OPEN

Probabilistic evaluation of shallow groundwater resources at a hypothetical carbon sequestration site

Zhenxue Dai[1], Elizabeth Keating[1], Diana Bacon[2], Hari Viswanathan[1], Philip Stauffer[1], Amy Jordan[1] & Rajesh Pawar[1]

[1]Los Alamos National Laboratory, Los Alamos, NM 87545, [2]Pacific Northwest National Laboratory, Richland, WA 99352.

Carbon sequestration in geologic reservoirs is an important approach for mitigating greenhouse gases emissions to the atmosphere. This study first develops an integrated Monte Carlo method for simulating CO_2 and brine leakage from carbon sequestration and subsequent geochemical interactions in shallow aquifers. Then, we estimate probability distributions of five risk proxies related to the likelihood and volume of changes in pH, total dissolved solids, and trace concentrations of lead, arsenic, and cadmium for two possible consequence thresholds. The results indicate that shallow groundwater resources may degrade locally around leakage points by reduced pH and increased total dissolved solids (TDS). The volumes of pH and TDS plumes are most sensitive to aquifer porosity, permeability, and CO_2 and brine leakage rates. The estimated plume size of pH change is the largest, while that of cadmium is the smallest among the risk proxies. Plume volume distributions of arsenic and lead are similar to those of TDS. The scientific results from this study provide substantial insight for understanding risks of deep fluids leaking into shallow aquifers, determining the area of review, and designing monitoring networks at carbon sequestration sites.

G eologic carbon sequestration (GCS) in deep reservoirs is considered to be a viable long-term solution for mitigating greenhouse gases in the atmosphere[1]. But there are substantial regulatory and technical challenges that need to be overcome before GCS can be widely deployed. A major regulatory concern that has garnered much research attention is the potential leakage of CO_2 and brine from deep storage reservoirs to shallow groundwater resources through preferential pathways, such as faults, wells, and local high-permeability zones in caprocks. The leaked CO_2 and brine may cause a decrease in groundwater pH, an increase in total dissolved solids (TDS), and a potential mobilization of trace metals from aquifer materials into groundwater[2–20]. While the potential risks to shallow aquifers at GCS operations are known, no defensible methodology for quantifying the risks has been reported to date[3,4].

As part of the National Risk Assessment Partnership (NRAP) project[21], we are developing approaches that can be used to quantify long-term risks at GCS sites. Current efforts are focused on analysis of risk proxies (i.e., impacts and their likelihoods) that will ultimately be combined with consequence analyses to create rigorous risk assessments. NRAP utilizes an Integrated Assessment Model (IAM) to predict long-term performance of GCS sites. The IAM is developed using a system modeling approach that captures the physical and geochemical processes that occur within and among various sub-components of the sequestration sites, e.g., from sequestration reservoir to leaky well, then to underground sources of drinking water (USDW) and/or the atmosphere. As an integrated model, the IAM simulates the coupling processes from the point of CO_2 injection to the interactions among CO_2, brine, groundwater, and aquifer materials in case of a leak. Coupling the complex sub-component models involves significant computational costs, since the inherent heterogeneities and uncertainties in geologic systems require modeling a large number of realizations to bracket and quantify uncertainties associated with likelihoods, impacts, and risks. To meet the requirements of modeling coupled processes and performing large number of realizations, NRAP is developing and using reduced order models which are computationally efficient versions of the full process models that capture the underlying physical and chemical processes at reduced computational costs[21].

This study demonstrates the developed methodology by using the highly heterogeneous unconfined Edwards aquifer as an example (Texas, USA). This aquifer was used because there is considerable data available, and because carbonate aquifers represent an important class of drinking water sources in the USA. Furthermore, the Edwards aquifer overlies potential sequestration sites including oil and gas reservoirs[23]. The Edwards aquifer is

Table 1 | Uncertain parameters and the ranges

#	Parameters	Min	Max
P1	Permeability variance	0.01	0.8
P2	Integral scale (km)	1.0	4.0
P3	Anisotropy (kx/kz)	1.0	50
P4	Mean permeability (m²)	10^{-14}	10^{-10}
P5	Mean Porosity (positively correlated to permeability)	0.05	0.35
P6	Max. CO_2 pressure (MPa)	10	40
P7	Max. CO_2 saturation	0.1	1.0
P8	Wellbore permeability (m²)	10^{-16}	10^{-12}
P9	Cl concentration (mol/l)	0.5	5.4

one of the most productive Karst aquifers in the world. It serves the diverse agricultural, industrial, recreational, and domestic needs of almost two million users in south central Texas[22]. The total groundwater withdrawal from the Edwards aquifer in 2000 was 2.8×10^6 m³/d of which withdrawal for public supply was about 1.6×10^6 m³/d. The Edwards aquifer consists of an underground layer of porous, fractured, honeycombed, water-bearing limestone rock with highly heterogeneous structures. In the recharge zone, the aquifer is unconfined. In the transition and discharge zones, it is an artesian aquifer confined by a layer of very low permeability sediment.

We select an unconfined zone of the Edwards aquifer as an example in order to evaluate the impact of the leaked CO_2 and brine on the USDW and the potential CO_2 discharge rate into the atmosphere (Note that other NRAP groups use the confined High Plain aquifer as an example to cover different types of shallow aquifers in the USA[21]). The study area is located north of San Antonio, where the unconfined carbonate aquifer has a total thickness between 100 and 200 m, with a mean thickness of 150 m. The groundwater flows from northwest to southeast with a hydraulic gradient around 0.00087 and the permeability structure of the unconfined carbonate is highly heterogeneous[24]. By reviewing the previous study of the permeability data[24–28], we summarize the statistical and spatial correlation parameters such as mean, variance, and integral scale in Table 1. The vertical integral scale is assumed to be linealy correlated to the horizontal integral scale with a factor of 0.01. The existing permeability data from Lindgren et al.[24–25] are used as pilot-point data for geostatistical simulations. A pilot-point-based Gaussian simulation method is modified from GSLIB[29] to a geotatistical model GEOST[30,31] for simulating permeability fields. PSUADE[32] and GEOST are coupled to generate nearly 500 realizations of the process model with Latin Hypercube sampling and geostatistical modeling for an integrated Monte Carlo simulation of CO_2 reactive transport in the unconfined carbonate aquifer.

For each realization the uncertain parameters are sampled first and then the heterogeneous permeability fields for the unconfined aquifer are simulated with the pilot-point-based Gaussian method. Figure 1 shows one realization of the generated permeability fields in three dimensions. The model y axis is set along the groundwater flow direction and the lateral flow rate is varied with the sampled permeability and porosity. The numerical model size is $5000 \times 8000 \times 150$ m with 64315 elements. The possible CO_2 and brine leakage point is set at the upstream with a coordinate of (2500, 7000, 0), where the grid is highly refined with minimum grid sizes Δx, Δy, Δz of 2, 4, and 1 m, respectively. Away from the leakage point, the numerical grids become coarse (Figure 1).

The maximum CO_2 pressure and saturation in the sequestration reservoirs and the wellbore permeability in the caprocks listed in Table 1 are three parameters we used to estimate the CO_2 and brine leakage rates from the deep reservoirs to the shallow aquifer with a reduced order model developed as part of NRAP[28]. The reduced order model computes CO_2 and brine leakage source terms using CO_2 pressure and saturation in sequestration reservoir and the wellbore permeability. One example of the CO_2 and brine leakage functions is shown in Figure 2, in which the CO_2 injection time and rate in the deep reservoirs were assumed to be 50 years and 5 million tons per year, respectively. When CO_2 and brine leaks into the unconfined carbonate aquifer, there could be very complex CO_2-induced geochemical reactions in the aquifer, including aqueous equilibrium reactions (or acid-base and complexation reactions), trace metal adsorption or desorption, and mineral dissolution or precipitation[11,15,33]. After reviewing all those reactions discussed by Bacon[33] and Zheng et al.[16] for carbonate aquifers such as Edwards aquifer, we incorporated the reactions, which have the most impact on the pH plume development, into our reactive transport simulations:

(1) CO_2 gas dissolution in water to reduce pH,

$$CO_{2(gas)} \Rightarrow CO_{2(aqueous)} + H_2O \Rightarrow HCO_3^- + H^+,$$
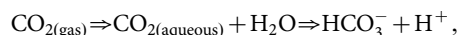
(2) Aqueous equilibrium reactions and the equilibrium constants at 25°C,

$$HCO_3^- + H^+ = H_2CO_3 \quad \log K = -6.3,$$

$$CO_3^{2-} + 2H^+ = H_2CO_3 \quad \log K = -16.6,$$
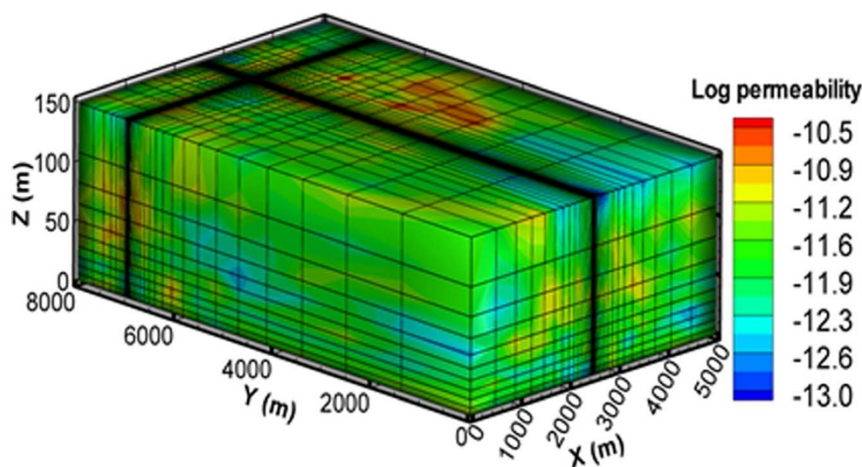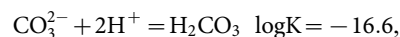
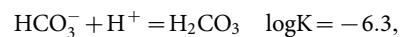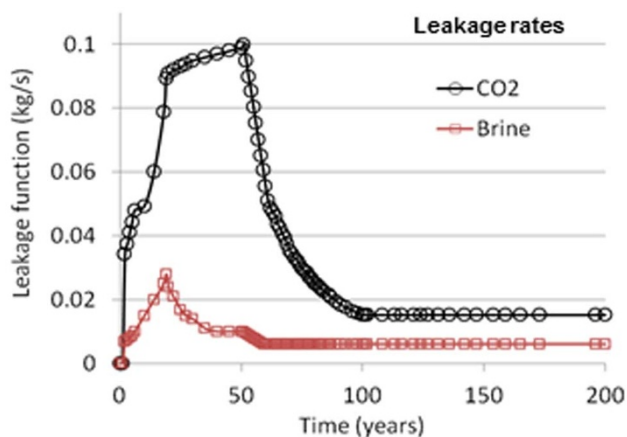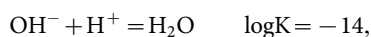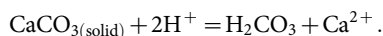**Figure 1 | The simulated heterogeneous permeability (logm²) field in a 3-D view.**

Figure 2 | The simulated CO_2 and brine leakage rates in a leaky well from one realization.

$$OH^- + H^+ = H_2O \qquad logK = -14,$$

(3)   Calcite kinetic dissolution or precipitation,

$$CaCO_{3(solid)} + 2H^+ = H_2CO_3 + Ca^{2+}.$$

Calcite is the most plentiful mineral in the carbonate aquifer and can significantly buffer the pH. The forward rate for the kinetically-controlled calcite dissolution or precipitation at 25°C is 1.5×10^{-6} mol/m²/s[11]. The specific surface area for calcite is 0.1 cm²/g. Dolomite is another common mineral in carbonate aquifers. In this study area the volume fraction of dolomite is very small relatively to that of calcite. We ignored dolomite in our geochemical calculations. We assumed that the brine is mainly made up of NaCl and treat its concentration as an uncertain variable, ranging from 0.5 to 5.4 mol/L. The geochemical analysis conducted by Bacon[33] and Last et al.[34] indicates that As, Cd, and Pb concentrations in the brine are proportional to the NaCl concentration. The proportional ratios are 3.16×10^{-7}, 3.16×10^{-8} and 1.0×10^{-5} on a molar basis for As, Cd, and Pb, respectively. We simulated the species Cl, Na, and trace metals (As, Cd, and Pb) from the leaked brine as conservative tracers by ignoring the sorption reactions of the trace metals. Since sorption processes generally reduce the concentrations of the trace metals, our approach represents the conservative or worst-case scenarios for the trace metals.

By examining the existing water quality data in the unconfined Edwards aquifer we establish baseline data sets and statistical protocols for determining statistically significant changes between background concentrations and predicted concentrations (after CO_2 and brine leakage)[33,34]. The background (or initial) concentrations of the major chemical components, as well as their EPA maximum contaminant levels (MCLs) and no-impact thresholds, are listed in Table 2[33,34]. Note that the no-impact thresholds are determined for pH, TDS, As, Cd, and Pb based on their standard deviations of the measured water quality data and are used to identify potential areas (or volumes) of contamination predicted by aquifer numerical models. This study quantitatively evaluates the following five volume-based risk proxies: 1) plume volume of pH less than the MCL or a no-impact threshold in the aquifer; 2) plume volumes of TDS larger than the MCL or a no-impact threshold; and 3) plume volumes of the three trace metals (e.g. As, Cd, and Pb) over their MCLs and no-impact thresholds. Additionally, we calculate the CO_2 discharge rate into the atmosphere (It should be noted that currently there is no consequence threshold used for CO_2 discharge rate to the atmosphere).

Table 2 | Initial values, no-impact thresholds, and MCLs for each component

Component	Initial values	No-impact thresholds	MCLs	Units
Arsenic	0.31	0.55	10	µg/L
Cadmium	0	0.04	5	µg/L
Lead	0.06	0.15	15	µg/L
pH	6.9	6.6	6.5	standard
TDS	330	420	500	mg/L

Results

The integrated MC simulation has been developed in this study for assessing the plume volumes of pH, TDS, and trace metals (e.g. As, Cd, and Pb), and CO_2 discharge rates into the atmosphere under a probability framework. In most of the MC realizations the shallow groundwater resources are degraded locally around leakage points by reduced pH, increased total dissolved solids, and trace metals mobilized from the injection reservoir. The simulated pH plumes are the largest of the risk proxies and extend from 100 to 1500 m along the flow direction. Perpendicularly to the flow direction pH plumes extend less than 500 m. Subsequent to the leakage free phase of CO_2 gas is formed, which preferentially migrates upwards due to buoyancy and leads to formation of pH plumes to the water table. The amount of CO_2 leaving the top of the aquifer ranges between 0 – 98% of the CO_2 leaking into the aquifer.

The global sensitivity results indicate that the volumes of pH less than the MCL and the no-impact threshold are most sensitive to aquifer porosity and CO_2 leakage rates, as well as the cumulative CO_2 mass and aquifer mean permeability. The volumes of TDS larger than the MCL and the no-impact threshold are most sensitive to the cumulative brine mass, brine leakage rate, and aquifer porosity. The CO_2 discharge rates (into the atmosphere) are positively correlated to the leakage rate from deep reservoirs.

For all of the output variables pH plume volumes (measured with MCL or no-impact thresholds) have the largest mean and median while the trace metal Cd plume volumes are the smallest. When measured with no-impact thresholds, the plume volume distributions of As and Pb are similar to those of TDS. Since the pH plume sizes are the largest among the five output variables, we will take it as one of the major risk proxies and use it as the reference to determine the area of review and to design the monitoring network in the unconfined shallow aquifer.

Discussion

The numerical results from the multi-phase reactive transport simulator FEHM[35] show that when CO_2 leaks into the aquifer, there is a small fraction of CO_2 gas (less 10%) dissolved into water, which leads immediately to a pH reduction at the leak points. The groundwater lateral flow transports the gaseous and dissolved liquid CO_2 downstream to extend the pH plumes. At the same time, the reduced pH enhances calcite dissolution to consume the increasing H^+, which buffers the pH. Because the leaked gaseous CO_2 has a much smaller density than water in the aquifer, the gaseous CO_2 also transports upwards by buoyancy, which leads to pH plumes to extend vertically until the water table (see Figures 3A, 3B and 3C). Finally the gaseous CO_2 leaves the top of the unconfined aquifer into the atmosphere. In most of the MC realizations, the simulated pH < 6.5 plumes extend from 100 to 1500 m along the flow direction. Perpendicularly to the flow direction pH plumes extend less than 500 m (Figure 3D).

Global sensitivity analysis. By using the Monte Carlo simulation results as the input for PSUADE[32], we conduct a global sensitivity analysis with the Main Effect method. At this step the three parameters that define the CO_2 and brine leakage functions (e.g. maximum CO_2 pressure and saturation in the squestration
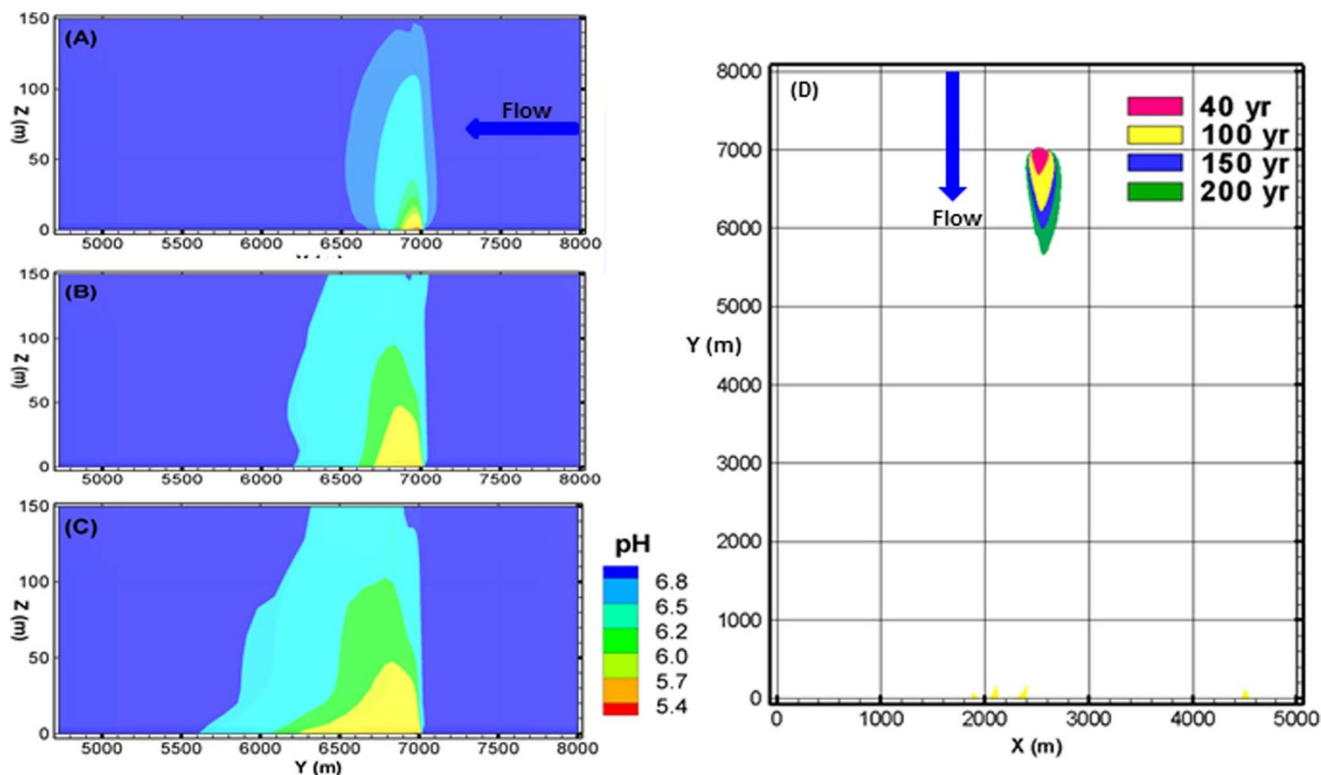
Figure 3 | Simulated pH plumes in the aquifer at different times from a YZ plane view at 40 (A), 100 (B), and 200 years (C), and a XY plane view (D) at four different times.

reservoirs and the wellbore permeability in the caprock) are converted to time-dependent leakage rates and the accumulated CO_2 mass computed from the leakage functions for each realization. The results plotted in Figure 4 show that the plume volumes measured with MCLs and no-impact thresholds have the similar sensitivity to the input parameters, but different risk proxies (e.g. pH and TDS volumes) are sensitive to different parameters. For example, the volumes of pH less than the MCL and no-impact threshold are most sensitive to aquifer porosity and CO_2 leakage rates, as well as the cumulative CO_2 mass and aquifer mean permeability (Figure 4A). Since the aquifer porosity is linearly correlated to the volumes of pH plumes, it is most sensitive to that parameter. The volumes of TDS larger than the MCL and no-impact threshold are most sensitive to the cumulative brine mass, brine leakage rate and aquifer porosity (Figure 4B). The sensitivities of the volumes of the tracer metals (As, Cd, and Pb) less than their

MCLs and thresholds are similar to those of the volumes of TDS. The CO_2 discharge rate (into the atmosphere) is positively correlated to the leakage rate from the deep reservoirs.

Probability measures for the risk proxies. By using the post-processing results of the original 460 (200 year) MC simulations, we conduct a statistical analysis of the output variables with their MCLs and no-impact thresholds. The 5th, 50th (median), and 95th percentiles of the outputs are computed at 20 different time steps. The statistical analysis results of the CO_2 discharge rate to the atmosphere are shown in Figure 5.

Since the MCLs and no-impact thresholds for pH and TDS are very close to each other (Table 2), the computed pH and TDS plume volumes with thresholds are slightly higher than their volumes with MCLs (Figures 5A and 5B). The MCLs and no-impact thresholds for trace metals (As, Cd, and Pb) are quite different and their percentile
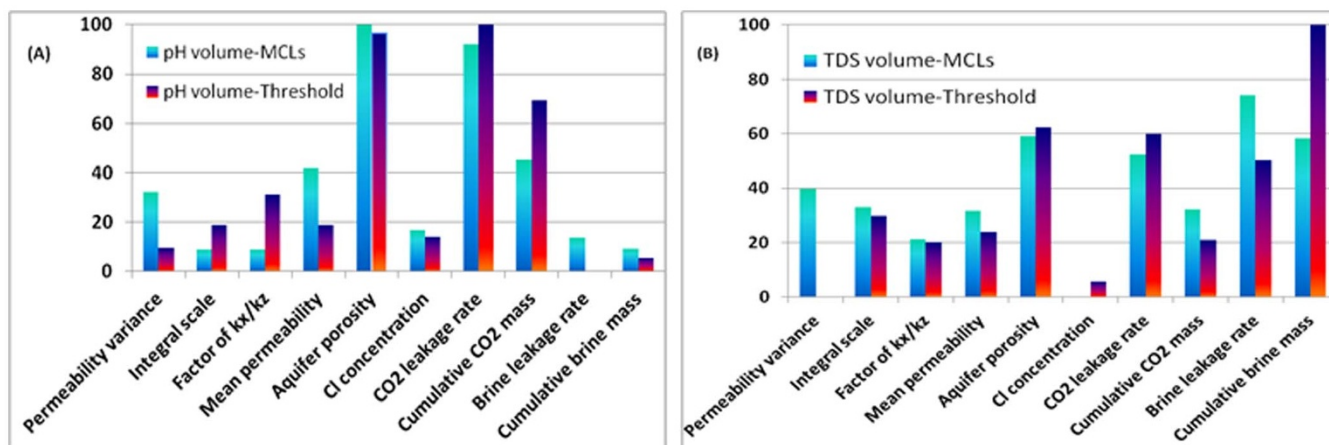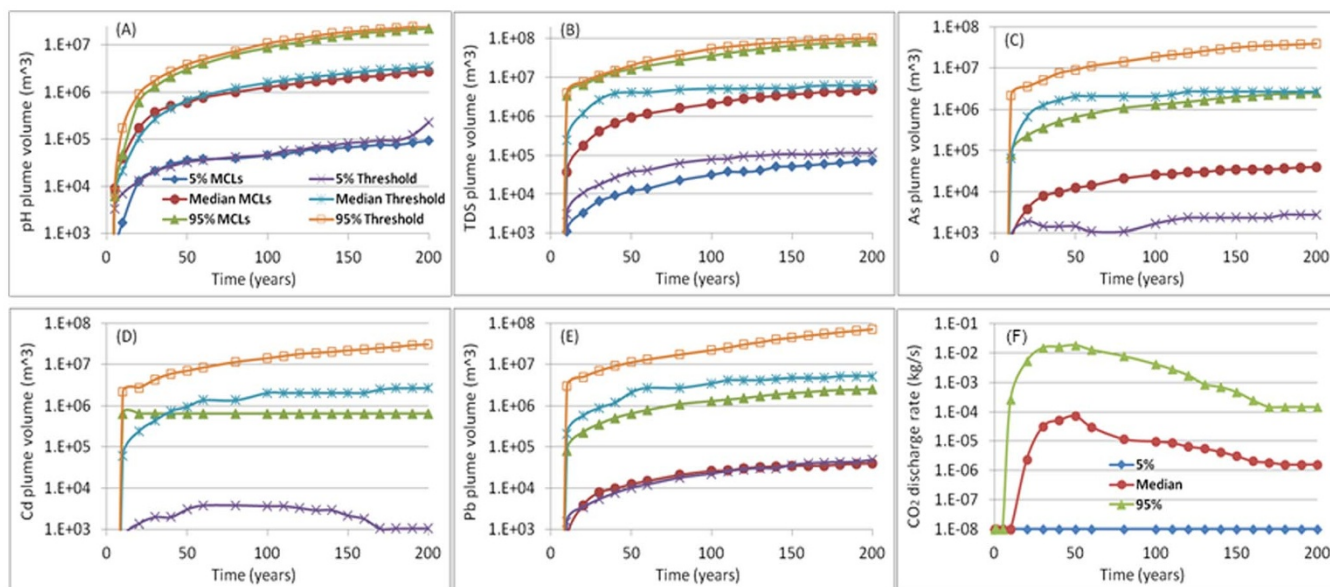

Figure 4 | Global sensitivity of the volumes of pH (A) and TDS plume volumes (B) to the 10 input parameters.

Figure 5 | The statistical analysis results for the 5 groundwater risk proxies and atmospheric leakage of CO_2 with their 5th, 50th (median), and 95th percentiles.

curves are also quite different (Figures 5C, 5D, and 5E). With high standard MCLs the plume volumes of the three trace metals are 0 at the 5th percentiles. The median plume volume of Cd is also 0 when measuring with its MCL (Figure 5D). The trace metal plume volumes measured with no-impact thresholds are significantly larger than those measured with MCLs. Surprisingly, the computed 5th, 50th (median), and 95th percentiles of the trace metal plume volumes with no-impact thresholds are very similar to the TDS volumes. Note that all these plume volumes still keep increasing slowly even though the CO_2 and brine leakage decreases after 50 years. The time-dependent trend of the of the CO_2 discharge rate to the atmosphere is similar to the CO_2 leakage rate to the aquifer, but the discharge rate has a larger tail at the late time, which means that when the leakage rate (into

aquifer) goes down after 50 years the gaseous CO_2 in the unconfined aquifer still keeps releasing into the atmosphere with a gradually reduced rate. The 5th percentile of the CO_2 discharge rate is 0 (a value of 10^{-8} in the log scale to represent 0 in Figure 5F).

Figure 6 shows histograms of the plume volume distributions of the 5 risk proxies at 200 years. Note that the log volume of 0 represents no plume. Among all these 5 risk proxies pH plume volumes (both with MCLs and thresholds) have the largest mean and median while the Cd plume volumes are the smallest. For As and Pb, 33% of the 460 realizations do not exceed the MCL threshold and 80% of the realizations do not exceed the MCL for Cd since Cd concentration in the leaked brine source is about one order of magnitude lower than those of As and Pb. When measured with no-impact thresholds, the
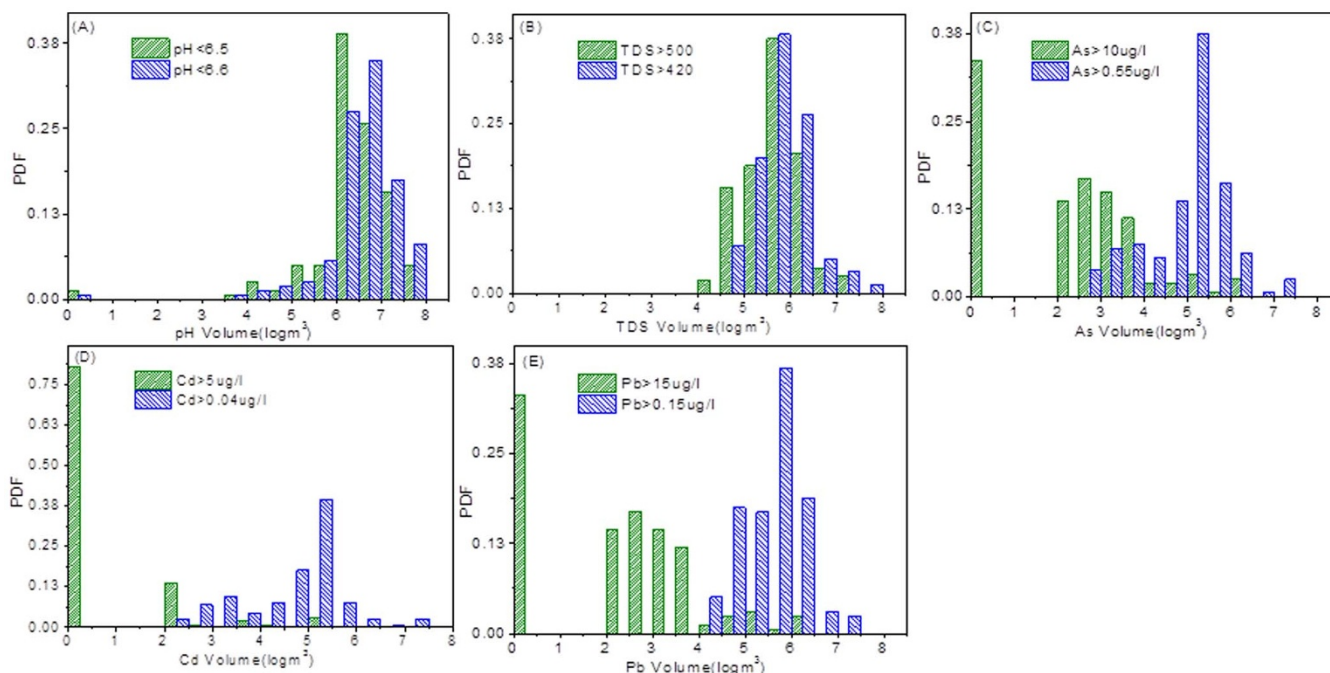
Figure 6 | The histograms of the computed plume volumes for pH, TDS, As, Cd and Pb measured with their MCLs and no-impact thresholds at 200 years.

plume volume distributions of As and Pb are similar to those of TDS. The volumes of the risk proxies measured with thresholds represent the potential areas (or volumes) of contamination (but not over the MCLs) in the shallow aquifer. Since the pH plume sizes are the largest among the five risk proxies, they will be used as the reference for determining the area of review and to design the monitoring network for the unconfined shallow aquifer.

Implications. An integrated Monte Carlo (MC) simulation under a probability framework has been developed in this study for assessing risk proxies (plume volumes and probabilities) related to shallow groundwater resources in the unconfined carbonate aquifer. Risk proxies are compared to both MCL and no-impact consequence thresholds. Differences in risk proxies between the two example consequence thresholds (MCL and no-impact) are significant, especially for the trace metals. This implies that careful consideration of consequence threshold definitions must be made.

Even though our study is for a hypothetical site with CO2 and brine leakage, results of our study provide substantial insights into the impact of the potential CO_2 and brine leakage on shallow groundwater quality for the range of parameters studied. Results presented here will be used to develop reduced order models for shallow aquifers which will be incorporated in IAMs and subsequently used in risk assessment calculations. When more site-specific aquifer parameters are obtained from laboratory-scale to field-scale tests, calculations of uncertain parameters must be revisited to refine estimates of all risk proxies. Finally, since the pH plume sizes are the largest among the five risk proxies, they will be used as the reference for determining the area of review and to design effective monitoring network for shallow aquifers.

Methods

Previous studies in the Edwards aquifer provide plentiful information about the ranges and distributions of the aquifer heterogeneous parameters, such as the aquifer thickness, hydraulic gradient, conductivity, and porosity[24–27]. Based on the statistics of these parameters, we develop a comprehensive methodology for conducting an integrated MC simulation of water and CO_2 flow and reactive transport in the unconfined aquifer. The integrated MC simulation sequentially uses the uncertainty quantification code PSUADE[32], the geostatistical modeling code GEOST[30,31] (modified from GSLIB[29]) and the multi-phase reactive transport simulator FEHM[35]. PSUADE is used to sample all of the uncertain parameters based on the statistical analysis results of the field data at or near the study area, conduct global sensitivity analysis of the output variables in relation to the input parameters, and evaluate the statistical distributions of the pH, TDS, and trace metal plume volumes. GEOST is used to analyze the heterogeneous structures of the Edwards aquifer and to simulate/generate the heterogeneous fields for permeability and porosity based on the heterogeneity parameters sampled from PSUADE.

The multi-phase reactive transport simulator FEHM is used to model CO_2 and water flow and reactive transport in the unconfined carbonate aquifer for each generated heterogeneous model for the unconfined carbonate aquifer. The Brooks and Corey[35,36] relative permeability is used for water/CO_2 multiphase flow simulations and the related coefficients are adopted from a literature review[35–38]. The numerical simulations always start from steady-state flow simulations as the initial conditions and then simulate the CO_2 multi-phase reactive transport in the aquifer for 200 years. In the transport model, it is assumed that the molecular diffusion coefficient is 10^{-9} m²/s and the longitudinal and transverse dispersivities are 800 m and 500 m, respectively. For each simulation a post-process is conducted to evaluate the four risk proxies, including the plume volumes of pH, TDS, and trace metals, and the CO_2 discharge rate into the atmosphere at 20 different times, which means that each risk factor will have 20 time-dependent values. During the MC simulations we keep checking the variance and mean of the risk proxies. After 300 MC runs we find that their variance and mean are approaching constants. Then, we stop our MC simulations after completing 460 runs when the MC simulations are considered to be converged. By taking into account the 20 times as an additional dimension, we obtain the total number of outputs of 9200 (realization number (460) times the number of time steps (20)). These data are re-organized in a PSUADE input format for global sensitivity and risk analysis conducted under a probability framework to evaluate quantitatively the chosen risk proxies.

In order to determine the key flow and transport parameters driving CO_2 transport behavior in the shallow aquifer, global sensitivity analysis techniques are used for investigating input-output sensitivities that are valid over the entire range of the parameter sampling. The Monte Carlo simulations provide 460 realizations of input transport parameter sets generated with Latin Hypercube sampling and geostatistical modeling. Each realization is propagated through the transport simulator to yield the

output response functions represented by the output variables. Global sensitivity analysis entails the comparison of the output distributions to each of the input parameter distributions and identifies the most sensitive parameters for each output variable. The main effect method[32] is used to quantify the uncertainty and sensitivity of the output variables to input parameters.

The main effect method is a variance-based analysis and it displays first-order Sobol' indices for the response surface built from the Monte Carlo simulations[32]. The essence of this analysis is the statistical measure called variance of conditional expectation. The variance-based analysis uses the following equation:

$$VCE(X_k) = \frac{100}{s} \sum_{j=1}^{s} (\overline{Y}_j - \overline{Y})^2 - \frac{1}{sr^2} \sum_{j=1}^{s} \sum_{i=1}^{r} (Y_{ij} - \overline{Y}_j)^2, \qquad (1)$$

where, VCE measures the variability in the conditional expected values of output variable Y when the input parameter X_k takes on different values. s is the number of distinct values of each input parameter and r is the number of replications. $N = sr$ is the sample size. The sensitivity of the output variable to the input parameters is quantified by Equation (1) and re-ranked with numbers from 0 to 100 to represent the importance of the input parameters[32].

1. Blondes, M. S. et al. National assessment of geologic carbon dioxide storage resources—Methodology implementation: U.S. Geological Survey Open-File Report 2013–1055, 26 p., http://pubs.usgs.gov/of/2013/1055/, 15/01/2014. (2013).
2. Apps, J., Zheng, L., Zhang, Y., Xu, T. & Birkholzer, J. Evaluation of potential changes in groundwater quality in response to CO2 leakage from deep geologic storage. Transp. in Porous Media 82, 215–246 (2010).
3. Harvey, O. R. et al. Geochemical Implications of Gas Leakage associated with Geologic CO2 Storage—A Qualitative Review. Environ. Sci. & Technol. 47, 23–36 (2012).
4. Pawar, R. J. et al. Overview of a CO2 sequestration field test in the West Pearl Queen reservoir, New Mexico. Environ. Geol. 13, 163–180 (2006).
5. Karamalidis, A. K. et al. Trace Metal Source Terms in Carbon Sequestration Environments. Environ. Sci. & Technol. 47, 322–329 (2012).
6. Kharaka, Y. et al. Changes in the chemistry of shallow groundwater related to the 2008 injection of CO2 at the ZERT field site, Bozeman, Montana. Environ. Earth Sci. 60, 273–284 (2010).
7. Little, M. G. & Jackson, R. B. Potential impacts of leakage from deep CO2 geosequestration on overlying freshwater aquifers. Environ. Sci. & Technol. 44, 9225–9232 (2010).
8. Lu, J., Partin, J., Hovorka, S. & Wong, C. Potential risks to freshwater resources as a result of leakage from CO2 geological storage: a batch-reaction experiment. Environ. Earth Sci. 60, 335–348 (2010).
9. Stauffer, P. H., Viswanathan, H. S., Pawar, R. J. & Guthrie, G. D. A system model for geologic sequestration of carbon dioxide. Environ. Sci. & Technol. 43, 565–570 (2008).
10. Stauffer, P. H. et al. Greening coal: breakthroughs and challenges in carbon capture and storage. Environ. Sci. & Technol. 45, 8597–8604 (2011).
11. Trautz, R. C. et al. Effect of Dissolved CO2 on a Shallow Groundwater System: A Controlled Release Field Experiment. Environ. Sci. & Technol. 47, 298–305 (2013).
12. Viswanathan, H. et al. Developing a robust geochemical and reactive transport model to evaluate possible sources of arsenic at the CO2 sequestration natural analog site in Chimayo, New Mexico. Int. J. Greenhouse Gas Control 10, 199–214 (2012).
13. Wang, S. & Jaffe, P. R. Dissolution of a mineral phase in potable aquifers due to CO2 releases from deep formations; effect of dissolution kinetics. Energy Conversion and Managem. 45, 2833–2848 (2004).
14. Wilkin, R. T. & DiGiulio, D. C. Geochemical Impacts to groundwater from geologic carbon sequestration: controls on pH and inorganic carbon concentrations from reaction path and kinetic modeling. Environ. Sci. & Technol. 44, 4821–4827 (2010).
15. Zheng, L. et al. Geochemical modeling of changes in shallow groundwater chemistry observed during the MSU-ZERT CO2 injection experiment. Int. J. Greenhouse Gas Control 7, 202–217 (2012).
16. Yang, C. et al. Single-Well Push-Pull Test for Assessing Potential Impacts of CO2 Leakage on Groundwater Quality in a Shallow Gulf Coast Aquifer in Cranfield, Mississippi. Int. J. Greenhouse Gas Control 18, 375–387 (2013).
17. Keating, E. et al. CO2 leakage impacts on shallow groundwater: Field-scale reactive-transport simulations informed by observations at a natural analog site. Applied Geochem. 30, 136–147 (2012).
18. Deng, H., Stauffer, P., Dai, Z., Jiao, Z. & Surdam, R. Simulation of industrial-scale CO2 storage: Multi-scale heterogeneity and its impacts on storage capacity, injectivity and leakage. Int. J. Greenhouse Gas Control 10, 397–418 (2012).
19. Dai, Z. et al. An integrated framework for optimizing CO2 sequestration and enhanced oil recovery. Environ. Sci. & Technol. Lett. DOI: 10.1021/ez4001033, 1, 49–54 (2014).
20. Zheng, L. et al. On modeling the potential impacts of CO2 sequestration on shallow groundwater: Transport of organics and co-injected H2S by supercritical CO2 to shallow aquifers. Int. J. Greenhouse Gas Control 14, 113–127 (2013).

21. NETL, National Risk Assessment Partnership, NRAP Project Summary, http://www.netl.doe.gov/publications/factsheets/rd/R%26D179.pdf, 15/01/2014. National Energy Technology Laboratory, (2011).

22. Eckhardt, G. The Edwards Aquifer Website, http://www.edwardsaquifer.net/gregg.html, 15/01/2014. (2011).

23. NETL, Carbon Sequestration Atlas of the United States and Canada — Fourth Edition (Atlas IV), http://www.netl.doe.gov/technologies/carbon_seq/refshelf/atlasIV/, 15/01/2014, National Energy Technology Laboratory, (2012).

24. Lindgren, R. J., Dutton, A. R., Hovorka, S. D., Worthington, S. R. & Painter, S. Conceptulization and Simulation of the Edwards Aquifer, San Antonio Region, Texas, Scientific Investigation Report 2004-5277, http://pubs.usgs.gov/sir/2004/5277/pdf/sir2004-5277.pdf, 15/01/2014, U.S. Geological Survey, 143p (2004).

25. Lindgren, R. J. Diffuse-Flow Conceptulization and Simulation of the Edwards Aquifer, San Antonio Region, Texas, Scientific Investigation Report 2006-5319, http://pubs.usgs.gov/sir/2006/5319/pdf/sir2006-5319.pdf, 15/01/2014, U.S. Geological Survey, 48p (2006).

26. Painter, S., Woodbury, A. & Jiang, Y. Transmissivity estimation for highly heterogeneous aquifers: comparison of three methods applied to the Edwards Aquifer, Texas, USA. *Hydrogeol. J.*, **15**, 315–331 (2006).

27. Hutchison, W. R. & Hill, M. E. Recalibration of the Edwards (Balcones Fault Zone) Aquifer – Barton Springs Segment – Groundwater Flow Model, State of Texas, http://www.twdb.state.tx.us/groundwater/models/alt/ebfz_b/EBFZ_B_Model_Recalibration_Report.pdf, 15/01/2014, U.S. Geological Survey, 115p (2011).

28. Viswanathan, H. *et al.* Development of a hybrid process and system model for the assessment of wellbore leakage at a geologic CO2 sequestration site. *Environ. Sci. & Technol.* **42**, 7280–7286 (2008).

29. Deutsch, C. V. & Journel, A. G. *GSLIB: Geostatistical Software Library and User's Guide* 380 p, (Oxford Univ. Press, New York, 1992).

30. Dai, Z., Wolfsberg, A., Lu, Z. & Ritzi, R. Representing Aquifer Architecture in Macrodispersivity Models with an Analytical Solution of the Transition Probability Matrix. *Geophys. Res. Lett.* **34**, L20406 (2007).

31. Harp, D. *et al.* Aquifer structure identification using stochastic inversion. *Geophys. Res. Lett.* **35**, L08404 (2008).

32. Tong, C. 2011. PSUADE User's Manual (Version 1.2.0), Technical Report, LLNL-SM-407882, Lawrence Livermore National Laboratory, Livermore, CA 94551-0808, May 2011).

33. Bacon, D. Reduced Order Model for the Geochemical Impacts of Carbon Dioxide, Brine, and Trace Metal Leakage into an Unconfined, Oxidizing Carbonate Aquifer, Version 2.1, NRAP Report, PNNL-22285, Pacific Northwest National Laboratory, Richland, Washington, (2013).

34. Last, G. V. *et al.* Determining Threshold Values for Identification of Contamination Predicted by Reduced Order Models, NRAP Report, PNNL-22077, Pacific Northwest National Laboratory, Richland, Washington, (2012).

35. Zyvoloski, G. A., Robinson, B. A., Dash, Z. V. & Trease, L. L. Summary of the Models and Methods for the (FEHM) Application -- A Finite-Element Heat- and Mass-Transfer Code. LA-13306-MS, Los Alamos National Laboratory, Los Alamos, New Mexico, (2011).

36. Brooks, R. H. & Corey, A. T. *Hydraulic properties of porous media: Hydrology Papers*, 24p, (Colorado State University, Fort Collins, 1964).

37. Pruess, K. & Garcia, J. Multiphase flow dynamics during CO_2 disposal into saline aquifers. *Environ. Geol.* **42**, 282–295 (2002).

38. Dai, Z., Samper, J., Wolfsberg, A. & Levitt, D. Identification of relative conductivity models for water flow and solute transport in unsaturated compacted bentonite. *Physics and Chem. of the Earth* **33**, S177–S185 (2008).

Author contributions

Z.D. designed and carried out the integrated MC simulations, performed data analysis and wrote the draft manuscript. E.K., R.P. and D.B. designed the numerical and geochemical models, discussed the results, reviewed and revised the manuscript. H.V., P.S. and A.J. discussed the results, and revised the manuscript.

Additional information