2002) and the African Rift Valley (Darling et al., 1996; Rango et al., 2013). Furthermore, numerous studies have found elevated As concentrations in pore-waters of volcanic rocks (Amini et al., 2008; Busico et al., 2018; Johannesson and Tang, 2009). Aquifers containing toxic concentrations of As throughout Latin America are commonly comprised of volcanic rocks and sediments which contain remnant geothermal heat (Bundschuh et al., 2012; Morales-Arredondo et al., 2018). Exhaustive reviews, however, of reports of arsenic occurrence in groundwater throughout Latin America indicate that few studies have analyzed long-term changes of As concentrations in heavily pumped, geothermally influenced aquifers (Bundschuh et al., 2012).

Groundwater is the primary source of water for municipal, agricultural and industrial demand in central Mexico. This groundwater is typically stored within the sediments and rocks of inter-montane basins with near-surface geothermal heat (Morales-Arredondo et al., 2018). These aquifers are being dewatered at alarming rates. Amongst other problems such as subsidence, the falling water table drives up drilling and electricity costs for pumping (Hoogesteger, 2018; Hoogesteger and Wester, 2017; Mexicano, 2017; Ortega-Guerrero, 2009; Wester et al., 2011). The present study is concerned with the deterioration in groundwater quality that accompanies the dewatering of geothermally influenced aquifers to serve agriculture. The objectives of this study are to: 1) estimate changes in temperature and groundwater chemistry across the Independence Basin between 1999 and 2016; 2) estimate changes in As concentrations at different spatial scales; and 3) deduce the hydrogeochemical processes driving As concentrations using multivariate regression and machine learning.

Geothermally influenced ground waters in volcanic rock and sediments commonly contain high dissolved As concentrations along with high temperature, pH and concentrations of fluoride (F), silica (Si), sodium (Na), lithium (Li), and boron (B) (Grassi et al., 2014; Guo et al., 2008; Morales-Arredondo et al., 2018; Munoz et al., 2015; Rango et al., 2013). Higher pore-water temperatures accelerate the hydrolysis of albite to kaolinite clay. This releases Si and Na and raises the pH (Rango et al., 2013). This reaction generates new cation exchange sites that favor calcium (Ca) over Na, which releases additional Na to pore waters. The rising pH moreover converts the charge of metal oxide surfaces from positive to negative, releasing oxyanions of As and molybdenum (Mo) (Dixit and Hering, 2003; Dzombak and Morel, 1990; Johannesson and Tang, 2009; Smedley et al., 2002). Lastly, the reduction in dissolved Ca concentrations through cation exchange on the newly produced clay promotes the dissolution of fluorite, releasing F to groundwater via the common ion effect (Apambire et al., 1997).

This sequence of reactions will be deduced from the joint, basin-wide analysis of: 1) changing temperature and chemistry between 1999 and 2016; 2) the multivariate structure of the physical and chemical properties in 2016; 3) calculated 2016 saturation indices; and 4) the spatial relationship between As concentrations and surface geology. Other potentially confounding processes that may locally influence well water temperature and chemistry were tested through spatial analysis. These include: 1) including mixing of shallow and deep geothermal waters by upwelling along faults (Morales-Arredondo et al., 2018); 2) the infiltration of fertilizers from fields; and 3) the infiltration of untreated sewage from cities.

2. Study area and methods

2.1. Description of study area

The study area is the so-called Independence Basin in the state of Guanajuato, Mexico (Ortega-Guerrero, 2009). According to water authorities, the official name is Upper Río Laja Watershed (Cuenca Alta del Río Laja). This study area was discussed in previous publications but is briefly reviewed here (de Leon et al., 2005; Knappett et al., 2018; Mahlknecht, 2003; Mahlknecht et al., 2006a, 2004; Ortega-Guerrero, 2009). The basin is roughly 100 km in diameter and ranges in elevation from 1900 m above sea level (masl) to 3000 masl. The flat central and eastern part of the basin is known as the “dry lake” (Laguna Seca) area. The western central part of the basin has 100 m of undulating relief created by arroyos and larger channels of the Rio Laja and its tributaries. The basin is surrounded on all sides by outcropping acidic, intermediate and basic volcanic rocks. The western side of the basin is bounded by rhyolite and ignimbrite of the Guanajuato Sierras. The southern side of the basin is bounded by basalt, andesite and pyroclastic material of the extinct Palo Huerfano volcano. The east and northern edges are comprised of alternating zones of ignimbrites, tuff and pyroclastic material and basalts (Fig. 1) (del Rio Varela et al., 2020). The central area of the basin is underlain by layers of clastic sedimentary rocks derived from the surrounding mountains. These fluvial deposits are interlayered with welded tuff, volcanic ash and basaltic lava flows (Consultores en Geologia, 1992; del Rio Varela et al., 2020; Ingenieros Civiles y Geologos Asociados, 1980). The basement rock is comprised of fractured, acidic volcanic rocks (Shepherd, 2018). Lastly, marine carbonate rocks outcrop in a few places throughout the basin.

Normal faults and graben structures are common in the basin and are part of the extensive Taxco-San Miguel de Allende system of faults which run through central Mexico in a general NNE-SSW direction (Alaniz-Alvarez and Nieto-Samaniego, 2007). Prominent cliffs throughout the basin demarcate the location of normal faults

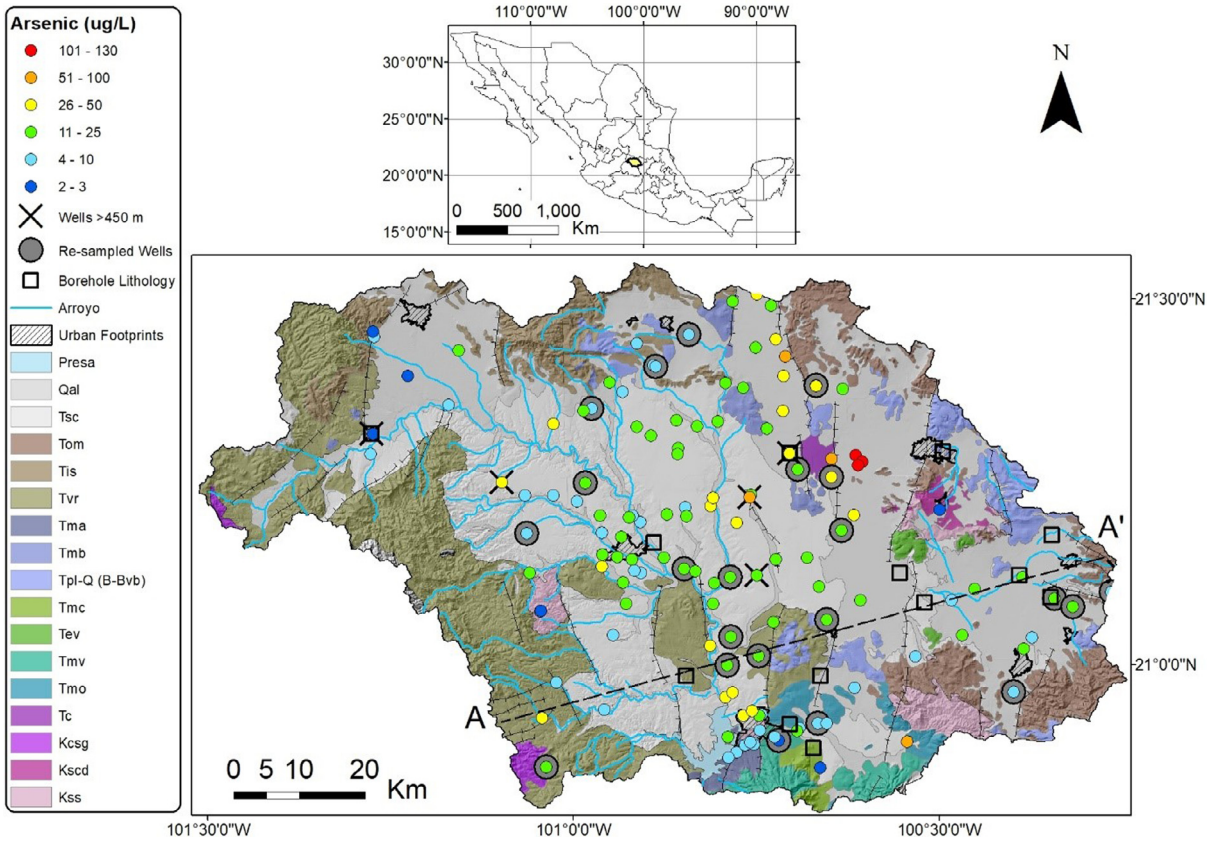

Fig. 1. Concentrations of arsenic in 137 wells sampled in 2016 in the Independence Basin including surficial geology type and the locations of major faults. Large, gray circles indicate re-sampled wells from the 1999 sampling campaign. Large, black X's indicate wells >450 deep. This plot includes the location of geologic cross-section A-A' (Fig. 2). Open, black squares indicate 13 wells for which borehole lithology is available. Surface geology is grouped into general classes of gray (clastic sedimentary), brown (acidic igneous), blue (basic igneous), green (intermediate igneous) and pink (carbonates). The symbols mean: Qal – Holocene deposits of sand, clay, and gravel; Tsc – Tertiary conglomerates, sandstone, and siltstone; Tom and Tis – Oligocene-Miocene ignimbrites, rhyolitic tuff or pumice; Tvr – Tertiary rhyolite and ignimbrites with ash, breccia and pyroclastic materal; Tma – Miocene andesites and dacites; Tmb – Miocene basalts; Tpl-Q– Pliocene-Pleistocene basalts and andesites; Tmc – Miocene andesites and dacites; Tev – Eocene-Oligocene andesites and dacites; Tmv – Miocene andesites with pyroclastics; Tmo – Miocene andesites; Tc – Tertiary siltstone, calcareous sandstone, limestone; Kcsg – Cretaceous carbonates, and clastic rocks of siltstone and sandstones of volcanic origin; Kscd – Cretaceous siltstone and sandstone; Kss – Cretaceous alternating limestone and loam; Tmc – andesites and dacites, and flows associate with lahars and breccia.

formed during extensional events over the last 30 million years (Nieto-Samaniego et al., 2007). These events caused the sinking of the basement rock and the infilling of grabens with thick layers of sediments from the eroding mountains. One of these grabens comprises the western half of the basin and hosts the drainage network of the Rio Laja (Fig. 2). The porous sandstone and conglomerate that infills this graben is known to state and federal water managers as the Rio Laja Basin Aquifer (CARL). The Laja River exits the basin in the southwest corner and flows into the Lerma-Chapala River system towards the Pacific Ocean.

The basin experiences an average annual temperature of 16 °C and receives 600 mm of rainfall per year with the majority falling from June to September. Maximum evaporation is 1828 mm. Water table maps and chemistry show that recharge to the aquifer occurs along the perimeter of the basin through faults, channels and through arroyos in the western half of the basin (de Leon et al., 2005; Knappett et al., 2018; Li et al., 2020). Agriculture consumes approximately 85% of all the groundwater extracted in this basin (Wester et al., 2011). The same aquifers pumped for irrigation are relied upon for a growing population of approximately 600,000 people (INEGI, 2019).

2.2. Field measurements and sampling

The locations of wells were surveyed using a handheld GPS (etrex 30, Garmin). Well construction information was obtained from the well caretaker or owner. When it was possible, static and dynamic water levels were measured using an electric water level meter (Model 102, Solinst Canada Ltd., Georgetown, Canada).

Prior to sampling, temperature, pH, specific conductance (SC) and Oxidative-Reductive Potential (ORP) were measured using a calibrated multi-meter (YSI Professional Plus©). The well was pumped until these parameters stabilized after which samples were taken. Alkalinity was measured by handheld titration with H_2SO_4 (Model AL-DT, HACH©). Three samples intended for chemical analyses in the laboratory were filtered through 0.45 μm nylon syringe filters (Model S25NY045, Simsii Inc., Irvine, CA) into pre-rinsed 20 mL High Density Polyethylene (HDPE) scintillation vials (LS Vial, Wheaton Industries Inc., Millville, NJ). Two vials were acidified to 0.1% nitric acid for cation and total elemental analyses. The nitric acid was distilled from 70% ACS grade nitric acid using an acid purification system (1000 DST, Savillex Corp., Eden Prairie, MN). One additional 20 mL glass vial was filled with filtered water for water isotopes ($\delta^{18}O$ and $\delta^{2}H$). All samples were stored at 4 °C prior to analysis. Most of the wells were chlorinated with an inline chlorination pump. To avoid chlorine, samples were collected upstream of the chlorine injection port. A total of 137 wells were sampled like this between 2015 and 2017. These were sampled over several field trips which occurred during the dry season months of January and March, and the wet season months of June and July. These are referred to as the "2016″ database. In addition, over this period, approximately 35 of the 137 wells were sampled several times (including the subset of 22 wells that were re-sampled between 1999 and 2016) to assess the short-term (2 year)

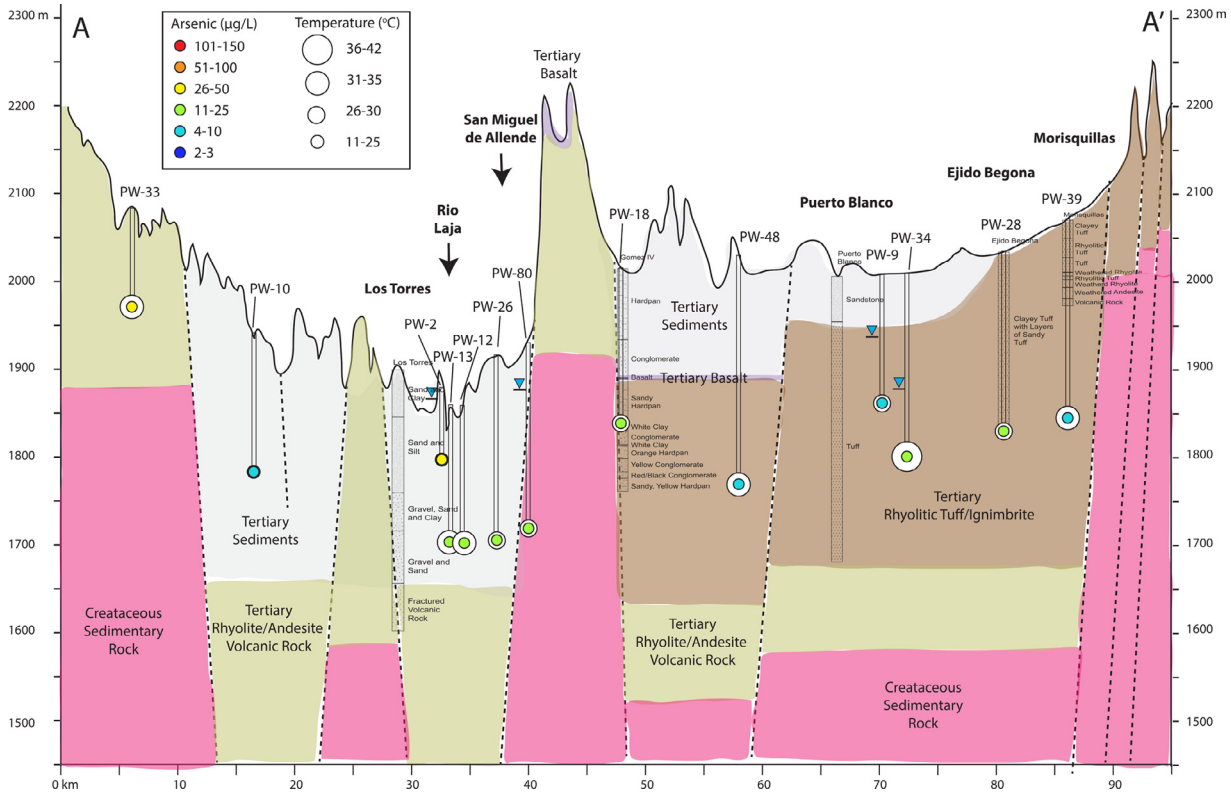

Fig. 2. Cross-section A-A' showing the general geology as constrained by surface geology maps and five borehole drill cuttings (Ingenieros Civiles y Geologos Asociados, 1980). Static water levels are represented by blue grad symbols. Arsenic concentration and temperature are shown for 13 wells that fall within 5 km from the cross-section.

stability of the chemistry and temperature in the wells. Therefore, for 137 wells, a total of 224 water samples were available to assess potential systematic effects of season (dry season vs. wet season) on the observed temperature and chemistry.

2.3. Laboratory analytical methods

A study completed in 1999 sampled a total of 246 wells. The samples were analyzed for a range of chemical parameters including major ions and the geogenic contaminants As, manganese (Mn) and selenium (Se), and the anthropogenic contaminant nitrate (NO_3–N) (Mahlknecht, 2003). All samples were filtered with a 0.45 µm syringe filter. Samples for cations were acidified to a pH of <2 using ultrapure nitric acid. Cations and anions were measured with an inductively-coupled plasma mass spectrometer (ICP-MS) and an ion chromatograph, respectively by Activation Laboratories Ltd. (Ancaster, Ontario). Alkalinity concentrations were determined in the field by titration with 1.5 N H_2SO_4 and the pH indicator Methyl orange (Mahlknecht et al., 2004). Water samples were analyzed for water isotopes at the Isotope Laboratory of the University of Waterloo, Canada.

The concentrations of major ions were compared between 1999 and 2016 in our previous study (Knappett et al., 2018), but here we introduce the results from an expanded dataset from 137 wells and includes trace elements which were not previously available. Samples taken from the 137 production wells in 2016 were analyzed for major cations (Na^+, K^+, Ca^{2+}, Mg^{2+}, Li^+) and anions (Cl^-, Br^-, NO_3^-, SO_4^{2-}, F^-) using an Ion Chromatograph at the Department of Geology & Geophysics at Texas A&M University (TAMU) (Dionex 500, Thermo-Fischer Scientific, Waltham, MA). Standards and blanks were measured every 15 samples to confirm accuracy. Charge balance calculations based on major ions resulted in 5th and 95th percentiles of −6 and +23%, respectively, with a median of +11%.

Trace elements were measured with ICP-MS (Element XR, Thermo Scientific, Waltham, MA) at the Radiogenic Isotope Laboratory at TAMU. Some elements measured in 2016 were not measured in 1999 so these elements could not be compared across time. These include boron (B), iron (Fe), strontium (Sr), cadmium (Cd), chromium (Cr), antimony (Sb), uranium (U), aluminum (Al), phosphorous (P), cobalt (Co), nickel (Ni), copper (Cu) and zinc (Zn).

Water isotopes $\delta^{18}O$ and $\delta^{2}H$ were analyzed on a Picarro cavity ring down system (Picarro Inc., Santa Clara, CA) within the Stable Isotope Geosciences Facility (SIGF) at TAMU. Based on running the known standards from the International Agency for Atomic Energy (IAEA), VSMOW2 and SLAP, the error was estimated to within ±0.1 ‰.

2.4. Geo-spatial analyses

Local or "neighborhood" concentrations of As were contoured along with temperature and chemical parameters measured on samples from 246 wells sampled in 1999. These were used to produce baseline values from which changes in 2016 could be calculated. The typical spatial extent of a "neighborhood" for each parameter was calculated with ordinary kriging (OK) (Montero et al., 2015). In contrast to other settings with high spatial variability in As concentrations over short distances (<100 m) (Aziz et al., 2017; van Geen et al., 2003), long lag distances of approximately 10 km were observed in temperature and most of the chemical parameters. Thus, aquifer chemistry and specifically As concentrations varied gradually over long distances in 1999. The uncertainty of interpolating 1999 concentrations was assessed by cross-validation as detailed in our previous study (Haslauer et al., 2016; Knappett et al., 2018). This error was reasonably small since the average spacing between the 1999 samples was much closer than the spatial lag distance.

This method of comparing the value of a parameter in a given well overlaid on the past 2-D interpolated concentration does not explicitly address the third dimension, depth. This issue is addressed here. The depths of the wells sampled in 1999 were not reported, but the depths for 102 out 137 wells were reported in 2016. Of these 102 wells with known depths, the vast majority had perforated casings that tapped the depth interval from 150 to 300 m below the surface. The water table is generally 100–150 m below surface. Therefore, most of the wells are pumping water from a similar depth range. This agrees with a large database of over 1000 wells from the state water authority (CEAG) in which practically all of the wells are installed to approximately 150 to 300 m below surface. It is assumed that wells had a similar depth range in 1999 because few farmers or communities are willing to spend extra money to drill deeper than they need to for a sustained water source. Six deep wells (>450 m), however, were included in the 2016 dataset which introduced the possibility of a depth bias in evaluating changing aquifer chemistry. Little relationship was observed, however, between well depth and temperature or chemical parameters. This implies that the long well screens are mixing vertically stratified aquifers in similar proportions (Izbicki et al., 2015; Mayo, 2010).

To test the hypothesis that geothermal waters rising up faults is driving As concentrations, the influence of proximity to known faults (< 1 km) on As concentrations was assessed using the non-parametric Kruskal-Wallis test. Overlay and proximity tools were implemented in ArcMap to extract the distance to the nearest mapped fault. The fault map was a published structural geologic and tectonic map of central Mexico that was modified based on extensive field observations of the basin by the authors (Nieto-Samaniego et al., 2007).

Since the mineralogy of the overlying rock, sediments and soil may influence As concentrations (Rango et al., 2013; Shamsudduha et al., 2015; Winkel et al., 2008), their potential role was assessed using the Kruskal-Wallis test combined with overlay and proximity tools in ArcMap. A digital surface geology map was created in ArcMap from 16 maps of the region of 1:250,000 and 1:50,000 scale published by the Mexican Geological Service (Servicio Geologico Mexicano). A soil 1:250,000 scale soil map created by the National Institute for Statistics, Geography and Information (INEGI, 2007) was used to assess spatial correlations between As concentrations and soil type.

To test whether human alteration of the land surface and vadose zone through disposal of untreated urban wastewater or the application of fertilizers and irrigation return flow from farm fields may be influencing As concentrations, land use immediately surrounding each well was defined as “urban”, “peri–urban”, “farming” or “rural” based on visual observation on Google Earth ™ images taken in 2016. The influence of the land use variable was then tested using the Kruskal-Wallis test.

2.5. Bivariate and multivariate statistical analyses

To identify possible mixing or reaction processes driving high As concentrations, covariates were assessed through bivariate regression with temperature and chemical parameters on the 2016 dataset. Next, PCA was performed on the 2016 dataset to visualize the collinearity between parameters. Collinearity manifests by clustering of parameters on the PC1 vs. PC2 loadings plot. To minimize the impact of outliers and linearize non-linear relationships between parameters, the PCA was performed on z-scores of both un-transformed and ln-transformed datasets, respectively.

To quantify the specific impact that each covariate had on As concentrations while controlling for the influence of other parameters, three types of step-wise linear regression (simple, robust and ln-transformed) were performed on the 2016 dataset. The performance of all models were assessed by their ability to predict As concentrations using the root-mean-squared-error (RMSE) and R^2. The goal was to develop a parsimonious model predicting As concentrations across the basin as a function of temperature and chemical parameters to elucidate geochemical and hydrological processes driving As concentrations. The point estimates of the slopes (b), standard errors (se) and significance level (p-value) for each statistically significant parameter were reported from the multiple regression models. All bivariate and multivariate modeling was performed using concentrations in units of mM so that stoichiometric relationships could be determined.

The direction and magnitude of changes (1999–2016) in neighborhood parameter values were compared to the slopes of statistically significant predictor variables in the multiple regression models of the 2016 dataset. If the multivariate analysis on the 2016 dataset indicated that parameter x was significantly and positively correlated to As concentrations, parameter x also should have increased between 1999 and 2016.

2.6. Random forest modeling

Random Forest (RF) regression models, a classification and regression tree (CART) model, were trained to analyze and account for non-linear relationships (Breiman, 2001; Breiman et al., 1984) between As concentrations and the measured hydro-chemical parameters. The RF is one of a growing number of machine learning algorithms that have been applied to understand and predict the distribution of anthropogenic and geogenic contaminants in groundwater (Bui et al., 2020; Smith et al., 2018). One features of these machine learning models is that even poorly correlated parameters can help make the predictions more accurate (Bui et al., 2020). Therefore, unlike the multiple regression models where only statistically significant parameters were retained in the final model, all parameters were used in the RF model. The RF algorithm has many advantages compared to single decision tree methods. It can handle the high dimensionality of the data set, it is relatively fast to train and robust against outliers, missing data and non-normal datasets which are typical of groundwater chemistry studies. The pre-modeling data treatment was the same as for the step-wise linear regression models. Random forest, compared to other decision tree models, grows multiple decision trees and the results are averaged over all grown trees with randomized predictor subsets for tree growing and node splitting. This can lead to more accurate and stable predictions. Here, up to 1000 trees were grown and maximal number of predictors to split a node were set 3/n (n = number of predictors). To prevent over-fitting, leave-one-group-out cross validation with 100 repetitions. For each repetition training (80%) and validation (20%) datasets were randomly selected for model training and validation, respectively. The final model was selected by one-standard error of the minimal RMSE of all trained models.

To analyze the impact of the predictors on As concentrations, variable importance was assessed by averaging the impurity decreases for all nodes over all trees with the Gini impurity metric, which defines the variable selection for node splitting (Louppe et al., 2013). Therefore, the higher the value of the mean decrease in Gini, the higher the importance of the variable in the model and consequently for predicting As concentrations. However, compared to linear regression models, the variable importance of random forest models does not account for the direction (positive or negative) of the relationship between the predictor and As concentrations. Therefore, partial dependence plots (PDP) were estimated to test effects between the predicted response variable and individual predictor variables by averaging the effect of all other predictors (Friedman, 2001) to visualize the direction of the effect in the trained random forest model.

Table 1
Summary of neighborhood change analysis (1999–2016) and results of multivariate modeling of 2016 dataset. Increasing and decreasing neighborhood values of a parameter are yellow and green, respectively. Parameters that correlate positively or negatively ($p<0.05$) with As concentrations are yellow and green, respectively. The terms b, se and p refer to slope, standard error, and significance, respectively. gray filled cells indicate slopes calculated in a multiple regression model that had the opposite sign to that found in bivariate models. RMSE and *k* refer to the Root-Mean-Square-Error and the number of parameters in each multiple regression model, respectively.

	Summary of Neigborhood Changes (1999-2016)				Bivariate Correlations of 2016 Samples						Multivariate Modeling of 2016 Samples									
					Robust Linear			ln-Transformed			Linear Regression			Robust Linear Regression			Ln-transformed Regression			Random Forest
	Units	Median 1999	Change 1999-2016	Δ (%)	b	se	p	b	se	p	b	se	p	b	se	p	b	se	p	Gini Decrease
Dependent Variable																				
Arsenic	mM	1.3E-04	2.4E-05	19																
Independent Variables																				
Temperature	°C	27.0	1.5	5	1.2E-05	2.2E-06	**0.00**	2.66	0.36	**0.00**							0.97	0.26	**0.00**	**7.0**
pH	-	7.3	0.3	4	8.9E-05	2.4E-05	**0.00**	9.07	0.94	**0.00**	1.2E-04	3.2E-05	**0.00**				2.09	0.70	**0.00**	**9.9**
SC	(µS/cm)	464.4	-2.7	-1	1.2E-07	7.0E-08	0.06	0.80	0.17	**0.00**	1.1E-06	1.6E-07	**0.00**							**7.1**
ORP	mV	na	na	na	-1.4E-07	1.6E-07	0.40	-0.37	0.08	**0.00**										1.1
HCO_3	mM	3.9600	-1.0100	-26	1.7E-05	9.6E-06	0.08	0.62	0.14	**0.00**										0.8
Na	mM	1.8700	0.5200	28	1.4E-05	3.6E-06	**0.00**	0.64	0.07	**0.00**	-2.8E-05	7.9E-06	**0.00**							**11.4**
K	mM	na	na	na	7.2E-05	4.6E-05	0.12	0.05	0.11	0.66	-1.7E-04	6.4E-05	**0.01**							2.3
Mg	mM	na	na	na	-1.5E-04	4.9E-05	**0.00**	-0.32	0.05	**0.00**	-3.0E-04	1.2E-04	**0.01**							**8.0**
Si	mM	1.2300	0.0990	8	7.7E-05	2.7E-05	**0.00**	0.53	0.20	**0.01**							0.32	0.13	**0.02**	5.8
Ca	mM	0.8300	-0.0790	-10	-5.0E-05	2.2E-05	**0.02**	-0.56	0.08	**0.00**	-4.1E-04	6.5E-05	**0.00**	-8.1E-05	2.7E-05	**0.00**				**13.6**
S	mM	0.2300	0.0390	17	2.0E-04	4.0E-05	**0.00**	0.48	0.07	**0.00**	2.5E-04	8.4E-05	**0.00**	1.4E-04	4.9E-05	**0.01**				**7.3**
Fe	mM	0.0017	-0.0016	-96	3.0E-02	5.4E-03	**0.00**	0.05	0.03	0.05	3.0E-02	6.8E-03	**0.00**	3.1E-02	5.3E-03	**0.00**				0.5
Mo	mM	na	na	na	2.9E+00	2.3E-01	**0.00**	0.65	0.05	**0.00**	2.2E+00	5.8E-01	**0.00**				0.24	0.06	**0.00**	**16.5**
Ba	mM	0.5100	0.1840	36	-1.1E-02	8.4E-03	0.18	-0.12	0.04	**0.00**										4.1
U	mM	na	na	na	1.7E-03	7.4E-01	1.00	0.10	0.07	0.14	-3.5E+00	9.9E-01	**0.00**							1.2
P	mM	na	na	na	-2.8E-02	5.0E-03	**0.00**	-0.19	0.04	**0.00**										4.5
Mn	mM	0.0200	-0.0110	-55	-5.1E-03	1.1E-02	0.64	-0.01	0.04	0.89										2.6
Sr	mM	na	na	na	-4.1E-03	4.0E-03	0.31	-0.16	0.08	**0.05**	3.5E-02	1.0E-02	**0.00**							5.0
Li	mM	0.0100	0.0027	27	2.8E-03	6.6E-04	**0.00**	0.35	0.05	**0.00**				1.7E-03	7.8E-04	**0.03**	0.07	0.04	0.09	**7.7**
F	mM	0.0368	0.0089	24	2.7E-03	9.8E-05	**0.00**	0.85	0.06	**0.00**				1.4E-03	1.7E-04	**0.00**	0.37	0.09	**0.00**	**24.6**
Cl	mM	0.2700	-0.0063	-2	2.1E-04	3.9E-05	**0.00**	0.39	0.09	**0.00**										4.7
Br	mM	0.0015	-0.0005	-33	3.7E-02	5.8E-03	**0.00**	0.26	0.05	**0.00**	-2.2E-02	8.5E-03	**0.01**							4.4
NO_3-N	mM	0.1260	-0.0100	-8	-5.6E-05	4.2E-05	0.19	-0.26	0.05	**0.00**							-0.14	0.03	**0.00**	**9.6**
$\delta^{18}O$	‰	-10.1	0.5	-4	9.8E-06	8.0E-06	0.22	0.75	0.56	0.19										-2.3
$\delta^{2}H$	‰	na	na	na	1.5E-06	1.4E-06	0.29	0.12	0.21	0.58	-6.5E-06	1.9E-06	**0.00**							2.3
									RMSE		1.5E-04			1.3E-04			1.5E-04			9.5E-05
									R^2		0.77			0.39			0.72			0.90
									k		13			5			6			25

2.7. Geochemical modeling

Saturation indices (SI's) for over 150 minerals were calculated using temperature and water chemistry data from the 2016 datasets using the software PHREEQC and the wateq4 database (Parkhurst and Appelo, 2013) . Saturation indices indicate the equilibrium state of water with respect to minerals. It is defined as SI = log (IAP/K) where IAP is the ion activity product and K is the equilibrium constant. If SI = 0, the solution is in equilibrium with the mineral phase and therefore highly influential on the dissolved concentrations of the constituent elements in that mineral phase. If SI > 0, groundwater is oversaturated with respect to the particular mineral, which means that the mineral phase may precipitate to achieve equilibrium. Finally, if SI < 0, the mineral will dissolve, if it is present in the aquifer.

3. Results

3.1. Changing neighborhood temperature and chemical parameters over time

The median percentage increase in 137 neighborhood concentrations from 1999 to 2016 of Ba, Na, Li, S, Si, temperature, and pH was 36, 28, 27, 17, 8, 5 and 4%, respectively (Table 1; Fig. S2). In contrast, the median percentage decrease in neighborhood concentrations of Fe, Mn, Br, HCO_3, Ca and NO_3–N was 97, 55, 33, 26, 10 and 8%, respectively. The median neighborhood temperature increase was 1.5 °C while pH increased by 0.3 units. The previously reported rise in $\delta^{18}O$ of 0.9 ‰ was maintained with the larger dataset used in the present study (Knappett et al., 2018). Significantly, neighborhood Cl concentrations and SC were stable between 1999 and 2016. To control for uncertainties arising from interpolating neighborhood, baseline concentrations and temperature in 1999, changes in parameter values were assessed using the smaller number of 22 re-sampled wells and approximately the same magnitude and direction of change was observed in the same parameters (Fig. S1).

3.2. Evidence for rising arsenic concentrations

Of all wells sampled in 1999 the median and 90th percentile As concentrations were 9 and 20 µg/L, respectively (Fig. 3a). In 2016, however, the median and 90th percentile As concentrations were 14 and 40 µg/L, respectively. In 1999, 19 (8%) and 93 (38%) of 246 wells exceeded the Mexican (25 µg/L) and WHO drinking water limits for As, respectively. In 2016, however, 29 (21%) and 88

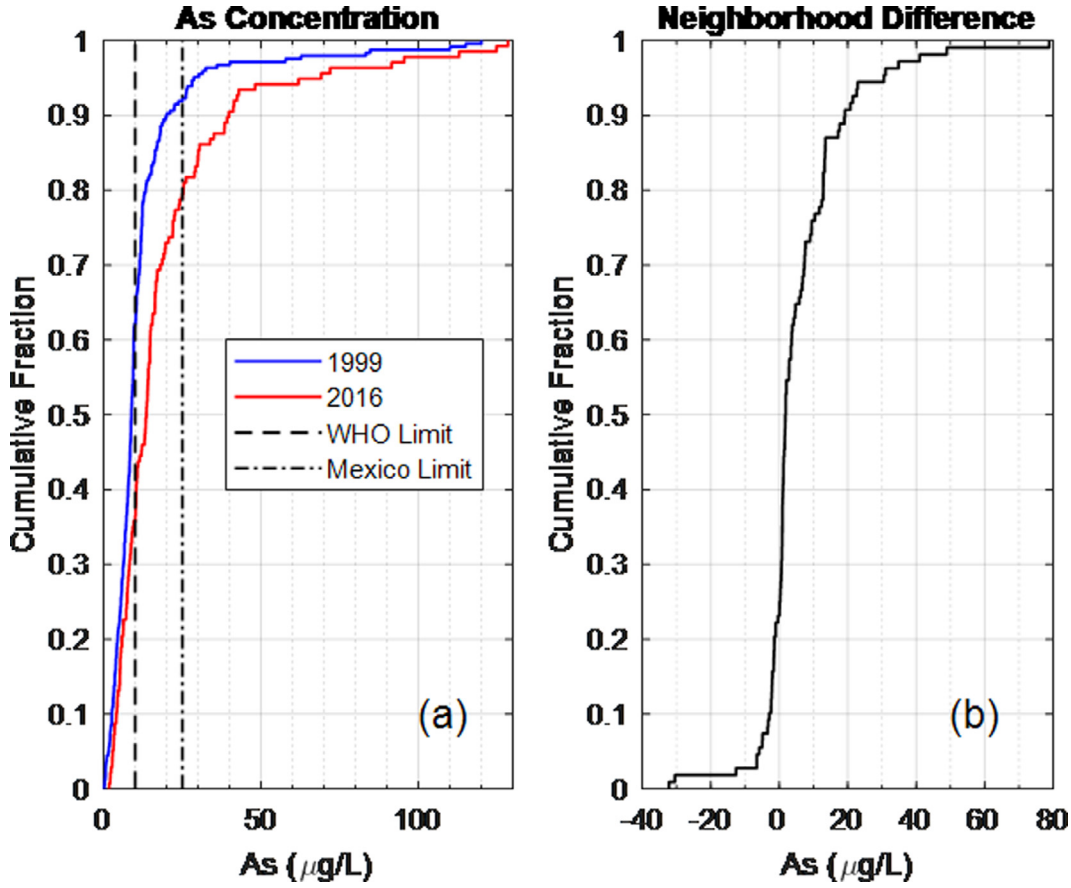

Fig. 3. Cumulative distribution curves of As concentrations from 1999 to 2016 (a), and of the changes in As concentrations at the neighborhood level (b). The values for panel b were derived from the analysis shown in Fig. 3. The median for 1999 and 2016 were 8.8 and 13.6 μg/L, respectively.

(64%) of 137 wells exceeded Mexican and WHO limits, respectively (Fig. 3a). Furthermore, 3 of the 22 wells that were re-sampled in 2016 increased by more than 10 μg/L compared to the 1999 value (Fig. 4b). Substantial increases in As concentrations were observed in certain neighborhoods (Fig. 4a,c). Thirty percent of wells sampled in 2016 contained As concentrations that exceeded their previous, baseline, neighborhood concentrations (1999) by at least 10 μg/L (Fig. 3b). An additional 45% of wells exhibited modest increases in As concentrations (0–10 μg/L).

The greatest increases in As concentrations were observed in two geographical areas. The first was an area that previously had high As concentrations in 1999, located in the northeast part of the basin called "Laguna Seca". The second area with the greatest increases were distributed throughout the western part of the basin or the CARL aquifer. In 1999, all wells sampled in the CARL aquifer contained low As concentrations near the WHO limit. In 2016, however, this region hosted many wells exceeding the WHO limit and some exceeding the Mexican limit (Fig. 1). No systematic variation in As concentrations was observed with season (Fig. S3) and samples were taken at different times of the year, so it is unlikely that seasonal fluctuations in As concentrations, possibly driven by more pumping to irrigate during the dry season, can explain the widespread increases.

3.3. Arsenic concentrations, land use, and geology

To assess the possible influence of geological composition and faulting, and human land use, on As concentrations, 2016 concentrations were compared to four land surface variables: proximity to known faults, surface geology type, soil type and land use type. The non-parametric Kruskal-test confirmed that wells located within 1 km of a known fault did not have higher As concentrations than wells located further away (Fig. S4). The vast majority of wells sampled in 2016 are installed on outcropping sedimentary deposits such as conglomerate and sandstone ($n = 113$) and, along with wells installed on acidic volcanic rock ($n = 7$), tend to have the highest As concentrations ($p<0.05$) (Fig. S5). The eight wells installed on outcropping basic volcanic rocks had the lowest As concentrations. Wells installed on outcropping carbonate rocks had intermediate As concentrations. Arsenic concentrations were systematically higher in wells underlying farm land and rural areas compared to urban areas ($p<0.05$) (Fig. S6). No systematic variation in As concentration was observed across the three major soil types in the basin (Fig. S7). In summary, wells installed in agricultural or rural areas underlain by sedimentary or acidic volcanic rocks, had the highest As concentrations.

3.4. Arsenic concentrations and well construction and use

To yield insight into the impact of flow processes within wells on rising As concentrations, As concentrations measured in 2016 were compared to well construction and operation variables including well depth, bottom elevation and pumping rate. Numerous reports (Consultores en Geologia, 1992; Ingenieros Civiles y Geologos Asociados, 1980) suggest that wells in this region are perforated over their entire length. This creates the possibility for cross-formational flow through these long well screens, especially in the six recently installed community wells ranging from 450 to 650 m in depth. These very deep wells typically have moderately high As concentrations (Fig. 5b), ranging from 20 to 60 μg/L. This finding

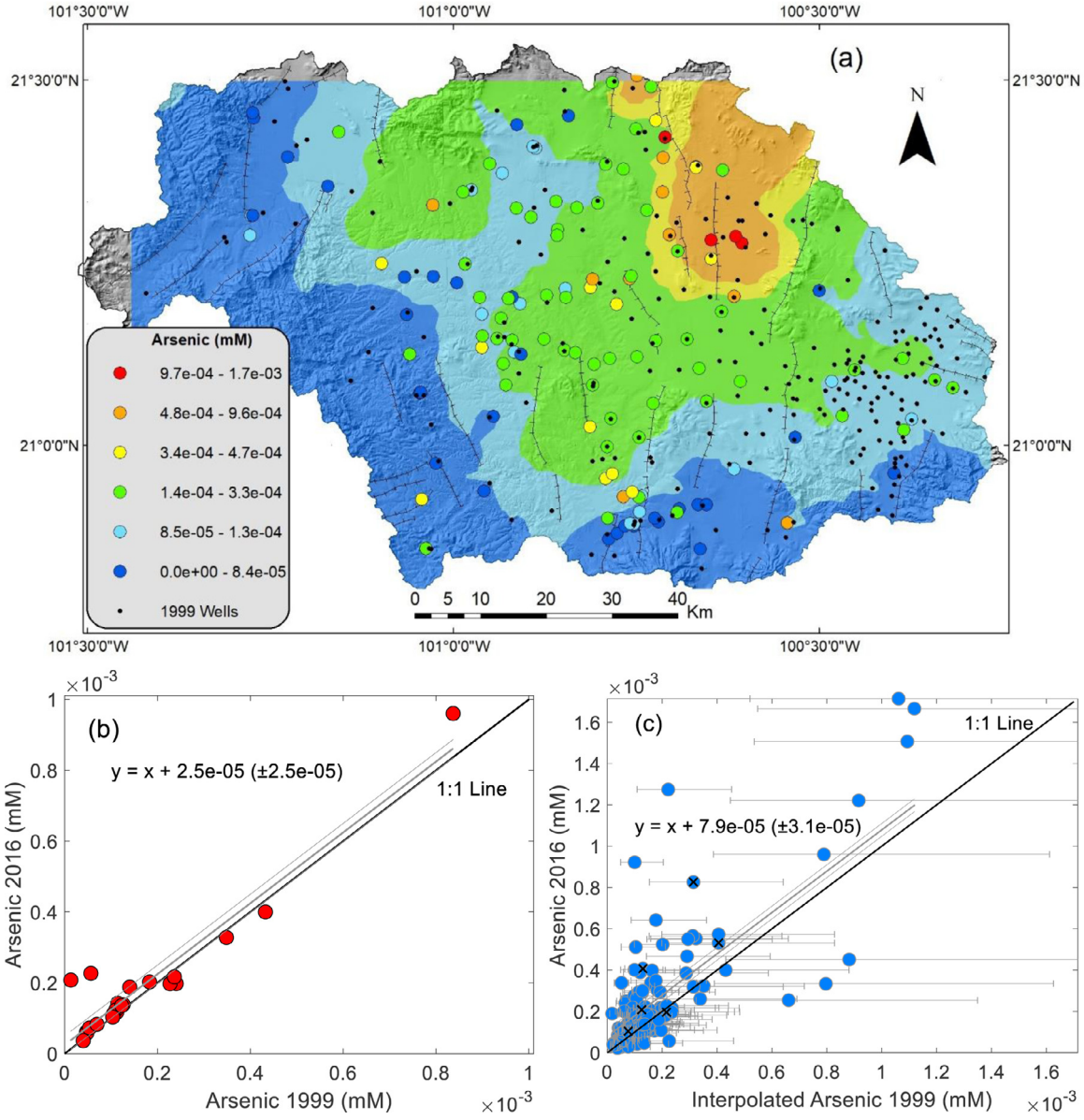

Fig. 4. (a) Comparison between neighborhood As concentrations in 1999 (contours) and 2016 (filled circles). Black dots indicate the locations of wells sampled in 1999. (b) Comparison between As concentrations measured in 22 re-sampled wells in 1999 and 2016. (c) Comparison between neighborhood As concentrations between 1999 and 2016. The horizontal error bars represent the estimated uncertainty from interpolating between points with known concentrations in 1999 assessed using cross-validation. Black X's mark 6 recently installed wells >450 m in depth. The WHO and Mexican drinking water limits for As are 1.3e-04 mM (10 µg/L) and 3.3e-04 mM (25 µg/L), respectively.

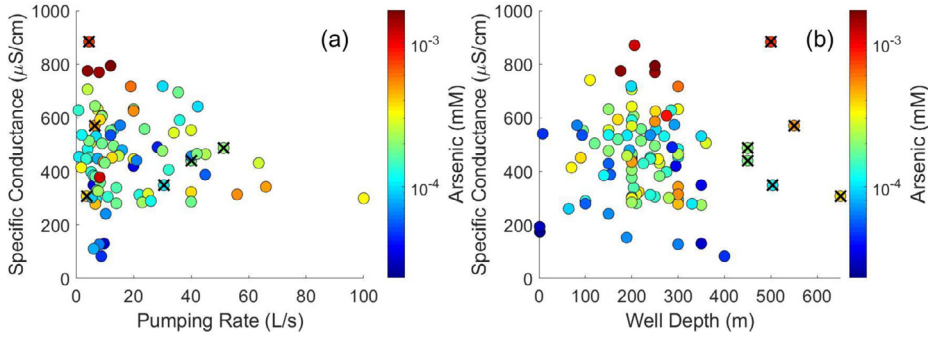

Fig. 5. Specific conductance of production well water plotted against pumping rate (L/s) (n = 74) and well depth (n = 102). The points are color coded for As concentration. Black X's mark 6 recently installed wells >450 m in depth. The WHO and Mexican drinking water limits for As are 1.3e-04 mM (10 µg/L) (green and above) and 3.3e-04 mM (25 µg/L) (yellow and above), respectively.

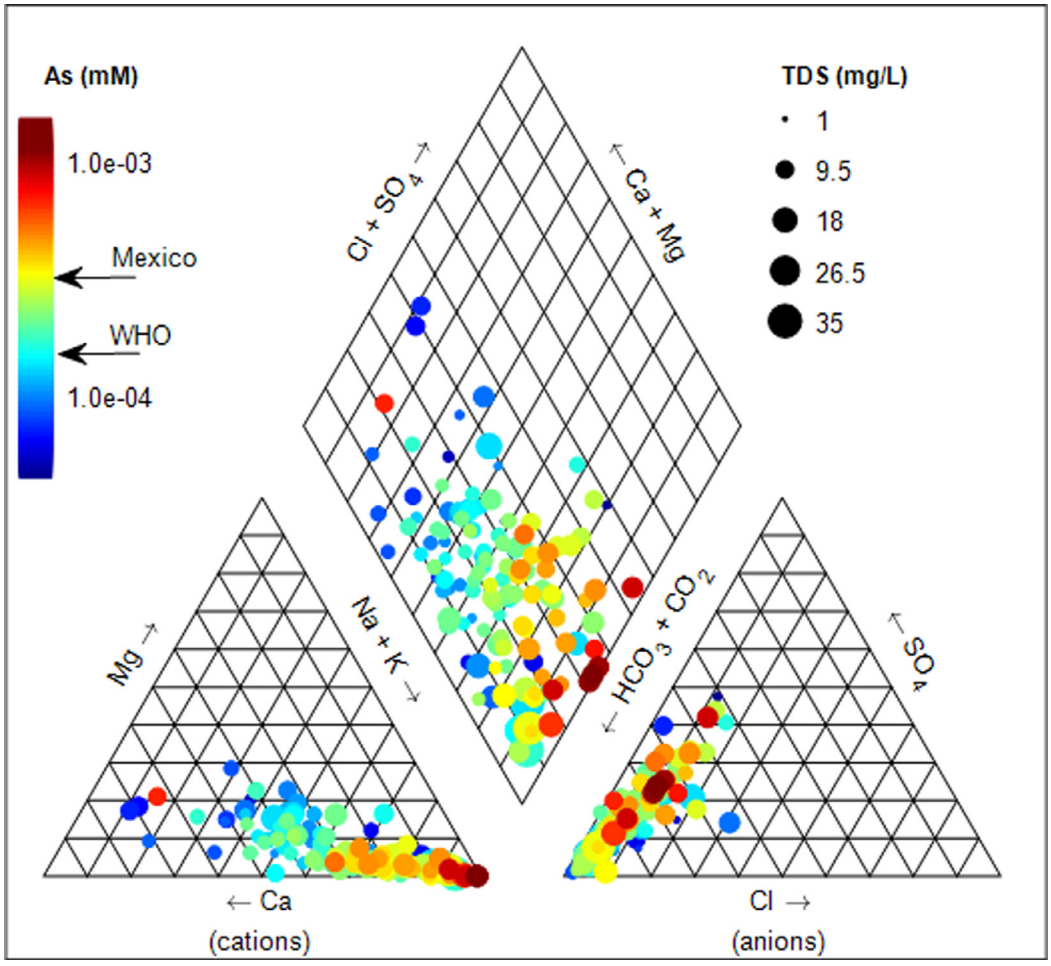

Fig. 6. Arsenic concentrations in well water sampled in 2016 as a function of dominant cation and anion composition. The Mexican and WHO drinking water limits for As are indicated on the color scale. Plot made using the PIED piper Matlab™ library (www.syr.edu/hydrology/pied_piper) (Russoniello and Lautz, 2019).

is likely because of dilution with low TDS, shallow groundwater across the long well screens. This is further evidenced by convergence to the basin's mean SC value in wells with the highest pumping rates where moderate As concentrations also are found (Fig. 5a). Wells with the highest pumping rates tended to range from 150 to 300 m depth, but there was a substantial range of pumping rates observed for wells in every depth interval (Fig. S8).

3.5. Arsenic concentrations and temperature and chemical parameters

The relationship between As concentrations and temperature, chemical parameters and water isotopes were analyzed to gain insight into the hydrological and geochemical processes that may be driving As concentrations higher in this intensively pumped aquifer system. A Piper Plot revealed that As concentrations increase smoothly as the dominant cation transitions from Ca-Mg to Na-K (Fig. 6).

Bivariate correlations were calculated between As concentrations and all potential co-variates measured in 2016 (Table 1). Both robust and ln-transformed (Fig. S9), bivariate regressions were performed to minimize the influence of localized processes. In the ln-transformed bivariate regression analysis, As concentrations correlated significantly and positively to temperature, pH, SC, HCO_3, Na, Si, S, Mo, Li, F, and Cl (Table 1).

The collinearity between variables were further investigated on the ln-transformed data using PCA (Fig. 7, Fig. S10). The strongest collinearity was observed between the variables SC, Na, HCO_3, Cl, S, and Mo which formed the first principal component (PC1) (Fig. 7). The variables As, F, temperature, pH and Li also clustered together, but were distinct from PC1. The variables Mg, Ca, Sr, Ba and NO_3–N comprised PC2.

The specific influence of each variable on As concentrations, while controlling for the effects of other covariates, was assessed using step-wise, multiple regression and RF modeling (Table 1). Three types of linear regression were used: regular, robust, and ln-transformed. The latter two models were performed to minimize the impact of outliers on the model. Regular multiple regression indicated many significant predictor variables of As concentrations, however many of these variables did not have robust bivariate relationships to As, or the bivariate relationship was of the opposite sign (Table 1). As assessed by R^2, the robust regression model performed poorly (R^2 = 0.39), under-predicting the concentrations of 5 wells with the highest As concentrations in the basin (Fig. 8b). The ln-transformed performed consistently across all As concentrations (Fig. 8c), was internally consistent since the bivariate relationships had the same sign as in the model, and was much more parsimonious compared to the first model. This regression model has six significant predictor variables: temperature, pH, Si, Mo, F and NO_3–N. Of these, only NO_3–N has a negative slope meaning that decreasing NO_3–N concentrations predicted increasing As concentrations when controlling for the effects of other predictor variables (Table 1).

The RF model performed the best (R^2 = 0.90) (Fig. 8d). This machine learning model confirmed the importance of the variables

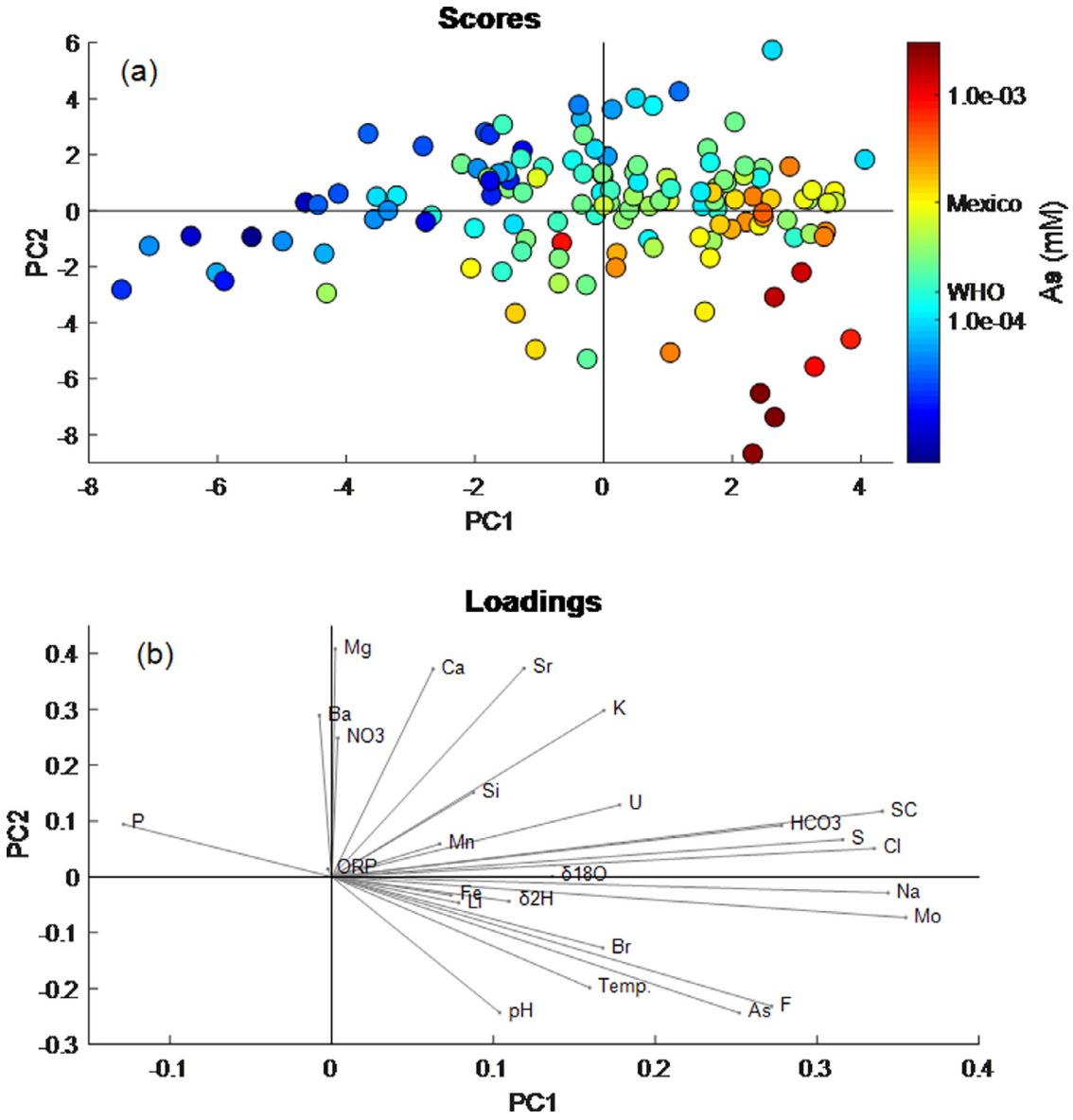

Fig. 7. Correlation structure of temperature, chemistry and water isotopes in 136 well water samples across the Basin in 2016 as assessed by Principal Components Analysis (PCA) on the z-scores of the ln-transformed values of each parameter (outlier well PW-7 was excluded). Following PCA, a varimax rotation was performed to optimize the correlations of parameters to the factors. The first 3 PC's explained approximately 60% of the total variability in the data set. (a) Scores of water samples along PC 1 and 2 coded for As concentrations. The Mexican and WHO drinking water limits for As are indicated on the color scale. (b) Loadings of parameters on PC1 and PC2.

identified by stepwise multiple regression but also detected the influence of Ca and Na, which may have been masked through collinearity in the ln-transformed multiple regression model. The partial dependence plots from the RF model show that Ca and Na had a negative and positive effect on As concentrations, respectively (Fig. S11, S12). The variables identified by RF modeling from most to least important for predicting As concentrations were F, Mo, Ca, Na, temperature, pH, NO_3–N and Li. The negative influence of Ca on As concentrations was detected by bivariate and two multiple regression models (Table 1; Fig. S9).

The spatial distribution of each of the eight As-predictor variables identified by the ln-transformed multiple regression and RF models reveal further insight into the prediction power of these variables on As concentrations (Fig. S13-S20). Wells emplaced upon outcropping bedrock on the higher elevation periphery of the basin tended to have cooler temperatures and low As concentrations (Fig. S13). The center of the basin exhibits zones with temperatures varying from 30 to over 40 °C with the highest temperature zones exhibiting the highest As concentrations. pH exhibits a similar pattern to temperature (Fig. S14). Fluoride concentrations also peak in wells with the warmest and most alkaline waters (Fig. S17) (Knappett et al., 2018). Molybdenum also closely correlates to As across the basin (Fig. S16). In contrast, Si concentrations tend to be highest in the center of the basin, and lowest around the periphery, but wells with intermediate (10–25 μg/L) As concentrations tend to have the highest Si concentrations (Fig. S15). The RF model did not rank Si highly as an important variable in predicting As concentrations (Table 1). A subset of 109 wells analyzed for the element boron (B) showed at strong bivariate correlation (R^2=0.4) with As (Fig. S21), however, because of the missing data, this element was not included in the multivariate analyses.

The largest clustering of sampled wells with high NO_3–N (>5 mg/L, >0.36 mM), and low As (<6 μg/L, <8e-05 mM), concentrations occurs around the city of San Miguel de Allende, however,

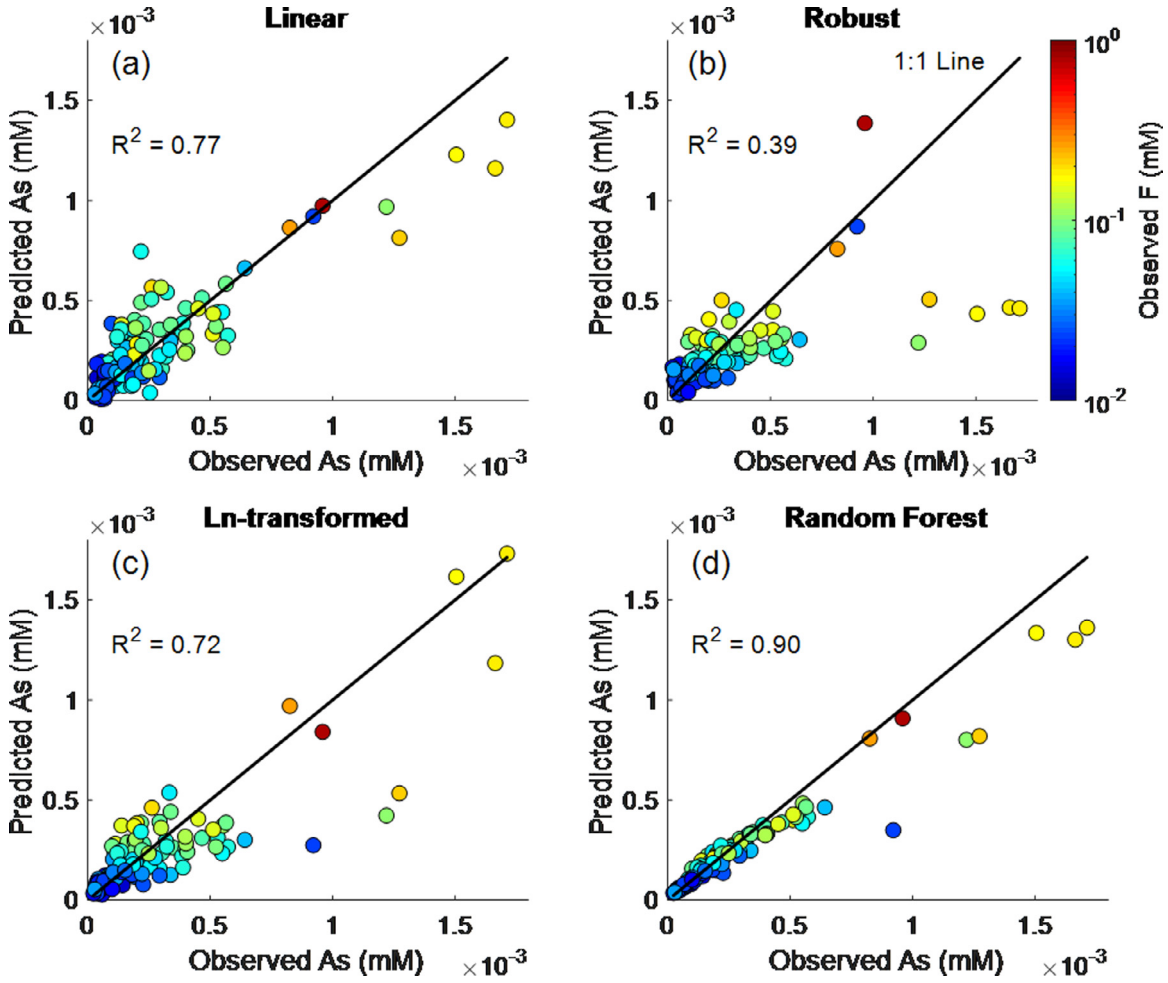

Fig. 8. Model performance of three stepwise multiple regression and one machine learning (RF) model. The points are color coded according to F concentrations. The WHO drinking water limit for F is 0.079 mM (1.5 mg/L) (green and above).

this is a common occurrence over the outer periphery of the basin (Fig. S18). In contrast, in the center of the basin many wells have intermediate concentrations of both As and NO_3–N (0.6–5.0 mg/L). And wells with the highest As concentrations (>50 µg/L, >4.7e-04 mM), have the least NO_3–N (<1.5 mg/L, <0.1 mM).

A subset of 53 wells tested for oxygen contained concentrations ranging from 3 to 7 mg/L, therefore the aquifer appears to be well oxidized. A subset of 45 wells contained DOC ranging from 0.5 to 7 mg/L. The presence of abundant oxygen and high ORP suggests this DOC is not bioavailable (Chapelle et al., 2013). The reason may be that the organic carbon is locked up in biofilms on grain surfaces that are mobilized by shear forces during pumping (Graham et al., 2015). Alternatively, wells screened across geochemically stratified aquifers may promote mixing between oxidizing and reducing agents (Izbicki et al., 2015; Mayo, 2010; McMahon and Chapelle, 2008) so the abundant oxygen and DOC may not come from the same parts of the aquifer system.

4. Discussion

4.1. Accelerated silicate hydrolysis from mixing with geothermal waters

The best multiple regression model explained 72% of the variability of As concentrations across the basin in 2016, whereas the RF machine learning model explained 90%. Many of the parameters that drive the As concentrations higher are common tracers of geothermal waters (Smedley and Kinniburgh, 2002). Hot, geothermal waters accelerate the dissolution of widely present silicate minerals such as albite releasing high concentrations of Na and Si. The mineral albite, was near saturation across the basin (Fig. S22). The transformation of albite to kaolinite also raises the SI for albite in pore waters while raising pH through the following hydrolysis reaction:

$$NaAlSi_3O_8 + 2H^+ + 9H_2O \rightarrow Na^+ + 4H_4SiO_4 + Al_2Si_2(OH)_4 \quad (1)$$

Studies in the MER suggested that the hydrolysis of albite drove increasing pH (Rango et al., 2013). Albite hydrolysis tends to proceed most quickly at acidic and basic pH's, and slowest at near-neutral (Chen and Brantley, 1997). This process is greatly accelerated by higher temperatures. Albite glass that is found in pyroclastic flows and tuffs (Gulbrandsen and Cressman, 1960) is weathered more easily than crystalline albite found in volcanic rocks like rhyolite (Hamilton et al., 2000). This rising pH converts iron oxide oxyanion binding sites from positive to negatively charged which releases oxyanions like As (and Mo) (Dixit and Hering, 2003; Johannesson and Tang, 2009; Smedley et al., 2002). Albite hydrolysis furthermore releases Na ions when it transforms to kaolinite (Fig. 9). The mineral kaolinite was supersaturated across much of the basin (Fig. S22). The formation of kaolinite and other clays increases the abundance of cation exchange sites which further depletes the aquifer of Ca and releases more Na (Rango et al., 2013).

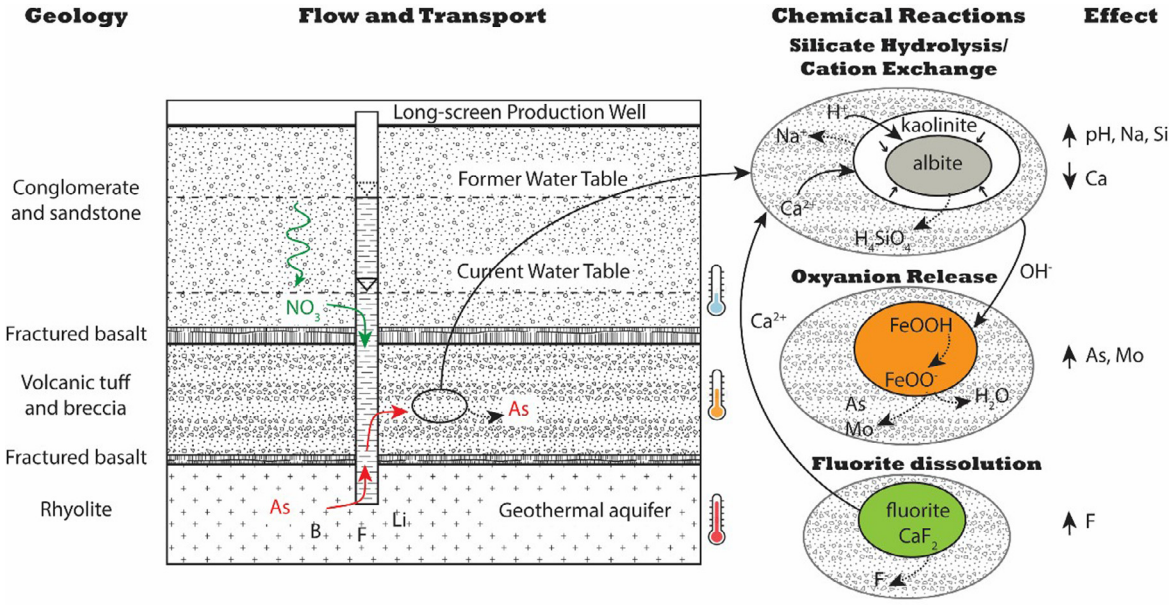

Fig. 9. Conceptual model of hydrogeochemical processes simultaneously driving As and F concentrations across the geothermally influenced inter-montane basin.

Cation exchange on kaolinite can be represented by the following formula:

$$Na^{+}\text{-}\ Al_2Si_2(OH)_4 + Ca^{2+} \rightarrow Na^{+} + Ca^{2+}\text{-}\ Al_2Si_2(OH)_4 \quad (2)$$

The mineral fluorite is consistently, slightly undersaturated (SI<0) across the basin suggesting this soluble mineral controls the concentrations of Ca and F ions. Therefore, decreasing Ca concentration through reactions 1–2, acts to dissolve the mineral fluorite through the common ion effect (Apambire et al., 1997):

$$CaF_{2(s)} \rightarrow Ca^{2+} + 2F^{-} \quad (3)$$

This explains the co-rising F concentrations between 1999 and 2016 and the tight correlation between As and F concentrations in the 2016 dataset (Table 1, Fig. 9).

Evidence of the internal consistency of the ln-transformed and RF models can be found by examining the bivariate associations. Both the regular linear regression and the robust regression models are limited in that they are seeking simple, straight-line relationships between As and each variable. Many variables like temperature and pH follow a positive exponential relationship with As, whereas other variables like Ca, Mg, and NO_3–N, have an inverse relationship. These are poorly fit with straight lines. Other researchers have found such negative correlations between Ca and Mg, and As and F, in the geothermally influenced sedimentary aquifers of the MER (Rango et al., 2013). These authors found that basaltic rock aquifers tended to have high concentrations of Ca and Mg, but low concentrations of As, F and Na. The basic volcanic rocks on the flanks of the Palo Huerfano volcano located at the southern edge of the basin host wells that are consistently low in As concentrations and high in Ca and Mg (Fig. 1, Fig. S20).

A similar multi-variate structure amongst chemical parameters, but not temperature, was observed in 769 wells within Orange County, North Carolina (Dinwiddie and Liu, 2018). In that study the source of As was attributed to the oxidation of pyrite within the clastic sedimentary rocks and the release of As from metal oxides at higher pH. The authors, furthermore, found that As concentrations increased with proximity to faults, which they attributed to the presence of more As-bearing sulfide minerals. In the present study area, pyrite was rarely found in a recent mineralogical study on surface outcrops and the exhaustive analyses of borehole cuttings from two deep (>450 m) wells, including one within the high As region in the eastern part of the basin (Fig. 1) (Shepherd, 2018). Furthermore, the oxidation of pyrite does not explain the apparent rising pH over time, or the highly significant impact that pH has on As concentrations in the 2016 dataset since pyrite oxidation drives pH down (Appleyard et al., 2006).

The significance of the parameters temperature, pH, and F along with Si and Na in the ln-transformed regression and RF models, respectively, and the tendency for the highest As wells to be found in felsic sedimentary or volcanic rocks, suggest that warm, geothermally influenced waters are driving As concentrations higher through hydrolysis of silicate minerals like albite glass found in volcanic tuff. High As wells, however, are not limited to those installed in volcanic tuff, which is suggested by surface mapping and driller's logs to be present in large quantities in the southeastern part of the basin (Fig. 1 and 2). Across the basin, the majority of wells are drawing water from the first 100–300 m which is typically comprised of clastic sedimentary rocks, sourced from the surrounding mountains. A subset of approximately 109 samples that were analyzed for boron (B) had a highly significant bivariate positive correlation with As (p<0.01). This further strengthens the argument for mixing with geothermal waters being the primary driver of As since both Li and B are considered a tracer of geothermal waters (Busico et al., 2018; Grassi et al., 2014; Guo et al., 2008; Morales-Arredondo et al., 2018; Munoz et al., 2015; Rango et al., 2013).

The slopes of the ln-transformed multiple regression model can be quantitatively interpreted (Table 1). For example, a 1% increase in temperature will produce a 1% increase in As concentrations. A 1% increase in pH will produce a 2% increase in As concentrations. A 1% increase in Si, Mo, and F concentrations will produce 0.32, 0.24 and 0.37% increases in As concentrations, respectively. In contrast, a 1% increase in NO_3–N will cause a 0.14% decrease in As concentrations.

Between 1999 and 2016 temperature, pH, and concentrations of Na, Si, F and Li typically increased in specific wells and at the neighborhood scale, further supporting the hypothesis that geothermal waters are mixing in increasing proportion with the freshly recharged, young groundwater, driving silicate hydrolysis and raising pH. The fact that Cl and NO_3–N concentrations re-

mained relatively stable, however, implies that changes to the composition of recharge water from irrigation return flow has yet to reach the water table (Mahlknecht et al., 2004). This agrees with the estimated 2,000 to >10,000 year old well water that was calculated using C^{14} age dating on water samples taken from 24 wells across the basin in 2000 (Mahlknecht et al., 2006a) and the absence of tritium in 21 of those wells. Prior to the drilling of production wells in the 1940′s the water table was close to the surface in many low-lying areas (de Leon et al., 2005). Between 1999 and 2016 the vadose zone thickness typically grew approximately 20 m across the basin (Knappett et al., 2018; Li et al., 2020). In 2016 the water table was typically found 50 to 200 m below the surface. Studies based on observed chloride mass accounting from rainfall and groundwater age dating have demonstrated that thick vadose zones in arid to semi-arid regions delay the arrival of recharge water by thousands of years (Phillips, 1994; Scanlon et al., 2003; Tyler et al., 1996). Together these results suggest that the growing vadose zone has protected the aquifer from contamination from land-use changes including fertilizer application. Irrigation return flow as a process for concentrating dissolved solids is therefore not driving As concentrations (Clark and Fritz, 1997). The rising $\delta^{18}O$ that we previously reported (Knappett et al., 2018) implies that the geothermal waters are more enriched in ^{18}O compared to more recently recharged water. This could be attributed to cooler paleoclimate conditions in this region at the time of recharge resulting in more depleted heavy isotopes (Clark and Fritz, 1997).

The close correlation between As and Mo oxyanions can be explained by their binding to positively charged metal oxides sites at neutral pH. When pH rises, these sites are converted to negatively charged releasing the oxyanions to pore waters (Rango et al., 2013; Smedley et al., 2002).

4.2. Nitrate indicates shallow aquifer water and predicts low arsenic as concentrations

Between 1999 and 2016 NO_3–N concentrations were relatively stable (Table 1). The most toxic concentrations of NO_3–N (>10 mg/L) underlie the city of San Miguel de Allende, founded in the 16th century on a permeable alluvial fan (Mahlknecht et al., 2006b). The source of this nitrate is debatable. It may be attributed to household latrines (onsite sanitation) which were traditionally installed in the basements of houses in San Miguel. Alternatively, in the 1990′s hog farms located near the city were suspected as the source and have since been moved away. Perhaps coincidentally, these city wells are installed on the basic igneous rocks along the flanks of the Palo Huerfano volcano which as mentioned above drive As concentrations lower. Nevertheless, across the basin, a robust inverse relationship was observed between As and NO_3–N concentrations. This relationship was significant in the ln-transformed multiple regression model and an important variable in the RF model, suggesting that even when controlling for the impacts of other processes and atypical regions like the urbanized slopes of the Palo Huerfano volcano, higher NO_3 concentrations were predictive of low As concentrations. This can be explained by mixing of shallow, recently recharged water (<100 years) with deeper, geothermally impacted, older groundwater that hasn't been impacted by the intensive growth in agriculture since 1940. A greater proportion of fresh, shallow groundwater in a well acts to dilute the older, more mineralized groundwater. Mixing of diverse water types through long well screens is likely an important process in this basin that impacts well water chemistry. This is supported by the observation that wells with high pumping rates (>40 L/s) tended to have intermediate SC values (Fig. 5a). In contrast, wells with the small pumping rates (<10 L/s) sample aquifer conditions that are specific to a narrower depth interval. Although this mixes waters contaminated by geothermal and anthropogenic (surface) processes, it also acts to dilute each. Therefore, water that has abundant NO_3–N suppresses As concentrations through dilution.

The presence of long-screened wells means that there is ample opportunity in this basin for cross-formational flow. Several thousand wells have been installed across the basin (Hoogesteger, 2018). The typical well depths are 100 to 300 m deep (Knappett et al., 2018), although municipal wells installed in the last 5 years tend to be much deeper up to 650 m. These long-screened wells have likely connected geochemically distinct aquifers and promoted mixing along these conduits, driven by hydraulic head differences between the layers (Izbicki et al., 2015; Mayo, 2010).

4.3. Limitations of this study

Although the evidence presented indicates rising or stable As concentrations, but almost never decreasing, as water tables fall, the long-term trends are not available owing to lack of regular resampling in the years between 1999 and 2016. Therefore, some of the variability in the changing As concentrations (dAs) in the 22 re-sampled well and 137 neighborhood comparison may be driven by short-term fluctuations. The cumulative distribution curves of As concentrations for the entire sampling campaigns of 1999 and 2016 suggest, however, the concentrations are rising, but this may have been influenced by a greater sampling density in the center of the basin in 2016, than in 1999 where many of the wells were located in the southeast corner of the Basin known as the Dr. Mora/San Jose Iturbide (DMSJ) groundwater management zone by the State and Federal water managers. Although it is impossible to capture long-term trends with only two snap shots of the basin, the situation appears quite serious since in some cases neighborhood concentrations have doubled over the 17 years. In spite of these limitations, this study describes one approach to cautiously utilize previous sampling campaigns in aquifer systems which are rarely monitored and publicly reported to suggest changing chemistry, develop hypotheses and design more robust experiments to test those hypotheses.

The direction of change in other variables over 17 years added strength to the conclusions of the multivariate modeling on the 2016 dataset (Fig. S2), but these changes are best regarded as tendencies wherein a parameter commonly both increased, stayed the same and decreased in different wells. Nevertheless, the match of the direction of change in parameters over time to the sign of the coefficients of those parameters in the regression and RF modeling suggest that the same processes driving As concentrations as water tables fall are operating over time and space. Land surface changes seem to have minimal impact on As concentrations. But the natural geology, shallow geothermal heat, natural recharge areas and aggressive pumping from long-screened wells are critically important. Until this aquifer system is geologically mapped using hundreds of borehole lithologies, and long-screen wells are profiled using down-hole sampling to map the aquifer chemistry and hydraulic properties, this aquifer cannot be confidently managed from the perspective of the future fate of geogenic contaminants like As and F. This study is not a replacement for these more conclusive and invasive methods to study aquifer systems, but rather shows one approach to monitor and infer what can be known by studying the geology, the multi-variate structure of the temperature and chemistry data sampled from long-screened wells.

4.4. Implications of rising arsenic concentrations with falling water tables

Arsenic concentrations rose over a time period when the average water table decline was approximately 20 m over 17 years

(Knappett et al., 2018; Li et al., 2020). In 1999, the area of the basin that exceeded the WHO drinking water limit was mostly confined to the eastern side of the basin in the Laguna Seca region. In 2016, however, while As concentrations increased in this more impacted region, the western half of the basin now hosts numerous wells that exceed the WHO drinking water limit. This study suggests that the primary process driving increasing As concentrations is mixing with geothermal waters, specifically the observed rising pH liberates As from metal oxide surfaces (Rango et al., 2013).

The geospatial analysis found that the highest As concentrations tended to be located in the most intensively farmed parts of the basin which are also most influenced by geothermal waters. Irrigation return flow, however, is not driving the rising As concentrations as evidenced by the stability of indicators of irrigation return (Cl and NO_3–N). This suggests that there is a long lag-time of at least several decades between changes to the land surface and the aquifer chemistry at 150–300 m depth (Mahlknecht et al., 2004). Therefore, the density of farms and irrigation pumping coincides with falling water tables driving the most mixing with geothermal waters in the basin center. Known faults were not found to play a major role across the basin, but may still have local importance.

5. Conclusions

The long-term consequences of over-exploiting aquifer systems in semi-arid, volcanic and geothermal regions is unknown owing to lack of long-term monitoring programs and process-level insights into what drives changes in the chemical composition of well water. This study helps to address each of these deficiencies by comparing the distribution of As concentrations across an intensively exploited basin over a period of 17 years and evaluating the local environment and chemical covariates that consistently drive As concentrations in both space and time.

- Between 1999 and 2016 neighborhood As concentrations increased by 19% in long-screened production wells across in an over-exploited, inter-montane aquifer system with near-surface geothermal heat.
- Parts of the aquifer system with high temperature and pH, and high As concentrations in 1999, had much higher As concentrations in 2016. This is reflected by the doubling of the 90th percentile As concentration from 20 µg/L in 1999 to 40 µg/L in 2016.
- Parts of the aquifer system in the western recharge area with low As concentrations in 1999, witnessed As concentrations rising over the WHO drinking water limit in many wells.
- Ln-transformed stepwise multiple regression indicated five statistically significant, positively correlated predictor variables of As: temperature, pH, Si, Mo, and F. And one negatively correlating predictor variable: NO_3–N.
- Random Forest modeling of As added Na (positive correlate) and Ca (negative correlate) to the list of important variables that predict As concentrations while confirming the importance of the others.
- The neighborhood values of the above-mentioned positive and negative predictor variables increased and decreased between 1999 and 2016, respectively. This suggests that these parameters impact As concentrations across space and time.
- These observations are consistent with three chemical reactions that are accelerated by rising groundwater temperatures: 1) silicate hydrolysis to kaolinite clay, raising pH; 2) conversion of metal-oxide sorption sites from positive to negative, releasing As oxyanions to pore-waters; and 3) stripping of Ca from pore waters through cation exchange with Na, thereby driving the release of F through dissolution of the mineral fluorite. The presence of long well screens allows for the possibility that this sequence of reactions are occurring in response to the transport of deeper hot, geothermal groundwater into shallower aquifers by flowing through the artificial open conduits but more work is needed to localize these reactions through characterizing water flow and transport.
- This study demonstrates a link between falling water tables and rising As concentrations across a 100 km wide basin. This relationship should be robustly quantified within over-exploited aquifer systems with known geogenic contaminants so that an accurate accounting of the costs and benefits of pumping that aquifer can be made.

The findings of this study suggests future research and management strategies which may help mitigate exposure to rising As and other geogenic contaminants. First, the hydrostratigraphy (geology and hydraulic properties) of the aquifer system needs to be systematically mapped through a system that publicly shares borehole lithology from newly installed wells. Second, the pore-water chemistry needs to be mapped at the sub-hydrostratigraphic unit scale to determine at which depth, and in what rock or sediment, the long-screened wells are drawing the most As, F and NO_3. Third, this information should be widely shared so that farmers and municipalities can use the new information to target currently safe units with short screened wells. Fourth, the water chemistry should be systematically monitored twice per year (dry and wet season) in key wells that are representative of a range of hydraulic, geologic and chemical conditions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to students from the University of Guanajuato for assisting in field work. The authors are grateful for the help of Ms. Bertha Chuc in obtaining hydrogeological reports from CONAGUA. This study benefited from conversations with Matthew Polizzotto, Lex van Geen, Brian Mailloux and Joel Podgorski. Funding was provided by Texas A&M University - CONACYT collaborative research grant program (grant number 2014–001 and 2017–034 s) and the Universidad de Guanajuato (grant number 004/2015 and 1013/2016). SM was supported by the Fulbright-Nehru Post-Doctoral Fellowship (USIEF).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2020.116257. The raw data used to create this study can be found at: http://www.hydroshare.org/resource/d43fad1af0c642b697650d2aa979f29d.

References

Alaniz-Alvarez, S.A. and Nieto-Samaniego, A.F. 2007. The Taxco-San Miguel de Allende fault system and the transmexican volcanic belt: two tectonic boundaries in Central Mexico active during the Cenozoic. Geological Society of America Special Paper 422.

Amini, M., Abbaspour, K.C., Berg, M., Winkel, L., Hug, S.J., Hoehn, E., Yang, H., Johnson, C.A., 2008. Statistical modeling of global geogenic arsenic contamination in groundwater. Environ Sci Technol. 42 (10), 3669–3675.

Apambire, W.B., Boyle, D.R., Michel, F.A., 1997. Geochemistry, genesis, and health implications of fluoriferous groundwaters in the upper regions of Ghana. Environ. Geol. 33 (1), 13–24.

Appleyard, S.J., Angeloni, J., Watkins, R., 2006. Arsenic-rich groundwater in an urban area experiencing drought and increasing population density, Perth, Australia. Appl. Geochem. 21 (1), 83–97.

Ayotte, J.D., Belaval, M., Olson, S.A., Burow, K.R., Flanagan, S.M., Hinkle, S.R., Lindsey, B.D., 2015. Factors affecting temporal variability of arsenic in groundwater used for drinking water supply in the United States. Sci. Total Environ. 505, 1370–1379.
Ayotte, J.D., Szabo, Z., Focazio, M.J., Eberts, S.M., 2011. Effects of human-induced alteration of groundwater flow on concentrations of naturally-occurring trace elements at water-supply wells. Appl. Geochem. 26 (5), 747–762.
Aziz, Z., Bostick, B.C., Zheng, Y., Huq, M.R., Rahman, M.M., Ahmed, K.M., van Geen, A., 2017. Evidence of decoupling between arsenic and phosphate in shallow groundwater of Bangladesh and potential implications. Appl. Geochem. 77, 167–177.
Breiman, L., 2001. Random forests. Mach Learn 45 (1), 5–32.
Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. Classification and Regression Trees. Taylor & Francis Group, L.L.C., Boca Raton, FL.
Bui, D.T., Khosravi, K., Karimi, M., Busico, G., Khozani, Z.S., Nguyen, H., Mastrocicco, M., Tedesco, D., Cuoco, E., Kazakis, N., 2020. Enhancing nitrate and strontium concentration prediction in groundwater by using new data mining algorithm. Sci. Total Environ. 715.
Bundschuh, J., Litter, M.I., Parvez, F., Roman-Ross, G., Nicolli, H.B., Jean, J.S., Liu, C.W., Lopez, D., Armienta, M.A., Guilherme, L.R.G., Cuevas, A.G., Cornejo, L., Cumbal, L., Toujaguez, R., 2012. One century of arsenic exposure in Latin America: a review of history and occurrence from 14 countries. Sci. Total Environ. 429, 2–35.
Busico, G., Cuoco, E., Kazakis, N., Colombani, N., Mastrocicco, M., Tedesco, D., Voudouris, K., 2018. Multivariate statistical analysis to characterize discriminate between anthropogenic and geogenic trace elements occurrence in the Campania Plain, Southern Italy. Environ. Pollut. 234, 260–269.
Chapelle, F.H., Bradley, P.M., Journey, C.A., McMahon, P.B., 2013. Assessing the Relative Bioavailability of DOC in Regional Groundwater Systems. Ground Water 51 (3), 363–372.
Chen, Y., Brantley, S.L., 1997. Temperature- and pH-dependence of albite dissolution rate at acid pH. Chem. Geol. 135 (3–4), 275–290.
Clark, I.D., Fritz, P., 1997. Environmental Isotopes in Hydrogeology. CRC Press.
Consultores en Geologia,, S.A., de, C.V., 1992. Modelo Matematico de La Cuenca Alta Del Rio de La Laja. Comision Nacional del Agua.
Darling, W.G., Gizaw, B., Arusei, M.K., 1996. Lake-groundwater relationships and fluid-rock interaction in the east African rift valley: isotopic evidence. J. Afr. Earth Sci. 22 (4), 423–431.
de Leon, I.N., Garfias-Soliz, J., Mahlknecht, J., 2005. Groundwater flow regime under natural conditions as inferred from past evidence and contemporary field observations in a semi-arid basin: cuenca de la Independencia, Guanajuato, Mexico. J. Arid. Environ. 63 (4), 756–771.
del Rio Varela, P., Nieto-Samaniego, A.F., Alaniz-Alvarez, S.A., Angeles-Moreno, E., Escalona-Alcazar, F.J., del Pilar-Martinez, A., 2020. Geology and structure of Guanajuato and Cordonices ranges, Mesa Central, Mexico. Boletin de la Sociedad Geologica Mexicana 72 (1).
Dinwiddie, E., Liu, X.M., 2018. Examining the Geologic Link of Arsenic Contamination in Groundwater in Orange County, North Carolina. Front Earth Sci. 6.
Dixit, S., Hering, J.G., 2003. Comparison of arsenic(V) and arsenic(III) sorption onto iron oxide minerals: implications for arsenic mobility. Environ. Sci. Technol. 37 (18), 4182–4189.
Dzombak, D.A. and Morel, F. (1990) Surface complexation modeling : hydrous ferric oxide.
Erban, L.E., Gorelick, S.M., Zebker, H.A., Fendorf, S., 2013. Release of arsenic to deep groundwater in the Mekong Delta, Vietnam, linked to pumping-induced land subsidence. P Natl. Acad. Sci. USA 110 (34), 13751–13756.
Fendorf, S., Michael, H.A., van Geen, A., 2010. Spatial and Temporal Variations of Groundwater Arsenic in South and Southeast Asia. Science 328 (5982), 1123–1127.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232.
Graham, P.W., Baker, A., Andersen, M.S., 2015. Dissolved Organic Carbon Mobilisation in a Groundwater System Stressed by Pumping. Sci. Rep.-Uk 5.
Grassi, S., Amadori, M., Pennisi, M., Cortecci, G., 2014. Identifying sources of B and As contamination in surface water and groundwater downstream of the Larderello geothermal - industrial area (Tuscany-Central Italy). J. Hydrol. 509, 66–82.
Gulbrandsen, R.A., Cressman, E.R., 1960. Analcime and Albite in Altered Jurassic Tuff in Idaho and Wyoming. J. Geol. 68 (4), 458–464.
Guo, Q.H., Wang, Y.X., Liu, W., 2008. B, As, and F contamination of river water due to wastewater discharge of the Yangbajing geothermal power plant, Tibet, China. Environ. Geol. 56 (1), 197–205.
Hamilton, J.P., Pantano, C.G., Brantley, S.L., 2000. Dissolution of albite glass and crystal. Geochim Cosmochim. Ac 64 (15), 2603–2615.
Harvey, C.F., Swartz, C.H., Badruzzaman, A.B.M., Keon-Blute, N., Yu, W., Ali, M.A., Jay, J., Beckie, R., Niedan, V., Brabander, D., Oates, P.M., Ashfaque, K.N., Islam, S., Hemond, H.F., Ahmed, M.F., 2005. Groundwater arsenic contamination on the Ganges Delta: biogeochemistry, hydrology, human perturbations, and human suffering on a large scale. C.R. Geosci. 337 (1–2), 285–296.
Haslauer, C.P., Heisserer, T., Bardossy, A., 2016. Including land use information for the spatial estimation of groundwater quality parameters-2. Interpolation methods, results, and comparison. J. Hydrol. 535, 699–709.
Hellegers, P.J.G.J., Perry, C.J., Al-Aulaqi, N., 2011. Incentives to Reduce Groundwater Consumption in Yemen. Irrig. Drain 60 (1), 93–102.
Hoogesteger, J., 2018. The Ostrich Politics of Groundwater Development and Neoliberal Regulation in Mexico. Water Altern. 11 (3), 552–571.
Hoogesteger, J., Wester, P., 2017. Regulating groundwater use: the challenges of policy implementation in Guanajuato, Central Mexico. Environ. Sci. Policy 77, 107–113.
INEGI, 2007. Conjunto De Datos Vectorial Edafologico Escala 1:250,000 Serie II. Continuo Nacional (Guanajuato). Instituto Nacional de Estadistica, Geografia e informatica.
INEGI 2019 2010 National Census.
Ingenieros Civiles y Geologos Asociados,, S.A., 1980. Estudio Geohydrologico en El Valle De San Jose Iturbide. Guanajuato.
Izbicki, J.A., Teague, N.F., Hatzinger, P.B., Bohlke, J.K., Sturchio, N.C., 2015. Groundwater movement, recharge, and perchlorate occurrence in a faulted alluvial aquifer in California (USA). Hydrogeol. J. 23 (3), 467–491.
Johannesson, K.H., Tang, J.W., 2009. Conservative behavior of arsenic and other oxyanion-forming trace elements in an oxic groundwater flow system. J. Hydrol. 378 (1–2), 13–28.
Knappett, P.S.K., Li, Y.M., Hernandez, H., Rodriguez, R., Aviles, M., Deng, C., Pina, V., Giardino, J.R., Mahlknecht, J., Datta, S., 2018. Changing recharge pathways within an intensively pumped aquifer with high fluoride concentrations in Central Mexico. Sci. Total Environ. 622, 1029–1045.
Li, Y., Hernandez, H.J., Aviles, M., Knappett, P.S.K., Giardino, J.R., Miranda, R., Puy, M.J., Padilla, F., Morales, J., 2020. Empirical Bayesian Kriging method to evaluate inter-annual water-table evolution in the Cuenca Alta del Río Laja aquifer, Guanajuato. México. J. Hydrol. 582 (124517).
Ling, H.B., Xu, H.L., Shi, W., Zhang, Q.Q., 2011. Regional climate change and its effects on the runoff of Manas River, Xinjiang. China. Environ. Earth Sci. 64 (8), 2203–2213.
Louppe, G., Wehenkel, L., Sutera, A., Geurts, P., 2013. Understanding Variable Importances in Forests of Randomized Trees. Curran Associates Inc., Red Hook, NY, United States, pp. 431–439.
Mahlknecht, J., 2003. Estimation of Recharge in the Indpendence aquifer, Central Mexico, By Combining Geochemical and Groundwater Flow Models. University of Agricultural and Life Sciences (BOKU).
Mahlknecht, J., Garfias-Solis, J., Aravena, R., Tesch, R., 2006a. Geochemical and isotopic investigations on groundwater residence time and flow in the Independence Basin, Mexico. J. Hydrol. 324 (1–4), 283–300.
Mahlknecht, J., Medina-Mejia, M.G., Garfias-Solis, J., Cano-Aguilera, I., 2006b. Intrinsic aquifer vulnerability assessment: validation by environmental tracers in San Miguel de Allende, Mexico. Environ. Geol. 51 (3), 477–491.
Mahlknecht, J., Schneider, J., Merkel, B., de Leon, I.N., Bernasconi, S., 2004. Groundwater recharge in a sedimentary basin in semi-arid Mexico. Hydrogeol. J. 12 (5), 511–530.
Mayo, A.L., 2010. Ambient well-bore mixing, aquifer cross-contamination, pumping stress, and water quality from long-screened wells: what is sampled and what is not? Hydrogeol. J. 18 (4), 823–837.
McMahon, P.B., Chapelle, F.H., 2008. Redox processes and water quality of selected principal aquifer systems. Ground Water 46 (2), 259–271.
Mexicano, S.G., 2017. Guanajuato, Servicio Geologico Mexicano. Pachucha, Hidalgo, Mexico.
Mihajlov, I., Mozumder, M.R.H., Bostick, B.C., Stute, M., Mailloux, B.J., Knappett, P.S.K., Choudhury, I., Ahmed, K.M., Schlosser, P., van Geen, A., 2020. Arsenic contaminantion of Bangladesh aquifers exacerbated by clay layers. Nat. Commun. 11 (2244).
Montero, J., Fernandez-Aviles, G., Mateu, J., 2015. Spatial and Spatio-Temporal Geostatistical Modeling and Kriging. John Wiley & Sons, Ltd., West Sussex, UK.
Morales-Arredondo, J.I., Esteller-Alberich, M.V., Hernandez, M.A.A., Martinez-Florentino, T.A.K., 2018. Characterizing the hydrogeochemistry of two low-temperature thermal systems in Central Mexico. J. Geochem. Explor. 185, 93–104.
Munoz, M.O., Bhattacharya, P., Sracek, O., Ramos, O.R., Aguirre, J.Q., Bundschuh, J., Maity, J.P., 2015. Arsenic and other trace elements in thermal springs and in cold waters from drinking water wells on the Bolivian Altiplano. J. S Am. Earth Sci. 60, 10–20.
Naseem, S., McArthur, J.M., 2018. Arsenic and other water-quality issues affecting groundwater, Indus alluvial plain, Pakistan. Hydrol. Process 32 (9), 1235–1253.
Nickson, R., McArthur, J., Burgess, W., Ahmed, K.M., Ravenscroft, P., Rahman, M., 1998. Arsenic poisoning of Bangladesh groundwater. Nature 395 (6700), 338.
Nieto-Samaniego, A.F., Alaniz-Alvarez, S.A. and Camprubi, A. (2007) Geology of Mexico: celebrating the Centenary of the Geological Society of Mexico: geological Society of America Special Paper. Alaniz-Alvarez, S.A. and Nieto-Samaniego, A.F. (eds), pp. 41–70, Geological Society of America.
Nordstrom, D.K., 2002. Public health - Worldwide occurrences of arsenic in ground water. ScienceScience 296 (5576), 2143–2145.
Ortega-Guerrero, M.A., 2009. Occurrence, distribution, hydrochemistry and origin of arsenic, fluoride and other trace elements dissolved in groundwater at basin scale in central Mexico. Rev. Mex. Cienc. Geol. 26 (1), 143–161.
Parkhurst, D.L., Appelo, C.A.J., 2013. User's Guide to PHREEQC Version 3 - A computer Program For speciation, batch-reaction, One-Dimensional transport, and Inverse Geochemical Calculations. United States Geological Survey, Denver, Colorado, p. 497.
Phillips, F.M., 1994. Environmental Tracers for Water-Movement in Desert Soils of the American Southwest. Soil Sci. Soc. Am. J. 58 (1), 15–24.
Rango, T., Vengosh, A., Dwyer, G., Bianchini, G., 2013. Mobilization of arsenic and other naturally occurring contaminants in groundwater of the Main Ethiopian Rift aquifers. Water Res. 47 (15), 5801–5818.
Russoniello, C.J., Lautz, L.K., 2019. Pay the PIED Piper: guidelines to Visualize Large Geochemical Datasets on Piper Diagrams. Groundwater.

Scanlon, B.R., Keese, K., Reedy, R.C., Simunek, J., Andraski, B.J., 2003. Variations in flow and transport in thick desert vadose zones in response to paleoclimatic forcing (0-90 kyr): field measurements, modeling, and uncertainties. Water Resour. Res. 39 (7).

Schaefer, M.V., Ying, S.C., Benner, S.G., Duan, Y.H., Wang, Y.X., Fendorf, S., 2016. Aquifer Arsenic Cycling Induced by Seasonal Hydrologic Changes within the Yangtze River Basin. Environ. Sci. Technol. 50 (7), 3521–3529.

Shamsudduha, M., Taylor, R.G., Chandler, R.E., 2015. A generalized regression model of arsenic variations in the shallow groundwater of Bangladesh. Water Resour. Res. 51 (1), 685–703.

Shepherd, F., 2018. Arsenic and Fluoride Contamination in the Independence Basin Aquifer System of Guanajuato. Mexico, Kansas State University, Manhattan, Kansas.

Smedley, P.L., Kinniburgh, D.G., 2002. A review of the source, behaviour and distribution of arsenic in natural waters. Appl. Geochem. 17 (5), 517–568.

Smedley, P.L., Nicolli, H.B., Macdonald, D.M.J., Barros, A.J., Tullio, J.O., 2002. Hydrogeochemistry of arsenic and other inorganic constituents in groundwaters from La Pampa. Argentina. Appl. Geochem. 17 (3), 259–284.

Smith, R., Knight, R., Fendorf, S., 2018. Overpumping leads to California groundwater arsenic threat. Nat. Commun. 9.

Tyler, S.W., Chapman, J.B., Conrad, S.H., Hammermeister, D.P., Blout, D.O., Miller, J.J., Sully, M.J., Ginanni, J.M., 1996. Soil-water flux in the southern Great Basin, United States: temporal and spatial variations over the last 120,000 years. Water Resour. Res. 32 (6), 1481–1499.

United Nations-Water. Integrated Monitoring Guide for Sustainable Development Goal 6 on Water and Sanitation: Targets and Global Indicators. July 14, 2017.

van Geen, A., Zheng, Y., Versteeg, R., Stute, M., Horneman, A., Dhar, R., Steckler, M., Gelman, A., Small, C., Ahsan, H., Graziano, J.H., Hussain, I., Ahmed, K.M., 2003. Spatial variability of arsenic in 6000 tube wells in a 25 km(2) area of Bangladesh. Water Resour. Res. 39 (5).

Wester, P., Minero, R.S., Hoogesteger, J., 2011. Assessment of the development of aquifer management councils (COTAS) for sustainable groundwater management in Guanajuato, Mexico. Hydrogeol. J. 19 (4), 889–899.

Winkel, L., Berg, M., Amini, M., Hug, S.J., Johnson, C.A., 2008. Predicting groundwater arsenic contamination in Southeast Asia from surface parameters. Nat. Geosci. 1 (8), 536–542.

Zhang, Z., Guo, H.M., Zhao, W.G., Liu, S., Cao, Y.S., Jia, Y.F., 2018. Influences of groundwater extraction on flow dynamics and arsenic levels in the western Hetao Basin, Inner Mongolia. China. Hydrogeol. J. 26 (5), 1499–1512.