LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-648252

Key Factors for Determining Risk of Groundwater Impacts Due to Leakage from Geologic Carbon Sequestration Reservoirs

S. Carroll, E. Keating, K. Mansoor, Z. Dai, Y. Sun, W. Trainor-Guitton, C. Brown, D. Bacon

January 6, 2014

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Key Factors for Determining Groundwater Impacts Due to Leakage from Geologic Carbon Sequestration Reservoirs

Susan A. Carroll[1], Elizabeth Keating[2], Kayyum Mansoor[1], Zhenxue Dai[2], Yunwei Sun[1], Whitney Trainor-Guitton[1], Chris Brown[3], Diana Bacon[3]

[1]Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550

[2]Los Alamos National Laboratory, Bikini Atoll Rd., SM 30, Los Alamos, NM 87545

[3]Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99354

Susan Carroll
925-423-5694 (phone)
925-423-0153 (fax)
carroll6@llnl.gov

Elizabeth Keating
505-665-6714 (phone)
505-665-8737 (fax)
ekeating@lanl.gov

Kayyum Mansoor
925 423 1170
925 423 5764 (fax)
mansoor1@llnl.gov

Zhenue Dai
505-665-6387 (phone)
505-665-8737 (fax)
daiz@lanl.gov

Yunwei Sun
925-422-1587 (phone)
925-423-0153 (fax)
sun4@llnl.gov

Whitney Trainor-Guitton
925-423-1825
925-423-5764 (fax)
trainorguitton@lln.gov

Chris Brown
509-371-7381 (phone)
509-371-7344 (fax)
Christopher.Brown@pnnl.gov

Diana Bacon
509-372-6132 (phone)
509-372-6089 (fax)
Diana.Bacon@pnnl.gov

ABSTRACT

The National Risk Assessment Partnership (NRAP) is developing a science-based toolset for the analysis of potential impacts to groundwater chemistry from CO_2 injection (www.netldoe.gov/nrap). The toolset adopts a stochastic approach in which predictions address uncertainties in shallow groundwater and leakage scenarios. It is derived from detailed physics and chemistry simulation results that are used to train more computationally efficient models, referred to here as reduced-order models (ROMs), for each component system. In particular, these tools can be used to help regulators and operators understand the expected sizes and longevity of plumes in pH, TDS, and dissolved metals that could result from a leakage of brine and/or CO_2 from a storage reservoir into aquifers. This information can inform, for example, decisions on monitoring strategies that are both effective and efficient. We have used this approach to develop predictive reduced-order models for two common types of reservoirs, but the approach could be used to develop a model for a specific aquifer or other common types of aquifers.

In this paper we describe potential impacts to groundwater quality due to CO_2 and brine leakage, discuss an approach to calculate thresholds under which "no impact" to groundwater occurs, describe the time scale for impact on groundwater, and discuss the probability of detecting a groundwater plume should leakage occur. To facilitate this, multi-phase flow and reactive transport simulations and emulations were developed for two classes of aquifers, considering uncertainty in leakage source terms and aquifer hydrogeology. We targeted an unconfined fractured carbonate aquifer based on the Edwards aquifer in Texas and a confined alluvium aquifer based on the High Plains Aquifer in Kansas, which share characteristics typical of many drinking water aquifers in the United States. The hypothetical leakage scenarios centered on the notion that wellbores are the most likely conduits for brine and CO_2 leaks. Leakage uncertainty was based on hypothetical injection of CO_2 for 50 years at a rate of 5 million tons per year into a depleted oil/gas reservoir with high permeability and, one or more wells provided leakage pathways from the storage reservoir to the overlying aquifer. This scenario corresponds to a storage site with historical oil/gas production and some poorly completed legacy wells that went undetected through site evaluation, operations, and post-closure.

For the aquifer systems and leakage scenarios studied here, CO_2 and brine leakage are *likely* to drive pH below and increase total dissolved solids (TDS) above the "no-impact thresholds;" and the subsequent plumes, although small, are likely to persist for long periods of time in the absence of remediation. In these scenarios, however, risk to human health may not be significant for two reasons. First, our simulated plume volumes are much smaller than the average inter-well spacing for these representative aquifers, so the impacted groundwater would be unlikely to be pumped for drinking water. Second, even within the impacted plume volumes little water exceeds the primary maximum contamination levels. These observations point to:

- The potential utility of the NRAP toolset to evaluate the risk of leakage and inform monitoring and corrective action plans of a potential site for long-term CO_2 storage by capturing storage reservoir, leakage pathway, and aquifer heterogeneity.
- The importance of establishing baseline groundwater chemistry that captures the pre-injection variability of underground sources of drinking water (USDW) above the

reservoir because the EPA has adopted a "no net degradation" policy towards the protection of groundwater resources.

- The need to test and develop spatially diverse monitoring techniques capable of detecting leakage early to employ effective mitigation strategies, and more importantly to add confidence to assessments used to evaluate the length of the post-injection site care. In our study, the probability of detecting plumes using existing wells to sample the groundwater chemistry was very low, because the plumes were relatively small in both aquifers.
- The need to develop methodologies that prevent and/or directly detect leakage prior to reaching USDWs, because our simulations predict that even small amounts of CO_2 and brine, when left unmitigated, can change USDW pH and TDS concentrations for long periods of time.

1. INTRODUCTION

Deep underground storage of CO_2 from stationary sources, such as power plants and industrial processes, is a promising strategy to limit the amount of CO_2 emitted into the atmosphere and to mitigate the effects of global climate change (IPCC, 2005; NETL, 2012). Long-term storage of CO_2 in deep underground reservoirs requires careful assessment of the reservoir integrity, well and fault susceptibility for potential leakage pathways, and consideration of the impact of leaks into the atmosphere or on shallow groundwater sources (Bachu, 2008; Herzog et al., 2003).

Potential impacts to groundwater quality are a focus for both state and federal regulatory agencies, because leakage of brine and/or CO_2 into groundwater resources and subsequent geochemical transformations may impact water quality. In the United States, the Class VI Rule sets minimum federal technical criteria that injection of supercritical CO_2 in geologic reservoirs are protective of underground sources of drinking water (USDW) that have less than 10,000 mg L^{-1} dissolved solids. The Class VI Rule and related documents are available at http://water.epa.gov/type/groundwater/uic/wells_sequestration.cfm.

The Class VI injection well permitting process requires baseline geochemical information on subsurface formations in the area of review (AoR) and the assessment of risk to water quality for all USDWs within the AoR prior to injection. Additionally, testing and monitoring for signs of leakage is required during the injection and post-injection phases above the confining zone and within overlying USDWs. The default period for post-injection site care (PISC) is currently set for 50 years, during which time operators are required to use periodic indirect (geophysical) and direct (well water) data to assess if USDWs have been or are likely to be compromised. Current guidelines on the duration of the PISC period are flexible and may be reduced if the operators can illustrate that CO_2 and brine are contained and USDWs are protected.

The National Risk Assessment Partnership (NRAP) is developing a science-based toolset for the analysis of potential impacts to groundwater chemistry from CO_2 injection (www.netldoe.gov/nrap). The toolset adopts a stochastic approach in which predictions address uncertainties in shallow groundwater and leakage scenarios. It is derived from detailed physics and chemistry simulation results that are used to train more computationally efficient models, referred to here as reduced-order models (ROMs), for each component of the system. In particular, these tools can be used to help regulators and operators understand the expected sizes and longevity of plumes in pH, TDS, and dissolved metals that could result from a leakage of brine and/or CO_2 from a storage reservoir into aquifers. This information can inform, for example, decisions on monitoring strategies that are both effective and efficient. We have used this approach to develop predictive reduced order models for two common types of reservoirs, but the approach could be used to develop a model for a specific aquifer or other common types of aquifers.

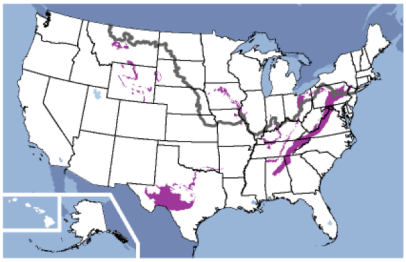
The objectives of this publication are four fold:

- Present summary findings that describe potential impacts to groundwater quality due to CO_2 and brine leakage

- Discuss an approach to calculate thresholds under which "no impact" to groundwater occurs
- Describe the time scale for impact on groundwater
- Discuss the probability of detecting a groundwater plume should leakage occur

To facilitate this, we compared the impact of CO_2 and brine leakage on groundwater quality within two distinct classes of shallow aquifer systems given the same CO_2 storage reservoir and leakage pathways. We focused on shallow USDWs because they are resources that are currently in use. We targeted two classes of aquifers, which share characteristics typical of many drinking water aquifers in the U.S.: an unconfined fractured carbonate aquifer and a confined alluvium aquifer (Figure 1).

a.



b.

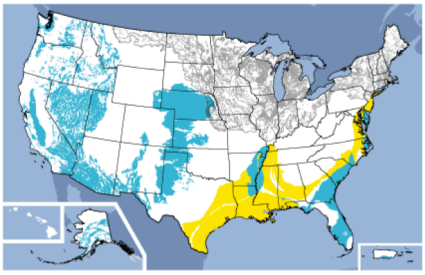


Figure 1. Locations of sand/carbonate (a, violet) and sand/gravel (b, cyan and yellow) shallow drinking waters mapped by the US Geological Survey. (http://water.usgs.gov/ogw/aquiferbasics/).

Multi-phase flow and reactive transport simulations and emulations were developed for these two classes of aquifers, considering uncertainty in leakage source terms and aquifer hydrogeology. The uncertain source term variables considered were: location and number of leaky wells, time-dependent brine/CO_2 leakage rates at each well, and total dissolved solids (TDS) and trace metal concentrations of the leaking brine. Two well-studied aquifers, the Edwards aquifer in Texas and the High Plains Aquifer in Kansas, were used to represent hydrogeologic characteristics of carbonate and alluvium aquifers, respectively. Uncertainty in hydrogeologic properties was considered, as well.

Each simulation provided a spatially explicit, temporal evolution of a shallow groundwater plume. Due to dissolution of CO_2 in groundwater and advective transport of brine, the plumes are lower in pH, and higher in TDS and trace metal concentrations relative to background conditions. Changes in trace metal concentration due to reactions, such as decreases due to adsorption or increases due to pH-related desorption or dissolution were ignored for the purpose of simplification. Two types of metrics were considered: the volume of the plume as defined by concentrations that exceed 1) drinking water standards or 2) "background" thresholds. The latter metric requires statistical analysis of ample background water chemistry sampling at the site.

We created hypothetical leakage scenarios centered on the premise that abandoned wellbores are the most likely conduits for brine and CO_2 leaks. Leakage uncertainty was based on hypothetical injection of CO_2 for 50 years at a rate of 5 million tons per year into a depleted oil/gas reservoir with high permeability, and one or more wells provided leakage pathways from the storage reservoir to the overlying aquifer. The simulations capture variability within the storage reservoir, leakage pathway, and aquifer heterogeneity.

2. METHODS

Ultimately, the NRAP toolset will allow a complete stochastic assessment of carbon storage sites using integrated assessment models (IAMs) that couple individual sub-models for potential storage reservoirs, leakage pathways (such as wellbores or fractures), and groundwater aquifers. Each of the underpinning stochastic models can be used separately to gain insights into the behavior of specific components in the storage-site system. Although the IAMs are still under development, we anticipate that a user will be able to develop and substitute site-specific sub-models as desired to assess various components of the storage-site system to plan the injection and post-injection site care activities. In the current work, we focus on the behavior of USDW aquifers, using leakage-scenario inputs developed separately from specific reservoir and wellbore sub-models.

Our analysis uses the results of between 500 and 700 high-fidelity reactive-transport simulations of the physical and chemical processes that are likely to change groundwater quality if CO_2 and/or brine were to leak from storage reservoirs to USDW aquifers, as well as statistical approximations generated by reduced-order models (ROMs) trained by detailed simulations. The simulated concentrations were used to quantify the size and location of the plumes relative to leakage sources (deep wellbores) and shallow groundwater receptors (drinking, agricultural, and industrial wells) to base discussions on monitoring and corrective action plans that are needed for the permitting of Class VI injection wells. The ROMs are needed to capture variability within the storage, leakage, and USDW aquifer systems through more thorough sampling of the parameter space and significantly faster simulation times to calculate the probability of a change in groundwater chemistry. We used an uncertainty quantification code called PSUADE (Tong, 2005, 2010) to establish sampling points for the reactive-transport simulations, to conduct parameter sensitivity analysis, and to train ROMs.

2.1 STORAGE RESERVOIR

The reservoir ROM is a look up table for the spatial and temporal distribution of CO_2 saturation and pressure as function of variable permeability, porosity, pore compressibility, and van Genuchten α and m for geologic layers in the storage and caprock formations (Table 1). It was derived from 200 simulations of CO_2 injection using a geologic model developed for a potential industrial-scale storage project in the Southern San Joaquin Basin near Kimberlina, California (Zhou and Birkholzer, 2011; Wainwright et al., 2012). The geological structure and hydrogeological parameters of various subsurface layers were determined from field data. The storage formation, based on field data from the Vedder sandstone, was divided into six sand and shale layers. The parallel version of TOUGH2 (Zhang et al., 2008) was used to simulate CO_2/brine migration and pressure buildup within the CO_2 storage formation and overlying/underlying formations. The simulation time includes an injection period of 50 years with an injection rate of 5 Mt per year, and a post-injection period of 150 years.

Table 1: Reference parameter values: horizontal permeability k_h, (±10) anisotropy ratio k_v/k_h, porosity Φ (±30%), pore compressibility β_p (±5), van Genuchten α (±5) and *m* (±30%).

	Caprock	Vedder Sand	Vedder Shale
K_h, *mD*	0.002	Depth Dependent	0.1
K_v/k_h	0.5	0.2	0.5
Φ	0.338	Depth Dependent	0.32
β_p, 10^{-10} Pa^{-1}	14.5	4.9	14.5
α, 10^{-5} Pa^{-1}	0.42	13	0.42
m	0.457	0.457	0.457

2.2 WELLBORE LEAKAGE

In all the scenarios considered in this study, abandoned legacy wells were presumed to be the most likely source of leakage. This scenario would be consistent with a storage site with legacy wells from previous oil and gas operations that were not identified or remediated during site characterization or through monitoring at the site. The intent was to allow leakage in order to understand how the aquifers would respond should a failure occur, recognizing that these assumptions would not be expected for sites using best practices and operated under current class VI regulations. We did select "plausible" leakage scenarios under these conditions, such that the volumes of leaked brines and CO_2 were physically realistic.

Leaky wells were assumed to fully penetrate the caprock and connect the storage reservoir and the shallow aquifer. To generate a range of plausible wellbore leakage scenarios, simplified ROMs for the sequestration reservoir and leaky wellbore were linked. The predicted leak rates were then applied as a CO_2/brine source at the base of the aquifer, as shown in Figure 2. The wellbore leakage ROM considered uncertainty in wellbore permeability and depth (Jordan et al., 2013). This ROM used input from the storage reservoir ROM to link reservoir pressure and CO_2/brine saturation to the leakage rates. In all cases, CO_2 injection was assumed to cease after the first 50 years.

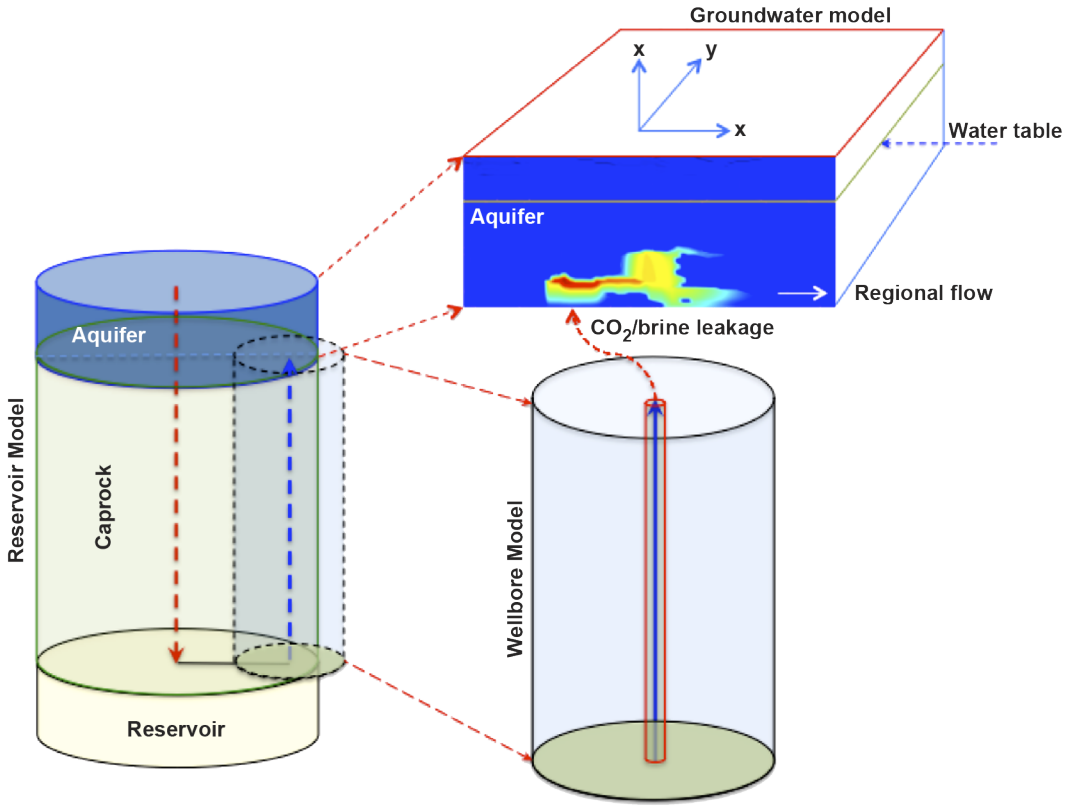


Figure 2: Schematic showing the links between reservoir, well leakage, and aquifer models using the High Plains Aquifer case study. Links between reservoir, well leakage, and the Edwards Aquifer case study are identical.

Assumptions about the location and number of leaky wells differed between the two simulation studies, but both studies considered a similar range of leakage rates and brine chemistry assuming wellbore permeability between 10^{-14} to 10^{-10} m^2 (Table 2). Typical CO_2 leakage rates were between 0.1 and 1 g s^{-1} and cumulatively, represent less than 0.4% of the total mass of CO_2 injected in the reservoir. The rates are similar to those measured at Mammoth Mountain and Crystal Geyser (Lewicki et al., 2007; Wilson et al., 2007). A typical time-varying leakage scenario is shown in Figure 3.

For the High Plains Aquifer study, the location of possible leaky wells was pre-determined using a database of 165 well locations. For each realization, between 1 and 5 leaky well location were selected randomly from those 48 wells located within a 5,000-meter radius of the injector. This percentage of wells (2–10%) spans the percentage (5%) of wells observed to have sustained casing pressure in the Canadian oil fields which was reported by Watson and Bachu (2009), who offered this as an expected rate should legacy wells in an oil/gas region be left unchecked.

For the Edwards aquifer study, only one leaky well was considered, with a fixed location. This assumption allowed very fine grid resolution at the location of the leak. And like the High Plains study, consideration of variation in reservoir and leaky wellbore properties allowed a large number of leakage rates to be considered.

We also considered uncertainty in brine chemistry. A brine chemistry database (www.natcarbviewer.org) was used to evaluate the possible range of sodium and chloride

concentrations in the brine, and an experimental study was used to evaluate the possible range of three trace metals (As, Pb, and Cd) (Karamalidis et al., 2013).

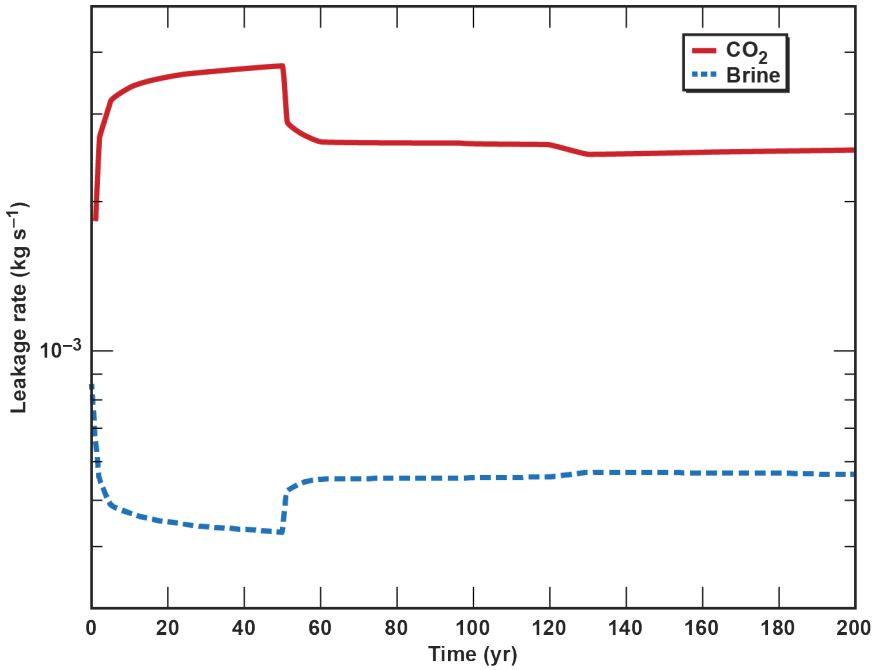


Figure 3: Example CO_2 and brine leakage rates as functions of time.

Table 2: Variable parameters and ranges sampled in the High Plains and Edwards studies

Parameter	High Plains		Edwards		Unit
	Minimum	Maximum	Minimum	Maximum	
CO_2 leakage rate[1]	0.0	168	.00001	311	[g s^{-1}]
Cumulative CO_2 mass[1]	0.0	995	.00124	1840	kton
Brine leakage rate[1]	0.0	56	.0018	36	[g s^{-1}]
Brine mass[1]	0.0	324	.0112	291	kton
NaCl	0.001	6.7	0.1	6.7	[mol L^{-1}]
Arsenic[2]	$10^{-7.98}$	$10^{-5.87}$	$10^{-7.76}$	$10^{-5.94}$	[mol L^{-1}]
Cadmium[2]	$10^{-8.87}$	$10^{-6.43}$	$10^{-8.76}$	$10^{-6.94}$	[mol L^{-1}]
Lead[2]	$10^{-8.12}$	$10^{-4.74}$	$10^{-8.02}$	$10^{-6.19}$	[mol L^{-1}]

[1]Time dependence of CO_2 and brine leakage rates and masses were calculated from variations in wellbore permeability (10^{-14} to 10^{-10} m^2).

[2]Trace metal concentrations were sampled independently from NaCl concentrations for the High Plains study. For the Edwards study trace metals were varied as a constant ratio of Cl (molar concentrations).

2.3 MULTI-PHASE REACTIVE TRANSPORT OF CO_2 AND BRINE IN AQUIFERS

The two classes of aquifers studied were an unconfined fractured carbonate aquifer based on the Edwards Aquifer, Texas, and a confined aquifer with variable lenses of high permeable sands and low permeable silts based on the High Plains Aquifer, Kansas. We simulated the response of representative portions of these aquifers to CO_2 and brine leakage through wells from a CO_2 storage reservoir using multi-phase and multi-component reactive transport codes and calculated changes in pH, TDS, As, Cd, and Pb concentrations under a wide range of hydrogeologic conditions.

Computer codes NUFT (Nitao, 1998; Hao et al., 2012) and FEHM (Zyvoloski et al., 2011) were used for the confined alluvium aquifer and the unconfined carbonate aquifer, respectively. Both codes are highly flexible for modeling non-isothermal, multi-phase flow and reactive transport and have been extensively verified and used for a variety of subsurface flow and transport problems, including nuclear waste disposal, groundwater remediation, CO_2 sequestration and hydrocarbon production.

The reactive transport simulations include a limited amount of chemistry to account for changes in groundwater pH due to CO_2 dissolution, as well as dissolved sodium and chloride as indicators of TDS. The dissolution of CO_2 in groundwater promotes the following sets of reactions:

$$CO_2(g) \leftrightarrow CO_2(aq) \tag{1}$$

$$CO_2(aq) + H_2O \leftrightarrow HCO_3^- + H^+ \tag{2}$$

$$CaCO_3(calcite) + H^+ \leftrightarrow Ca^{2+} + HCO_3^- \tag{3}$$

These reactions promote the acidification of the system, which is then buffered by calcite dissolution. We did not include reactions that might affect trace metal concentrations in the aquifer, such as decreases due to adsorption or increases due to pH-related desorption were ignored for the purpose of simplification. The trace metal plumes described below, therefore, only describe the fate of trace metals originating in the brine.

2.3.1 Unconfined Carbonate Aquifer

We studied the impact of possible leakage from sequestration reservoirs on water quality in carbonate aquifers, because a large percentage of the U.S. drinking water supply is derived from carbonate aquifers. To guide our numerical model construction, we selected a particularly well-characterized example: the carbonate Edwards Aquifer located in south-central Texas (Figure 4). This aquifer covers an area of more than 10^5 km^2 (Painter et al., 2007). We focused on an unconfined portion of the aquifer near San Antonio. The San Antonio segment is one of the most productive karst aquifers in the world, and is the sole source water supply for more than 2 million people (Musgrove et al., 2010). Water levels and groundwater chemistry data from USGS reports for the San Antonio area were used to establish the local hydrologic gradient and background chemistry (Lindgren et al., 2004; Musgrove et al., 2010). The aquifer is composed of carbonate rocks of the Georgetown Formation and the Edwards Group (or their stratigraphic equivalents), which range in thickness from 121–152 m.

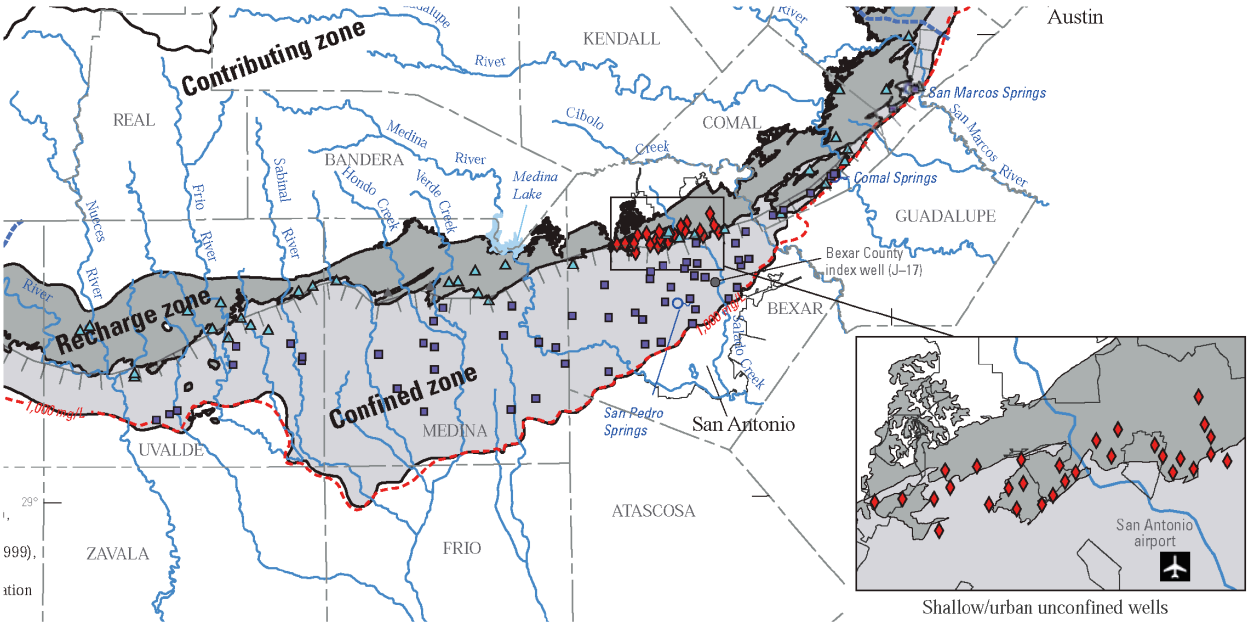


Figure 4: Location of unconfined portion of the Edwards Aquifer in south-central Texas.

The numerical model consisted of hydrostatic lateral boundaries, a water table equal to the local hydrologic gradient of 7.5 Pa m^{-1} and a thickness of 150 m, which matched the upper end of the aquifer's observed thickness. The lateral extent of the model, 8,000 km × 5,000 km, was selected pragmatically to be as small as possible (to allow very small grid blocks), yet much larger than any simulated groundwater plume. The computational mesh had variable grid spacing consisting of small cells near the well ($\Delta x = \Delta y = 9$ m, $\Delta z = 6$ m), gradually increasing to larger cells in the far field ($\Delta x = 200$ m, $\Delta y = 300$ m, $\Delta z = 20$ m) to capture CO_2 buoyancy physics. Using the range of model parameters described in Table 1, plumes never approached the lateral boundaries of the model.

Aquifer heterogeneity in the Edwards is controlled by large and unpredictable variations in karst features (Lindgren, 2006). We assumed random Gaussian variations in permeability, using mean, variance, and correlation lengths determined for this aquifer by Painter et al. (2007) and Lindgren (2006). Stochastic fields of heterogeneous permeability were generated using the pilot point method and random Gaussian interpolation (Deutsch and Journel, 1992; Dai et al., 2007; Harp et al., 2008). All nodes were assumed to have anisotropic intrinsic permeability. Models allowed porosity to vary spatially along with permeability:

$$k = a\emptyset^b \tag{4}$$

where, k [m^2] is permeability, $\emptyset$ is porosity, and a and b are coefficients (a=4.84×10^{-10} and b=3) (Bernabe et al., 2003; Deng et al., 2012). Ranges for uncertain rock parameters listed in Table 3 represent our current understanding of system variability and could be redefined over an alternate range to better describe characterization data from another site.

346 **Table 3: Uncertain parameters and their ranges for unconfined carbonate aquifer simulation**

	Parameter		Minimum	Maximum	Unit
1	Permeability	Variance	0.017	0.79	[km^2]
2		Correlation length	1	3.95	[km]
3		Anisotropy	1.1	49.1	[-]
4		Mean	-13.5	-10.6	$Log_{10}[m^2]$
5	Mean porosity		0.05	0.34	[-]

347

348 ### 2.3.2 Confined Alluvium Aquifer

349 The High Plains Aquifer is representative of a sedimentary aquifer that might overlay a CO_2
350 storage reservoir. The aquifer, also known as the Ogallala aquifer, is one of the largest aquifers
351 in the world covering about 450,000 km^2 and spanning eight states in the Great Plains (Figure 5).
352 The aquifer accounts for approximately 27% of all irrigated land in the United States and about
353 30% of all groundwater used for irrigation (USGS, 2011). It is comprised mainly of
354 unconsolidated or partly consolidated silt, sand, gravel and clay rock debris deposited in the late
355 Miocene to early Pliocene period when the Rocky Mountains were tectonically active (Gutentag
356 et al., 1984).

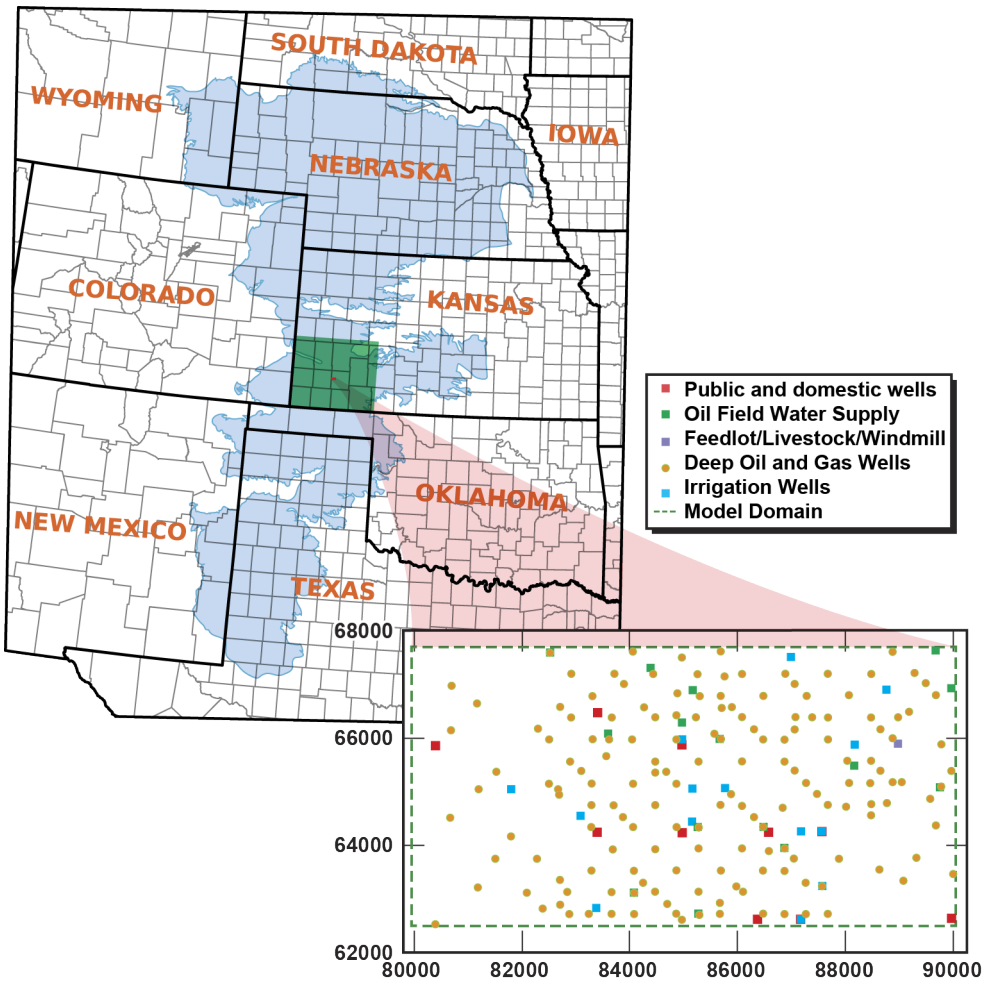


Figure 5: Location of alluvial High Plains Aquifer in Kansas.

The top 300 m of the aquifer was used to develop a lithological model for the geological realizations of the numerical model. Lithologic descriptions were obtained from the Kansas Geological Survey's Water Well Completion Record (WWC5) database (KGS, 2011). All simulations were based on a 10 km × 5 km domain that lies primarily in Haskel County and was selected due to the relatively higher density of lithological picks needed to represent the model depth of 240 m. A total of 468 lithological picks from 48 domestic, feedlot, irrigation, public water supply and oil field water supply wells were used to develop the geostatistically derived indicator models using the TPROGS software (Carle, 1999). Correlation lengths were derived from the transition probability approach. The correlation lengths in the x- and y-direction varied uniformly from 200–2500 m and the correlation length in *z*-direction varied uniformly from 0.50–25.0 m. A total of 1,000 conditional geostatistical realizations were developed based on randomly selected material-volume fraction and correlation lengths using the PSUADE uncertainty quantification software package (Tong, 2005, 2010).

The 3-D numerical model domain captured the unsaturated and saturated zones of the heterogeneous High Plains Aquifer. The model domain extended to 10,000 m × 5,000 m × 240 m with 1 to 5 leakage sources placed at 198 m depth at known well locations. The orthogonal

numerical grid contained fixed cell widths in the x (Δx = 100.0 m), y (Δy = 100.0 m), and z (Δz =4.8 m) directions. The grid dimensions were 100, 50, and 50 nodes in the x, y, and z directions, respectively, for a total of 250,000 nodes. Isothermal conditions were assumed with a generic temperature of 17°C in the entire domain. The uppermost portion of the model was set as atmospheric allowing for both saturated and unsaturated conditions. The east (minimum x) and west (maximum x) model boundaries were fixed in time. Hydrostatic-pressure gradients were achieved by changing the gravity vector. No-flow boundaries were assumed at the southern (minimum y), northern (maximum y) and bottom (minimum z) boundaries. A constant-pressure boundary condition was set on ground surface and at the aquifer bottom to maintain a hydrostatic initial condition with saturated and unsaturated zones. Regional groundwater flow was maintained by a 0.3% hydraulic gradient. Since the regional groundwater flow of the Great Plains Aquifer in southwestern Kansas flows eastward, the mesh is structured to accommodate flow in the predominant x-direction. The leakage source term was estimated from reservoir and wellbore ROMs (Jordan et al., 2013; Wainwright et al., 2012). Each simulation was executed for ~20–60 hours using the high performance computing facility at Lawrence Livermore National Laboratory (LNNL).

Ranges for uncertain rock parameters are listed in Table 4. Physical parameters, including porosity, density, permeability and van Genuchten unsaturated parameters, were taken from the USDA Rosetta database (Schaap et al., 2001).

Table 4: Uncertain parameters and their ranges for alluvium aquifer simulation.

	[1]Parameter	Minimum	Maximum	Unit
1	Sand volume fraction of aquifer	0.35	0.65	[-]
2	Correlation length of aquifer in x	200.0	2500.0	[m]
3	Correlation length of aquifer in z	0.50	25.0	[m]
4	Sand permeability of aquifer	-13	-10	$Log_{10}[m^2]$
5	Clay permeability of aquifer	-18	-15	$Log_{10}[m^2]$

[1]Mean values for correlation length in y = 1350 m, sand and shale porosity of 0.38 and 0.47, sand and shale van Genuchten m of 0.66 and 0.19, sand and clay van Genuchten $\alpha = 5.6234 \times 10^{-5}$ and 1.5136×10^{-5} Pa^{-1}, and CO_2 diffusivity 10^{-9} $m^2 s^{-1}$)

2.4 IMPACT THRESHOLDS

The simulations were used to calculate the volume of groundwater within the shallow aquifers that exceeds certain water quality thresholds. We considered two thresholds in this study, as defined in Table 5 and developed by Last et al. (2013): “no impact” and maximum contaminant level (MCL) thresholds. The no-impact thresholds represent the lowest detectable concentrations above the background water chemistry that could be used to quantify a change in groundwater chemistry due to CO_2 or brine leakage, and were calculated as the 95%-confidence, 95%-coverage tolerance limit from data sets specific to each aquifer type. A key feature of the data presented in Table 5 is that the no-impact thresholds are much closer to the initial water chemistry in the carbonate aquifer case than the sands aquifer case. This may reflect differences in site-specific data used to define the initial model chemistry and the data used to estimate the no-impact thresholds. For the carbonate aquifer, the background thresholds were based on temporal data within or immediately adjacent to the model domain for the unconfined portion of

412 the Edwards Aquifer (Musgrove et al., 2010). For the confined alluvium aquifer, the background
413 thresholds were based on a 2010 USGS groundwater survey of 30 wells within the High Plains
414 Aquifer from an area outside of the lithologic model site. The high no-impact threshold for the
415 High Plains Aquifer reflects spatial and temporal variability sampled by the survey. It was
416 necessary to use these data because spatial and temporal data were not available from within the
417 model domain.

418 The MCL threshold refers to concentrations that exceed primary or secondary maximum
419 contaminant levels designated by the U.S. EPA (2009). Primary drinking water standards are for
420 trace metals (such as As, Cd, Cr, Cu, and Pb, among others) and are legally enforced for the
421 protection of public health by limiting the levels of contaminants in drinking water. Secondary
422 drinking water standards (which include standards for Fe, Mn, and Zn) are non-enforceable
423 guidelines regulating contaminants that may cause cosmetic or aesthetic effects in drinking
424 water.

425 **Table 5: Initial aquifer concentrations used in the simulations, no-impact and MCL (EPA, 2009) thresholds**
426 **reported in Last et al. (2013).**

Analyte	Unconfined Carbonate Aquifer (Edwards Acquifier)		Confined Unconsolidated Sands Aquifer (High Plains Acquifer)		U.S. EPA Regulatory Standard
	Initial Model	No-Impact Threshold[a]	Initial Model	No-Impact Threshold[a]	MCL Threshold[a]
pH	6.9	6.6	7.6	6.625[f]	6.5
Total Dissolved Solids	330	420 mg L^{-1}	570 mg L^{-1}	1300 mg L^{-1}[b]	500 mg L^{-1}[c]
Arsenic	0.31	0.55 µg L^{-1}	1.5 µg L^{-1}	9.3 µg L^{-1}	10 µg L^{-1}
Cadmium	0.00	0.04 µg L^{-1}	0.059 µg L^{-1}	0.25 µg L^{-1}	5 µg L^{-1}
Lead	0.06	0.15 µg L^{-1}	.086 µg L^{-1}	0.63 µg L^{-1}	15 µg L^{-1}

427 (a) 95%-confidence, 95%-coverage tolerance limit based on log values except for pH, which is already a log value.
428 (b) Threshold value exceeds regulatory standard, however using the regulatory standard may result in widespread
429 false positives under field conditions.
430 (c) Value is about 0.5 pH units lower than no-impact threshold estimated by Last et al. (2013) because ROMs at
431 higher threshold produced non-physical results.

432

3. CHANGES TO GROUNDWATER QUALITY

CO_2 and brine leakage into shallow aquifer resources can change the groundwater chemistry to values above the no-impact and MCL thresholds. pH and TDS plume distributions over the 200-year simulation period were, in large part, controlled by the distinct lithology of the respective aquifers, as is illustrated for single realizations for the High Plains and Edwards aquifers in Figures 6 and 7. This is especially true for the pH plume because it is tied to the transport of CO_2 gas in the aquifer systems through chemical solubility (Equations 1–3). The unconfined nature of the carbonate Edwards Aquifer allows buoyant CO_2 gas to transport vertically from the leakage source term to the atmosphere with some advection in the direction of groundwater flow. Once the plume reaches the water table, the flux rate of CO_2 across the water table rapidly reaches steady-state and matches the flux of CO_2 from the leaking wellbore (Figure 8). In contrast, variable lenses of permeable sands and impermeable shale, characteristic of the High Plains Aquifer, limit the vertical transport of CO_2 gas and yield plumes that are largely relegated to the lower permeable sand units within the aquifer, where only a small fraction of the CO_2 leaked into the aquifer is transported to the vadose zone above the water table (0.01 and 0.1%).

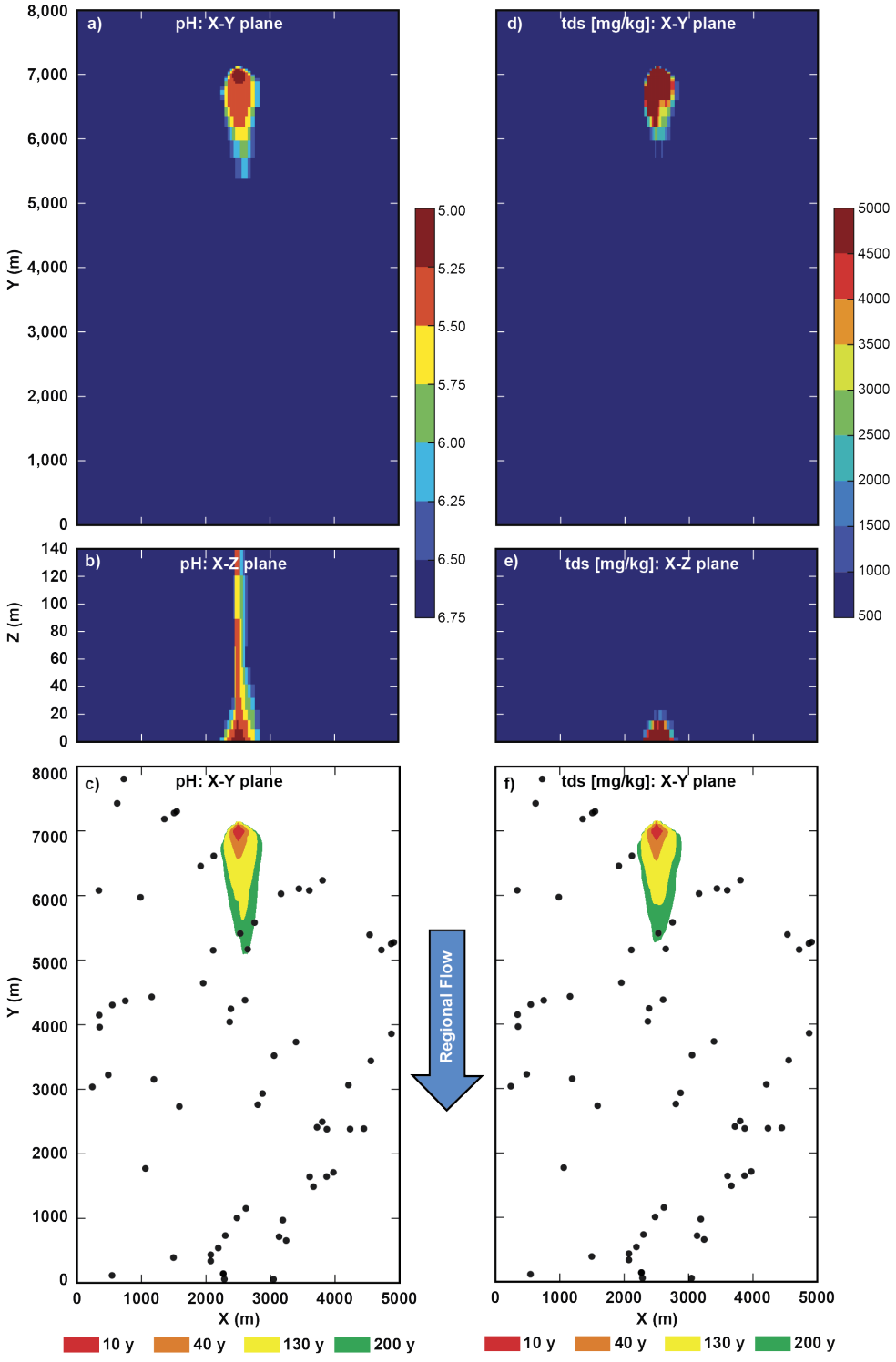


Figure 6: Color contour plots of pH and TDS at 200 years in plan view (XY) and cross section (XZ) (a,b,d,e) and the no-impact thresholds projected against shallow well locations (black dots) for a single simulation of the Edwards Aquifer (c,f). Groundwater flow is in the Y direction (North to South).

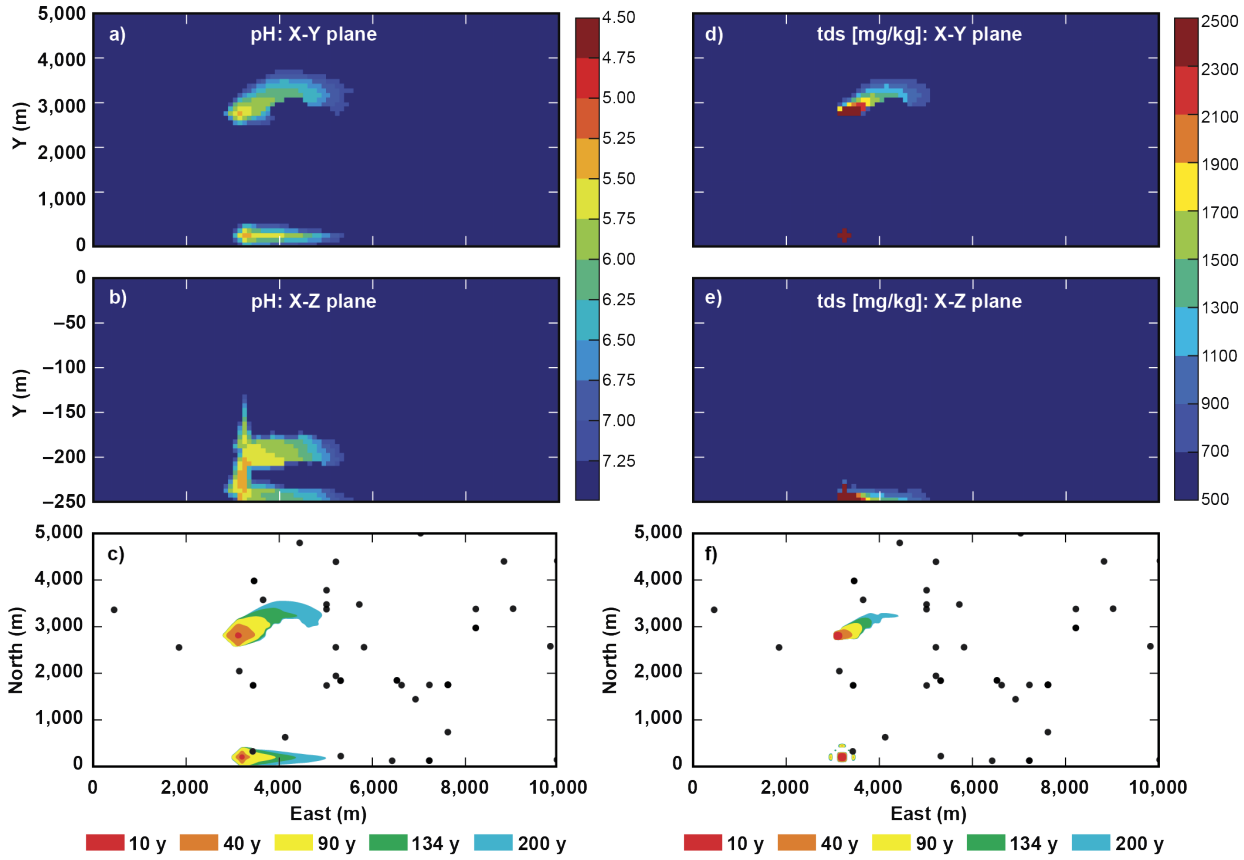


Figure 7: Color contour plots of pH and TDS at 200 years in plan view (XY) and cross section (XZ) (a,b,d,e) and the no-impact thresholds projected against shallow wellbore locations (black dots) for a single simulation of the High Plains Aquifer (c,f). Groundwater flow is in the X direction (East).

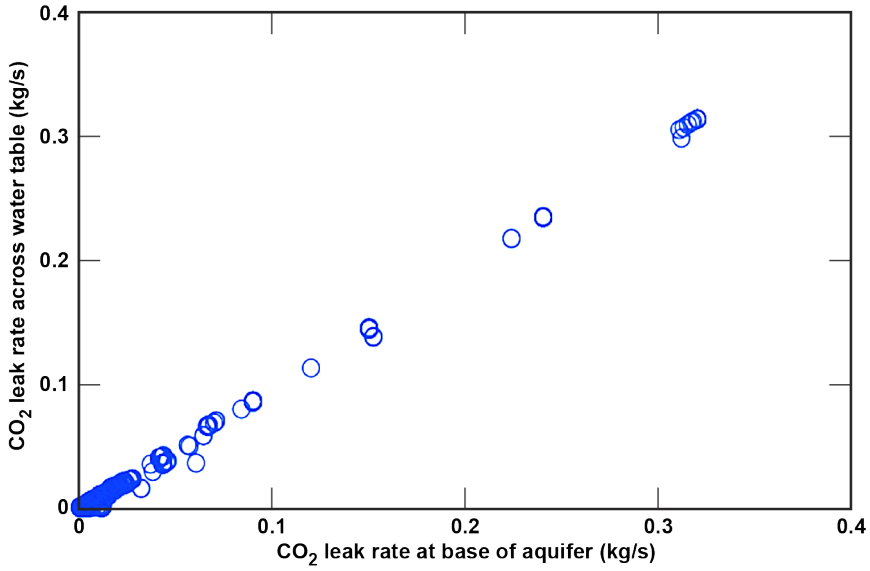


Figure 8: Correlation of CO_2 leakage rate in and out of the water table for the Edwards Aquifer.

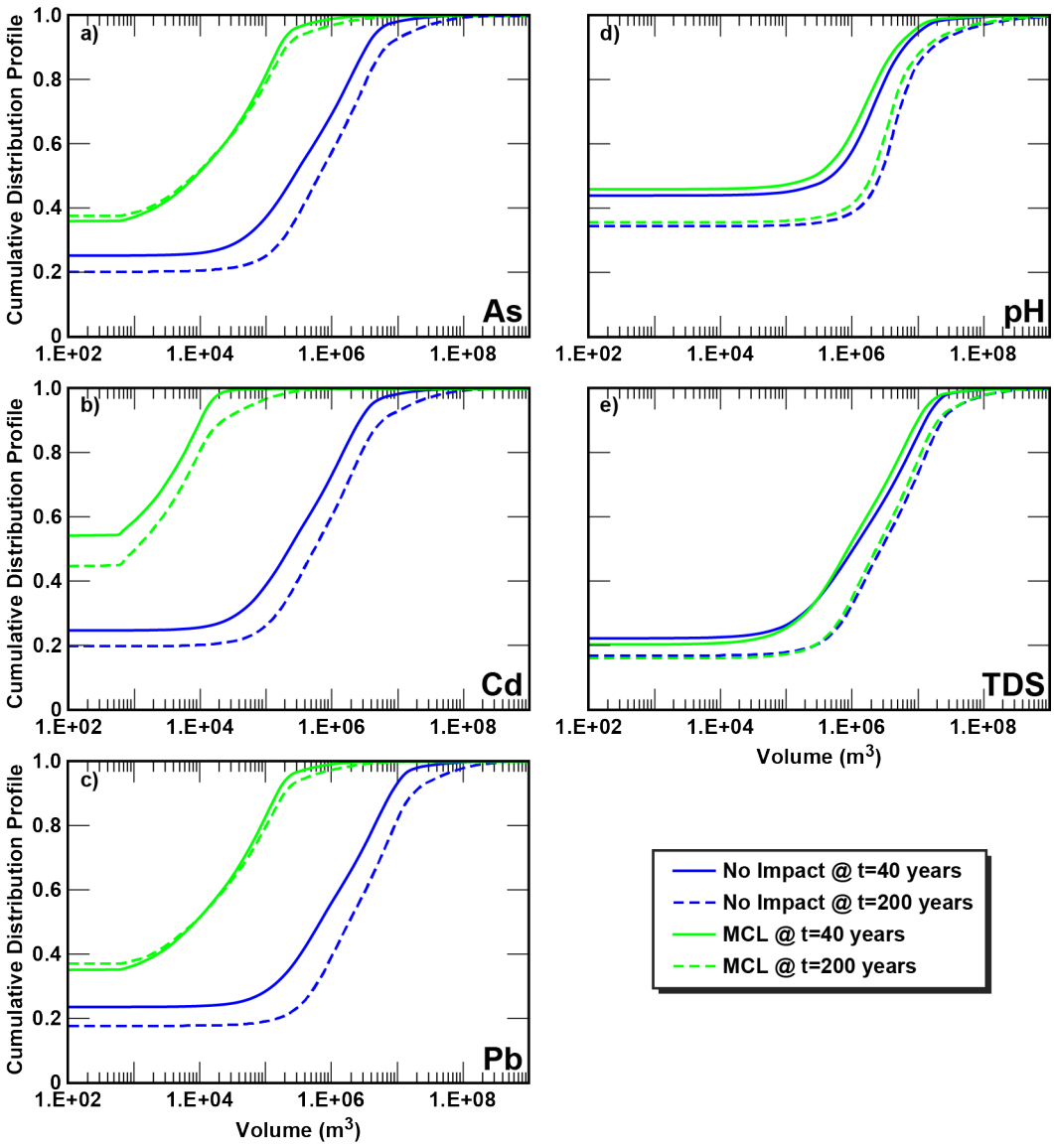


Figure 9: Unconfined carbonate (Edwards) aquifer: Comparison of plume volumes after 200 years of wellbore leakage for the MCL and no-impact thresholds for (a) As, (b) Cd, (c) Pb, (d) pH, and (e) TDS.

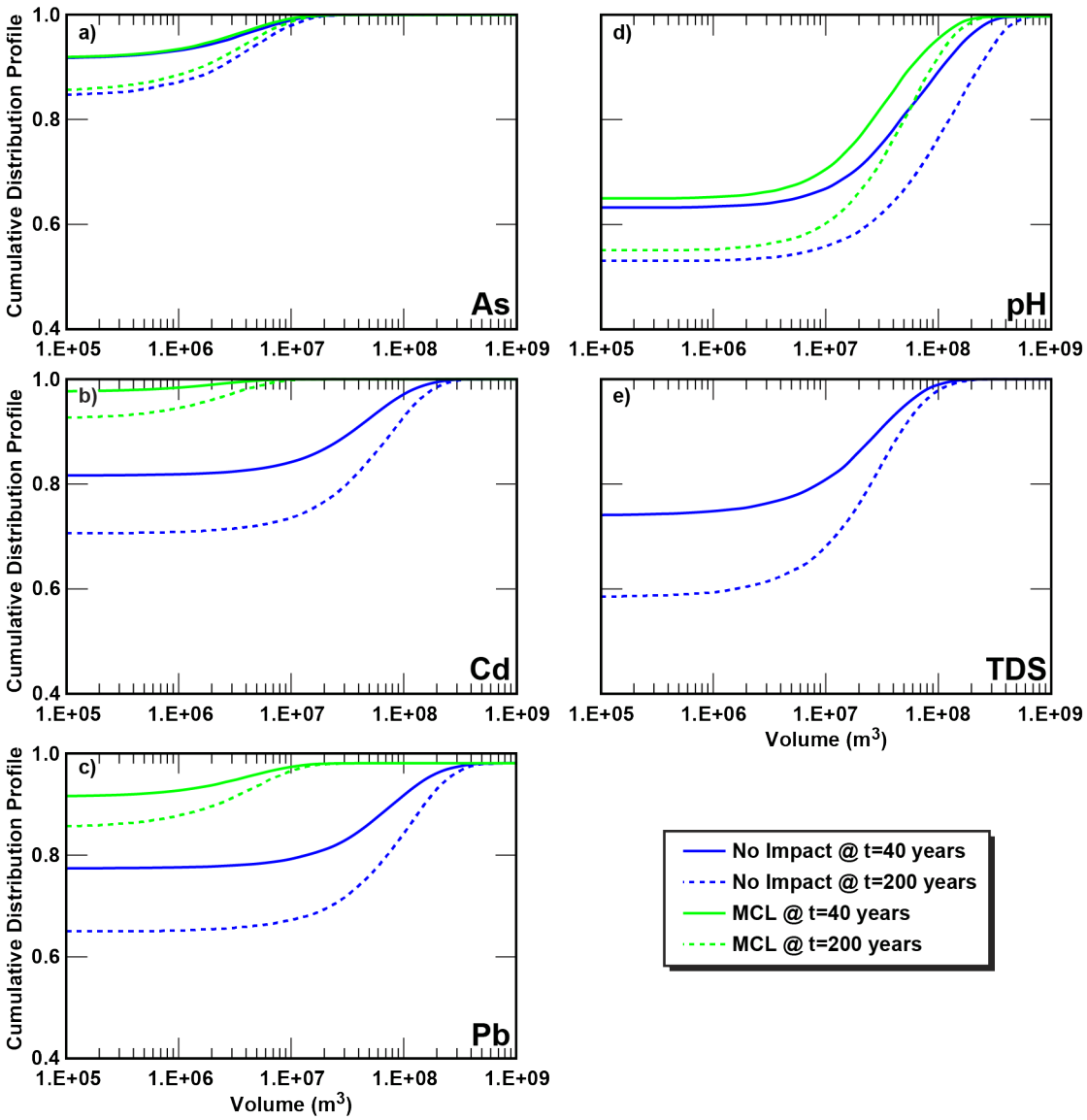


Figure 10: Confined alluvium (High Plains) aquifer: Comparison of plume volumes after 200 years of wellbore leakage for the MCL and no-impact thresholds for (a) As, (b) Cd, (c) Pb, (d) pH, and (e) TDS.

Images, such as those shown in Figures 6 and 7, highlight the role that aquifer characteristics play on the spatial and temporal distribution of groundwater plumes, but only for a single realization. In total, more than 500 and 700 detailed reactive transport simulations were performed to fully capture knowledge gaps and natural system variability for the Edwards and High Plains aquifers, respectively. The results measure our current understanding of the contribution of CO_2 and brine leakage on groundwater quality given the inherent uncertainty in the storage reservoir, leakage pathways, and dilute aquifer for these model systems. Figures 9 and 10 plot the cumulative distribution of emulated plume volumes at representative time intervals for CO_2 injection (40 years) and post injection (200 years) for pH and TDS thresholds for the unconfined carbonate (Edwards) and confined alluvium (High Plains) aquifers when exposed to the same leakage scenarios (MCL TDS for the alluvium aquifer was below the baseline value and is not plotted). The unconfined aquifer plot extends to smaller plume volumes (10^2 m^3) because of the smaller grid size and consequent ability to resolve smaller plumes.

We use the results and probability of occurrence ranges to forecast the likelihood that leakage will impact groundwater quality over 200 years (Figure 11). We use 10^5 m^3 as the lowest volume threshold so that the results of the two models can be compared, regardless of grid size. The probability of occurrence is shown against the no-impact and the MCL thresholds for each aquifer. There is a higher probability of exceeding the no-impact threshold for the unconfined carbonate aquifer than the alluvium aquifer because the thresholds in these examples are much lower and closer to the initial model chemistry for the Edwards aquifer than for the High Plains aquifer. Leakage is *likely* to result in a statistically significant change of the trace metal concentrations pH and TDS for the Edwards example. Whereas, leakage is only *likely* to cause a statistically significant change to groundwater pH for the High Plains example; changes in TDS and Pb have an *even chance of occurring*, and changes in Cd and As concentrations are *unlikely*. Forecasts of groundwater quality measured against no-impact thresholds are site specific and cannot be transferred to similar aquifer sites, because the site threshold depends on spatial and temporal variability as well as the absolute concentration.

Groundwater impacts to unconfined carbonate and confined alluvium aquifers are comparable when measured against MCL thresholds, with leakage *likely* to change pH and TDS concentrations above the thresholds. Of importance for this MCL-based metric is that probability of occurrence for trace metal impacts is *unlikely* to *extremely unlikely* to occur. Furthermore, detailed geochemical modeling of the aquifers that included inputs from both the reservoir and USDW showed trace metal concentrations could be reduced by uptake onto the aquifer sediments (Bianchi et al., 2013; Bacon, 2013).

The likelihood ranges are useful, because they forecast the gross performance of the storage system, but they do not convey information on the size or the evolution of the plume with time. We refer the reader back to the emulated volumes shown in Figures 9 and 10 to discuss the time scale of groundwater impacts if leakage were to occur. Recall that all emulated groundwater plumes result from the injection of 5 million tons of CO_2 per year for 50 years in which leakage is allowed to occur in up to 10% of the wells with variable permeability (10^{-14} to 10^{-10} m^2) with no option to mitigate the leak if it were detected. The pH plumes continue to increase because of buoyancy driven CO_2 transport and because smaller more acidic pH plumes are diluted through natural recharge and dispersion towards the more neutral thresholds, as are the TDS plumes. The emulations show a 10-fold increase in plume volume between the injection and post injection periods, on average, from 40 to 200 years. Because impacts to shallow groundwater chemistry,

as measured changes in pH and TDS above pre-injection values, can be sustained for long periods of time, it is important to detect and mitigate leakage sources as early as possible.

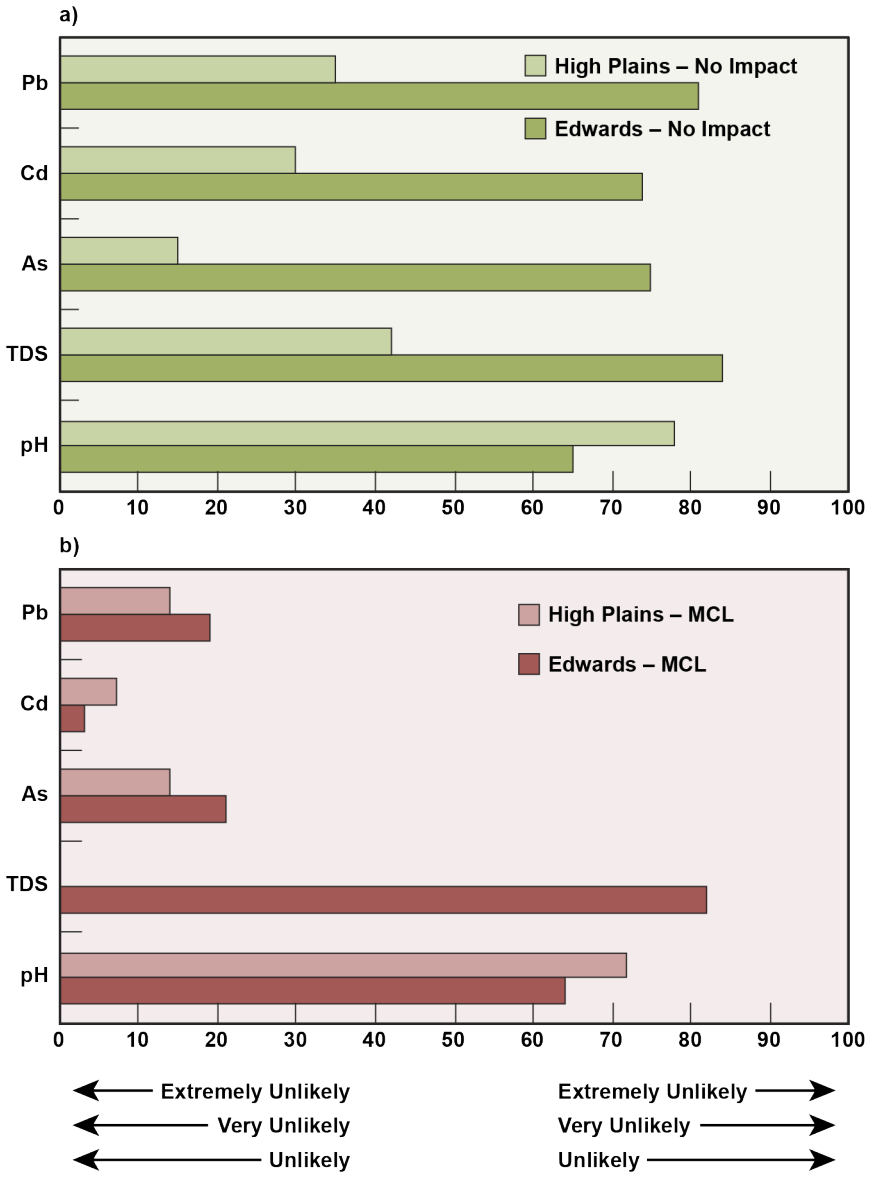


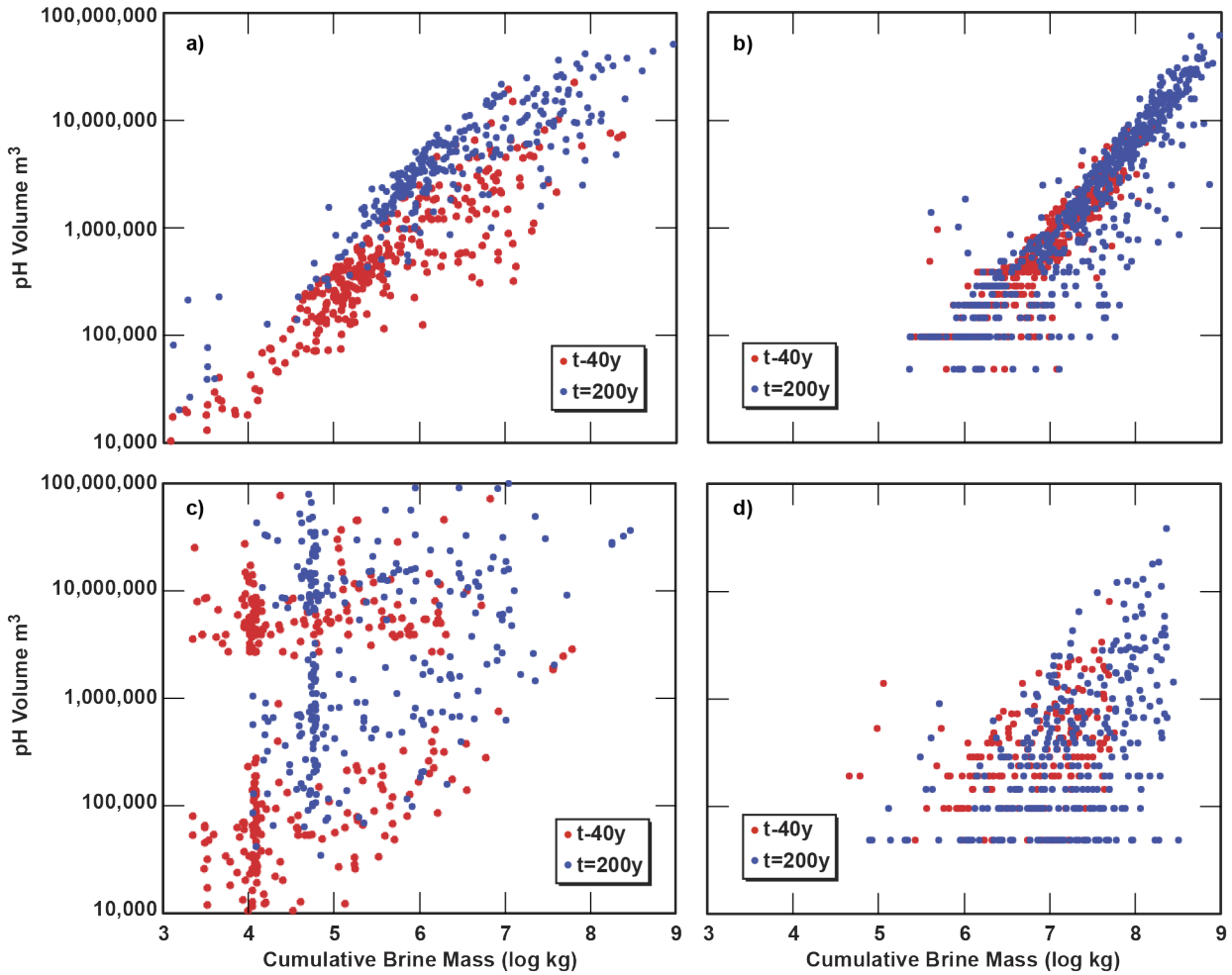
Figure 11: Comparison of the probability of occurrence of emulated plumes for the leakage scenarios investigated for the High Plains alluvium and Edwards carbonate aquifers that exceed the no-impact (a) and MCL (b) thresholds for pH, TDS, As, Cd, and Pb for volumes greater than 10^5 m^3 against the likelihood ranges for expressing the probability of occurrence.

Figure 12 plots pH and TDS plume volumes for no-impact thresholds against the cumulative mass of CO_2 and brine leaked into the unconfined carbonate and confined alluvium aquifers at 40 and 200 years after the initiation of CO_2 injection. Plume volume is largely dependent on the mass of CO_2 or brine that leaks into the aquifer (where TDS concentration is also important). Up to one million tons of CO_2 and brine leaked into the aquifers and produced plume volumes as large as 100 million cubic meters ($10^8 m^3$). In both the Edwards and High Plains models, the amount of leaked CO_2 and brine comprised a very small fraction (≤ 0.4%) of the CO_2 injected

529 into the storage reservoir (250 million tons) and reservoir brine in the area of review even after
530 200 years of simulation.

531 We use the data in Figure 12 to estimate leakage bounds that do not result in a measureable
532 change in the groundwater composition (defining plume volume > $10^5 m^3$). The lower bound for
533 pH plume volumes is about 100–1,000 tons of CO_2 for the unconfined carbonate or confined
534 alluvium aquifer examples. Similar lower bounds on CO_2 leakage for the two different aquifers
535 can be explained by buffering capacity of carbonate minerals in both systems and by the nearly
536 identical no-impact thresholds for each system. There is a more marked difference for the lower
537 bounds for brine leakage for the two aquifers, largely because the no-impact thresholds differ
538 significantly. The simulated results indicate that leakage as small as 1–10 tons could result in a
539 measurable change in the carbonate aquifer with a $TDS_{no\ impact\ threshold}$ = 420 mg L^{-1}. Whereas the
540 lower bound for the alluvial aquifer was about 100–1,000 tons of brine because this particular
541 aquifer has a higher no-impact threshold ($TDS_{no\ impact\ threshold}$ = 1300 mg L^{-1}). Establishing a given
542 aquifer's leakage tolerance requires a thorough assessment of the pre-injection chemistry at the
543 site that accounts for variability of current land use practices. In some cases, a high-density of
544 data may be available within the model domain, as was the case for the Edwards aquifer used as
545 the basis for the unconfined aquifer in this study. However this was not the case for the High
546 Plains aquifer where the no-impact threshold was based on data collected over a very large
547 region and consequently sampled greater variability.

548



549

550 **Figure 12: pH (a, b) and TDS (c, d) no-impact plume volumes plotted versus cumulative mass of CO_2 and**
551 **brine leaked into the unconfined carbonate aquifer (a, c) and the confined alluvium aquifer (b, d), where red**
552 **and blue symbols indicated plume volumes 40 and 200 years after CO_2 injection has started.**

553

4. DETECTION OF GROUNDWATER PLUMES

Figures 9, 10, and 12 suggest that relatively small amounts of CO_2 and brine leaked from the storage reservoir can result in a measureable change to the shallow aquifer chemistry. Despite the nominally large range of plume volumes, our computations show the probability of detecting the plumes using the available shallow water infrastructure is *extremely* to *very unlikely* over the entire 200-year period (Table 6). This determination was made using the simulated results because it allowed individual leakage source points to be compared with individual shallow groundwater receptors. Our conceptual model contains 165 deep wells, of which 49 serve as potential leakage source terms because they penetrate the area of review. Actual shallow groundwater receptors consist of 128 drinking wells in the carbonate aquifer; and 48 drinking, agricultural, and industrial wells for the confined alluvial aquifer. Receptor density is about 2.6 wells per km for the carbonate aquifer's model domain and 1 well per km for the alluvial aquifer's model domain. The analysis assumes that shallow wells are screened from the top to the bottom of the aquifer could detect a plume at any depth.

Table 6: Percent probability that any of the shallow aquifer wells will contain groundwater above the no-impact thresholds over the 200-year period. Actual shallow well locations are specific to the Edwards and High Plains areas shown in Figure 5 and 6. Deep well locations are the same for both systems.

Aqueous Component	Unconfined Carbonate Aquifer (Edwards)	Confined Alluvium Aquifer (High Plains)
pH	4.3%	9.6%
Total Dissolved Solids	3.8%	5.0%
Arsenic	1.7%	1.7%
Cadmium	1.4%	0.6%
Lead	2.8%	0.7%

Clearly, to increase the likelihood of detecting changes in groundwater chemistry a much higher density of shallow wells would be needed. Moreover, groundwater sampling is unlikely to be a reliable early leak detection strategy. Alternatives should be considered, including geophysical techniques such as electrical resistance (ER) data that samples regions in between monitoring wells using surface arrays of electrodes. Trainor-Guitton et al. (2013) computed an ER sensitivity index for a suite of groundwater simulations to assess ER's sensitivity to plume and non-plume results, where the sensitivity index is a mean log ratio of the electrical response at two different times. In other words, the ratio is the electrical response at a time after CO_2 injection scaled by the electrical response at time = 0 (before CO_2 injection). In general, Trainor-Guitton et al. found that the sensitivity of electrical resistivity depends on both the aspect ratio (the plume's dimension versus depth) and plume's TDS concentration. The study demonstrates the trade-off introduced when using a geophysical technique: it provides better areal coverage (between wells) without the expense of drilling boreholes, but there is a possibility of "false negatives" or "false positives" of plume occurrence because the groundwater is not sampled directly. This is demonstrated in Figure 13, where plume length (plume defined as TDS $\geq$ 1500 mg L^{-1}) is plotted versus the ER sensitivity index. Ambiguity in the remote sensing data exists because samples with no plume (green) yielded a positive index. Both plume and non-plume simulations can produce the same ER sensitivity as seen in the area between the dashed lines corresponding to the 50^{th} and 95^{th} percentile of ER sensitivity for non-plumes. A reasonable

threshold for detectability would be for all plumes with ER sensitivity ≥ 0.0012 (to the right of the 95th percentile - cyan line) corresponding to plumes between 100 and 3,000 m in length. Although this technique may not be able to resolve the diffuse boundary defined by the no-impact threshold, it would be able to detect more concentrated brines within the plume. Once detected, monitoring wells could be drilled to target the plume and assess the ability to mitigate the leak and the need for corrective actions.

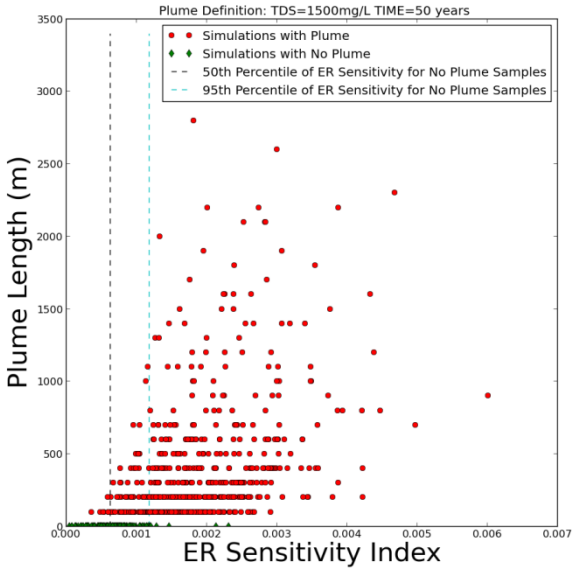


Figure 13: Plot of ER sensitivity index versus the plume length. Vertical lines represent the 50th and 95th percentile of the ER sensitivity index for simulations with non-plumes. ER will identify with high likelihood plumes that are ≥1,000 m in length (all samples to the right of the cyan vertical line). The samples between these two lines represent the most ambiguity in the ER signal.

Above-zone pressure measurements have also been suggested as an effective means to detecting leakage because pressure signals travel fast and can be collected continuously at relatively low cost. Leakage simulations into shallow groundwater suggest that leakage rates comparable to those studied here can lead to small changes in down-hole pressure (1–5 psi) and can be detected at wellbore spacing between 100 and 500 m away from the leaking well (Sun et al., 2013; Sun and Nicot, 2013). Continuous pressure testing of monitoring wells may provide early detection of leakage into shallow groundwater.

5. IMPLICATIONS FOR MONITORING

The U.S. EPA Class VI well permitting process for CO_2 storage requires that the area of review and risk to overlying USDW resources be assessed. Simulation and emulation studies that capture the storage reservoir, leakage pathways, and aquifer heterogeneity can be used to evaluate effective monitoring strategies of a potential storage site. Identification of possible leakage rates, coupled with predictions of plume volumes, can be used to identify potential monitoring and corrective action strategies should leakage from the storage reservoir occur. In the case of our two aquifer examples, the models assumed a fixed 50-year injection period, variable wellbore leakage pathways to either an unconfined carbonate aquifer or a confined alluvium aquifer, and that no corrective actions were made to the leaking wells for the 200-year simulation period. In the discussion that follows, we refer to results measured against the no-impact thresholds, as they represent the earliest point at which a detectable change in groundwater quality can be measured in the aquifer systems studied here.

The U.S. EPA has adopted a no net degradation policy for managing groundwater resources. Therefore, it is extremely important to establish a given USDW's baseline chemistry, as this baseline data can be used to develop no-impact threshold values for the site. The no-impact thresholds calculated as part of this study were demonstrated to be site specific. Key differences in the calculated values between the two sites were due to a combination of aquifer properties, as well as by the availability of existing spatial and temporal groundwater data on which the no-impact threshold was based. In the case of the Edwards aquifer, sufficient data were available from wells located within the model domain. However this was not the case for the High Plains aquifer, where the no-impact threshold values were based on data collected over a very large region. If no-impact thresholds are used to define plumes, aquifers with substantial temporal and spatial variability in water quality will have smaller plumes that will be more difficult to detect. Despite vertical transport of CO_2 out of the unconfined carbonate aquifer, the probability of impact to groundwater quality is higher than for the unconfined carbonate aquifer because the pre-injection chemistry is lower and the natural variability is smaller.

Although CO_2 and brine leakage are *likely* to drive pH below and increase TDS above the no-impact thresholds for both aquifers evaluated, the size of the plumes is small relative to spacing of the current network of wells in both the unconfined (2.6 wells/km^2) and confined (1 well/km^2) aquifers. There is a very low probability that the plumes would intersect USDW wells in the two study areas and in other areas with similar receptor density, based on our initial simulations and current understanding of parameters for both shallow aquifer systems. This result points to the need to test and develop spatially diverse, yet robust, monitoring techniques capable of detecting leakage early, which can be used to add confidence to data generated through typical groundwater assessments.

Some period of post-injection site care is required. Our simulations predict that even small amounts of CO_2 and brine, when left unmitigated, can change USDW pH and TDS concentrations for long-periods of time. The difficulty is deciding the time period, because it could take many years to directly observe the impacted waters in monitoring wells. The focus of this study was on using simulations to predict potential impacts within shallow USDWs, not on identifying methods for early leakage detection. Future efforts will focus on understanding how early detection and mitigation of leaks impacts plume volume and the time required for the aquifer chemistry to rebound to pre-leakage conditions.

6. REFERENCES

Bachu, S. Legal and regulatory challenges in the implementation of CO2 geological storage: An Alberta and Canadian perspective. *International J. Greenhouse Gas Control* **2008,** *2*, 259–273. doi: 10.1016/j.ijggc.2007.12.003

Bernabe, Y.; Mok, U.; Evans, B. Permeability-porosity relationships in rock subjected to various evolution processes. *Pure and Applied Geophysics* **2003,** *160*, 937–960.

Carle, S. F. *T-PROG: Transitional probability geostatistical software users' guide*; University of California: Davis, CA, 1999.

Carroll, S.A.; Hao, Y.; Aines, R. Geochemical detection of carbon dioxide in dilute aquifers. *Geochemical Transactions* **2009,** *10*.

Dai, Z.; Wolfsberg, A.; Lu, Z.; Ritzi, R. Representing Aquifer Architecture in Macrodispersivity Models with an Analytical Solution of the Transition Probability Matrix. *Geophysical Research Letters* **2007,** *34*, L20406. doi:10.1029/2007GL031608

Deng, H.; Stauffer, P. H.; Dai, Z.; Jiao, Z.; Surdam, R. C.; Simulation of industrial-scale CO_2 storage: Multi-scale heterogeneity and its impacts on storage capacity, injectivity and leakage. *Int. J. Greenhouse Gas Control* **2012,** *10*, 397–418.

Deutsch, C. V.; Journel, A. G. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford Univ. Press: New York, 1992, p 340.

U.S. Department of Energy. *DOE Carbon Sequestration Technology Roadmap and Program Plan*; U.S. Department of Energy and National Energy Technology Laboratory: Washington DC, 2007.

Gutentag, E.; Heimes, F.; Krothe, N.; Luckey, R.; Weeks, J. *Geohydrology of the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming*; Technical Report Professional Paper, 1400–B; U.S. Geological Survey, 1984.

Hao, Y.; Sun, Y.; Nitao, J. Overview of NUFT: a versatile numerical model for simulating flow and reactive transport in porous media. In *Groundwater Reactive Transport Models;* Zhang et al., Eds.; Bentham Science Publishers, 2012; p 213–240.

Harp, D.; Dai, Z.; Wolfsberg, A.; Vrugt, J.; Robinson, B.; Vesselinov, V. Aquifer structure identification using stochastic inversion. *Geophysical Research Letters* **2008,** *35*, L08404. doi:10.1029/2008GL033585

Herzog, H.; Calderia, K.; Reilly, J. An issue of permanence: assessing the effectiveness of temporary carbon storage. *Climate Change* **2003,** *59*, 293–310. doi:10.1023/A:1024801618900

Intergovernmental Panel on Climate Change (IPCC); Cambridge University Press, 2005.

Jordan, A. B.; Stauffer, P. H.; Harp, D. R.; Carey, J. W.; Pawar, R. J. *A Method for Predicting CO_2 and Brine Leakage from Geologic Sequestration along Cemented Wellbores in System-Level Models*. Los Alamos National Laboratory Rep. LA-UR-13-29243; 2013

Kansas Geological Survey (KGS), Wells, logs, core, and other databases. http://www.kgs.ku.edu/PRS/petroDB.html (accessed on 9/30/2011).

693 Karamalidis, A. K.; Torres, S. G.; Hakala, A. J.; Shao, H.; Cantrell, K. J.; Carroll, S. A. Trace
694 Metal Source terms in carbon sequestration environments. *Environ. Sci. Technol.* **2013,**
695 *47*, 322–329. DOI: 10.1021/es304832m

696 Last, G. V.; Brown, C. F.; Murray, C. J.; Bacon, D. H.; Qafoku, M. P.; Sharma, M.; Jordan, P. D.
697 *Determining threshold values for identification of contamination predicted by reduced*
698 *order models*; Pacific Northwest National Laboratory, 2013.

699 Lewicki JL, Birkholzer J, Tsang C-F: Natural andindustrial ana- logues for leakage of CO_2 from
700 storage reservoirs: identification of features, events, and processes and lessons learned.
701 *Environ Geol* 2007, **52:**457-467.

702 Lindgren, R. J.; Dutton, A. R.; Hovorka, S. D.; Worthington, S. R.; Painter, S. Conceptulization
703 and Simulation of the Edwards Aquifer, San Antonio Region, Texas; Scientific
704 Investigation Report 2004-5277; U.S. Geological Survey, 2004; p 143.

705 Lindgren, R. J. Diffuse-Flow Conceptulization and Simulation of the Edwards Aquifer, San
706 Antonio Region, Texas; Scientific Investigation Report 2006-5319; U.S. Geological
707 Survey, 2006; p 48.

708 Musgrove, M.; Stern, L. A.; Banner, J. L. Spring water geochemistry at Honey Creek State
709 Natural Area, Central Texas: implications for surface water and groundwater interaction
710 in a karst aquifer. *Journal of Hydrology* **2010,** *388*, 144–156.

711 National Energy Technology Laboratory (NETL). The United States 2012 Carbon Utilization
712 and Storage Atlas, 4th Ed., 2012. www.netl.doe.gov

713 National Carbon Sequestration Database and Geographic Information System (NATCARB).
714 www.natcarb.org (accessed Nov 2013).

715 National Risk Assessment Partnership (NRAP), 2012. www.netl.doe.gov/nrap (accessed Sept
716 2013).

717 Nitao, J. J. *User's manual for the USNT module of the NUFT code, version 2 (NP-phase, NC-*
718 *component, thermal)*; UCRL-MA-130653; Lawrence Livermore National Laboratory,
719 1998.

720 Painter, S. L.; Woodbury, A. D.; Jiang, Y. Transmissivity estimation for highly heterogeneous
721 aquifers: comparison of three methods applied to the Edwards Aquifer, Texas, USA.
722 *Hydrogeol. J.* **2007,** *15*, 315–331.

723 Schaap, M.; Leij, F.; van Genuchten, M. Rosetta: A computer program for estimating soil
724 hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* **2001,** *251*, 163–
725 176.

726 Sun, A. Y.; Zeidouni, M.; Nicot, J.-P.; Lu, Z.; Zhnag, D. Assessing leakage detectability at
727 geologic CO_2 sequestration sites using the probabilistic collocation method. *Advances in*
728 *Water Resources* **2013,** *56*, 49–60. dx.doi.org/10.1016/j.advwatres.2012.11.017.

729 Sun, A. Y.; Nicot, J.-P. Inversion of pressure anomaly data for detecting leakage at geologic
730 carbon sequestration sites. *Advances in Water Resources* **2013,** *44*, 20–29.
731 Dx.doi.org/10.1016/j.advwatres.2012.04.006.

732 Tong, C. *PSUADE User's Manual*; LLNL-SM-407882; Lawrence Livermore National
733 Laboratory, 2005.

734 Tong, C. Self–validated variance–based methods for sensitivity analysis of model outputs.
735 *Reliab. Eng. Syst. Safe.* **2010,** *95*, 301–309.

736 Tainor-Guitton, W. J.; Ramirez, A.; Yang, X.; Mansoor, K.; Sun, Y.; Carroll, S. A. Value of
737 information methodology for assessing the ability of electrical resistivity to detect
738 CO_2/brine leakage into a shallow aquifer. *Int. J. Greenh. Gas Control* **2013,** *10*, 101–113.

739 U.S. EPA. Code of Federal Regulations, 40 CFR Part 141 and 143, EPA 816-F-09-0004, 2009.

740 U.S. EPA. http://water.epa.gov/type/groundwater/uic/wells_sequestration.cfm (accessed Sept
741 2013).

742 U.S. Geological Survey (USGS). *National water–quality assessment (nawqa) program: High*
743 *plains regional ground water study*, 2011.
744 http://co.water.usgs.gov/nawqa/hpgw/HPGW_home.html accessed on 9/1/2011.

745 Wainwright, H.; Finsterle, S.; Zhou, Q.; Birkholzer, J. *Modeling the Performance of Large-Scale*
746 *CO_2 Storage Systems: A Comparison of Different Sensitivity Analysis Methods*; NRAP-
747 TRS-III-002-2012; NRAP Technical Report Series; U.S. Department of Energy, National
748 Energy Technology Laboratory: Morgantown, WV, 2012; p 24.

749 Watson, T.; Bachu, S. Evaluation of the potential for gas and CO_2 leakage along wellbores. *SPE*
750 *Drilling & Completion* **2009,** *24*, 115–126.

751 Wilson EJ, Friedmann SJ, Pollak MF: Research for deployment: Incorporating risk, regulation,
752 and liability for carbon cap- ture and sequestration. *Environ Sci Technol* 2007, **41:**5945-
753 5952.

754 Zhang, K.; Wu, Y.S.; Pruess, K. *User's Guide for TOUGH2-MP – A massively parallel version*
755 *of the TOUGH2 code*; Report LBNL-315E; Lawrence Berkeley National Laboratory, Berkeley,
756 CA, 2008.

757 Zhou, Q.; Birkholzer, J.T. On scale and magnitude of pressure build-up induced by large-scale
758 geologic storage of CO2. *Greenhouse Gases: Science and Technology* **2011,** *1*, 11-20

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

762

ACKNOWLEDGEMENTS

This work was completed as part of National Risk Assessment Partnership (NRAP) project. Support for this project came from the DOE Office of Fossil Energy's Crosscutting Research program. The authors wish to acknowledge Robert Romanosky (NETL Strategic Center for Coal) and Regis Conrad (DOE Office of Fossil Energy) for programmatic guidance, direction, and support.

This work has been reviewed by Lawrence Livermore National Laboratory (LLNL-TR-648252), Los Alamos National Laboratory (LA-UR-14-20115), and Pacific Northwest National Laboratory (PNNL-23073) and approved for public release.

The authors are grateful for the technical reviews provided by Kirk Cantrell (PNNL). Carroll, Mansoor, and Sun also acknowledge Lawrence Livermore National Laboratories high performance computing facility that enabled physics-based simulations of the High Plains Aquifer.

Disclaimer