

Carnegie Mellon University

CARNEGIE INSTITUTE OF TECHNOLOGY

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF Doctor of Philosophy

TITLE Statistical Approaches to Quantifying Uncertainty of
Monitoring and Performance of Geologic CO₂ Storage Sites
on CO₂ Transport Through Fractured Caprocks

PRESENTED BY Nicholas A. Azzolina

ACCEPTED BY THE DEPARTMENTS OF

Civil and Environmental Engineering

Mitchell J. Small April 30, 2015
CO-ADVISOR, MAJOR PROFESSOR DATE

David Nakles April 30, 2015
CO-ADVISOR, MAJOR PROFESSOR DATE

David A. Dzombak May 1, 2015
DEPARTMENT HEAD DATE

APPROVED BY THE COLLEGE COUNCIL

Vijayakumar Bhagavatula May 5, 2015
DEAN DATE

**Statistical Approaches to Quantifying Uncertainty of
Monitoring and Performance at Geologic CO₂ Storage Sites**

Nicholas A. Azzolina

A.B. in Geology, Princeton University, Princeton, New Jersey
M.S. in Hydrogeology, Syracuse University, Syracuse, New York

A DISSERTATION
SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS
FOR THE DEGREE OF

Doctor of Philosophy

in

Environmental Management and Science

Carnegie Mellon University

Pittsburgh, PA

May 2015

Copyright © 2015

THESIS COMMITTEE

Professor Mitchell J. Small, Advisor

Departments of Civil and Environmental Engineering & Engineering and Public Policy

Carnegie Mellon University

Professor David V. Nakles, Co-Advisor

Department of Civil and Environmental Engineering

Carnegie Mellon University

Professor David A. Dzombak

Department of Civil and Environmental Engineering

Carnegie Mellon University

Dr. Grant S. Bromhal

United States Department of Energy

National Energy Technology Laboratory

To Julie, Sophia, Dominick and Luka

ABSTRACT

Geologic carbon dioxide (CO₂) storage is one approach for mitigating concentrations of CO₂ in the atmosphere that are caused by stationary anthropogenic inputs. Injecting CO₂ into the subsurface for long-term storage is an “engineered-natural system”. This engineered-natural system is complex, with potential interactions during CO₂ injection between CO₂ and other reservoir fluids and various components of the geologic system. The National Risk Assessment Partnership (NRAP) is an initiative within DOE’s Office of Fossil Energy that is improving the fundamental understanding of the complex science behind engineered-natural systems and is developing the risk assessment tools that are needed for safe, permanent geologic CO₂ storage. The NRAP technical approach entails an iterative modeling process that integrates component models into a system model which may then be used to provide quantitative assessments of potential risks and to design monitoring protocols that will effectively monitor risks at a geologic CO₂ storage project. A theme throughout all phases of the NRAP approach is *quantifying uncertainty and variability*. The focus of this dissertation is to contribute statistical methods and/or approaches for quantifying uncertainty and variability with respect to both monitoring and performance at geologic CO₂ storage sites. These methods are intended for future use by NRAP or other geologic CO₂ storage practitioners and may be incorporated into broader modeling approaches. However, the results and contributions from this work extend beyond geologic CO₂ storage and apply to other subsurface engineered-natural systems.

This dissertation examines three topic areas which have been organized into three separate chapters: (Chapter 2) evaluating the effectiveness of deep subsurface pressure

monitoring for brine leakage detection; (Chapter 3) quantifying the benefit of wellbore leakage potential estimates for prioritizing long-term monitoring; and (Chapter 4) quantifying CO₂ storage associated with CO₂ enhanced oil recovery. Key results for each of these topic areas and new contributions to the state-of-the-science are summarized below.

Evaluating the effectiveness of deep subsurface pressure monitoring for brine leakage detection. The work in Chapter 2 highlights the importance of uncertainty in the design of a reliable pressure monitoring network for geologic CO₂ storage sites. Subsurface pressure monitoring is attractive within the geologic CO₂ storage system, because the pressure pulse from CO₂ injection will cover a significantly broader portion of the system than the CO₂ plume. While several authors have investigated the use of pressure monitoring at geologic CO₂ storage sites and the sensitivity of pressure output to model inputs, uncertainty analysis has not been extensively conducted to assess detection sensitivity of pressure monitoring throughout the reservoir system as a function of space and time. Uncertainty analysis adds important information beyond sensitivity analysis, and helps define the probability at which monitoring during and post-injection will differentiate leakage of a given size from the background variability. The results suggest that measurements of the absolute change in pressure within the target injection aquifer would not be able to distinguish small leakage rates from baseline conditions, and that only large leakage rates would be discriminated with sufficient statistical power. Combining measurements, for example by taking the ratio of formation pressure in the aquifer above the cap rock to the target injection aquifer, provides better statistical power for distinguishing smaller leakage rates at earlier times in the injection program. Thus one

important contribution of this work is the idea that an optimized monitoring system which integrates multiple measurements could provide improved leakage detection sensitivity. In addition, the work in Chapter 2 shows that detection sensitivity for pressure is a function of lateral distance from the injection wells (x), vertical position within the storage system (z), and time (t), or $\sim f(x, z, t)$. Therefore, design of an adequate monitoring network for subsurface pressure should account for this time-space variability to ensure that the monitoring system performs to the necessary design criteria. There is unlikely to be a “one-size fits all” monitoring program for geologic CO₂ storage sites. However, a critical insight from the work in Chapter 2 is that the final monitoring program must be adaptive and dynamic, changing over the project life-cycle to adjust to changing subsurface conditions in response to CO₂ injection.

Quantifying the benefit of wellbore leakage potential estimates for prioritizing long-term monitoring. The work in Chapter 3 demonstrates that accurate prior knowledge about the probability of well leakage adds measurable value to the ability to detect a leaking well during the monitoring program, and that the loss in detection ability due to imperfect knowledge of the leakage probability can be quantified. The CO₂ injection phase for geologic CO₂ storage is likely to span decades and the post-injection site care would extend approximately 50 additional years beyond that timeline. Thus, sampling all wells within the Area of Review (AoR) at fixed frequency (e.g., annually) may be cost-prohibitive, especially for sites with many legacy wells. Optimizing a well sampling program using prior knowledge about the probability of well leakage (e.g., pre-injection well integrity surveys) would yield a more cost-effective monitoring program. Power curves were developed as a function of (1) the number of

leaking wells within the Area of Review; (2) the sampling design (random or judgmental, choosing first the wells with the highest deep leakage potential scores); (3) the number of wells included in the monitoring sampling plan; and (4) the relationship between a well's leakage potential score and its relative probability of leakage. Cases where the deep well leakage potential scores are fully or partially informative of the relative leakage probability are compared to a non-informative base case in which leakage is equiprobable across all wells in the Area of Review. The solution approach uses a stepwise procedure that was implemented using open-source (free) software; therefore, the process is reproducible and easily transferable to other modeling and simulation platforms. The results from Chapter 3 provide a practical example of how a field operator might incorporate prior knowledge about the probability of well leakage into their sampling plan design and thus quantitatively evaluate different plans to satisfy the site-specific monitoring objectives.

Quantifying CO₂ storage associated with CO₂ enhanced oil recovery. The work in Chapter 4 analyzes a database of 31 existing CO₂ enhanced oil recovery (EOR) projects to better understand the CO₂ retention, incremental oil recovery, and net CO₂ utilization of these oil fields. The results of this study showed that across all 31 sites, the 10th, 50th (median), and 90th percentile values for the three factors at 300% hydrocarbon pore volume (HCPV) were: CO₂ retention: 23.1%, 48.3%, and 61.8% retained; incremental oil recovery: 5.3%, 12.2%, and 21.5% OOIP (original oil in place); and net CO₂ utilization: 4.8, 8.7, and 10.5 Mscf/STB (stock-tank barrel). These results can be used to estimate the potential range of expected performance for similar candidate oil fields that are not currently undergoing CO₂ injection, including estimates of the

associated CO₂ storage potential of these candidate fields. In addition, the work in Chapter 4 employs a novel approach that incorporates nonlinear functions to quantify uncertainty of the CO₂ retention, incremental oil recovery, and net CO₂ utilization response with a handful of parameters, providing insight into the behavior of the reservoir across the entire timeline of the CO₂ flood. This work showed that a four-parameter log-logistic function was able to accurately describe the overall shape of the CO₂ retention and incremental oil recovery curves across the 31 sites, and that a two-parameter simple asymptotic function was able to accurately describe the overall shape of the net CO₂ utilization curves across the 16 sites for which net CO₂ utilization values could be calculated. These functions provide a practical approach for operators to summarize key performance metrics of their CO₂ EOR field data and forecast the performance into the future. In addition, these functions also provide useful screening tools for broad regional assessments of the CO₂ storage resource associated with candidate CO₂ EOR fields. Lastly, these analytical approaches lend themselves to Monte Carlo simulation and therefore rapid uncertainty quantification as compared to complex three-dimensional numerical simulation models.

ACKNOWLEDGEMENTS

I would like to thank my academic advisor Prof. Mitchell J. Small for his guidance, patience with an *In Absentia* ABD student, and insightful direction along the way. Never have I met someone who in less than five minutes can put his finger on the critical assumptions, develop a mathematical solution entirely from memory and fundamentals, and then make me laugh with a joke. It was truly a pleasure working with you and I very much enjoyed our phone calls and face-to-face meetings over the years. Thank you for imparting onto me a deeper appreciation for statistics.

I would also like to thank my academic co-advisor Prof. David V. Nakles for encouraging me to pursue my Ph.D. and providing sage advice from time-to-time. Your tireless support through each of the stages of the Ph.D. is a major reason that I finished. I will never forget your kindness and generosity, and I cannot possibly put into words the level of respect and admiration that I have for you. Thank you.

I would also like to thank my thesis committee members, Prof. David A. Dzombak and Dr. Grant S. Bromhal, for serving on my committee and for their comments, guidance and suggestions. I feel privileged to have had such a distinguished thesis committee and I am very grateful for your time and input.

I have also had the great pleasure to work with one graduate student in particular who reviewed each of my papers, served as a sounding board for many ideas, and helped me to sharpen my statistical pencil. Thank you Mr. Clint Noack – your friendship is something that I truly value and one of the greatest gifts from my time at Carnegie Mellon University.

This work was funded by the United States Department of Energy National Energy Technology Laboratory (NETL). I would like to express my deep gratitude for NETL's investment in my research and education. Your financial support made this degree possible.

The Energy and Environmental Research Center (EERC) in Grand Forks, North Dakota, in particular John Harju, Ed Steadman and Charlie Gorecki, very generously shared data with me and helped to support my work through the United States Department of Energy National Energy Technology Laboratory under Award no. DE-FE0009114. The EERC is a major part of this dissertation and I am very thankful for the support that you have given to me over the past five years.

And last but certainly not least, none of this would have been possible had my wife, Julie, not believed in me and encouraged me to "go for it". Your love is a source of strength and I appreciate the many sacrifices you have made over the past five years so that I could pursue this degree. Without you by my side none of this would have happened. Our three kids – Sophia, Dominick and Luka – are very lucky to have such an extraordinary mom. And I am one lucky guy to have you as my wife.

Table of Contents

Chapter 1: Introduction	1
1.1 Geologic CO ₂ Storage to Mitigate Atmospheric Inputs of CO ₂	1
1.2 National Risk Assessment Partnership Approach.....	2
1.3 Research Objectives.....	3
1.4 Dissertation Overview	5
1.5 References.....	8
Chapter 2: Effectiveness of subsurface pressure monitoring for brine leakage detection in an uncertain CO₂ sequestration system.....	9
Abstract.....	9
2.1 Introduction.....	10
2.2 Methods.....	16
2.2.1 <i>Reduced-Order Model</i>	16
2.2.2 <i>First-Order Uncertainty Analysis</i>	20
2.2.3 <i>Statistical Power Curves</i>	28
2.3 Results.....	28
2.3.1 <i>Model Nominal Response</i>	28
2.3.2 <i>Uncertainty in Pressure Response</i>	33
2.3.3 <i>Power Curves</i>	49
2.4 Sensitivity to Model Assumptions	50
2.5 Conclusions.....	55
2.6 References.....	56
Chapter 3: Quantifying the benefit of wellbore leakage potential estimates for prioritizing long-term MVA well sampling at a CO₂ storage site	60
Abstract.....	60
3.1 Introduction.....	61

3.2	Methods.....	66
3.2.1	<i>Generic Geologic CO₂ Storage Site</i>	66
3.2.2	<i>Well Leakage Potential Scores</i>	66
3.2.3	<i>Well Leakage and Well Sampling Scenarios</i>	70
3.3	Results.....	83
3.3.1	<i>Statistical Power Curves</i>	83
3.3.2	<i>Sensitivity to Wellbore Leakage Potential Accuracy</i>	89
3.3.3	<i>Sensitivity to Partial Wellbore Leakage Potential Scores</i>	91
3.4	Implications for Long-Term Monitoring at CO ₂ Storage Sites	95
3.5	Implicit Assumptions of the Statistical Simulation Approach.....	97
3.5.1	<i>Monitoring as a Discrete Outcome</i>	98
3.5.2	<i>Time Independence</i>	99
3.5.3	<i>Emphasis on Buoyant CO₂</i>	102
3.5.4	<i>Geospatial Information</i>	103
3.5.5	<i>Accuracy of the Deep Well Leakage Potential Score</i>	104
3.6	Conclusions.....	104
3.7	References.....	106

Chapter 4: CO₂ storage associated with CO₂ enhanced oil recovery: A statistical analysis of historical operations..... 110

	Abstract.....	110
4.1	Introduction.....	111
4.2	Methods.....	116
4.2.1	<i>Industry Data Set and Parameters</i>	116
4.2.1	<i>Dimensionless Variables</i>	117
4.2.2	<i>CO₂ Retention</i>	120
4.2.3	<i>Net CO₂ Utilization Factors</i>	120
4.2.4	<i>Nonlinear Regression</i>	123
4.3	Results.....	129

4.3.1	<i>Patterns in CO₂ Retention, Incremental Oil Recovery Factor, and Net CO₂ Utilization</i>	129
4.3.2	<i>Nonlinear Curve Fits of CO₂ Retention</i>	130
4.3.3	<i>Nonlinear Curve Fits of Incremental Oil Recovery</i>	134
4.3.4	<i>Nonlinear Curve Fits of Net CO₂ Utilization</i>	137
4.3.5	<i>Uncertainty Quantification – Percentile Estimates from 0% to 300% HCPVI</i> 141	
4.4	Discussion	143
4.4.1	<i>Comparison to Other Published Studies</i>	143
4.4.2	<i>Predictions of CO₂ Storage</i>	145
4.4.3	<i>CO₂ Retention versus CO₂ Storage</i>	147
4.5	Conclusions	150
4.6	References	152
Chapter 5:	Conclusions and Future Research Directions	155
5.1	Conclusions	155
5.1.1	<i>Chapter 2: Effectiveness of subsurface pressure monitoring for brine leakage detection in an uncertain CO₂ sequestration system</i>	155
5.1.2	<i>Chapter 3: Quantifying the benefit of wellbore leakage potential estimates for prioritizing long-term MVA well sampling at a CO₂ storage site</i>	156
5.1.3	<i>Chapter 4: CO₂ storage associated with CO₂ enhanced oil recovery: A statistical analysis of historical operations</i>	157
5.2	New Contributions	158
5.2.1	<i>Integration of Analytical Model and First-Order Uncertainty Analysis to Answer Fundamental Questions</i>	158
5.2.2	<i>Simulation Approach for Integrating Wellbore Leakage Risk and Monitoring</i> 159	
5.2.3	<i>Novel Approach for Fitting CO₂ EOR Performance Metrics</i>	160
5.3	Recommendations for Future Research	161
5.4	References	165

Appendix A: LBNL Model Input File for the Baseline Model.....	167
Appendix B: Biased Urn R Code	174

List of Tables

Table 2-1. LBNL Model input parameters for the baseline model that were changed by $\pm dx$ through a series of 18 separate modeling runs as part of the FOUA process. In the leaky scenarios, HCONP was modified by a factor of 2, 3, 4, 5, 10, 20, 30, 40, 50, and 100 for each of the 18 modeling runs.....	27
Table 3-1. Well attributes and weighting scores used to derive the deep well leakage potential scores for 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada.....	68
Table 3-2. Matrix of six well leakage and well sampling scenarios included in this work.	72
Table 3-3. Combinations of well attributes for Scenarios 6a and 6b.....	81
Table 3-4. Input parameters used in the analytical model of Nordbotten et al. (2005) and the semi-analytical model of Vilarrasa et al. (2013) to estimate the radial extent of the CO ₂ plume during a 50-years CO ₂ injection period.....	101
Table 4-1. Fitted parameters A, C, and D and their standard errors for the four-parameter log-logistic model used to fit CO ₂ retention. The column showing the maximum value for %HCPV is the length of the measured data record from the Industry Data Set. Parameter B was constant in the four-parameter log-logistic model (100 for CO ₂ retention) and is therefore not shown in the table. Cells with an “---” indicate that insufficient data were available to calculate estimates.	133
Table 4-2. Fitted parameters A, C, and D and their standard errors for the four-parameter log-logistic model used to fit incremental oil recovery. The column showing the maximum value for %HCPV is the length of the measured data record from the Industry Data Set. Parameter B was constant in the four-parameter log-logistic model (0 for incremental oil recovery) and is therefore not shown in the table. Cells with an “---” indicate that insufficient data were available to calculate estimates.	136
Table 4-3. Fitted parameters A and B and their standard errors for the two-parameter asymptotic function used to model net CO ₂ utilization. The column showing the maximum value for %HCPV is the length of the measured data record from the Industry Data Set. Cells with an “---” indicate that insufficient data were available to calculate estimates.	140
Table 4-4. Multi-site empirical percentile estimates for CO ₂ retention, incremental oil recovery factor, and net CO ₂ utilization as a function of total CO ₂ + H ₂ O injected (50%, 100%, 150%, 200%, 250%, and 300% HCPV). The modeled responses from each of the sites (31 sites for CO ₂ retention and incremental oil recovery and 16 sites for net CO ₂ utilization) were combined into a single data set to generate empirical percentile estimates across the sites.....	143

List of Figures

Figure 1-1. Simplified diagram illustrating key components of a geologic CO ₂ storage site: storage reservoir (Reservoirs), geologic structural features such as faults and fractures (Seismicity), cap rock overlying the storage reservoir (Seals), wellbores that are present within the storage site area (Wellbores), near-surface receptors such as underground sources of drinking water, the soil vadose zone, and ultimately the surface (Receptors) (Source: DOE, 2015).	2
Figure 2-1. Model domain profile. Aquifer 1 is the target CO ₂ storage reservoir, and is overlain by an aquitard/aquifer sequence to simulate a sedimentary basin. Aquitard 1 is the primary cap rock. Aquifers 5-8 are considered underground sources of drinking water (USDW), while Aquifers 1-4 are saline. Model input parameters for each unit are summarized in Figure S1 in the ESM (adapted from Cihan et al., 2012).	18
Figure 2-2. Overview showing the model domain profile and values for the input parameters for each geologic unit. Input parameter values were taken from Birkholzer et al. (2009) and Cihan et al. (2012).	19
Figure 2-3. Profiles from the baseline model output showing the pressure build-up above hydrostatic pressure after 30 years of CO ₂ injection at distances of 1 km (blue line), 5 km (green line) and 15 km from the CO ₂ injection wells. Based on these results, which show that pressure does not propagate into the shallow geologic units, the first-order uncertainty analysis focused on the two bottom-most aquifer units: A01 and A02.	20
Figure 2-4. Pressure build-up above hydrostatic pressure in Aquifer 1 (bottom) and Aquifer 2 (top) as a function of time and distance from the CO ₂ injection wells. Baseline (blue), 5x (green), 10x (orange), and 100x (red) show the relative change in pressure build-up as a function of the cap rock permeability.	30
Figure 2-5. Pressure build-up above hydrostatic pressure in Aquifer 1 (bottom) and Aquifer 2 (top) as a function of distance from the CO ₂ injection wells at a specified time. Baseline (blue), 5x (green), 10x (orange), and 100x (red) show the relative change in pressure build-up as a function of the cap rock permeability.....	31
Figure 2-6. Comparison of the ratio of pressure build-up in Aquifer 2/Aquifer 1 (A02/A01) as a function of time since injection and lateral distance from the injection wells. Throughout the injection period, the A02/A01 ratio baseline (blue) < 10x (green) < 100x (red) scenarios, suggesting that this ratio would provide greater detection sensitivity than measurements of absolute change in pressure within either unit.	33
Figure 2-7. Cumulative distribution functions for baseline, 2x, 5x, 10x, 20x, 50x, and 100x scenarios within Aquifer 1 as a function of distance from the injection wells at one year (bottom), five years (middle), and 10 years (top) after injection.	36
Figure 2-8. Comparison of the nominal response at 1km from the injection wells for baseline bounded by the 1st and 99th percentiles to the 10x leakage scenario (top) and 100x leakage scenario (bottom).	38
Figure 2-9. Comparison of the nominal response at 1 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario	

(green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).	39
Figure 2-10. Comparison of the nominal response at 2 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).	40
Figure 2-11. Comparison of the nominal response at 3 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).	41
Figure 2-12. Cumulative distribution functions for the ratio A_{02}/A_{01} for baseline, 2x, 5x, 10x, 20x, 50x, and 100x scenarios as a function of distance from the injection wells at one year (bottom), five years (middle), and 10 years (top) after injection.	43
Figure 2-13. Comparison of the nominal response at 1 km from the injection wells for baseline bounded by the 1st and 99th percentiles to the 10x leakage scenario (top) and 100x leakage scenario (bottom).	45
Figure 2-14. Comparison of the nominal response at 1 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).	46
Figure 2-15. Comparison of the nominal response at 2 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).	47
Figure 2-16. Comparison of the nominal response at 3 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).	48
Figure 2-17. Power curves showing the probability of detecting pressure less than baseline conditions in Aquifer 1 (left column) or an A_{02}/A_{01} ratio greater than baseline conditions (right column) at years 1, 5, and 10 of the injection program. The x-axis shows the factor at which the baseline permeability of the cap rock was multiplied for the leakage scenario, i.e., $x=10$ refers to the 10x leakage scenario. ..	50
Figure 2-18. Comparisons of the statistical power curves for three different scenarios of input parameter uncertainty: (1) less uncertain (10-fold case) [dotted line]; (2) nominal uncertainty (100-fold case) [solid line]; and (3) highly uncertain (1000-fold case) [dashed line] for the 1-km (left column), 2-km (middle column) and 3-km (right column) lateral distances at 1 year (top row), 5 years (middle row), and 10 years (bottom row) following injection.	54

Figure 3-1. Cumulative distribution of the deep well leakage potential scores for 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada border (gray open circles) [left] and our random subsample of 100 wells used in this work [right]. The blue dashed line in both panels represents the fitted cumulative distribution function of a 3-parameter gamma function to the full population of 826 legacy oil and gas wells. Many of the deep well leakage potential

scores were identical; therefore, a small jitter of 0.025 was added to the x - and y -directions of this plot to help visualize the number of points. 70

Figure 3-2. Fitted 3-parameter gamma cumulative distribution functions (cdfs, blue dashed line) to the two-attribute partial deep well leakage potential scores for 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada border (gray open circles) (Glazewski et al., 2013). A small jitter of 0.015 was added to the x - and y -directions of this plot to help visualize the number of points used to fit the cdfs. The combinations of two well attributes are shown for each panel..... 82

Figure 3-3. Fitted 3-parameter gamma cumulative distribution functions (cdfs, blue dashed line) to the three-attribute partial deep well leakage potential scores for 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada border (gray open circles) (Glazewski et al., 2013). A small jitter of 0.015 was added to the x - and y -directions of this plot to help visualize the number of points used to fit the cdfs. The combinations of three well attributes are shown for each panel..... 83

Figure 3-4. Probability of at least one leaky well sampled (POLWS) for Scenario 1 (equiprobable well leakage/random well sampling, [left]), Scenario 2 (equiprobable well leakage/score-based well sampling, [middle]), and Scenario 3 (score-based well leakage/random well sampling [right])..... 85

Figure 3-5. Probability that all leaky wells are sampled (PALWS) for Scenario 1 (equiprobable well leakage/random well sampling, [left]), Scenario 2 (equiprobable well leakage/score-based well sampling, [middle]), and Scenario 3 (score-based well leakage/random well sampling [right])..... 85

Figure 3-6. Expected fraction of leaky wells sampled (EFLWS) for Scenario 1 (equiprobable well leakage/random well sampling, [left]), Scenario 2 (equiprobable well leakage/score-based well sampling, [middle]), and Scenario 3 (score-based well leakage/random well sampling [right])..... 86

Figure 3-7. Probability of at least one leaky well sampled (POLWS, left column); probability that all leaky wells are sampled (PALWS, middle column); and expected fraction of leaky wells sampled (EFLWS, right column) for Scenario 1 (equiprobable well leakage/random well sampling, [top]) and Scenario 4 (score-based well leakage/score-based well sampling, [bottom]). 87

Figure 3-8. Ratio of the expected fraction of leaky wells sampled (EFLWS) for Scenario 4 (score-based well leakage/score-based sampling) / Scenario 1 (equiprobable well leakage/random well sampling) as a function of the number of wells sampled (n_S). 89

Figure 3-9. Ratios of the expected fraction of leaky wells sampled (EFLWS) for Scenario 4 / Scenario 5a (small accuracy loss, blue line) or Scenario 4 / Scenario 5b (large accuracy loss, dashed red line) as a function of n_S for $n_L = 1$ (top left), 2 (top right), 5 (bottom left), and 10 (bottom right) wells. Note: when $n_L = 1$, POLWS (probability at least one leaky well sampled) = PALWS (probability all leaky wells sampled) = EFLWS. 91

Figure 3-10. Ratio of EFLWS (Scenario 4) / EFLWS (Scenario 6a), where Scenario 6a utilizes only two of the four well attributes to develop the well sampling. The six combinations of two attributes include: (1) no. of fracture and acid treatments (upper left); (2) no. of fracture treatments and abandonment type (upper middle); (3) no. of

fracture treatments and completions (upper right); (4) no. of acid treatments and abandonment type (lower left); (5) no. of acid treatments and completions (lower middle); and (6) no. of completions and abandonment type (lower right).	94
Figure 3-11. Ratio of EFLWS (Scenario 4) / EFLWS (Scenario 6b), where Scenario 6b utilizes only three of the four well attributes to develop the well sampling. The four combinations of three attributes include: (1) no. of fracture treatments, no. of acid treatments, and abandonment type (upper left); (2) no. of fracture treatments, acid treatments, and completions (upper right); (3) no. of fracture treatments, no. of completions, and abandonment type (lower left); and (4) no. of acid treatments, no. of completions, and abandonment type (lower right).	95
Figure 3-12. Profile from the injector ($x=0$) outward of a hypothetical CO ₂ -brine interface at 1, 5, 10, 20, 30, 40, and 50 years into the CO ₂ injection period using the input parameters for the aquifer, brine, and CO ₂ shown in Table 3-4 and the semi-analytical model of Vilarrasa et al. (2013). The solution at 50 years for the same input parameters using the analytical model of Nordbotten et al. (2005) is shown for comparison (red dashed line).	102
Figure 4-1. Schematic of fluid flows in a CO ₂ EOR project. <i>qCO₂pur</i> refers to purchased CO ₂ ; <i>qCO₂rec</i> refers to recycled CO ₂ ; <i>qCO₂inj</i> refers to total injected CO ₂ , which is comprised of both purchased CO ₂ and recycled CO ₂ ; <i>qCO₂seq</i> refers to the sequestered CO ₂ that is permanently stored in the subsurface and not available for recycle; <i>qwatinj</i> refers to injected water; and <i>qoilprf</i> refers to incremental oil production (Source: van 't Veld et al., 2013).	114
Figure 4-2. Scatterplots of x = cumulative total CO ₂ + H ₂ O injected and y = CO ₂ retention for each of the 31 CO ₂ EOR sites included in the Industry Data Set. The red horizontal line across each panel represents 50% CO ₂ retention.	125
Figure 4-3. Scatterplots of x = cumulative total CO ₂ + H ₂ O injected and y = incremental oil recovery for each of the 31 CO ₂ EOR sites included in the Industry Data Set. The red horizontal line across each panel represents 10% OOIP.	126
Figure 4-4. Scatterplots of x = cumulative total CO ₂ + H ₂ O injected and y = net CO ₂ utilization factor for 16 of the 31 CO ₂ EOR sites included in the Industry Data Set. The red horizontal line across each panel represents 10 Mscf/STB. Note: y -axis is presented on a log-scale.	128
Figure 4-5. Fits of CO ₂ retention four-parameter log-logistic functions to six representative sites from the Industry Data Set. Blue circles represent observed data and the red line represents the fitted response from the four-parameter log-logistic model, Equation 4-8.	132
Figure 4-6. Fits of incremental oil recovery four-parameter log-logistic functions to six representative sites from the Industry Data Set. Blue circles represent observed data and the red line represents the fitted response from the four-parameter log-logistic model, Equation 4-8.	135
Figure 4-7. Fits of net CO ₂ utilization two-parameter asymptotic functions to six representative sites from the Industry Data Set. Blue circles represent observed data and the red line represents the fitted response from the two-parameter asymptotic model, Equation 4-9.	139

Figure 4-8. Percentile estimates generated from multi-site nonlinear regression fits for CO₂ retention (top), incremental oil recovery (middle), and net CO₂ utilization (bottom). The light gray shaded regions bounded by dashed lines represent the 10th to 90th percentiles; the dark gray shaded regions bounded by dotted lines represent the 25th to 75th percentiles (interquartile range); and the P₅₀ estimate (solid black line) represents the central value of the distribution. 142

Figure 4-9. Monte Carlo simulation results of the volume of CO₂ stored as a function of OOIP, incremental oil recovery factor, and net CO₂ utilization factor using Equation 4-10 and an OOIP of 500,000,000 bbl. The y-axis shows the frequency out of 10,000 simulations. The red vertical lines at 13.1 and 42.4 Mt show the P₁₀ and P₉₀ estimates, respectively. The blue vertical line at 25.3 shows the P₅₀ (median). 147

Chapter 1: Introduction

1.1 Geologic CO₂ Storage to Mitigate Atmospheric Inputs of CO₂

In their recent *Fifth Assessment Synthesis Report*, the Intergovernmental Panel on Climate Change (IPCC) asserts that the warming of Earth's climate system is real and observable, and that this warming is directly attributable to anthropogenic inputs of carbon dioxide (CO₂) and other greenhouse gases to the atmosphere, largely as a byproduct of fossil fuel combustion (IPCC, 2014). In an effort to mitigate concentrations of CO₂ in the atmosphere that are caused by stationary anthropogenic inputs, the United States Department of Energy (DOE) is pursuing geologic CO₂ storage as one approach in a portfolio of greenhouse gas reduction strategies. Through the Carbon Sequestration Program, the DOE is working with seven Regional Carbon Sequestration Partnerships (RCSPs) to identify feasible sites within the U.S. and portions of Canada for geologic storage of CO₂ (DOE, 2012).

Geologic CO₂ storage generally involves (1) separating CO₂ from an industrial process, (2) transporting the CO₂ to a storage location, and (3) injecting and sequestering the CO₂ in a geologic reservoir for long-term isolation from the atmosphere (IPCC, 2005). Injecting CO₂ into the subsurface for long-term storage is what DOE refers to as an “engineered-natural system” (DOE, 2015). This engineered-natural system is complex, with potential interactions during CO₂ injection between CO₂ and other reservoir fluids and various components of the geologic system. These components include the storage reservoir, structural features like faults and fractures, seal formations (the cap rock and other low-permeability layers which overly the storage reservoir), wellbores that are present within the storage site area, and near-surface receptors like underground sources

of drinking water (USDW), the soil vadose zone, and ultimately the surface (Figure 1-1). Understanding the interactions between CO₂ and other reservoir fluids and these components of the geologic system in response to CO₂ injection is paramount to accurately modeling the fate and transport processes in the subsurface and ultimately to quantifying the potential risks over the project life-cycle of the geologic CO₂ storage site.

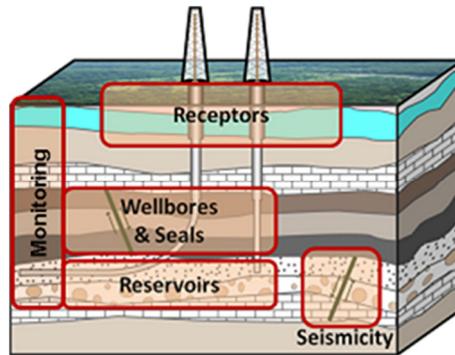


Figure 1-1. Simplified diagram illustrating key components of a geologic CO₂ storage site: storage reservoir (Reservoirs), geologic structural features such as faults and fractures (Seismicity), cap rock overlying the storage reservoir (Seals), wellbores that are present within the storage site area (Wellbores), near-surface receptors such as underground sources of drinking water, the soil vadose zone, and ultimately the surface (Receptors) (Source: DOE, 2015).

1.2 National Risk Assessment Partnership Approach

The National Risk Assessment Partnership (NRAP) is an initiative within DOE's Office of Fossil Energy that is improving the fundamental understanding of the complex science behind engineered-natural systems and is developing the risk assessment tools that are needed for safe, permanent geologic CO₂ storage (DOE, 2015). The NRAP technical approach for developing the risk assessment tools for geologic CO₂ storage entails an iterative modeling approach that integrates component models into a system model which may then be used to provide quantitative assessments of potential risks and to design monitoring protocols that will effectively monitor risks at a geologic CO₂ storage project. A theme throughout all phases of the NRAP approach is *quantifying uncertainty and variability*. For example, uncertainty exists in both the random and

systematic errors used to measure key geologic and fluid parameters, which in turn influence the subsequent simulation of the fate and transport of CO₂ and other reservoir fluids within the modeled system. However, there is also environmental variability over space for these parameters. This variability is inherent to the system and unlike uncertainty it cannot be reduced through additional measurements (Gilbert, 1987). Thus modeling of a complex engineered-natural system during CO₂ injection is not a *deterministic* (i.e., single answer) computation, but is instead a *probabilistic* (stochastic) representation of reality, where uncertainty in the system must be propagated throughout the model to produce a statistical distribution of outcomes rather than a single value (Note: In this dissertation, the term “uncertainty” is used to capture the aggregate effect of both uncertainty and variability.). The focus of this dissertation is to contribute statistical methods and/or approaches for quantifying uncertainty with respect to both monitoring and performance at geologic CO₂ storage sites. These methods are intended for future use by NRAP or other geologic CO₂ storage practitioners and may be incorporated into broader modeling approaches, thus improving the state-of-the-science with respect to developing the risk assessment tools that are needed for safe, permanent geologic CO₂ storage. However, the results and contributions from this work extend beyond geologic CO₂ storage and apply to other subsurface engineered-natural systems.

1.3 Research Objectives

As noted above, the primary objective of this dissertation is to provide statistical methods for quantifying uncertainty related to monitoring and performance at geologic CO₂ storage sites. The specific objectives were as follows:

- **Objective 1** – To estimate the statistical power of a pressure monitoring system to detect leakage of brine through the cap rock as a function of (1) spatial orientation of the pressure monitoring (i.e., the lateral distance from the injection wells and vertical orientation within the reservoir); (2) time (i.e., years since CO₂ injection); and (3) the specified leakage rate of brine through the cap rock.
- **Objective 2** – To estimate the statistical power of selecting one or more leaky wells from among a population of legacy wells as a function of (1) the number of leaking wells within the Area of Review; (2) the sampling design (random or judgmental); (3) the number of wells included in the monitoring sampling plan; and (4) the relationship between a well’s leakage potential score and its relative probability of leakage.
- **Objective 3** – To quantify the uncertainty associated with the CO₂ retention, incremental oil recovery, and net CO₂ utilization for CO₂ enhanced oil recovery (EOR) sites using historical reservoir performance data from 31 CO₂ EOR projects located within the continental United States and heavily dominated by West Texas carbonate floods. Objective 3 is particularly important, as CO₂ EOR is likely to be the primary means of geologic CO₂ storage during the early stages of commercial implementation given the lack of a national policy and the viability of the current business case (Peridas, 2008; Leach et al., 2011).

Each of these specific objectives contributes to the primary objective by illustrating the application of statistical tools and models to different components related to monitoring and performance at geologic CO₂ storage sites. For example, pressure monitoring and

wellbore sampling are principal elements of the monitoring associated with geologic CO₂ storage sites. Thus Objectives 1 and 2 provide important results that may be used to directly answer key monitoring questions for these types of sites. Objective 3 is related to performance. CO₂ retention, incremental oil recovery, and net CO₂ utilization are three factors that significantly influence the long-term performance and economic viability of CO₂ EOR projects. Therefore, Objective 3 provides key information about the expected values and uncertainties associated with these three performance metrics for CO₂ EOR projects.

1.4 Dissertation Overview

This dissertation consists of five chapters. The main content of the thesis is presented in Chapter 2 to 4, which are materials that have been published in peer-reviewed journals.

Chapter 1 provides a general introduction to the dissertation, including the background of the NRAP approach for developing the risk assessment tools that are needed for safe, permanent geologic CO₂ storage; a statement of the research objectives; and a brief overview of the content of this dissertation.

Chapter 2 evaluates the detection sensitivity of deep subsurface pressure monitoring within an uncertain CO₂ sequestration system by linking the output of an analytical reduced-order model and first-order uncertainty analysis. This material was written by Nicholas A. Azzolina and co-authored by Mitchell J. Small, David V. Nakles,

and Grant S. Bromhal, and was published in the peer-reviewed journal, *Stochastic Environmental Research and Risk Assessment*.¹

Chapter 3 uses probabilistic methods to simulate a hypothetical geologic CO₂ storage site in a depleted oil and gas field where the large number of legacy wells would make it cost-prohibitive to sample all wells for all measurements as part of the post-injection site care. This material was written by Nicholas A. Azzolina and co-authored by Mitchell J. Small, David V. Nakles, Kyle A. Glazewski, Wesley D. Peck, Charlie D. Gorecki, Grant S. Bromhal, and Robert M. Dilmore, and was published in the peer-reviewed journal, *Environmental Science and Technology*.²

Chapter 4 analyzes a database of 31 existing CO₂ EOR projects that was compiled for the estimation of oil reserves to better understand the CO₂ retention, incremental oil recovery, and net CO₂ utilization of these oil fields. As noted above, CO₂ EOR is likely to be the primary means of geologic CO₂ storage during the early stages of commercial implementation given the lack of a national policy and the viability of the current business case. This material was written by Nicholas A. Azzolina and co-authored by David V. Nakles, Charlie D. Gorecki, Wesley D. Peck, Scott C. Ayash, L. Steven Melzer, and Sumon Chatterjee, and was published in the peer-reviewed journal, *International Journal of Greenhouse Gas Control*.³

¹ Azzolina, N.A.; Small, M.J.; Nakles, D.V.; and Bromhal, G.S. (2014) Effectiveness of subsurface pressure monitoring for brine leakage detection in an uncertain CO₂ sequestration system. *Stochastic Environmental Research and Risk Assessment*, 28:895-909.

² Azzolina, N.A.; Small, M.J.; Nakles, D.V.; Glazewski, K.A.; Peck, W.D.; Gorecki, C.; Bromhal, G.S.; and Dilmore, R.M. (2015) Quantifying the benefit of wellbore leakage potential estimates for prioritizing long-term MVA well sampling at a CO₂ storage site. *Environmental Science and Technology*, 49 (2): 1215-1224.

³ Azzolina, N.A.; Nakles, D.V.; Gorecki, C.D.; Peck, W.D.; Ayash, S.C.; Melzer, L.S.; and Chatterjee, S. (2015). CO₂ storage associated with CO₂ enhanced oil recovery: A statistical analysis of historical operations. *International Journal of Greenhouse Gas Control*, 37:384-397.

Chapter 5 summarizes the major conclusions of this dissertation, highlights key contributions to the state-of-the science, and provides suggestions for ongoing research and future work associated with the monitoring and modeling at geologic CO₂ storage sites.

1.5 References

DOE (United States Department of Energy) (2012) *The 2012 United States Carbon Utilization and Storage Atlas – Fourth Edition (Atlas IV)*; U.S. Department of Energy, National Energy Technology Laboratory.

DOE (2015) *National Risk Assessment Partnership*, Retrieved from <https://edx.netl.doe.gov/nrap/>

Gilbert, R.O. (1987) *Statistical Methods for Environmental Pollution Monitoring*, John Wiley & Sons, Inc., New York, 320 p.

IPCC (Intergovernmental Panel on Climate Change) (2005). *IPCC Special Report on Carbon Dioxide Capture and Storage*. Prepared by Working Group III of the Intergovernmental Panel on Climate Change [Metz, B., O. Davidson, H. C. de Coninck, M. Loos, and L. A. Meyer (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 442 pp.

IPCC (2014) *Summary for Policymakers. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects*. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32.

Leach, A.; C.F. Mason; and K. van 't Veld. (2011) Co-optimization of Enhanced Oil Recovery and Carbon Sequestration. *Resource & Energy Economics*, 33(4): pp 893-912.

Peridas, G. (2008) *Written testimony submitted to the Natural Resources Committee, Subcommittee on Energy and Mineral Resources*, Unites States House of Representatives hearing on "Spinning Straw Into Black Gold: Enhanced Oil Recovery Using Carbon Dioxide," June 12, 2008.

Chapter 2: Effectiveness of subsurface pressure monitoring for brine leakage detection in an uncertain CO₂ sequestration system⁴

Abstract

This work evaluates the detection sensitivity of deep subsurface pressure monitoring within an uncertain CO₂ sequestration system by linking the output of an analytical reduced-order model and first-order uncertainty analysis. A baseline (non-leaky) modeling run was compared against 10 different leakage scenarios, where the cap rock permeability was increased by factors of 2 to 100 (cap rock permeability from 10⁻³ to 10⁻¹ millidarcy). The uncertainty variance outputs were used to develop percentile estimates and detection sensitivity for pressure throughout the deep subsurface as a function of space (lateral distance from the injection wells and vertical orientation within the reservoir) and time (years since injection), or $P(x, z, t)$. Conditional probabilities were computed for combinations of x , z , and t , which were then used to generate power curves for detecting leakage scenarios. The results suggest that measurements of the absolute change in pressure within the target injection aquifer would not be able to distinguish small leakage rates (i.e., less than 50x baseline) from baseline conditions, and that only large leakage rates (i.e., >100x baseline) would be discriminated with sufficient statistical power (>99%). Combining measurements, for example by taking the ratio of formation pressure in Aquifer 2/Aquifer 1, provides better statistical power for distinguishing

⁴ Chapter 2 was the basis for the peer-reviewed publication. In the publication, several figures were presented in the Electronic Supplementary Material. However, in this dissertation the figures have been renumbered in consecutive numerical order and are embedded into the main text. Azzolina, N.A.; Small, M.J.; Nakles, D.V.; and Bromhal, G.S. (2014) Effectiveness of subsurface pressure monitoring for brine leakage detection in an uncertain CO₂ sequestration system. *Stochastic Environmental Research and Risk Assessment*, 28:895-909.

smaller leakage rates at earlier times in the injection program. Detection sensitivity for pressure is a function of space and time. Therefore, design of an adequate monitoring network for subsurface pressure should account for this space-time variability to ensure that the monitoring system performs to the necessary design criteria, e.g., specific false-negative and false-positive rates.

2.1 Introduction

In an effort to mitigate concentrations of carbon dioxide (CO₂) in the atmosphere that are caused by stationary anthropogenic inputs, the United States Department of Energy (DOE) is pursuing carbon capture and sequestration (CCS) as one approach in a portfolio of greenhouse gas (GHG) reduction strategies. CCS involves (1) separating CO₂ from an industrial process, (2) transporting the CO₂ to a storage location, and (3) injecting and sequestering the CO₂ in a geologic reservoir for long-term isolation from the atmosphere (IPCC, 2005). Through the Carbon Sequestration Program, the DOE is working with seven Regional Carbon Sequestration Partnerships (RCSPs) to identify feasible sites within the U.S. and portions of Canada for large-scale (i.e., one million tonnes of CO₂ or greater) CO₂ geologic sequestration (GS) (DOE, 2011).

The injection and long-term storage of CO₂ into geologic reservoirs is not without risks. In general, these risks include leakage of stored CO₂ or displaced formation water brine into the near-surface or surface environment or physical impacts resulting from modifying the subsurface stress regime. Several potential failure modes have been identified, including: (1) salinity impacts to freshwater underground sources of drinking water (USDW) from displaced formation water brines (Damen et al., 2006); (2)

mobilization of metals in USDW as a function of leaking CO₂ entering the USDW, lowering the aquifer pH, and modifying the aquifer conditions such that metals enter the aqueous phase (Siirila et al., 2010; Little and Jackson, 2010; Wilkin and Digulio, 2010; Apps et al., 2011); (3) induced seismicity caused by injection of CO₂ altering the subsurface stress conditions (Damen et al., 2006); and (4) CO₂ leakage impacting sensitive ecological receptors in the near-surface and surface environment (U.S. EPA 2008). These and other risks are analyzed and managed through combinations of site characterization, simulation models, and monitoring. Site characterization data include well logs, seismic data, geologic cores, and other measurements used to delineate the stratigraphy and subsurface properties of the storage system. Simulation models are then constructed from the site characterization data and used to predict the fate and transport of CO₂ and other reservoir fluids in the subsurface or to model the geomechanical integrity of the reservoir and overlying seal formation in response to injection. Lastly, monitoring is used throughout the project life-cycle (baseline monitoring prior to injection, operational monitoring during injection, and post-operational monitoring following the injection period) to collect measurements and, based on the results of those measurements, infer whether or not one or more system failures is occurring.

Monitoring will be an important component of making CO₂ GS an acceptable method for atmospheric CO₂ mitigation. Many different types of monitoring technologies are available, and each provides information with respect to making inferences about the probability of CO₂ leakage or other types of storage system failures. Ultimately, an integrated, model-based monitoring network which incorporates tailored combinations of technologies will provide the appropriate balance between risk reduction and cost

effectiveness (Seto and McRae, 2011; Yang et al., 2012). The final configuration of monitoring technologies at any given CO₂ GS site will be site-specific, including the appropriate number of technologies positioned at adequate sampling density and collecting information at the appropriate frequency (i.e., continuous, daily, weekly, monthly, or annually) to provide sufficient detection of leakage. However, heuristic studies of generic sites provide insight into how monitoring plans may be developed, and which technologies may provide sufficient detection sensitivity at lower cost.

In general, monitoring technologies may be classified by three different zones: (1) atmospheric, (2) near-surface, and (3) deep subsurface (DOE, 2009). While atmospheric and near-surface technologies provide information with respect to leakage at the surface or near-surface environments, they may be insufficient to provide detection early enough within the reservoir system to prevent impacts such as contamination of USDWs. In contrast, deep subsurface monitoring provides a sentinel location within the system, which could allow an operator to make adjustments during injection, thereby mitigating impacts to USDW or other near-surface and surface environments. One particular type of subsurface measurement that may provide information with respect to leakage is formation pressure (DOE, 2009). Pressure monitoring is attractive when there are a discrete number of monitoring wells within the system, because the pressure pulse from injection will cover a significantly broader portion of the system than the CO₂ plume. For example, the results of stochastic simulations using a semi-analytical model showed that for a 100 km² domain with 10 monitoring wells, the pressure pulse due to leakage after 50 years of injection covered 98% of the project area, while the CO₂ plume only covered 3.5% of the domain (Nogues et al., 2011).

While several authors have investigated the use of pressure monitoring at CO₂ GS sites (e.g., Nicot et al., 2009; Nogues et al., 2011) and the sensitivity of pressure output to model inputs (Birkholzer et al., 2009; Cihan et al., 2011; 2012), uncertainty analysis has not been extensively conducted to assess detection sensitivity of pressure monitoring throughout the reservoir system as a function of space and time. Uncertainty analysis adds important information beyond sensitivity analysis, and helps define the probability at which monitoring during and post-injection will differentiate leakage of a given size from the background variability. This work evaluates the detection sensitivity of deep subsurface pressure monitoring within an uncertain CO₂ sequestration system by using an analytical reduced-order model (ROM) and first-order uncertainty analysis (FOUA). This work shows the value of a first-order approach for estimating uncertainty, and the importance of uncertainty in the design of a reliable monitoring network. Probabilistic approaches that incorporate uncertainty are important for designing monitoring networks, and may be used to select the best network that minimizes total cost (Bierkens, 2006; Kim and Lee, 2007).

Modeling options for CO₂ GS sites range from more simplified analytical approaches (e.g., Nordbotten et al., 2004; Cihan et al., 2012) to fully coupled, numerical three-dimensional (3D) models that include multi-phase flow, multi-component transport with complex phase partitioning, coupled geomechanical and geochemical reactions (Court et al., 2011). Many different types of numerical models have been developed. For example, DOE (2009) lists over 20 numerical codes that have been developed by either national laboratories or commercial vendors. Most of these numerical models involve solving the mass balance equations within the model grid, and their complexity requires

the use of massively parallel supercomputers (Court et al., 2011). Simpler 3D models exist, where one or more interactions are de-coupled, thereby simplifying the governing equations and expediting computation. Models such as TOUGH2 (Pruess, 2005), ECLIPSE (Exploration Consultants Limited Implicit Program for Simulation Engineering) (Schlumberger, 2010), and codes developed at the national laboratories such as Pacific Northwest National Laboratory's (PNNL's) STOMP (Subsurface Transport Over Multiple Phases) (White, 1997) or Lawrence Livermore National Laboratory's (LLNS's) NUFT (Nonisothermal, Unsaturated Flow and Transport with Chemistry) (Nitao, 1998) are examples of simplified 3D models (Court et al., 2011). Lastly, vertically-averaged two-dimensional (2D) models integrate the full 3D equations over the direction perpendicular to the top and bottom boundaries of the formation of interest (often assumed to be the vertical direction) (Court et al., 2011). Due to the computational complexity of fully-coupled or partially de-coupled 3D models, it is generally not practical to conduct multiple realizations of these numerical models to assess uncertainty, as each model run may take tens of hours (or longer). Therefore, an open research question in this field is how to evaluate uncertainty given the modeling limitations. This work uses a recently-developed ROM for simulating reservoir injection and pressure that was developed by Lawrence Berkley National Lab (LBNL Model, Cihan et al., 2011). The LBNL Model is further described in the Methods section. The approach for evaluating detection sensitivity builds upon methods established by Yang et al. (2011a, b; 2012), and the process includes four primary steps: (1) Characterize the uncertainty and variability present in the expected value for pressure with near-zero reservoir leakage (i.e., baseline conditions); (2) Determine critical values of pressure

beyond which (above or below, depending upon whether leakage causes an increase or a decrease in pressure) the operator can reject the null hypothesis (H_0 : No leakage above baseline conditions); (3) Simulate the change in pressure in response to different leakage rates, including the variability and uncertainty of pressure at each leakage rate; and (4) Considering this prediction with its variability and uncertainty, compute the probability that the leak-impacted pressure will be beyond the critical value, allowing rejection of H_0 , and therefore inference of a leak. The above process yields estimates of the statistical power of the pressure monitoring system for the given leakage rate. In the development of power curves, we selected as critical values the 99th percentile (P_{99}) of the baseline pressure (when leakage causes an increase in pressure) or the 1st percentile (P_{01}) of the baseline pressure (when leakage causes a decrease in pressure). We used the FOUA variance outputs to develop percentile estimates and detection sensitivity for pressure throughout the deep subsurface as a function of space (lateral distance from the injection wells and vertical orientation within the reservoir) and time (years since injection). We defined leakage in terms of the additional brine flux that would pass through the primary cap rock as a function of increasing the value of the cap rock permeability by factors of two through 100 above baseline. Therefore, "leakage" is meant to imply that the cap rock transmits fluids at a greater rate than was originally anticipated during the site characterization (feasibility study) phase. This can be due to improper geologic characterization, or induced changes in the cap rock as a function of CO_2 injection (e.g., geochemical or geomechanical failures). In a sense, the system is "leaking" fluids to overlying strata, and this study is asking whether or not small "leakages" will be detected by changes in pressure in the subsurface. The ability to detect

these leakages would then be an input into the broader site decision framework for assessing potential brine impacts to USDW as CO₂ injection continues.

2.2 Methods

2.2.1 Reduced-Order Model

The LBNL Model has been extensively described in Cihan et al. (2011) and (2012). A brief overview of the solution approach is described below. The analytical solution assumes single-phase flow in a multilayered system of aquifers and aquitards with both diffuse brine leakage (i.e., flux through aquitards) and focused brine leakage (i.e., flow through wells) (Cihan et al., 2012). The LBNL Model is a single-phase model; therefore, multi-phase processes are not incorporated into the solution. However, injection of a single-phase fluid (brine) with an equivalent volume of CO₂ injected compared well with the numerical model, TOUGH2-ECO2N, and provided accurate results for pressures beyond the CO₂ plume and brine leakage zone (Birkholzer et al., 2009; Cihan et al., 2011; 2012). These results show that multi-phase processes inside the CO₂ plume may be assumed negligible for prediction of far-field pressure build-up. However, the analytical solution does over-predict pressure build-up within the CO₂ plume, and therefore this study restricts the modeling outputs to distances of greater than 1 km from the injection location (Cihan et al., 2012). All aquifers and aquitards are assumed homogeneous, with uniform thickness and infinite radial extent. Fluid flow is horizontal in the aquifers and vertical in the aquitards. Leaky wells are represented as Darcy-type flow pathways, and the user may specify well X- and Y-coordinates, the hydraulic conductivity of well-aquifer segments, the hydraulic conductivity of well-

aquitard segments, and the radius of the well (Cihan et al., 2012). The equations of horizontal groundwater flow in the aquifers are coupled to the vertical-flow equations in the aquitards and the flow-continuity equations in the leaky wells. The governing partial differential equations for single-phase flow in aquifers and aquitards are transformed into the Laplace domain, and the resulting coupled system of ordinary differential equations are solved using the eigenvalue analysis method (Cihan et al., 2012). The LBNL Model is compiled in FORTRAN, and the model output includes pressure build-up in the different geologic units with time.

To evaluate reservoir conditions that had already been published and validated using the numerical TOUGH2/ECO2N simulator (Prues et al., 1999; Pruess, 2005), parameters for the baseline scenario were taken from Cihan et al. (2012). The model includes a primary injection aquifer (1180 m to 1120 m), overlying cap rock (1120 m to 1020 m), and alternating units of aquifers and aquitards above the cap rock up to the land surface (Figure 2-1). Details on the model input parameters for each unit are provided in Figure 2-2 and Appendix A provides the input file for the baseline model. CO₂ was injected into a zone of 50 m radial extent within the primary injection aquifer, representing several distributed wells. The original model simulation included 30 years injection at an annual rate of 1.52 million tonnes of CO₂ followed by a 70-year post-injection period, for a total of 100 years (Cihan et al., 2012). The original baseline model pressure output was computed assuming a low-permeability (0.1 millidarcy (mD)) cap rock and outputs were recorded at multiple lateral distances from the injection wells (x), vertical intervals (z), and times since injection (t), or $P(x, z, t)$. A review of the pressure distribution throughout the system showed that significant pressure build-up did not

occur in many of the overlying units, even after 30 years of injection (Figure 2-3). Therefore, subsequent modeling and the FOUA focused on uncertainty within the three units representing the deepest zones within the system: the target injection aquifer (Aquifer 1), primary cap rock (Aquitard 1), and overlying aquifer (Aquifer 2). These three units are located between 1180 m to 960 m below ground surface (Figure 2-1). In addition, to focus on the period of greatest system sensitivity, FOUA output times were limited to the 30-year injection period (30 time steps in total), and did not include the 70-year post-injection pressure decline.

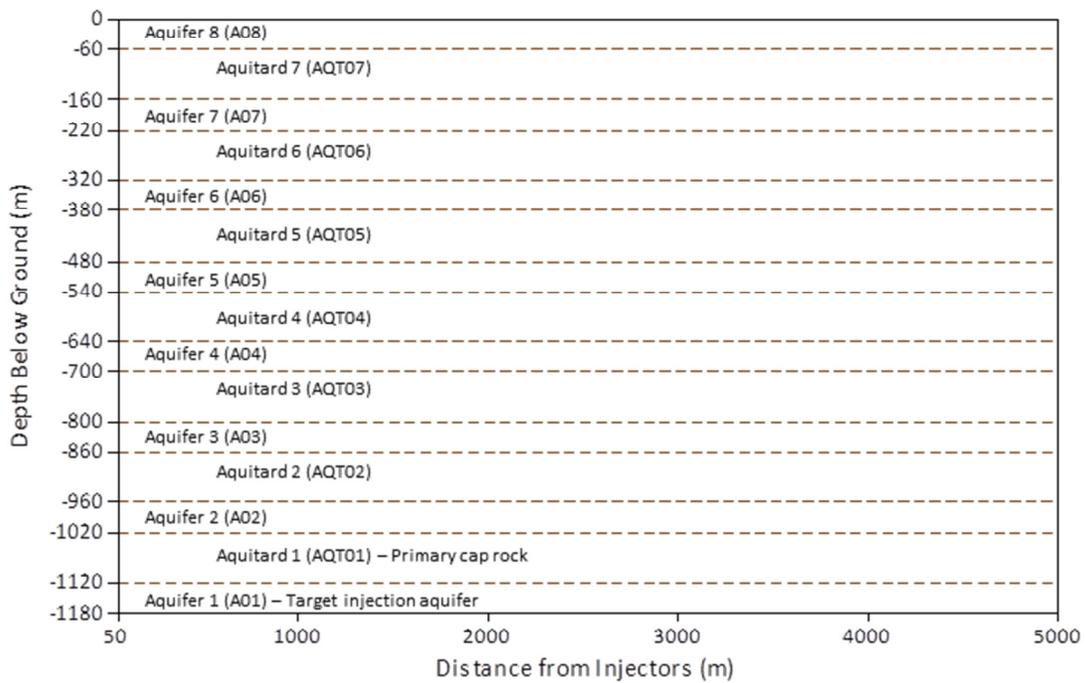


Figure 2-1. Model domain profile. Aquifer 1 is the target CO₂ storage reservoir, and is overlain by an aquitard/aquifer sequence to simulate a sedimentary basin. Aquitard 1 is the primary cap rock. Aquifers 5-8 are considered underground sources of drinking water (USDW), while Aquifers 1-4 are saline. Model input parameters for each unit are summarized in Figure S1 in the ESM (adapted from Cihan et al., 2012).

Unit Code	Hydrostatic Pressure (kPa)	Temp (°C)	Salt Mass Fraction	Density (kg/m ³)	Viscosity (Pa s)	Permeability (mD)	Porosity (%)	Hydraulic Cond. (m/d)	Brine Compress. (Pa ⁻¹)	Pore Compress. (Pa ⁻¹)	Specific Storage (m ⁻¹)
A08	500	15.0	0	999.28	1.12E-03	101	0.20	7.6E-02	4.7E-10	4.5E-10	1.8E-06
AQT07	1,375	17.5	0	999.34	1.08E-03	0.001	0.15	7.8E-07	4.7E-10	9.0E-10	2.0E-06
A07	2,125	18.8	0	999.37	1.04E-03	101	0.20	8.1E-02	4.6E-10	4.5E-10	1.8E-06
AQT06	3,000	20.3	0	999.37	9.99E-04	0.001	0.15	8.5E-07	4.6E-10	9.0E-10	2.0E-06
A06	3,750	21.8	0	999.36	9.58E-04	101	0.20	8.8E-02	4.6E-10	4.5E-10	1.8E-06
AQT05	4,500	23.5	0	999.31	9.22E-04	0.001	0.15	9.2E-07	4.5E-10	9.0E-10	2.0E-06
A05	5,250	25.5	0	999.26	8.89E-04	101	0.20	9.5E-02	4.5E-10	4.5E-10	1.8E-06
AQT04	6,000	27.0	0.011	1004.78	8.70E-04	0.001	0.15	9.8E-07	4.4E-10	9.0E-10	2.0E-06
A04	6,750	28.5	0.023	1013.32	8.59E-04	101	0.20	1.0E-01	4.3E-10	4.5E-10	1.7E-06
AQT03	7,500	30.0	0.038	1021.50	8.49E-04	0.001	0.15	1.0E-06	4.1E-10	9.0E-10	2.0E-06
A03	8,300	31.5	0.053	1030.10	8.42E-04	101	0.20	1.0E-01	4.0E-10	4.5E-10	1.7E-06
AQT02	9,050	33.0	0.069	1046.52	8.54E-04	0.001	0.15	1.0E-06	3.8E-10	9.0E-10	2.0E-06
A02	9,875	34.5	0.098	1062.46	8.69E-04	101	0.20	1.0E-01	3.7E-10	4.5E-10	1.7E-06
AQT01	10,750	36.0	0.120	1078.41	8.90E-04	0.001	0.15	1.0E-06	3.5E-10	9.0E-10	2.0E-06
A01	11,625	38.0	0.150	1095.63	9.16E-04	101	0.20	1.0E-01	3.4E-10	4.5E-10	1.7E-06

Figure 2-2. Overview showing the model domain profile and values for the input parameters for each geologic unit. Input parameter values were taken from Birkholzer et al. (2009) and Cihan et al. (2012).

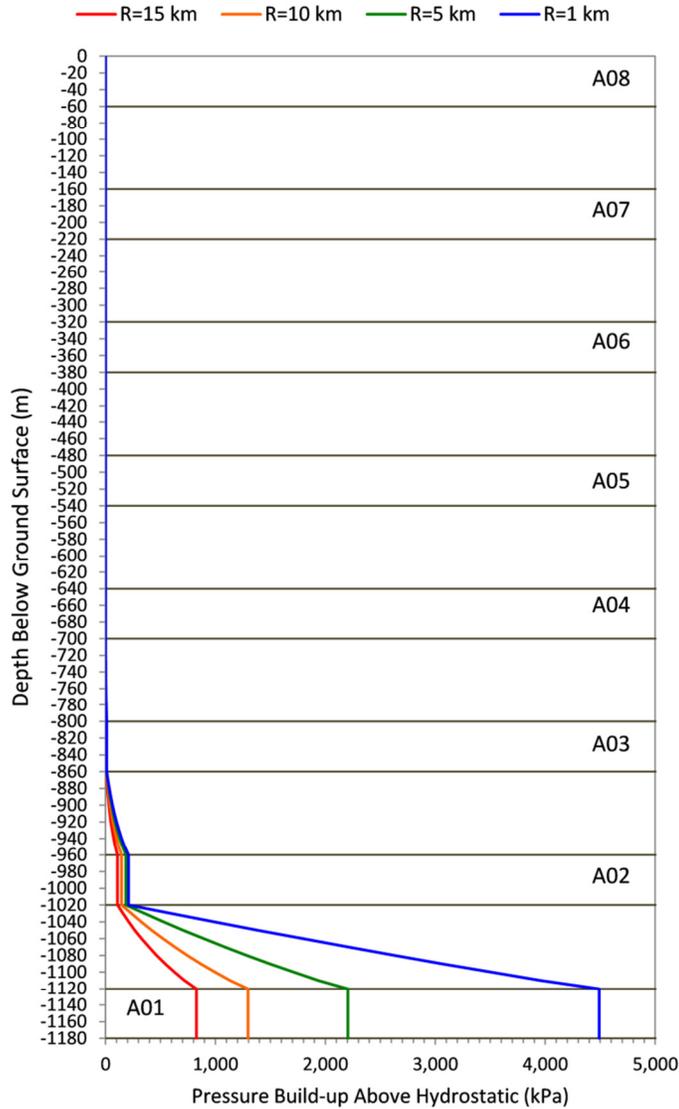


Figure 2-3. Profiles from the baseline model output showing the pressure build-up above hydrostatic pressure after 30 years of CO₂ injection at distances of 1 km (blue line), 5 km (green line) and 15 km from the CO₂ injection wells. Based on these results, which show that pressure does not propagate into the shallow geologic units, the first-order uncertainty analysis focused on the two bottom-most aquifer units: A01 and A02.

2.2.2 First-Order Uncertainty Analysis

Table 2-1 summarizes the nine principal uncertain inputs to the model, which included the hydraulic conductivity of Aquifers 1 (HCONX1) and 2 (HCONX2) and Aquitard 1 (HCONP1), the specific storage of Aquifers 1 (SS1) and 2 (SS2) and Aquitard

1 (SSP1), and the thickness of Aquifers 1 (BAQ1) and 2 (BAQ2) and Aquitard 1 (BAQP1).

In hydrogeologic systems, one source of uncertainty is the stochastic nature of the subsurface deposits and the resulting spatial field of hydraulic conductivity, which is a function fluid density, fluid viscosity, and the reservoir permeability as described in Equation 2-1 (Dagan, 1987; Rubin, 2003; Ezekwe, 2011).

$$k = \kappa \frac{\rho g}{\mu} \quad (2-1)$$

Where:

k is hydraulic conductivity (m d^{-1}),

κ is permeability (m^2),

ρ is the fluid density (kg m^{-3}),

g is the acceleration due to gravity (9.81 m s^{-2}), and

μ is the fluid viscosity ($\text{kg m}^{-1}\text{s}^{-1}$).

For the calculation of hydraulic conductivity, fluid density and viscosity were considered constant within an aquifer or aquitard layer and were calculated on-line using the CREWES Fluid Property Calculator (CREWES, 2012). Formation salinity, temperature, and pressure profiles used in the computation of fluid properties were taken from Birkholzer et al. (2009). Permeability was assumed to follow a lognormal distribution. This assumption is based on the fact that data from geologic cores have shown that permeability follows a lognormal distribution (Ezekwe, 2011). In addition,

estimations of correlation functions for reservoir heterogeneity across hydrogeologic systems has shown from well test data that permeability is more closely approximated by a lognormal distribution (Gelhar, 1993; Yortsos, 2000). The median values (mean of the log-transformed permeability) were set equal to the values published in Cihan et al. (2012). We assumed that the median value was known to within plus-or-minus an order-of-magnitude, and therefore established a 100-fold variability between the 0.1th and 99.9th percentiles of the lognormal distribution to solve for the parameters (a and b) and estimate the variance in permeability.

Specific storage was calculated according to Equation 2-2 and was also assumed to follow a lognormal distribution.

$$S_s = \rho g \phi (\alpha + \beta) \quad (2-2)$$

Where:

S_s is specific storage (m^{-1}),

ϕ is porosity (unitless),

α is the brine compressibility (Pa^{-1}), and

β is the pore compressibility (Pa^{-1}).

The median porosity was set at 20% for aquifers and 15% for aquitards (Birkholzer et al., 2009). The median brine compressibility was set at $3.9 \times 10^{-10} Pa^{-1}$ (Zhou et al., 2007) and the median pore compressibility was set at $4.5 \times 10^{-10} Pa^{-1}$ for aquifers and $9.0 \times 10^{-10} Pa^{-1}$ for aquitards (Birkholzer et al., 2009). As pore

compressibility could reasonably range from $4.5 \times 10^{-9} \text{ Pa}^{-1}$ to $4.5 \times 10^{-11} \text{ Pa}^{-1}$ (Zhou et al., 2007), we again assumed that the median value for specific storage was known to within plus-or-minus an order-of-magnitude, and therefore established a 100-fold variability between the 0.1th and 99.9th percentiles of the lognormal distribution to solve for the parameters and estimate the variance in specific storage.

Lastly, the median aquifer thicknesses were set to 60 m and the median aquitard thicknesses were set to 100 m (Birkholzer et al., 2009). These were assumed to follow a uniform distribution with minimum and maximum values set at 10 m below or above the median value.

The distributions assigned to these parameters attempt to account for both uncertainty and variability. For example, uncertainty exists in both the random and systematic errors used to measure these parameters, as well as in the assumptions used to estimate the parameters from theoretical models, such as estimating permeability from well test data (Horne, 1995; Ramaswami et al., 2005). However, there is also environmental variability over space for these parameters. This variability is inherent to the system and unlike uncertainty it cannot be reduced through additional measurements (Gilbert, 1987; Ramaswami et al., 2005).

The FOUA of the baseline model was conducted by changing each of the variables by a “delta” of $\pm 5\%$ through 18 separate modeling runs and computing the variance-covariance matrix of model outputs using methods summarized by Ramaswami et al. (2005), including the uncertain variance in $\ln(\text{pressure})$ at each (x, z, t) . FOUA uses a Taylor series approximation to represent the model and estimates the variance in an output variable from this approximation. As the Taylor series is truncated after the first

term, it is known as “first-order” uncertainty analysis (Ramaswami et al., 2005). For a model that can be represented as $y = f(x)$, where $x = \{x_1, x_2, \dots, x_n\}^T$ is a vector of uncertain parameters or model inputs and the function f is continuously differentiable in the neighborhood about x , then the Taylor series expansion for y about its nominal value, y^* is (Ramaswami et al., 2005):

$$y = y^* + \sum_{i=1}^n (x_i - x_i^*) \left[\frac{\partial y}{\partial x_i} \right]_{x^*} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (x_i - x_i^*)(x_j - x_j^*) \left[\frac{\partial^2 y}{\partial x_i \partial x_j} \right]_{x^*} + HOT \quad (2-3)$$

If $x_i^* = E[x_i]$ is the expected value of the parameter x_i , then the variance of y , $E[(y - y^*)^2]$, resulting from uncertainty in the x_i values, can be approximately evaluated from the first-order term of the Taylor series as (Ramaswami et al., 2005):

$$E[(y - y^*)^2] \cong E \left[\left(\sum_{i=1}^n (x_i - x_i^*) \left[\frac{\partial y}{\partial x_i} \right]_{x^*} \right)^2 \right] = \sum_{i=1}^n \sum_{j=1}^n Cov[x_i, x_j] \left[\frac{\partial y}{\partial x_i} \right]_{x^*} \left[\frac{\partial y}{\partial x_j} \right]_{x^*} \quad (2-4)$$

If the x_i are independent random variables ($Cov[x_i, x_j] = 0$ for $i \neq j$) then Equation 2-4 reduced to (Ramaswami et al., 2005):

$$E[(y - y^*)^2] \cong \sum_{i=1}^n Var[x_i] \left(\left[\frac{\partial y}{\partial x_i} \right]_{x^*} \right)^2 \quad (2-5)$$

The individual terms of the sum on the right of Equation 2-5 represent the contributions of each of the n parameters to the variance in the output variable. The nine inputs were considered independent for the initial FOUA work.

The first-order derivative shown in Equations 2-3 to 2-5 was estimated numerically by central differencing using $\Delta x = 5\%$ of x^* (Ramaswami et al., 2005):

$$\frac{\partial y}{\partial x_i} \cong \frac{\Delta y}{\Delta x_i} = \frac{y|_{x^*, x_i^* + \Delta x} - y|_{x^*, x_i^* - \Delta x}}{2\Delta x} \quad (2-6)$$

A matrix solution approach was used to evaluate the a covariance matrix of the model inputs. For n_x model inputs, $x_i, i = 1, \dots, n_x$ and n_y model outputs, $y_k, k = 1, \dots, n_y$; the n_y by n_x model sensitivity matrix, A_{yx} , is given by (Ramaswami et al., 2005):

$$[A_{yx}] = \left[\frac{\partial y_k}{\partial x_i} \right]_{x^*} \quad (2-7)$$

where x^* denotes that the local sensitivity coefficients are computed with all inputs at their nominal values. The $n_x \times n_x$ covariance matrix of the inputs, S_x , is defined as (Ramaswami et al., 2005):

$$[S_x] = [Cov[x_i, x_j]] \quad (2-8)$$

The $n_y \times n_y$ covariance matrix of the outputs, S_y , is then computed as (Ramaswami et al., 2005):

$$[S_y] = [A_{yx}] \times [S_x] \times [A_{yx}]^T \quad (2-9)$$

where $[]^T$ denotes the matrix transpose operation.

The outputs from the LBNL Model were imported into Minitab® Statistical Software (Minitab, 2010) and the computations of the LBNL Model output variance were conducted in four steps: (1) Populate the $[S_x]$ matrix of the input variables (the diagonals were equal to the variance and the off-diagonal elements were all set to zero due to independence); (2) Compute the first-order derivative for each x , z , and t , using Equation 2-6 and thereby populate the $[A_{yx}]$ matrix; (3) Transpose the $[A_{yx}]$ matrix; and (4) Calculate the $[S_y]$ matrix according to Equation 2-7. These four steps then provided a mean and variance estimate for each x , z , and t based on the variable inputs.

FOUAs of the leakage scenarios were computed using the same approach described above, by repeating the 18 different modeling runs but with the initial HCONP of the cap rock increasing by 2x, 3x, 4x, 5x, 10x, 20x, 30x, 40x, 50x, and 100x baseline to emulate diffuse leakage through the seal formation (Table 2-1). The changes in HCONP equate to the median permeability of the cap rock changing from a baseline value of 10^{-3} mD to the leakiest scenario of 10^{-1} mD.

The results of the FOUA provide information with respect to the distribution of pressure responses for each scenario, which thereby permits the quantification of the probability that a leak of a given size is detected (also referred to as the statistical power of the measurement) given the current pre-injection uncertainty in subsurface properties.

Table 2-1. LBNL Model input parameters for the baseline model that were changed by $\pm dx$ through a series of 18 separate modeling runs as part of the FOUA process. In the leaky scenarios, HCONP was modified by a factor of 2, 3, 4, 5, 10, 20, 30, 40, 50, and 100 for each of the 18 modeling runs

Run No.	Input Variable ^a	Unit	Unit Name	x^*	Units	Distribution	Parameters ^b					Range	dx
							Mean	Variance	Std Dev	a	b	b/a	
1	HCONX	Aquifer 1	Target inject aquifer	1.01E-01	m/d	lognormal	-2.29	0.56	0.75	1.0E-02	1.0E+00	100	5.1E-03
2	HCONX	Aquifer 2	Saline aquifer	1.04E-01	m/d	lognormal	-2.27	0.56	0.75	1.0E-02	1.0E+00	100	5.2E-03
3	SS	Aquifer 1	Target inject aquifer	1.69E-06	1/m	lognormal	-13.29	0.56	0.75	1.7E-07	1.7E-05	100	8.5E-08
4	SS	Aquifer 2	Saline aquifer	1.70E-06	1/m	lognormal	-13.28	0.56	0.75	1.7E-07	1.7E-05	100	8.5E-08
5	BAQ	Aquifer 1	Target inject aquifer	60	m	uniform	60	33.3	5.8	50	70	1.4	3
6	BAQ	Aquifer 2	Saline aquifer	60	m	uniform	60	33.3	5.8	50	70	1.4	3
7	HCONP	Aquitard 1	Primary cap rock	1.03E-06	m/d	lognormal	-13.79	0.56	0.75	1.0E-07	1.0E-05	100	5.1E-08
8	SSP	Aquitard 1	Primary cap rock	1.99E-06	1/m	lognormal	-13.13	0.56	0.75	2.0E-07	2.0E-05	100	9.9E-08
9	BAQT	Aquitard 1	Primary cap rock	100	m	uniform	100	33.3	5.8	90	110	1.2	5

^aInput variable definitions:

HCONX = hydraulic conductivity, x-direction (aquifers)

SS = specific storage (aquifers)

ANSR = anisotropy (aquifers only)

BAQ = aquifer thickness

HCONP = hydraulic conductivity, z-direction (aquitards)

BAQT = aquitard thickness

^b a and b :

For lognormal distributions, a = 0.1th percentile and b = 99.9th percentile.

For uniform distributions, a = minimum and b = maximum.

2.2.3 Statistical Power Curves

The cumulative distribution functions (CDFs) were calculated for the different $P(x, z, t)$ combinations in the baseline model using the mean and variance from the FOUA output and an assumed lognormal distribution. Conditional probabilities for the different leakage scenarios were derived by taking the inverse of the P_{01} and P_{99} from the baseline model using the mean and variance computed for each leaky scenario. The objective is to calculate the probability that each leaky scenario (HCONP = 2x, 3x, etc.) yields a pressure output beyond the P_{99} (or below the P_{01}) of the non-leaky (baseline) estimate. For example,

$$Prob[P_{x,z,t} > Y | HCONP = 2x] = F^{-1}(X, \mu 2x, \sigma 2x) \quad (2-10)$$

where X is the P_{99} value from the baseline model at that unique $P(x, z, t)$ and $\mu 2x$ and $\sigma 2x$ are the mean and standard deviation, respectively, of the 2x-leaky scenario at that same unique $P(x, z, t)$. Conditional probabilities were computed for all combinations of x, z, t , and HCONP of the cap rock, which were then used to generate power curves for selected x, z , and t values.

2.3 Results

2.3.1 Model Nominal Response

Increasing the permeability of the cap rock produces a direction of change that is to be expected, namely the pressure build-up decreases over time relative to the baseline

scenario within Aquifer 1 (the injection reservoir) and, in contrast, pressure build-up over time in Aquifer 2 increases relative to the baseline scenario (Figures 2-4 and 2-5). The magnitude of the change in pressure build-up is both a function of the aquifer (i.e., change Aquifer 1 > Aquifer 2) and lateral distance from the injection wells (i.e., the change is greatest closest to the injection wells and decreases as the radius increases). For example, after five years of injection the difference in pressure between (baseline – 10x) and (baseline – 100x) scenarios within Aquifer 1 are 515 and 1270 kilopascals (kPa), respectively, at 1 km. At a distance of 10 km, the differences are reduced to 234 and 347 kPa, respectively, at the same point in time.

Pressure build-up within Aquifer 3 in the Baseline model was negligible during the first decade of injection. At a 1-km distance from the injection wells, pressure build-up above 14 kPa was not observed until 32 years after the start of injection, which is after injection has stopped and into the post-injection long-term monitoring phase. In contrast, the 10x and 100x leakage scenarios showed pressure build-up in Aquifer 3 above 14 kPa at years 14 and 10, respectively (Figures 2-4 and 2-5). A pressure build-up of 14 kPa is within the range of the published pressure accuracy for a down-hole quartz pressure gauge (± 14 kPa) (Schlumberger, 2012). This result implies that detection of pressure increase above hydrostatic within Aquifer 3 during the early phases of injection would provide evidence of leakage. However, even with a relatively permeable cap rock, this leakage impact is not evident within the first decade of injection.

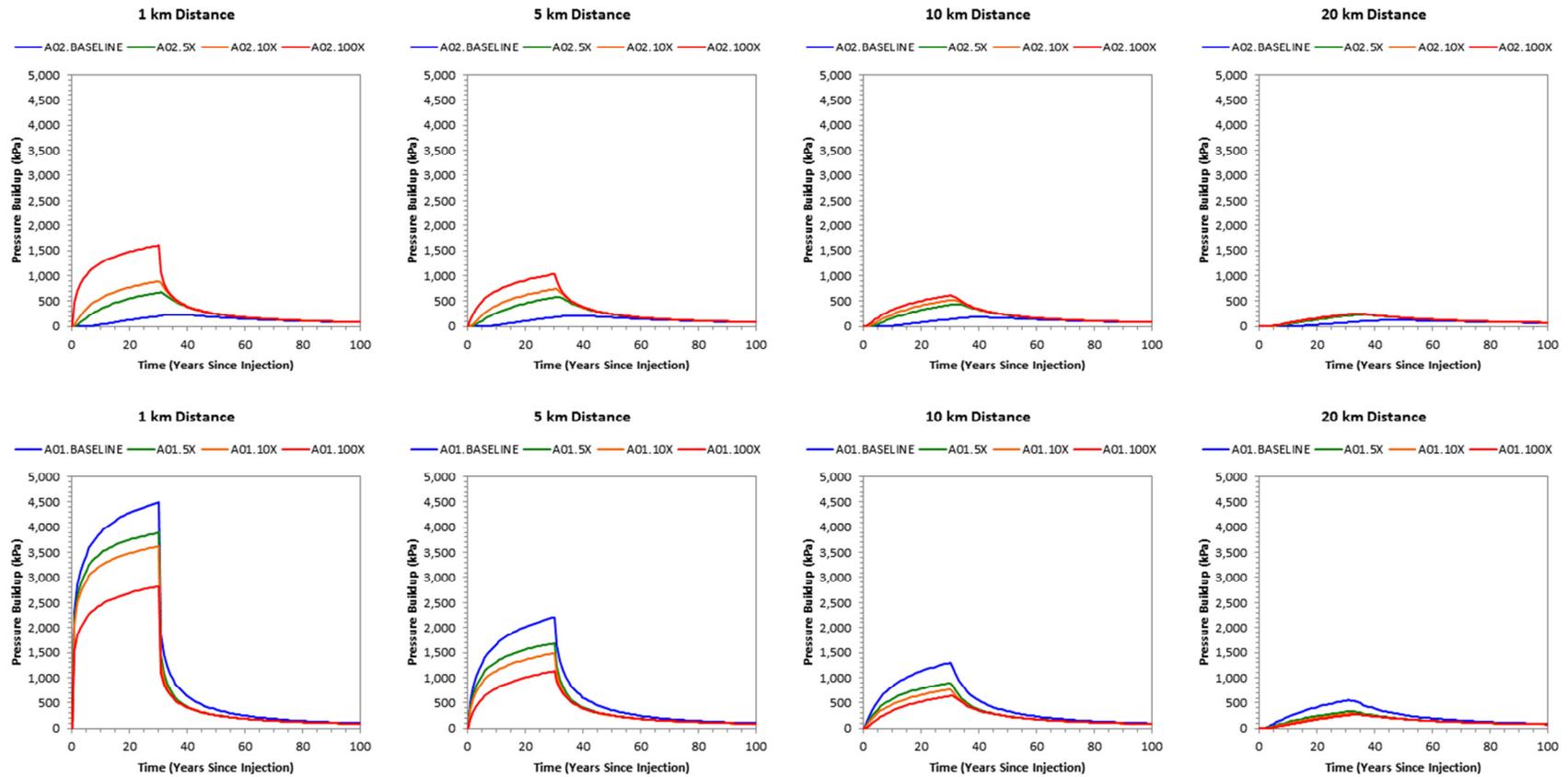


Figure 2-4. Pressure build-up above hydrostatic pressure in Aquifer 1 (bottom) and Aquifer 2 (top) as a function of time and distance from the CO₂ injection wells. Baseline (blue), 5x (green), 10x (orange), and 100x (red) show the relative change in pressure build-up as a function of the cap rock permeability.

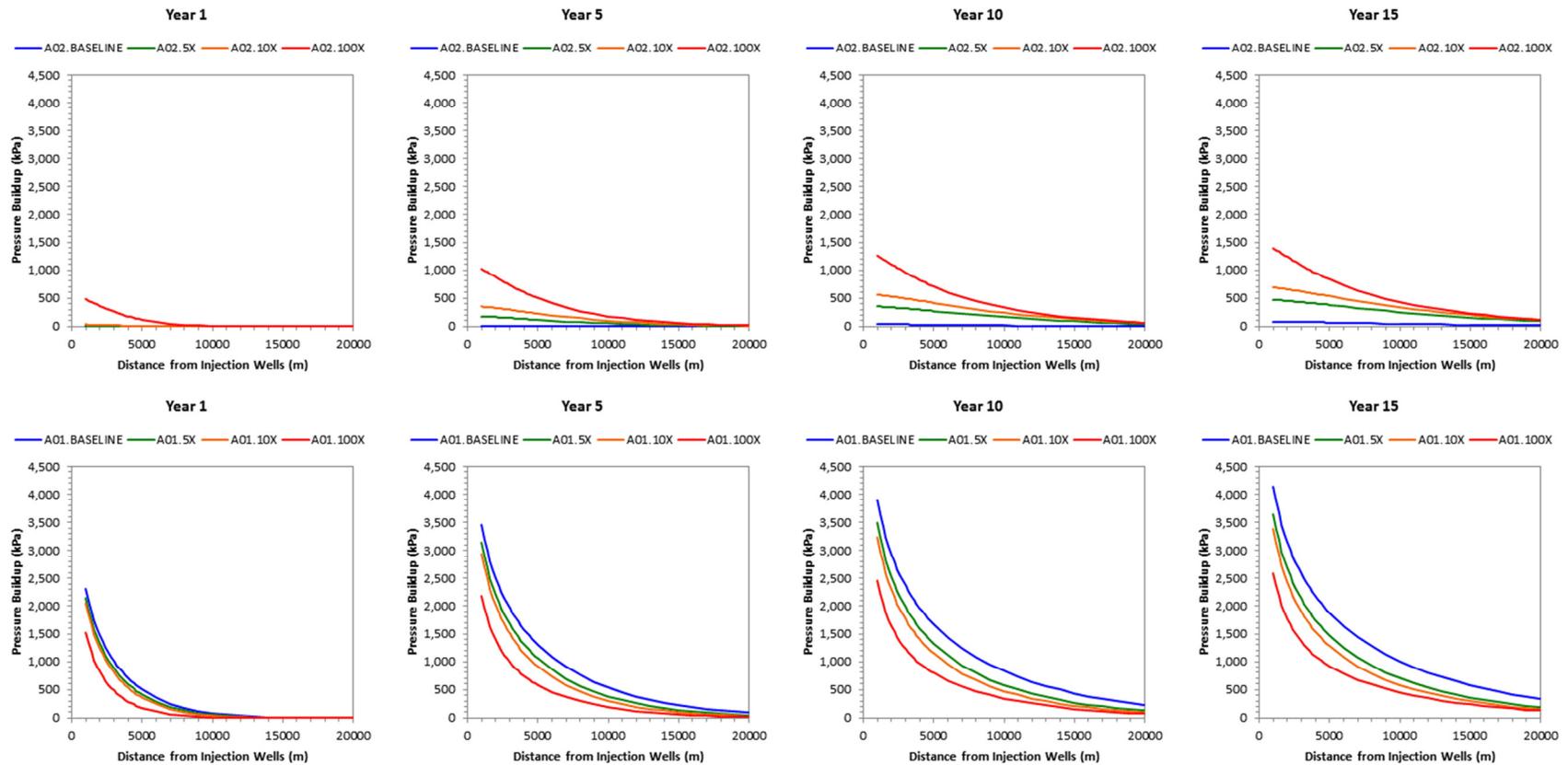


Figure 2-5. Pressure build-up above hydrostatic pressure in Aquifer 1 (bottom) and Aquifer 2 (top) as a function of distance from the CO₂ injection wells at a specified time. Baseline (blue), 5x (green), 10x (orange), and 100x (red) show the relative change in pressure build-up as a function of the cap rock permeability.

The ratio of pressure build-up in Aquifer 2/Aquifer 1 (A_{02}/A_{01}) was also evaluated to leverage the inverse correlation between pressure decrease in Aquifer 1 relative to baseline and pressure increase in Aquifer 2 relative to baseline as a function of leakage rate. Presumably, the difference in the ratio would be greater than the absolute change in either aquifer, thereby driving detection sensitivity lower. Comparisons of the A_{02}/A_{01} ratio over time and distance from the injection wells shows that the ratio differentiates between the baseline and leakage scenarios; however, the degree of separation is a function of both time and distance (Figure 2-6). At early time, when pressure build-up in Aquifer 2 is less pronounced, the A_{02}/A_{01} ratio is driven almost entirely by the pressure decrease in Aquifer 1. However, as time progresses the A_{02}/A_{01} ratio increases more rapidly until reaching a near-asymptote after ten years of injection.

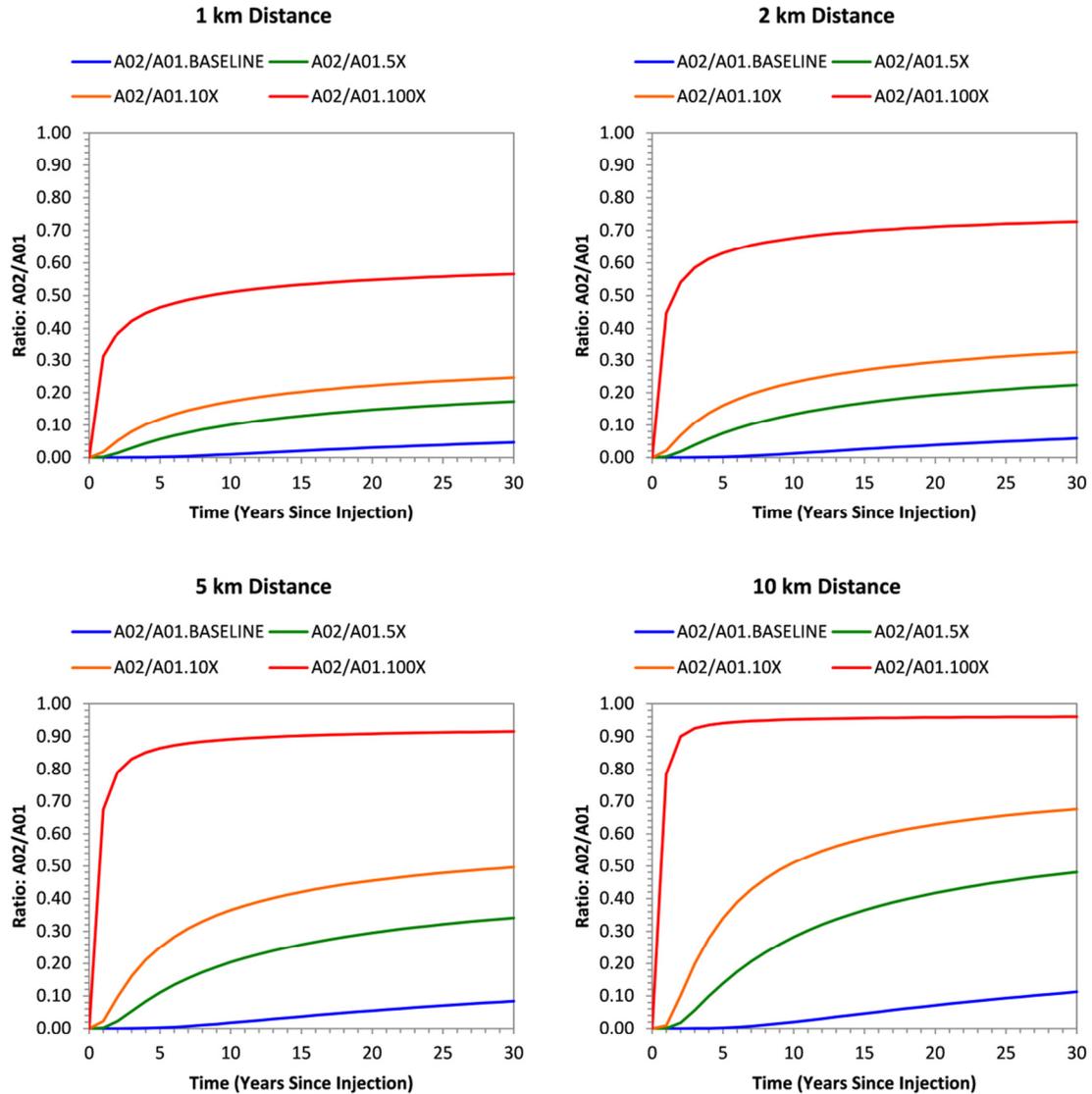


Figure 2-6. Comparison of the ratio of pressure build-up in Aquifer 2/Aquifer 1 (A02/A01) as a function of time since injection and lateral distance from the injection wells. Throughout the injection period, the A02/A01 ratio baseline (blue) < 10x (green) < 100x (red) scenarios, suggesting that this ratio would provide greater detection sensitivity than measurements of absolute change in pressure within either unit.

2.3.2 Uncertainty in Pressure Response

The nominal model output pressures, $P^*(x, z, t)$, described above indicate the general (and perhaps most likely) behavior of the system, but do not yet account for uncertainty in the model input parameters arising from the uncertainty in characteristics

like reservoir or aquitard hydraulic conductivity, specific storage, and thickness, which all contribute to the pressure response. The following comparisons take into account the uncertainty variance in pressure output and the consequence of this uncertainty on differentiating a leakage scenario from baseline conditions.

The magnitude of the variability in $P(x, z, t)$ within an aquifer was a function of leakage rate, distance from the injection wells, and time. For example, the baseline model output $\ln(\text{pressures})$ varied more than those of the leaky scenarios, with variance baseline $> 2x > \dots > 100x$. The uncertainty was greatest at the 1-km distance, and least at the 20-km distance. Lastly, the pressure uncertainty was lowest at early times, and increased throughout the 30-year injection period, i.e., variance 30 yr $>$ 20 yr $>$... $>$ 1 yr.

Aquifer 1 Pressure

Overlays of the CDFs for pressure in Aquifer 1 for different model scenarios show substantial overlap among the pressure distribution for the baseline scenario and the 2x through 50x leakage scenarios (Figure 2-7). For example, the P_{01} for baseline in Aquifer 1 at 1 km after one year of injection was 13,430 kPa (formation pressure). The $\text{Prob}[P(1 \text{ km}, A01, 1 \text{ yr}) < P_{01}]$ for the 2x, 5x, 10x, 20x, 50x, and 100x scenarios were 0.016, 0.036, 0.118, 0.218, 0.754, and 0.994, respectively. Therefore, a pressure decrease within Aquifer 1 resulting from higher leakages through the cap rock than presumed under the baseline scenario lack sufficient statistical power for differentiation from baseline for the 2x through 50x scenarios. In contrast, a leakage rate of 100x produced probability distributions that were significantly lower than baseline (Figure 2-7), with resulting power exceeding 0.99.

Figure 2-7 illustrates several trends with respect to the interplay among x , z , t , and detection sensitivity within Aquifer 1. First, the overlap in baseline and leakage-impacted CDFs decreases with time, and as injection continues the ability to discriminate smaller leakage sizes from baseline improves. For example, at 1 km during years 1, 5, 10, and 15, the probabilities that the Aquifer 1 pressure in the 30x scenario is less than P_{01} from baseline were 0.404, 0.763, 0.853, and 0.891, respectively. Therefore, at 15 years into the injection program a leakage scenario of 30x could be differentiated from baseline with approximately 90% confidence, as opposed to only 40% confidence within the first year of injection. A second trend is the relationship between uncertain variance and distance from the injection wells. Figure 2-7 shows nearly vertical CDFs at 3 km in the first year, because this region does not experience significant pressure build-up, and therefore the formation pressure at this location and time is nearly hydrostatic. Conversely, after 10 years of injection, there is nearly 1000 kPa uncertainty in the formation pressure at this location. These two trends highlight the importance of time and distance in the design of an optimal monitoring network configuration; the detection sensitivity is not a static value at any given point within the storage reservoir.

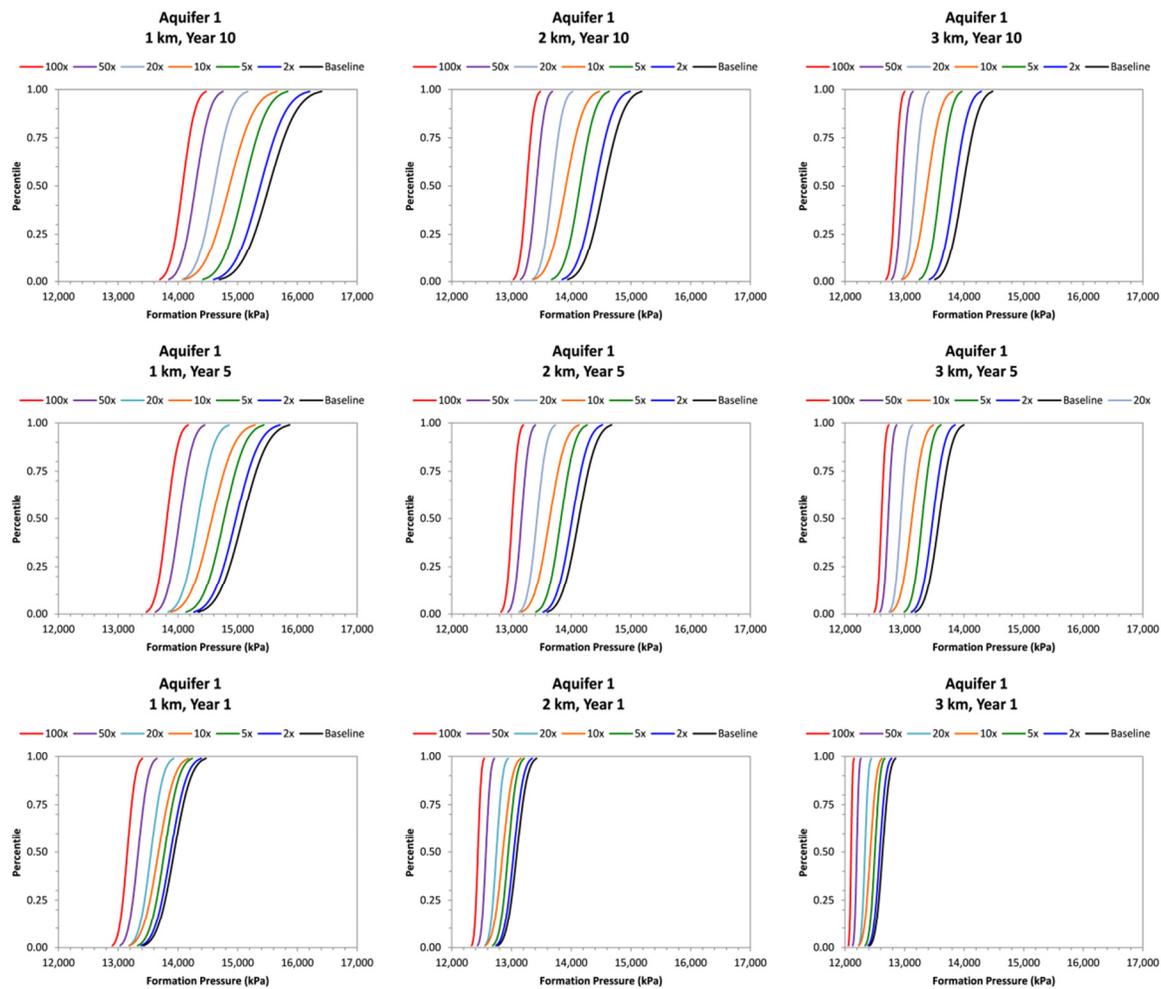


Figure 2-7. Cumulative distribution functions for baseline, 2x, 5x, 10x, 20x, 50x, and 100x scenarios within Aquifer 1 as a function of distance from the injection wells at one year (bottom), five years (middle), and 10 years (top) after injection.

The nominal, P_{01} , and P_{99} responses for Aquifer 1 at 1 km for the baseline, 10x, and 100x scenarios are shown for the 30-year injection period in Figure 2-8. Additional comparisons are shown in Figures 2-9, 2-10, and 2-11. The time-series plots illustrate several of the aforementioned trends. For example, the width of the interval from the P_{01} to the P_{99} is widest for baseline and smallest for the 100x scenario. In addition, at no time during the 30-year period at 1 km, 2 km, or 3 km from the injection wells are the 5x or 10x scenarios clearly distinguishable from the baseline scenario. At 1 km, the 50x scenario is not distinguishable from baseline, but at 2 km and 3 km there is sufficient separation between the baseline P_{01} and the 50x distribution to indicate likely inference of leakage. Lastly, at all times and distances, the 100x scenario is clearly different from baseline, so detection is very probable.

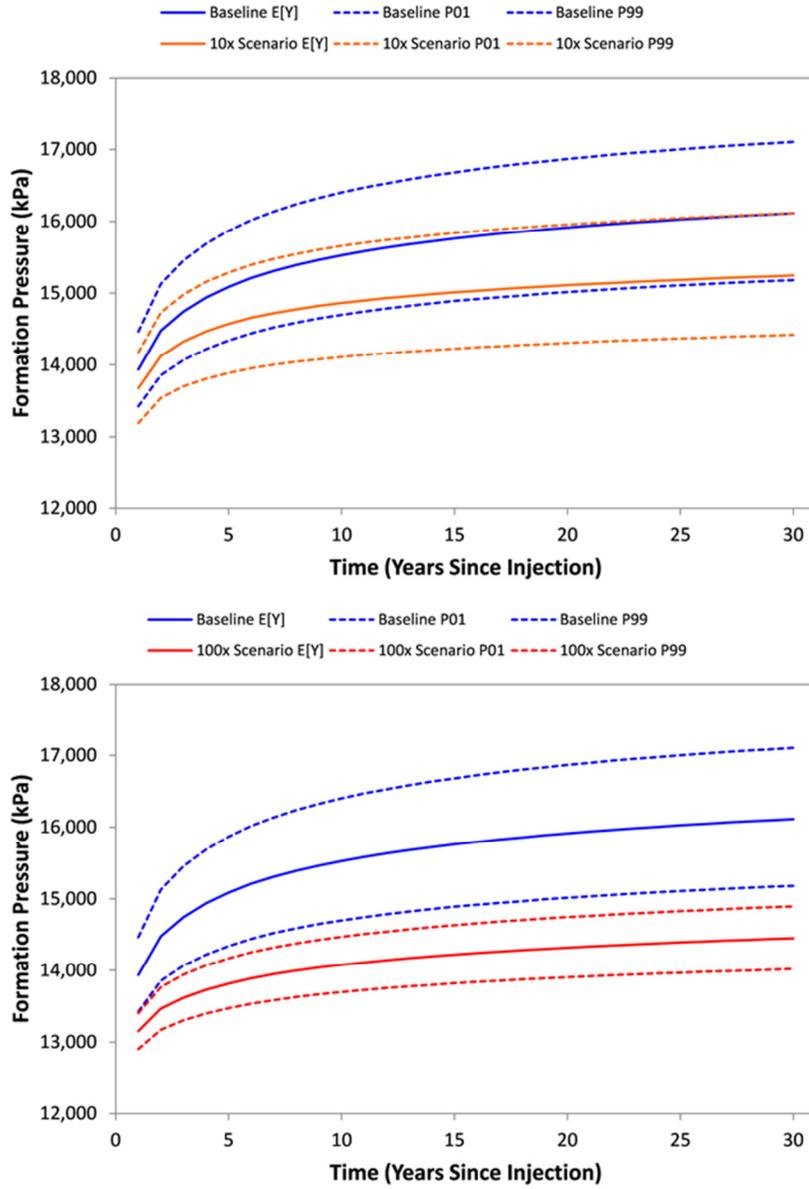


Figure 2-8. Comparison of the nominal response at 1km from the injection wells for baseline bounded by the 1st and 99th percentiles to the 10x leakage scenario (top) and 100x leakage scenario (bottom).

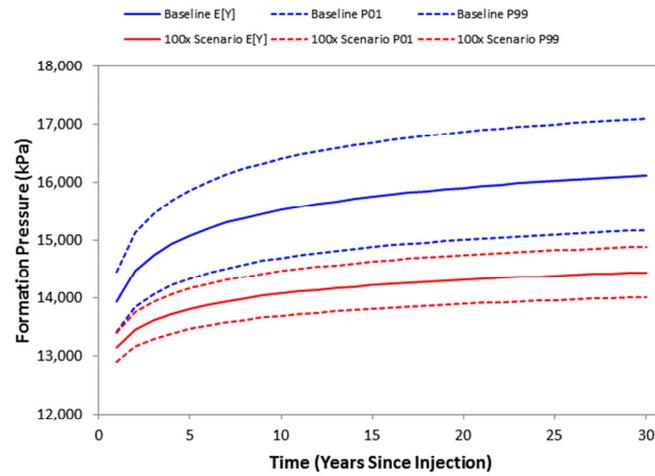
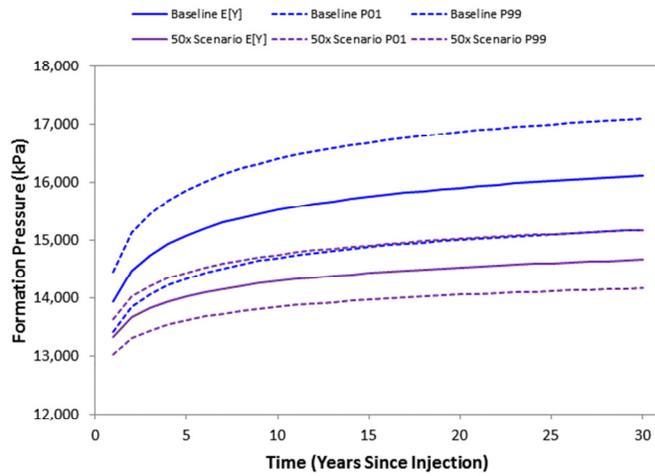
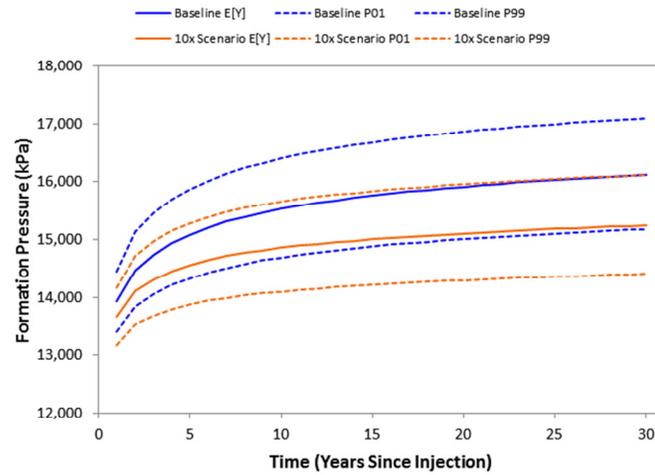
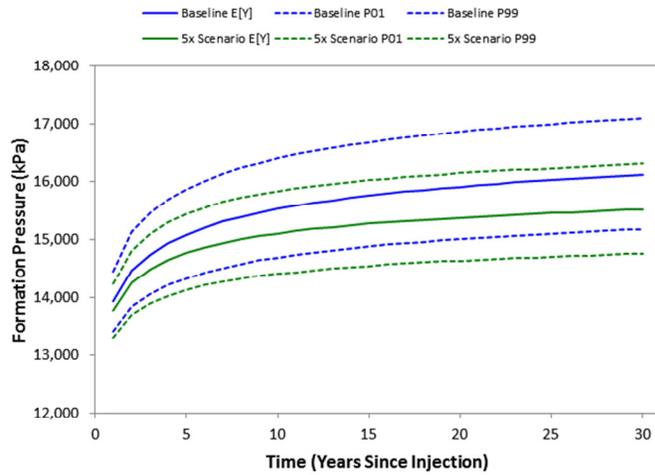


Figure 2-9. Comparison of the nominal response at 1 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).

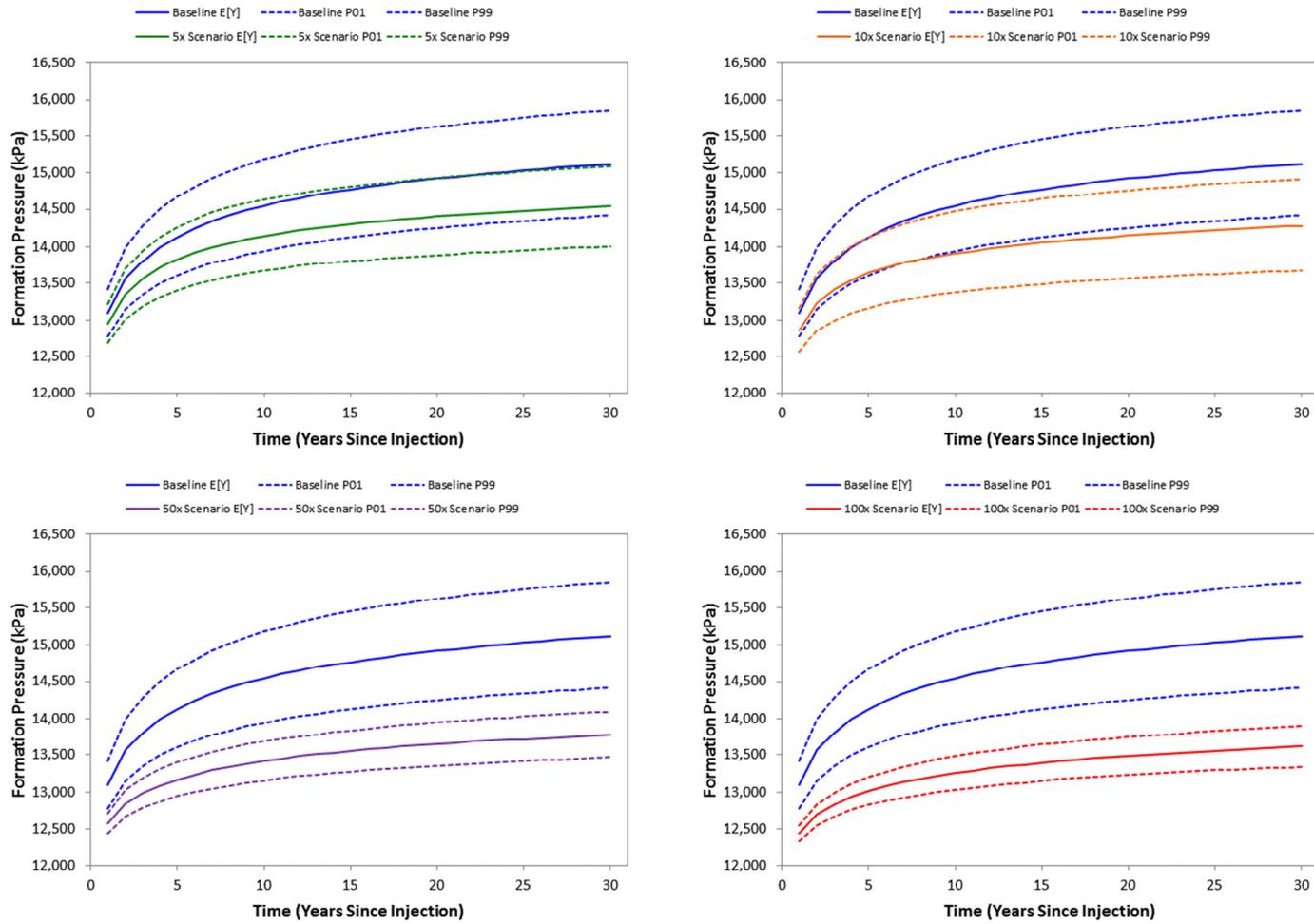


Figure 2-10. Comparison of the nominal response at 2 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).

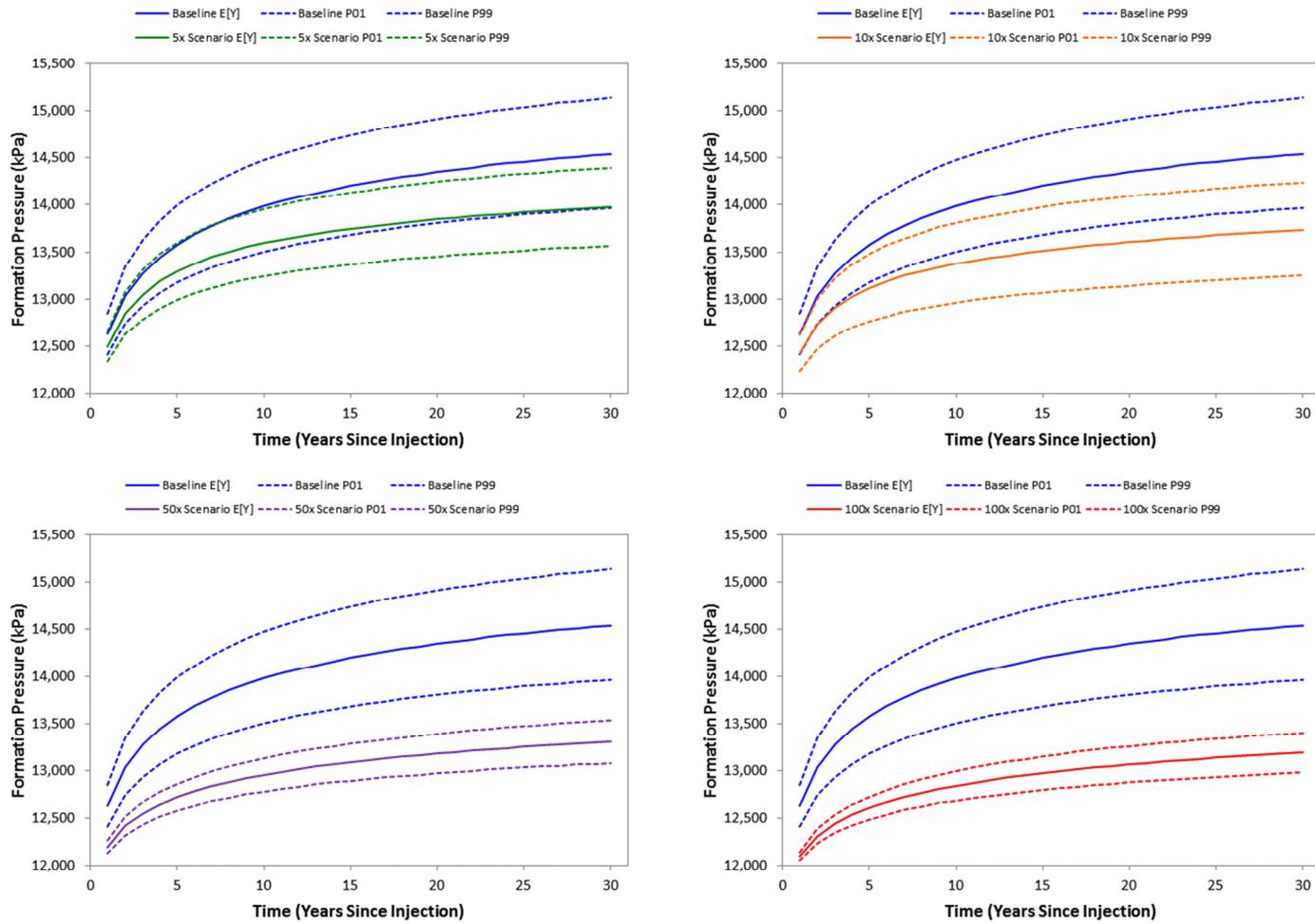


Figure 2-11. Comparison of the nominal response at 3 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).

Ratio Aquifer 2/Aquifer 1 (A02/A01)

The A02/A01 ratio increases as a function of leakage rate, as the pressure in A02 increases and the pressure in A01 decreases relative to baseline. Therefore, comparisons were made to the P_{99} A02/A01 ratio value for baseline (in contrast to the P_{01} for baseline pressure in Aquifer 1). At early times, the A02/A01 ratio is effectively the hydrostatic formation pressure in A02 divided by the decreasing formation pressure in A01. As a result, initially the ratio does not provide added differentiation among leakage scenarios beyond the pressure decrease in A01. However, with time, as the difference in pressure between A01 and A02 increases, the A02/A01 ratio provides more powerful detection of leakage (i.e., greater statistical power).

Overlays of the CDFs for A02/A01 for different model scenarios show substantial overlap between the baseline scenario and the 2x through 20x leakage scenarios in the first year; however, the separation improves for years 5 and 10 (Figure 2-12). The $\text{Prob}[P(1 \text{ km}, A02/A01, 1 \text{ yr}) > P_{99}]$ for the 2x, 5x, 10x, 20x, 50x, and 100x scenarios were 0.016, 0.038, 0.168, 0.605, >0.999, and >0.999, respectively. However, at five years the probabilities were 0.026, 0.244, 0.819, >0.999, >0.999, and >0.999, respectively. Therefore, after five years of injection, the ratio A02/A01 at 1 km is able to differentiate the 20x-leakage scenario from baseline. This is in contrast to the measurement of the absolute change in formation pressure within Aquifer 1 alone, which was not able to distinguish the 20x scenario even after 15 years. These results suggest that combining measurements from two different aquifers provides greater statistical power than the measurement of formation pressure at any one single location.

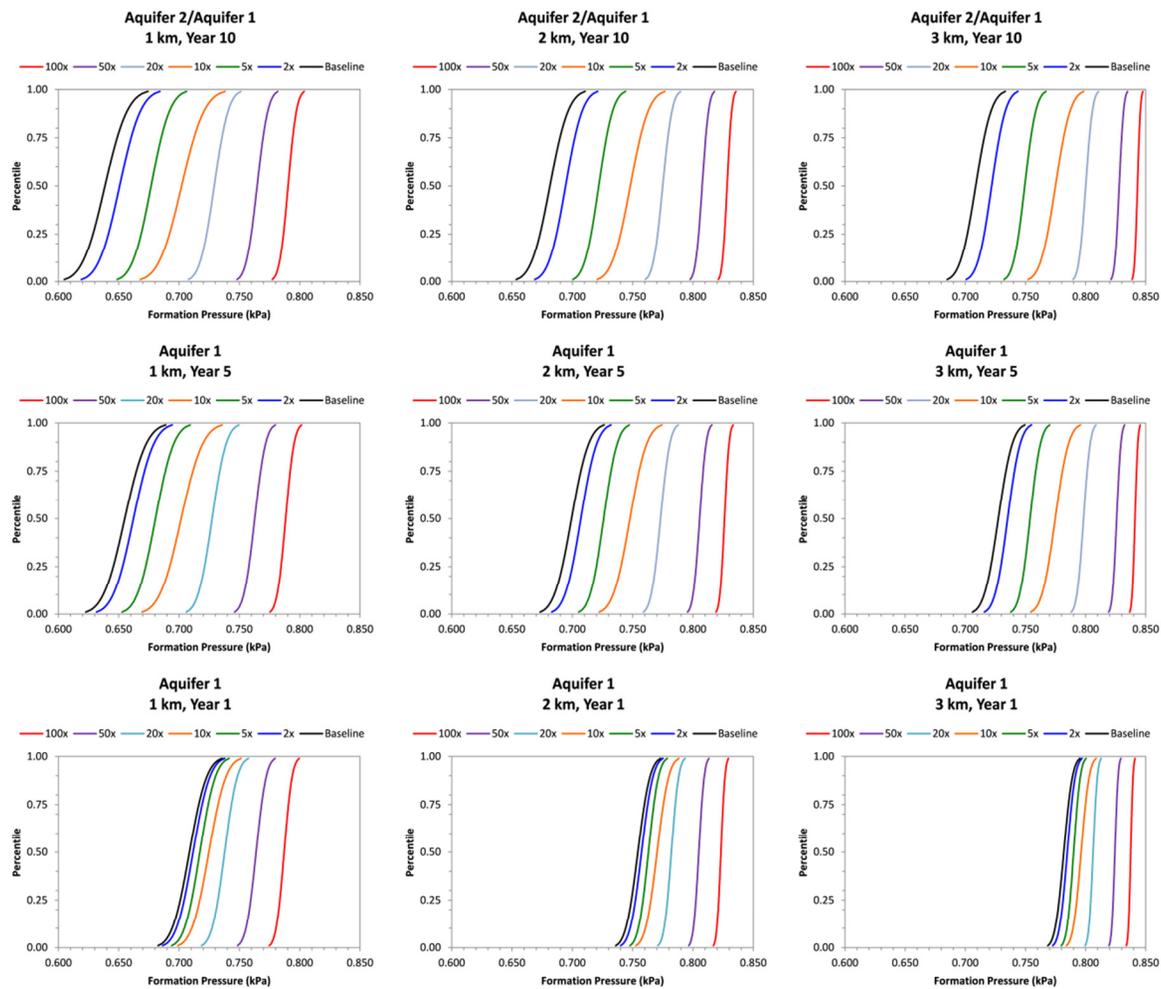


Figure 2-12. Cumulative distribution functions for the ratio A_{02}/A_{01} for baseline, 2x, 5x, 10x, 20x, 50x, and 100x scenarios as a function of distance from the injection wells at one year (bottom), five years (middle), and 10 years (top) after injection.

The nominal, P_{01} , and P_{99} responses for A02/A01 at 1 km for the baseline, 10x, and 100x scenarios are shown for the 30-year injection period in Figure 2-13. Additional comparisons are shown in Figures 2-14, 2-15, and 2-16. The time-series plots show that the ratio A02/A01 is more sensitive for detecting leakage than the measurement of the absolute change in pressure in a single aquifer. For example, at approximately 15 years, the A02/A01 ratio at 1 km distinguishes between the 10x and baseline scenarios, whereas the measurement of pressure decrease at 1 km in Aquifer 1 could not distinguish the 10x scenario from baseline at any point during the 30-year injection period (Figure 2-8).

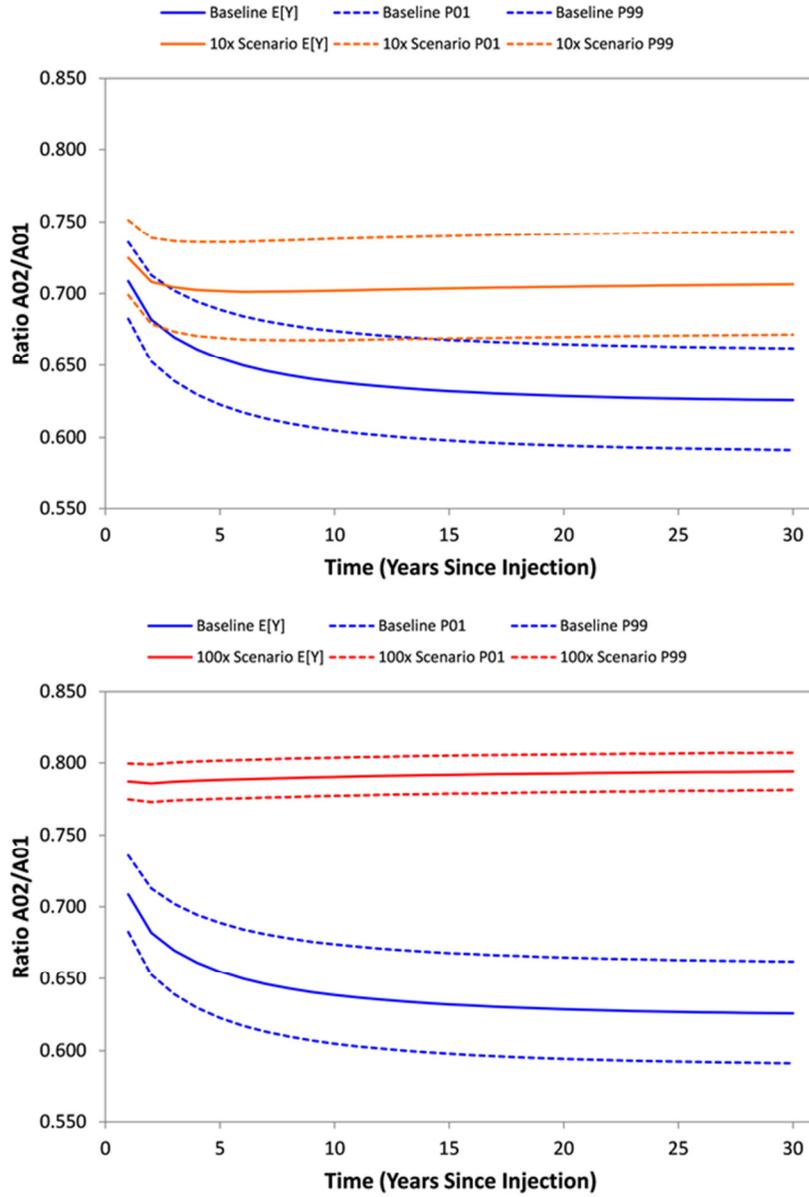


Figure 2-13. Comparison of the nominal response at 1 km from the injection wells for baseline bounded by the 1st and 99th percentiles to the 10x leakage scenario (top) and 100x leakage scenario (bottom).

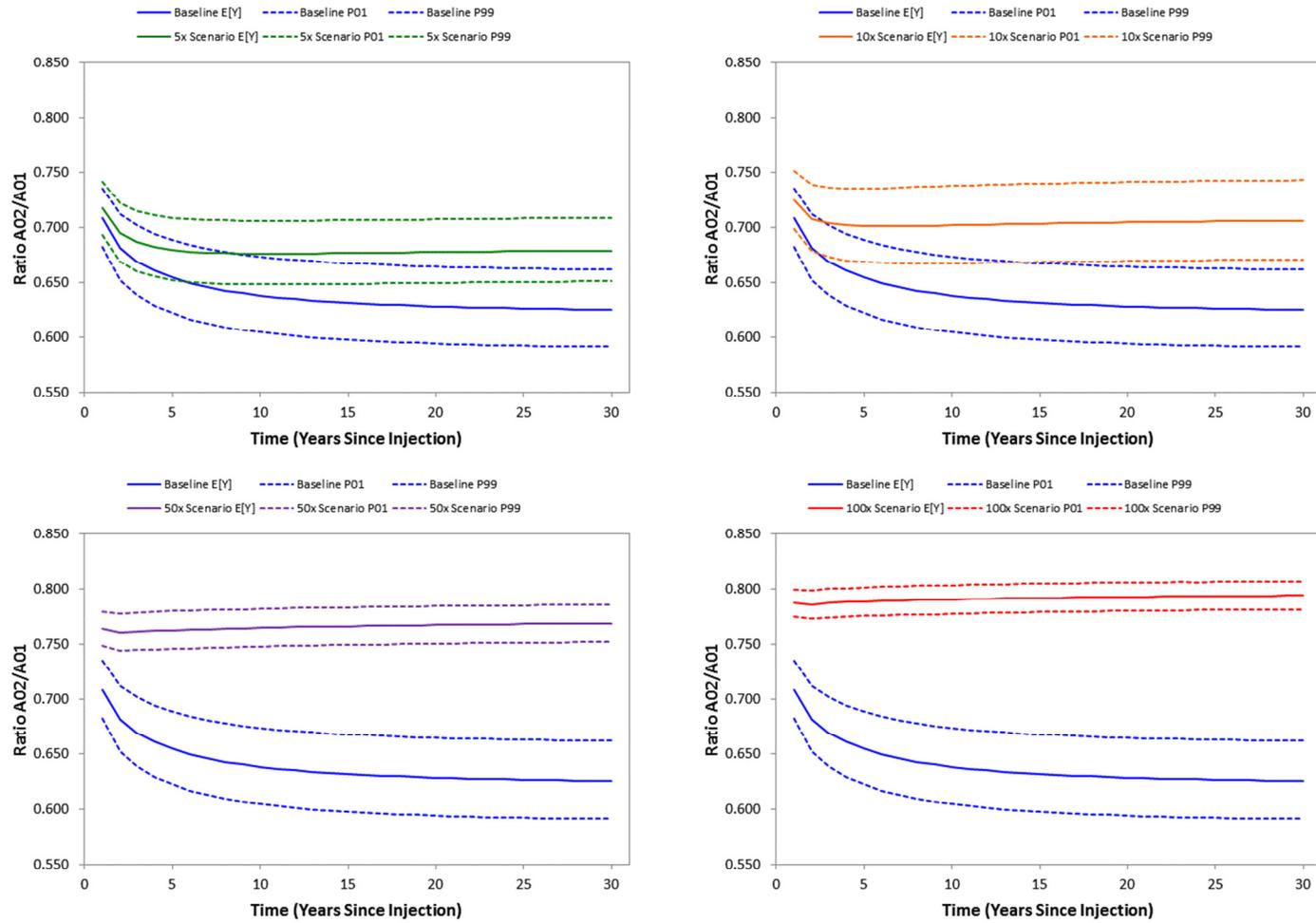


Figure 2-14. Comparison of the nominal response at 1 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).

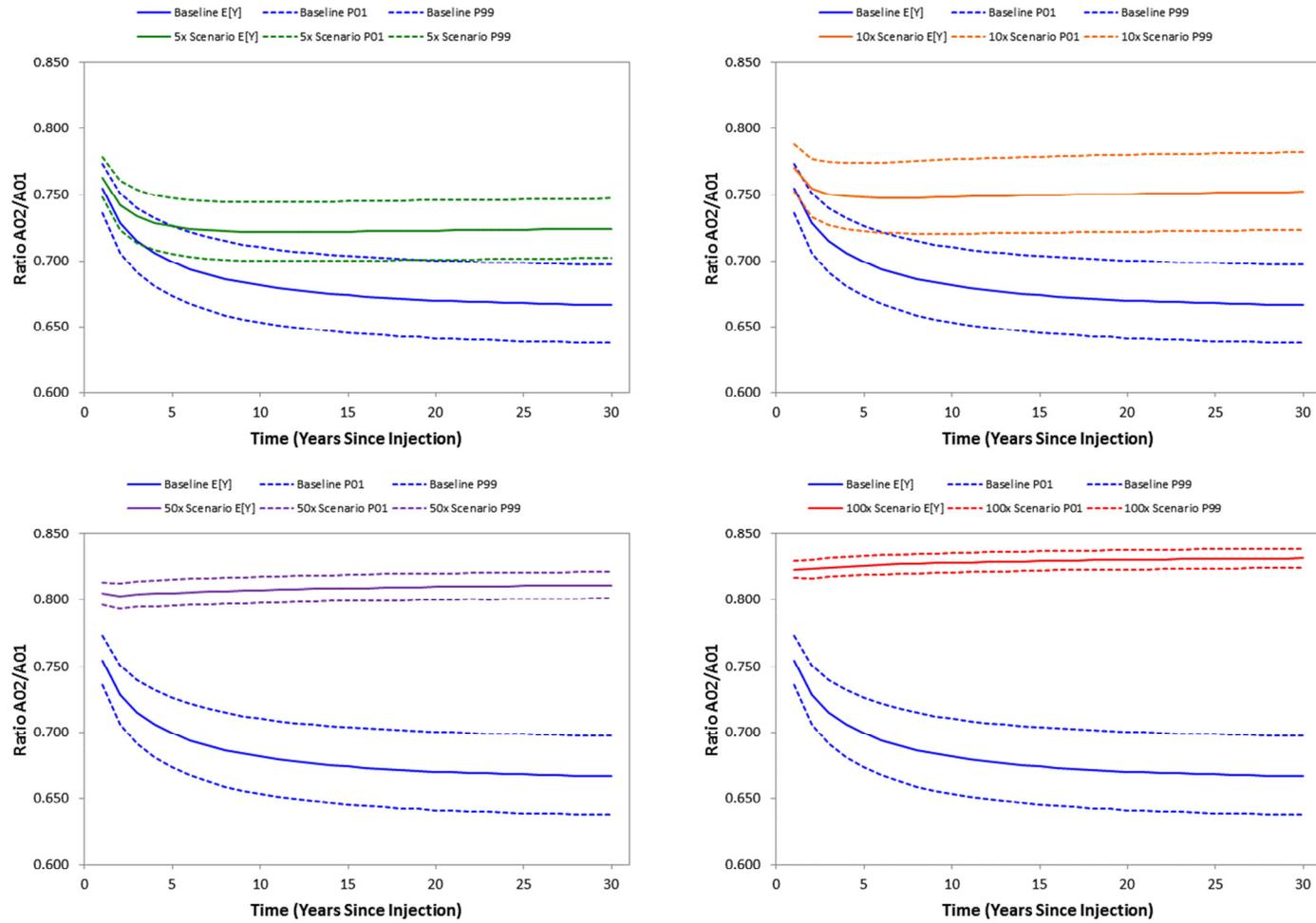


Figure 2-15. Comparison of the nominal response at 2 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).

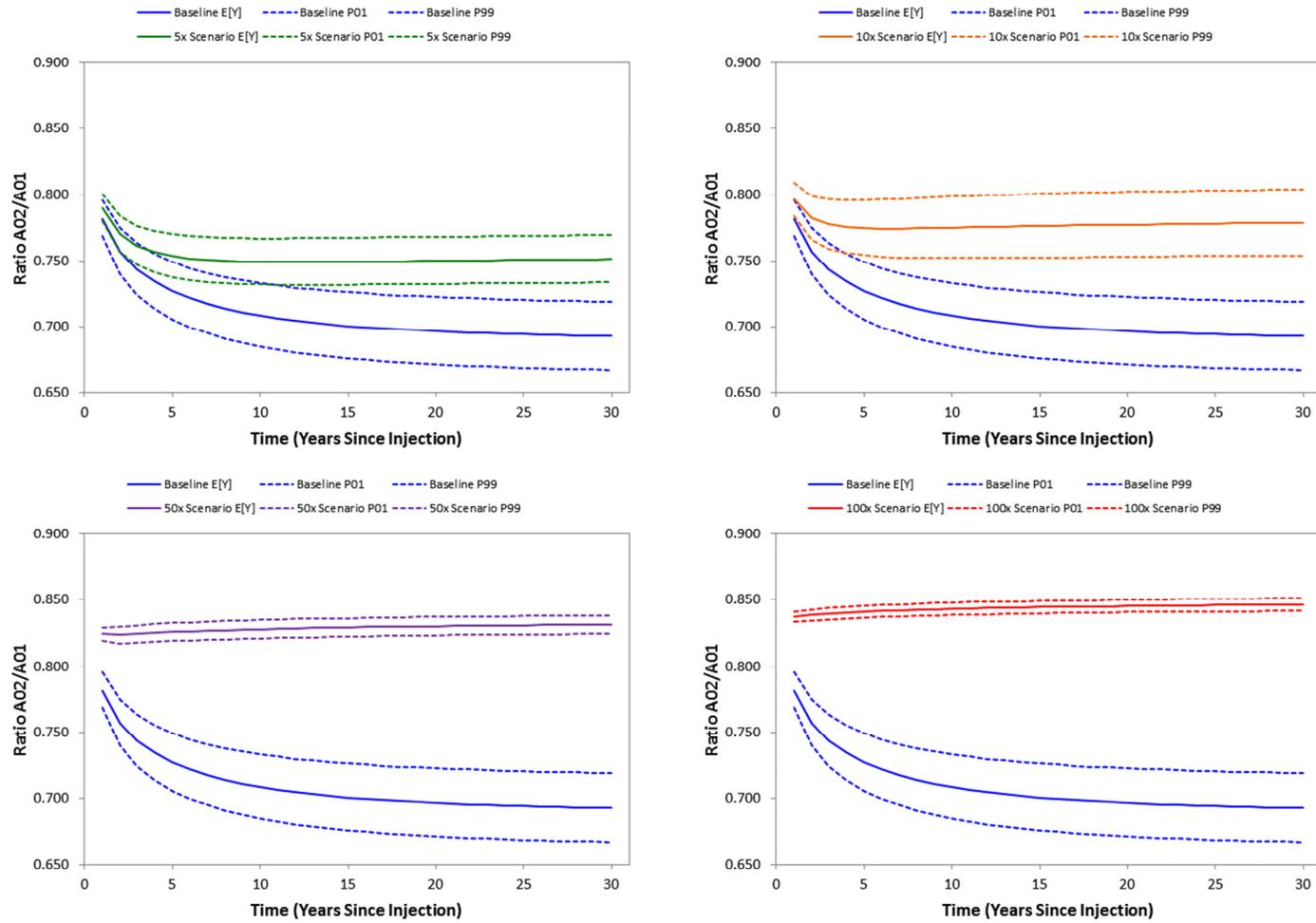


Figure 2-16. Comparison of the nominal response at 3 km from the injection wells for baseline (blue) bounded by the 1st and 99th percentiles to the 5x leakage scenario (green, upper left), 10x leakage scenario (orange, upper right), 50x leakage scenario (purple, lower left), and 100x leakage scenario (red, lower right).

2.3.3 Power Curves

Power curves showing the probability of detecting either (1) pressure less than baseline conditions for Aquifer 1 or (2) a ratio A_{02}/A_{01} greater than baseline conditions are shown in Figure 2-17. The comparisons show that in the first year, both measurements perform similarly for detecting leakage of a specific magnitude. However, as time progresses, the A_{02}/A_{01} ratio outperforms the measurement of pressure only in Aquifer 1. For example, at five years the A_{02}/A_{01} ratio provides greater than 95% probability of detecting a 10x leakage scenario at 2-5 km from the injection wells, whereas the measurement of pressure solely within Aquifer 1 has less than 90% probability for all but the 5-km distance. At 10 years, the A_{02}/A_{01} ratio can distinguish a 5x leakage scenario from baseline at 3-5 km with greater than 95% probability, while the measurement of pressure solely within Aquifer 1 has less than 50% probability of detecting 5x leakage at all distances.

The power curves show that the power for detection increases with further distance from the point of injection (maximum power at 5 km). This result is counterintuitive, since one in general expects measurements near the point of injection to be more sensitive. However, unlike direct measurements of CO_2 , which would likely increase sensitivity closer to the injection wells, pressure changes are more variable near the injection wells than further away. Beyond a certain distance, the pressure remains at or near hydrostatic through time, and therefore at such distances there is little information to be gained from monitoring. Between this distance and the injector lies an optimal point which varies through time where detection sensitivity to changing pressure conditions is greatest.

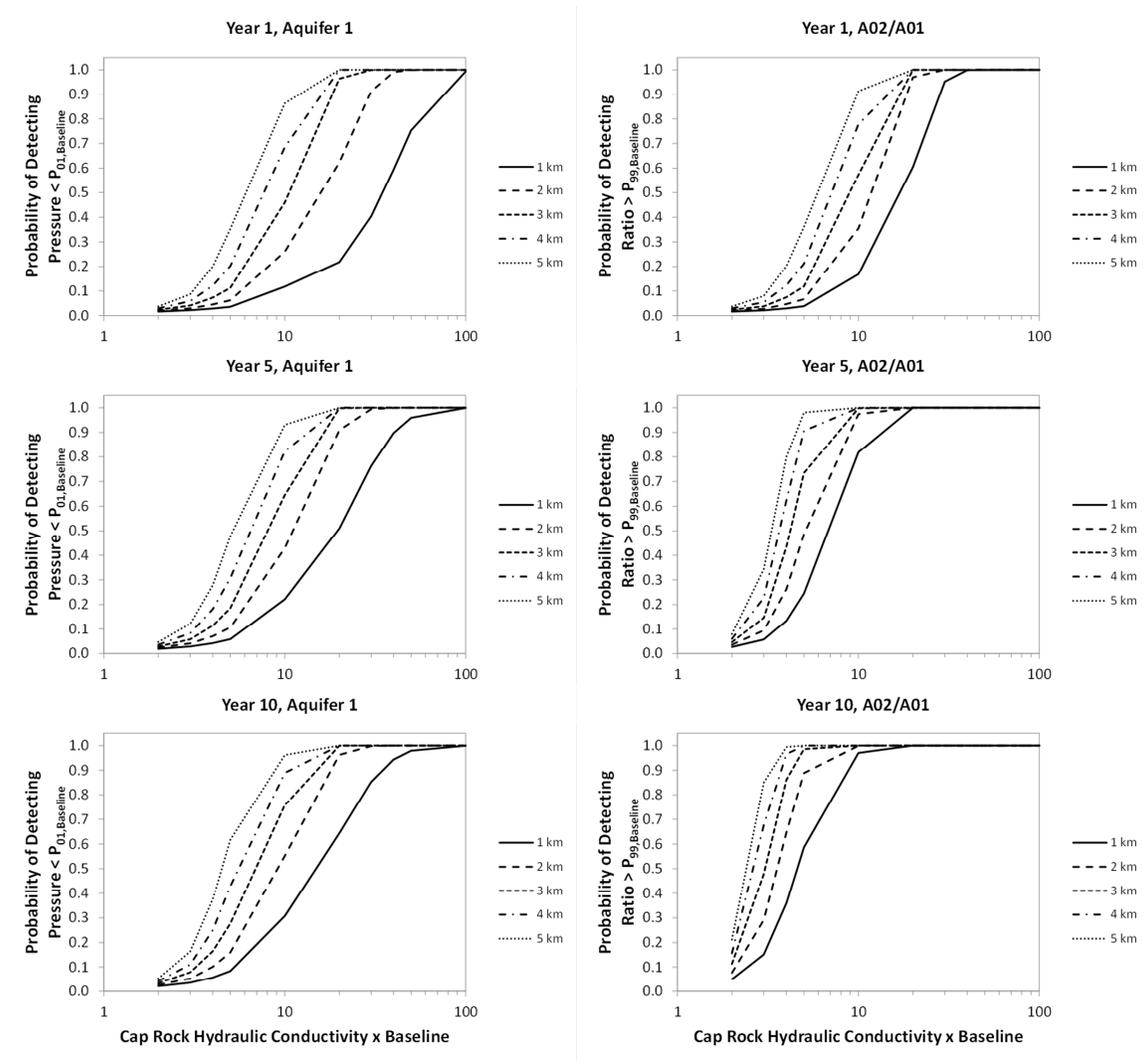


Figure 2-17. Power curves showing the probability of detecting pressure less than baseline conditions in Aquifer 1 (left column) or an A02/A01 ratio greater than baseline conditions (right column) at years 1, 5, and 10 of the injection program. The x-axis shows the factor at which the baseline permeability of the cap rock was multiplied for the leakage scenario, i.e., $x=10$ refers to the 10x leakage scenario.

2.4 Sensitivity to Model Assumptions

A number of assumptions were made in executing this work in addition to those assumptions already incorporated into the solution approach for the LBNL Model itself (e.g., single-phase, homogeneous aquifers and aquitards of constant thickness, etc.). First, this work assumes a lognormal distribution for pressure, and utilizes the nominal value as the lognormal mean and the second moment (variance) from the FOIA process to derive

the lognormal standard deviation. The lognormal distribution is justifiable due to the multiplicative effects incorporated into the analytical solutions of the LBNL Model, as the product of many independent, identically distributed, positive random variables has approximately a log-normal distribution. In addition, many of the input parameters are themselves log-normally distributed; therefore, the output distribution is also lognormal (Limpert et al., 2001). For example, the equation used for the diffuse leakage rate (specific discharge) from an aquifer to an overlying aquitard (w_i) as part of the governing equations used to develop the analytical solution in the LBNL Model is (Cihan et al., 2011):

$$w_i = -\frac{K_i}{B_i} \frac{\partial s_i}{\partial z_{Di}} \Big|_{z_{Di}=0} \quad (2-11)$$

Where K_i is the hydraulic conductivity, B_i is the aquifer thickness, s_i is the specific storativity of the aquifer, and z_{Di} is the dimensionless local vertical coordinate. As noted above in the Methods section, the hydraulic conductivity is lognormal and therefore w_i will be log-normally distributed. Other governing equations used in the LBNL Model have a similar form. In addition to more appropriately fitting the output distribution, the lognormal distribution prevents the output pressure values from going below zero. In contrast, the use of the normal distribution, particularly for outputs with significant variance, permits the lower percentile values (e.g., P_{01}) to be negative. This pressure response is not feasible, as CO_2 -equivalent brine volumes are being injected into the system and the resulting change in pressure must therefore be positive.

A second assumption was the uncertainty in the input parameters, which directly affects the S_x matrix (Equation 2-8) and hence, the output covariance matrix, S_y (Equation 2-9). The nominal case assumed that the median values for hydraulic conductivity and specific storage were known to within plus-or-minus an order-of-magnitude, and therefore established a 100-fold variability between the 0.1th and 99.9th percentiles of the lognormal distribution. Aquifer and aquitard thickness were assumed, somewhat arbitrarily, to be known to within plus-or-minus 10 m. To explore the sensitivity of the statistical power curves to these assumptions for input parameter uncertainty, two end-member scenarios were modeled and the resulting statistical power curves were compared. The first end-member represents a less uncertain system, with only a 10-fold variability between the 0.1th and 99.9th percentiles of the lognormal distribution for hydraulic conductivity and specific storage, and aquifer and aquitard thickness were assumed to be known within plus-or-minus 5 m. The second end-member represents a more uncertain system, with a 1000-fold variability between the 0.1th and 99.9th percentiles of the lognormal distribution for hydraulic conductivity and specific storage, and aquifer and aquitard thickness were assumed to be known within plus-or-minus 20 m. Comparisons of the statistical power curves for the three different input parameter uncertainty scenarios: (1) less uncertain (10-fold case); (2) nominal uncertainty (100-fold case); and (3) highly uncertain (1000-fold case) are shown in Figure 2-18 for the 1-, 2- and 3-km lateral distances at 1, 5, and 10 years following injection. The results show, as expected, that the less uncertain model has greater statistical power and the more uncertain model has less. However, the results are somewhat robust and the 10-fold differences in the input parameter uncertainty do not result in 10-fold difference in

statistical power. For example, after five years of injection and 1 km from the injection wells, the probability of detecting pressure less than baseline conditions for Aquifer 1 under the 100-fold uncertainty model is 0.221 for the 10x scenario. For the 10-fold uncertainty and 1000-fold uncertainty models, the probabilities are 0.413 and 0.105, respectively. Therefore, for this particular example, the 10-fold difference in the input parameter uncertainty results in a doubling (or halving) of the statistical power. The magnitude of the change in statistical power is a function of time since injection and lateral distance from the injection well. For example, after five years of injection and 3 km from the injection wells, the probability of detecting pressure less than baseline conditions for Aquifer 1 under the 100-fold uncertainty model is 0.645 for the 10x scenario. For the 10-fold uncertainty and 1000-fold uncertainty models, the probabilities are 0.973 and 0.442, respectively. Therefore, at this greater distance from the injection wells, the 10-fold difference in the input parameter uncertainty results in a 1.5 times greater (or less) statistical power.

Lastly, the FOUA process assumes a local linear approximation for the model output to variations in the input parameters. A CO₂ storage system may exhibit nonlinear behavior, in which case the FOUA estimates would be inadequate. A Monte Carlo analysis of uncertainty would be better-suited for estimating nonlinear behavior, and for providing estimates of the full probability distribution function for the model outputs (Ramaswami et al., 2005). However, a Monte Carlo approach would be computationally more intensive, and would likely require thousands of simulations.

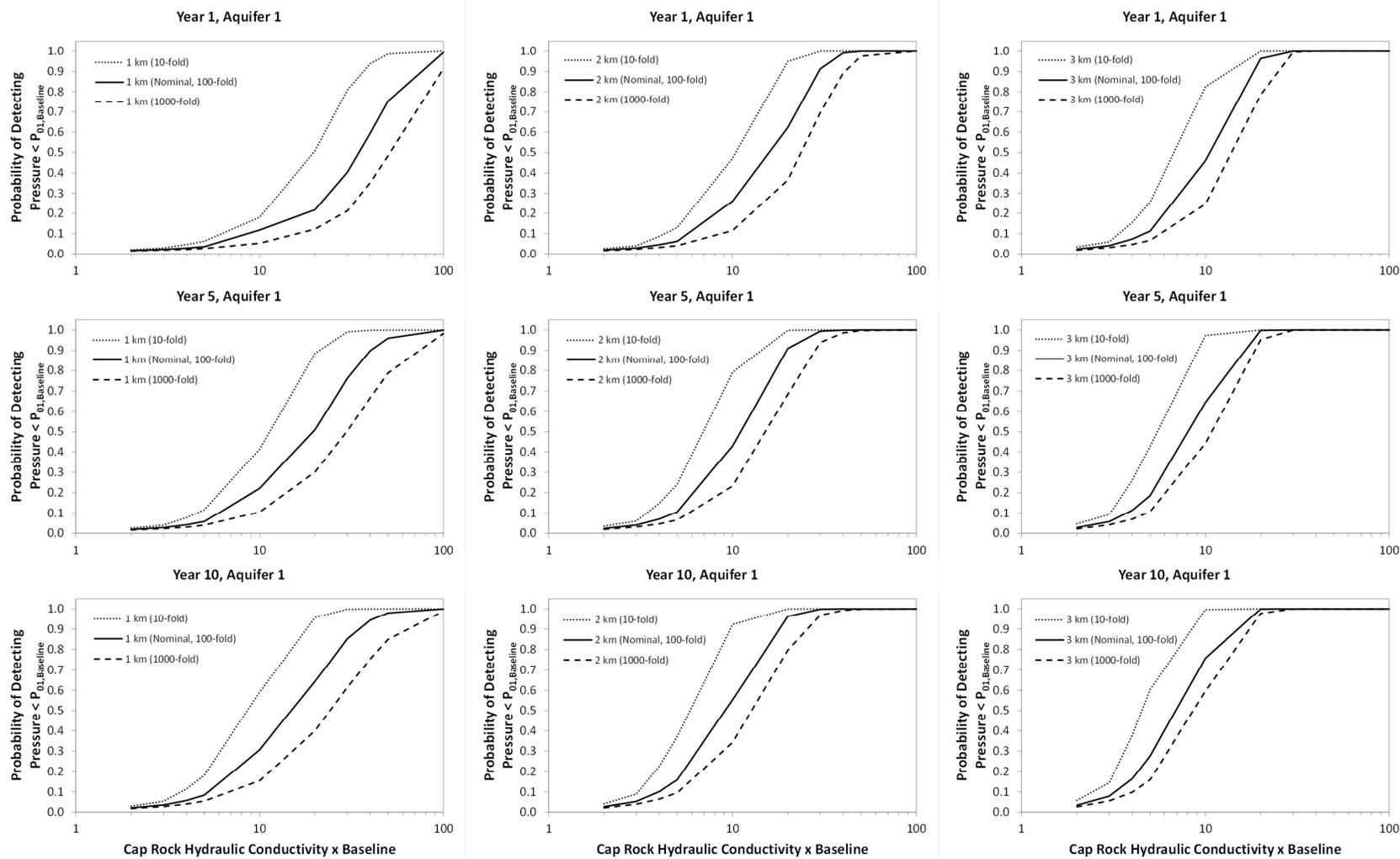


Figure 2-18. Comparisons of the statistical power curves for three different scenarios of input parameter uncertainty: (1) less uncertain (10-fold case) [dotted line]; (2) nominal uncertainty (100-fold case) [solid line]; and (3) highly uncertain (1000-fold case) [dashed line] for the 1-km (left column), 2-km (middle column) and 3-km (right column) lateral distances at 1 year (top row), 5 years (middle row), and 10 years (bottom row) following injection..

2.5 Conclusions

This work shows the value of a first-order approach for estimating uncertainty, and the importance of uncertainty in the design of a reliable monitoring network. The results suggest that measurements of the absolute change in pressure within the target injection aquifer would not be able to distinguish small leakage rates (i.e., 2x to 50x) from baseline conditions, and that only large leakage rates (i.e., >100x) would be discriminated with sufficient statistical power (probability $\geq 99\%$), for the monitoring strategy assumed here. Combining measurements, for example by taking the ratio of formation pressure in Aquifer 2/Aquifer 1, provides better statistical power for distinguishing smaller leakage rates at earlier times in the injection program. Detection sensitivity for pressure is a function of lateral distance from the injection wells, vertical position within the system, and time i.e., (x, z, t) . Therefore, design of an adequate monitoring network for subsurface pressure should account for this time-space variability to ensure that the monitoring system performs to the necessary design criteria, e.g., a specific false-negative and false-positive rate. The FOUA process demonstrated here required 10 nominal model runs and $180 \pm \Delta x$ model runs, for a total of 190 model runs to generate the first-order uncertainty assessment. The total simulation time was less than three hours (i.e., less than one minute per run); however, the subsequent data handling and FOUA matrix computations were more time-intensive.

Acknowledgements

This study was funded by the U.S. Department of Energy National Energy Technology Laboratory. The authors would like to thank Dr. Abdullah Cihan of the Ernest Orlando Lawrence Berkeley National Laboratory for sharing both the FORTRAN model and support for running the model over different scenarios.

2.6 References

Apps, J.A.; Zheng, L.; Spycher, N.; Birkholzer, J.T.; Kharaka, Y.; Thordsen, J.; Kakourosb, E.; and Trautz, R. (2011) Transient changes in shallow groundwater chemistry during the MSU ZERT CO₂ injection experiment. *Energy Procedia*, 4, 3231-3238.

Bierkens, M.F.P. (2006) Designing a monitoring network for detecting groundwater pollution with stochastic simulation and a cost model. *Stochastic Environmental Research and Risk Assessment*, 20(5):335-351.

Birkholzer, J.T.; Zhou, Q.; and Tsang, C.F. (2009) Large-scale impact of CO₂ storage in deep saline aquifers: a sensitivity study on the pressure response in stratified systems. *International Journal of Greenhouse Gas Control*, 3(2):181–194.

Cihan, A.; Zhou, Q.; and Birkholzer, J.T. (2011) Analytical solutions for pressure perturbation and fluid leakage through aquitards and wells in multilayered aquifer systems. *Water Resources Research*, 47:W10504. DOI:10.1029/2011WR010721.

Cihan, A.; Birkholzer, J.T.; and Zhou, Q. (2012) Pressure buildup and brine migration during CO₂ storage in multilayered aquifers. *Ground Water*, 51(2):252–267.

Court, B.; Celia, M.A.; Nordbotten, J.M.; Dobossy, M.; Elliot, T.R.; and Bandilla, K. (2011) Modeling Options to Answer Practical Questions for CO₂ Sequestration Operations. Princeton-Bergen Series on Carbon Storage, Retrieved March 22, 2013, from <http://arks.princeton.edu/ark:/88435/dsp01rf55z769f>

CREWES (2012) CREWES Fluid Property Calculator, Accessed March 1, 2013, from <http://www.crewes.org/ResearchLinks/ExplorerPrograms/FIProp/FluidProp.htm>

Dagan, G. (1987) Theory of solute transport by groundwater. *Annual Review of Fluid Mechanics*, 19:183-215.

Damen, K., Faaij, A., and Turkenburg, W. (2006). Health, safety, and environmental risks of underground CO₂ storage – Overview of mechanisms and current knowledge. *Climate Change*, 74 (1-3): 289-318.

DOE (United States Department of Energy) (2009) *Best Practices for: Monitoring, Verification, and Accounting of CO₂ Stored in Deep Geologic Formations*. U.S. Department of Energy, DOE/NETL-311/081508. January.

DOE (2011) *Best Practices for: Risk Analysis and Simulation for Geologic Storage of CO₂*. U.S. Department of Energy, DOE/NETL-2011/1459. March.

Ezekwe, N. (2011) *Petroleum Reservoir Engineering Practice*. Pearson Education, Inc., Boston.

Gelhar, L.W. (1993) *Stochastic Subsurface Hydrology*. Englewood Cliffs, New Jersey, Prentice-Hall, 390 p.

Gilbert, R.O. (1987) *Statistical Methods for Environmental Pollution Monitoring*, John Wiley & Sons, Inc., New York, 320 p.

Horne, R.N. (1995) *Modern Well Test Analysis: A Computer-Aided Approach, Second Edition*. Petroway, Inc., 257 p.

IPCC (2005) *IPCC Special Report on Carbon Dioxide Capture and Storage. Prepared by Working Group III of the Intergovernmental Panel on Climate Change* [Metz, B., O. Davidson, H. C. de Coninck, M. Loos, and L. A. Meyer (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 442 pp.

Kim, K.H. and Lee, K.K. (2007) Optimization of groundwater-monitoring networks for identification of the distribution of a contaminant plume. *Stochastic Environmental Research and Risk Assessment*, 21(6):785-794.

Limpert, E.; Stahel, W.A.; and Abbt, M. (2001) Log-normal distributions across the sciences: Keys and clues. *BioScience*, 51(5):341-352.

Little, M.G. and Jackson, R.B. (2010) Potential impacts of leakage from deep CO₂ geosequestration on overlying freshwater aquifers. *Environmental Science and Technology*, 44, 9225-9232.

Minitab 14 Statistical Software (2010) [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com).

Nicot, J.P.; Oldenburg, C.M.; Bryant, S.L.; and Hovorka, S.D. (2009) Pressure perturbations from geologic carbon sequestration: Area-of-review boundaries and borehole leakage driving forces. *Energy Procedia*, 1(1):47-54.

Nitao, J.J. (1998) *Reference Manual for the NUFT Flow and Transport Code, Version 3.0*. Lawrence Livermore National Laboratory, UCRL-MA-130651-REV-1.

Nogues, J.P.; Nordbotten, J.M.; and Celia, M.A. (2011) Detecting leakage of brine or CO₂ through abandoned wells in a geological sequestration operation using pressure monitoring wells. *Energy Procedia*, 4:3620-3627.

Nordbotten, J.; Celia, M.; and Bachu, S. (2004) Analytical solutions for leakage rates through abandoned wells. *Water Resources Research*, 40(4):.

Pruess, K. (2005) *ECO2N: A TOUGH2 Fluid Property Module for Mixtures of Water, NaCl, and CO₂*. Report LBNL-57952. Lawrence Berkeley National Laboratory, Berkeley, CA, USA. August.

Pruess, K.; Oldenburg, C.M.; and Moridis, G. (1999) *TOUGH2 User's Guide, Version 2.0*. Report LBNL-43134. Lawrence Berkeley National Laboratory, Berkeley, CA, USA. September.

Ramaswami, A.; Milford, J.B.; and Small, M.J. (2005) *Integrated Environmental Modeling: Pollutant Transport, Fate, and Risk in the Environment*. John Wiley & Sons, Inc. New Jersey. Chapter 14: Tools for Evaluation, Analysis, and Optimization of Environmental Models.

Rubin, Y. (2003) *Applied Stochastic Hydrogeology*. Oxford University Press, New York.

Schlumberger (2010) *ECLIPSE Technical Description*.

Schlumberger (2012) *WellWatcher Quartz: NLQG, NMQG, NPQG, NHQG Multidrop Pressure and Temperature Gauges*. Downloaded on October 1, 2012 from: http://www.slb.com/~media/Files/completions/product_sheets/quartzgauge_ps.pdf.

Seto, C.J. and McRae, G.J. (2011) Reducing risk in basin scale CO₂ sequestration: A framework for integrated monitoring design. *Environmental Science and Technology*, 45:845–859.

Siirila, E.R.; Navarre-Sitchler, A.K.; Maxwell, R.M.; and McCray, J.E. (2010) A quantitative methodology to assess the risks to human health from CO₂ leakage into groundwater. *Advances in Water Resources*, 36:146-164.

U.S. EPA (2008) *Vulnerability Evaluation Framework for Geologic Sequestration of Carbon Dioxide: Technical Support Document*. U.S. Environmental Protection Agency, EPA430-R-08-009, July 10, 2008.

White, M.D. and Oostrom, M. (1997) *STOMP: Subsurface Transport Over Multiple Phases*. Pacific Northwest National Laboratory Report PNNL-11218, Richland, WA, October 1997.

Wilkin, R.T. and DiGiulio, D.C. (2010) Geochemical impacts to groundwater from geologic carbon sequestration: Controls on pH and inorganic carbon concentrations from reaction path and kinetic modeling. *Environmental Science and Technology*, 44:4821-4827.

Yang, Y.M.; Small, M.J.; Ogretim, E.O.; Gray, D.D.; Bromhal, G.S.; Strazisar, B.R.; and Wells, A.W. (2011a) Probabilistic design of a near-surface CO₂ leak detection system. *Environmental Science and Technology*, 45:6380–6387.

Yang, Y.M.; Small, M.J.; Junker, B.; Bromhal, G.S.; Strazisar, B.R.; and Wells, A.W. (2011b) Bayesian hierarchical models for soil CO₂ flux and leak detection at geologic sequestration sites. *Environmental Earth Sciences*, 64(3):787-798.

Yang, Y.M.; Small, M.J.; Ogretim, E.O.; Gray, D.D.; Wells, A.W.; Bromhal, G.S.; and Strazisar, B.R. (2012) A Bayesian belief network (BBN) for combining evidence from multiple CO₂ leak detection technologies. *Greenhouse Gases: Science and Technology*, 2(3):185-199.

Yortsos, Y.C. (2000) Permeability variogram from pressure transients of multiple wells, in Zhang, D., and Winter, C.L., eds., *Theory, Modeling, and Field Investigation in Hydrogeology: A Special Volume in Honor of Shlomo P. Neuman's 60th Birthday*: Boulder, Colorado, Geological Society of America Special Paper 348, p.19-23.

Zhou, Q.; Birkholzer, J.T.; Rutqvist, J.; and Tsang, C.F. (2007) *Sensitivity Study of CO₂ Storage Capacity in Brine Aquifers with Closed Boundaries: Dependence on Hydrogeologic Properties*. Sixth Annual Conference on Carbon Capture and Sequestration, U.S. DOE-NETL, May 7-10, 2007.

Chapter 3: Quantifying the benefit of wellbore leakage potential estimates for prioritizing long-term MVA well sampling at a CO₂ storage site⁵

Abstract

This work uses probabilistic methods to simulate a hypothetical geologic CO₂ storage site in a depleted oil and gas field where the large number of legacy wells would make it cost-prohibitive to sample all wells for all measurements as part of the post-injection site care. Deep well leakage potential scores were assigned to the wells using a random subsample of 100 wells from a detailed study of 826 legacy wells that penetrate the basal Cambrian formation on the U.S. side of the U.S./Canadian border. Analytical solutions and Monte Carlo simulations were used to quantify the statistical power of selecting a leaking well. Power curves were developed as a function of (1) the number of leaking wells within the Area of Review; (2) the sampling design (random or judgmental, choosing first the wells with the highest deep leakage potential scores); (3) the number of wells included in the monitoring sampling plan; and (4) the relationship between a well's leakage potential score and its relative probability of leakage. Cases where the deep well leakage potential scores are fully or partially informative of the relative leakage probability are compared to a non-informative base case in which leakage is equiprobable across all wells in the Area of Review. The results show that accurate prior knowledge

⁵ Chapter 3 was the basis for the peer-reviewed publication. In the publication, several tables and figures were presented in the Supporting Information. However, in this dissertation the tables and figures have been renumbered in consecutive numerical order and are embedded into the main text. Azzolina, N.A.; Small, M.J.; Nakles, D.V.; Glazewski, K.A.; Peck, W.D.; Gorecki, C.; Bromhal, G.S.; and Dilmore, R.M. (2015) Quantifying the benefit of wellbore leakage potential estimates for prioritizing long-term MVA well sampling at a CO₂ storage site. *Environmental Science and Technology*, 49 (2): 1215-1224.

about the probability of well leakage adds measurable value to the ability to detect a leaking well during the monitoring program, and that the loss in detection ability due to imperfect knowledge of the leakage probability can be quantified. This work underscores the importance of a data-driven, risk-based monitoring program that incorporates uncertainty quantification into long-term monitoring sampling plans at geologic CO₂ storage sites.

3.1 Introduction

In an effort to mitigate concentrations of carbon dioxide (CO₂) in the atmosphere that are caused by stationary anthropogenic inputs, the United States Department of Energy (DOE) is pursuing carbon capture and sequestration (CCS) as one approach in a portfolio of greenhouse gas (GHG) reduction strategies. CCS involves (1) separating CO₂ from an industrial process, (2) transporting the CO₂ to a storage location, and (3) injecting and sequestering the CO₂ in a geologic reservoir for long-term isolation from the atmosphere (IPCC, 2005). Through the Carbon Sequestration Program, the DOE is working with seven Regional Carbon Sequestration Partnerships (RCSPs) to identify feasible sites within the U.S. and portions of Canada for large-scale (i.e., one million tonnes of CO₂ or greater) CO₂ geologic sequestration (U.S. DOE, 2011). The DOE is pursuing three primary types of geologic systems for long-term CO₂ storage: (1) depleted oil and gas fields; (2) unconventional formations such as gas shales, coal seams, and basalts; and (3) saline formations (U.S. DOE, 2012)

One of the potential risks associated with the injection and long-term storage of CO₂ into geologic reservoirs is leakage of stored CO₂ from geologic containment and into

the near-surface or surface environment. A potential leakage pathway in depleted oil and gas fields is associated with legacy exploration and production wells (Celia et al., 2007; 2011; Nogues et al., 2011) These legacy wells provide a potential conduit through low-permeability cap rock formations that would otherwise act as a seal to retain CO₂ in the storage reservoir. Extensive work has been conducted in Alberta, Canada over the past decade to assess the potential CO₂ leakage risk of legacy wells by drawing inferences from well completion and abandonment information. This work has, in part, been performed as part of the DOE Regional Partnership – Plains CO₂ Reduction (PCOR) Partnership. For example, Watson and Bachu (2007) compiled data from the Alberta Energy Resources Conservation Board on well leakage at the surface as surface-casing vent flow (SCVF) through wellbore annuli and soil gas migration (GM) outside casing. Their analysis compared these data on SCVF/GM well leaks against available well completion and abandonment factors. Based on this assessment, Watson and Bachu (2008) developed a scoring methodology to score well leakage potential according to a matrix of factors.

A well leakage potential scoring approach like the one developed by Watson and Bachu (2008) provides a quantitative means for ranking the increased probability of CO₂ leakage at a specific well because of SCVF and/or GM. Applying this scoring methodology to the legacy wells that are located within a particular region provides a screening-level risk assessment approach for identifying potential geologic CO₂ storage sites – areas with a high incidence of high-ranking wells would represent locations that are not favorable to long-term geologic storage of CO₂, while areas with a low incidence of high-ranking wells may be suitable future CO₂ injection and storage. In addition, once

a geologic CO₂ storage site has been identified, then such a well ranking approach also informs the monitoring, verification, and accounting (MVA) sampling plan for the site, as higher-ranking wells would take priority over lower-ranking wells.

This work investigates the use of a deep well leakage potential scoring approach like the one developed by Watson and Bachu (2008) to inform the CO₂ monitoring plan for legacy oil and gas wells at a simulated geologic CO₂ storage site. Previous studies have evaluated the effectiveness of leakage monitoring using specific measurements and their detection sensitivity, based on the underlying uncertainty and variability present in the baseline measurements (Yang et al., 2011a; 2011b; 2012; Azzolina et al., 2014) In contrast, this work assumes that a leakage event at a particular location (in this case, a legacy exploration or production well) will be detected, but only if the particular location is included in the sampling plan. The probability of detection is then determined by the sample size (relative to the total number of candidate locations/wells at the site) and the extent to which locations are properly prioritized based on their potential for leakage. Monitoring CO₂ leakage through legacy wells at a geologic CO₂ storage site would occur across the project phases: pre-injection (baseline), during injection (operational), and post-injection site care (long-term monitoring). Different approaches to sampling wells are available, from down-hole geophysical methods (e.g., vertical seismic profiles, neutron logs, and passive seismic) to collecting discrete fluid samples for geochemical analysis (e.g., carbon isotopes, alkalinity, and pH) (U.S. DOE, 2009; Hitchon, 2012). The full breadth of sampling methods is beyond the scope of this study. However, a site-specific sampling plan would need to specify not only the measurements to be made at each well, but the specific wells at which the measurements would be conducted.

Different methods for selecting the specific wells to include in the monitoring program are the focus of this work. For example, the most conservative approach would be to conduct all measurements on all wells within the Area of Review (AoR). An AoR is the region surrounding the geologic CO₂ storage project that may be impacted by the injection activity, which is based on simulation modeling that accounts for the separate-phase CO₂ plume extent and extent of storage formation pressure increase (U.S. EPA, 2011). While this approach ensures complete coverage, it may be cost-prohibitive and is neither data-driven nor risk-based. Alternatively, the number of wells and specific wells could be chosen randomly across the AoR or according to a judgmental sampling design which incorporates prior knowledge of the increased probability of CO₂ leakage at specific wells based on the leakage factors discussed above (Gilbert, 19987; U.S. EPA, 2002).

The objective of this work is to quantify the statistical power of selecting a set of leaking wells using judgmental sampling informed by estimated deep well leakage potential scores, under alternative assumptions regarding the extent to which the deep well leakage potential scores are reflective of actual leakage probabilities. A non-informative base case is first considered in which the occurrence of leakage across all wells is equiprobable, independent of the deep well leakage potential score. A second fully-informative case is next considered in which well leakage probabilities are assumed to be directly proportional to the deep well leakage potential scores. A third set of cases representing partially-informative deep well leakage potential-score information is then simulated assuming that the deep well leakage potential scores provide some information regarding corresponding well leakage probabilities, but with an error that is

parameterized by either a small or large degree of imprecision. Finally, a fourth set of cases is examined where only two or three of the four well attributes used to derive the deep well leakage potential score are used to inform the sampling, thus the sampler has incomplete information about the true probability of leakage. In all cases sampling is sequential, based on the deep well leakage potential score, that is, proceeding from highest to lowest deep well leakage potential score (though not necessary from the highest to the lowest well leakage probability).

This work evaluates several probabilistic simulations of a hypothetical geologic CO₂ storage site in a depleted oil and gas field where exhaustive sampling of a large number of legacy wells would be cost-prohibitive as part of the MVA sampling plan. This work does not use numerical simulation models to simulate CO₂ injection, subsurface fate and transport of CO₂ phases, and subsequent migration of CO₂ into abandoned wells. Instead, the sampling problem is reduced to different combinations of well leakage and well sampling, and the probabilistic simulations quantify the statistical power of selecting a set of leaking wells for each combination. The statistical modeling is aimed at late-stage operational monitoring phase and into the post-injection site care phase. Deep well leakage potential scores were assigned to the wells using a random subsample from a detailed study of 826 legacy wells that penetrate the basal Cambrian formation on the U.S. side of the U.S./Canadian border (Glazewski et al., 2013). Analytical solutions and Monte Carlo simulations were used to quantify the statistical power of selecting a leaking well. Power curves quantifying the probability that a leaky well was selected were developed as a function of (1) the number of leaking wells within the AoR; (2) the sampling design (random or judgmental); (3) the number of wells

sampled as part of the MVA plan; and (4) the failure model for the leaking wells (equiprobable; with leakage probabilities fully related to the deep well leakage potential score; or with leakage probabilities and leakage potential scores only partially related, i.e., with error). The results of this work underscore the importance of a data-driven, risk-based monitoring program that incorporates uncertainty quantification into MVA sampling plans at geologic CO₂ storage sites.

3.2 Methods

3.2.1 Generic Geologic CO₂ Storage Site

A generic geologic CO₂ storage site was defined using an AoR with $n=100$ legacy wells. This generic site and monitoring scenarios are intended to represent the late-stage operational monitoring phase and into the post-injection site care phase. Monitoring during the early-stage operational phase would focus resources closer to the CO₂ injection wells. In contrast, monitoring in the later stages of CO₂ injection and into the post-injection site care phase would need to sample a broader AoR based on the separate-phase CO₂ plume extent and extent of storage formation pressure increase within the subsurface. This broader radial extent and the need to economize well selection for monitoring within the AoR help to motivate the need for this work.

3.2.2 Well Leakage Potential Scores

This work focuses on deep well leakage potential scores. Deep well leakage is defined as leakage (cross-flow) from a target production zone or CO₂ injection zone back into the wellbore (or outside the casing) where it moves upward into an adjacent

permeable zone (productive zone or aquifer) (Bachu et al., 2012). Thus detecting deep well leakage is one component of the MVA sampling program to proactively identify leakage deep in the subsurface in an attempt to mitigate near-surface impacts. A detailed study of 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada border incorporated four well attributes to derive a deep well leakage potential score: (1) number of fracture treatments; (2) number of acid treatments; (3) number of completions (perforations); and (4) well abandonment type (Glazewski et al., 2013). Information was collected for each of the 826 wells penetrating the basal Cambrian system across the states of Montana, North Dakota, and South Dakota. Well data were collected from each state agency as follows: Montana Board of Oil and Gas (MTBOG), North Dakota Industrial Commission (NDIC) Oil and Gas Division, and South Dakota Department of Environment and Natural Resources (SD DENR).

Fracture treatments are designed to open cracks in the formation to allow oil or natural gas to flow back to the well. Matrix-acidizing treatments are used either to stimulate a well to greater-than-ideal matrix reservoir flow or to remove skin damage. Both fracture and acid treatments are executed at high pressure and therefore may contribute to degradation of local hydraulic isolation (Bachu et al., 2012). Well completions (perforations) provide a communication pathway through the production casing in order to access the formation. However, perforations could compromise well integrity as the casing may not be sealed when squeezing cement into the perforations during abandonment. Thus as the number of perforations increase, so does the potential risk of leakage. Lastly, wells in the study area were typically abandoned by using either cement plugs or cast iron bridge plugs. Cement plugs are generally considered more

reliable, while bridge plugs have been found to be more prone to corrosion and seal failure in the presence of CO₂ (Bachu et al., 2012). Cement/additive type was omitted from the scoring, as the well files in the basal Cambrian study area lacked sufficient cement and additive data to evaluate wellbore integrity. The scoring system is shown in Table 3-1.

Table 3-1. Well attributes and weighting scores used to derive the deep well leakage potential scores for 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada.

Well Attribute	Weighting Score
Number of fracture treatments	
0	1
1	1.5
>1	2
Number of acid treatments	
0	1
1	1.1
2	1.2
>2	1.5
Number of completions (perforations)	
0	1
1	1.5
>1	2
Abandonment plug type	
Cement	1
Unknown (abandoned)	2
Active well	2
Bridge plug	3

The basal Cambrian study showed that the deep well leakage potential scores reflected a right-skewed distribution, with most wells having relatively low risk scores and fewer wells having relatively high- to very-high scores. The deep well leakage risk scores were determined to fit a 3-parameter gamma distribution, with probability density function (pdf) (NIST/SEMATECH, 2014):

$$f(x|a, b, \theta) = \frac{\left(\frac{x-\theta}{b}\right)^{a-1} \exp\left(-\frac{x-\theta}{b}\right)}{b\Gamma(a)} \quad (3-1)$$

Where:

$f(x|a, b, \theta)$ = the pdf of the 3-parameter gamma distribution;

x = the deep well leakage risk score ($x \geq \theta$);

a = the shape parameter ($a > 0$);

b = the scale parameter ($b > 0$);

θ = the threshold parameter; and

Γ = the gamma function, which has the formula:

$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dt \quad (3-2)$$

The cumulative distribution function (cdf) of the deep well leakage risk score, $F(x|a, b, \theta)$, is computed using numerical methods or from commands available within most mathematics/statistics packages, such as Excel, Minitab, and Matlab. The 3-parameter gamma distribution was fit using the empirical cdf in Minitab 17 Statistical Software (Minitab Statistical Software, 2010), resulting in fitted parameters of $a=0.2513$, $b=5.899$, and $\theta=0.99$. The empirical and fitted 3-parameter gamma cdfs are compared for the 826 legacy oil and gas well deep leakage potential scores in Figure 3-1.

A random subsample of $n=100$ wells was collected from this group of $n=826$ wells to populate the deep well leakage potential scores used in this work (the total score based on all four attributes and each of the four component scores), and these values are assumed to apply to a set of $n=100$ wells in the AoR. The $n=100$ subsample was

determined to be representative of the larger well population. A comparison of the fitted cumulative distribution function to the population of 826 wells and the random subsample of 100 wells is shown in Figure 3-1.

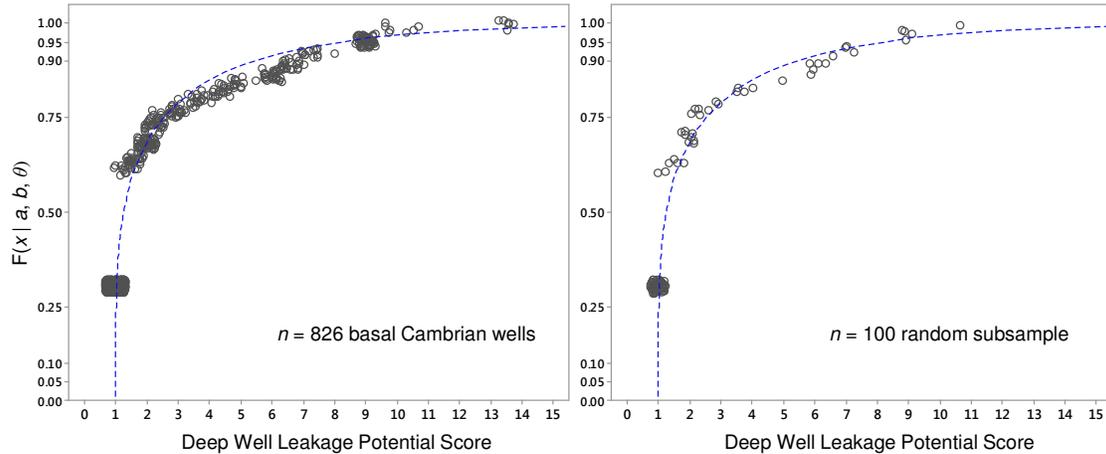


Figure 3-1. Cumulative distribution of the deep well leakage potential scores for 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada border (gray open circles) [left] and our random subsample of 100 wells used in this work [right]. The blue dashed line in both panels represents the fitted cumulative distribution function of a 3-parameter gamma function to the full population of 826 legacy oil and gas wells. Many of the deep well leakage potential scores were identical; therefore, a small jitter of 0.025 was added to the x - and y -directions of this plot to help visualize the number of points.

3.2.3 Well Leakage and Well Sampling Scenarios

Six different scenarios (Scenarios 1 through 6) were evaluated to assess different well leakage and well sampling combinations (Table 3-2). Two different failure modes were considered for well leakage: (1) equiprobable well leakage and (2) well leakage proportional to the deep well leakage potential score. Equiprobable well leakage would equate to some unknown failure mode independent of the well attributes or discrete geologic pathway such that all wells are equally likely to be leaking (e.g., site-wide seal formation failure). Well leakage proportional to the deep well leakage potential score implies that some combination of the number of fracture treatments, number of acidizing

treatments, number of completions, and/or well abandonment type are predictive of well leakage, thus wells with higher deep leakage potential scores leak first. Two different sampling plans were considered: (1) random sampling and (2) sequential sampling based on the deep well leakage potential scores. Random sampling represents the simplest sampling plan, and assumes that wells are randomly selected from the $n=100$ possible wells, thus each well is equally likely to be included in any given sampling event. Sequential sampling based on the deep well leakage potential scores selects wells from highest to lowest leakage potential score in sequential order. These two failure modes and two sampling plans yield four different leakage/sampling combinations, which have been designated Scenarios 1-4. Two additional evaluations have been included to look at the change in Scenario 4 when error is introduced into the deep well leakage potential score (Scenarios 5a and 5b) or when only a subset of the well attributes used to derive the deep well leakage potential score is known (Scenarios 6a and 6b) (Table 3-2).

Table 3-2. Matrix of six well leakage and well sampling scenarios included in this work.

Scenario	Well Leakage	Well Sampling	Risk Score for Sampling	Calculation Method
1	Equiprobable	Random	NA	Analytical - hypergeometric distribution
2	Equiprobable	Sequential	Fully informative	Leakage = random w/o replacement; Sampling = sequential
3	Score	Random	NA	Leakage = mwnchypg distribution; Sampling = random w/o replacement
4	Score	Sequential	Fully informative	Leakage = mwnchypg distribution; Sampling = sequential
5a	Score	Sequential	Partially informative (small error)	Leakage = mwnchypg distribution; Sampling = sequential
5b	Score	Sequential	Partially informative (large error)	Leakage = mwnchypg distribution; Sampling = sequential
6a	Score	Sequential	Partially informative (2 of 4 well attributes)	Leakage = mwnchypg distribution; Sampling = sequential
6b	Score	Sequential	Partially informative (3 of 4 well attributes)	Leakage = mwnchypg distribution; Sampling = sequential

- 1) NA = Not Applicable.
- 2) Score = proportional to the deep well leakage potential score.
- 3) Sequential well sampling was from highest-to-lowest deep well leakage potential score.
- 4) The “fully informative” leakage potential scores used to simulate the sampling matrix in Scenarios 2 and 4 assume that both the “true” increased probability of well leakage and the information used to develop the well sampling plan are identical. In contrast, the “partially informative” leakage potential scores used to simulate the sampling matrices in Scenarios 5a and 5b reflect either a small or large degradation, respectively, in the accuracy in the sampler’s beliefs about the true increased probability of well leakage. The partially informative leakage potential scores used to simulate the sampling matrices in Scenarios 6a and 6b reflect the use of only two or three of the four well attributes, respectively, that were used to derive the deep well leakage potential score.
- 5) The “mwnchypg distribution” used in Scenarios 3, 4, 5, and 6 is the multivariate Wallenius’ hypergeometric distribution.

Scenario 1 Calculations

In Scenario 1, all wells are equally likely to be leaking (uniform distribution of well leakage risk scores) and wells are sampled randomly. Scenario 1 is a discrete probability distribution where the probability of k successes ($k = n_{LS}$) in n draws ($n = n_s$) without replacement from a finite population of size N ($N = n$) containing exactly K successes ($K = n_L$), and is therefore described by the hypergeometric distribution (Equation 3-3) (Jaynes, 2003; Berkopec, 2007):

$$p = F(n_{LS}|n_s, n_L, n) = \sum_{i=0}^x \frac{\binom{n_L}{i} \binom{n-n_L}{n_s-i}}{\binom{n}{n_s}} \quad (3-3)$$

where p is the cumulative probability of sampling up to n_{LS} of a possible number of leaky wells (n_L), in a sample of n_s wells without replacement from a group of n wells; and $F(x|n_s, n_L, n) =$ the cdf of the hypergeometric distribution.

The hypergeometric distributions for Scenario 1 were calculated for $n_L = 1, 2, 5,$ and 10 wells and $n_s = 0$ to 100 using the “HYPGEOM.DIST” function in Microsoft Excel[®]. Scenario 1 represents the random, uninformed base case against which Scenarios 2, 3, 4, 5, and 6 were compared. Three outputs were used to compare scenarios: (1) the probability that at least one leaky well was sampled (POLWS); (2) the probability that all leaky wells were sampled (PALWS); and (3) the expected fraction of leaky wells sampled (EFLWS).

The output POLWS was quantified by setting $n_{LS}=1$ in Equation 3-3 and solving for $n_L = 1, 2, 5,$ and 10 wells. The output PALWS was quantified by the probability mass

function (pmf) of the hypergeometric distribution at $n_L = 1, 2, 5,$ and 10 wells. The expected value of a hypergeometric distribution is defined as (Equation 3-4):

$$E[X] = n_{LS} = n_s \frac{n_L}{n} \quad (3-4)$$

Scenario 2 Calculations

In Scenario 2, all wells are equally likely to be leaking and wells are sampled sequentially from the highest deep well leakage potential score to the lowest. The POLWS, PALWS, and EFLWS for Scenario A2 were computed through Monte Carlo simulation for $n_L = 1, 2, 5,$ and 10 wells and $n_S = 0$ to 100 . Each combination was run for 1000 realizations.

The first step in the simulation process was to generate a 1000-by- n_L matrix, $[\mathbf{A}_{nL}]$, of leaky wells, which was populated with random integers from 1 to n without replacement. Thus each row of matrix $[\mathbf{A}_{nL}]$ represents a unique simulation and contains the index of the equiprobable leaky wells. The second step in the simulation process was to generate a 1000-by- n_S matrix, $[\mathbf{B}_{nS}]$, of sampled wells, which was populated according to deep well leakage potential scores of the random subsample of 100 wells. The highest deep well leakage potential scores (highest rank) were sampled first, then the sample selection moved down the ranks to the lowest deep well leakage potential score. In the event of wells with tied ranks the wells were selected at random without replacement.

Logical functions were used to compare the indices of leaky wells and sampled wells to determine if one or more leaky wells were included in the sample. The POLWS was the number of times *one or more leaky wells* were sampled ($n_{LS} \geq 1$) out of 1000 for

a given n_S . The PALWS was the number of times *all leaky wells* were sampled ($n_L = n_{LS}$) out of 1000 for a given n_S . Lastly, the EFLWS was the *fraction of leaky wells* sampled (n_{LS} / n_L) out of 1000 for a given n_S .

Scenario 3 Calculations

In Scenario 3, each well has a leakage probability that is proportional to their deep well leakage potential score, but this information is ignored and the wells are sampled randomly. The POLWS, PALWS, and EFLWS for Scenario 3 were computed through Monte Carlo simulation for $n_L = 1, 2, 5, \text{ and } 10$ wells and $n_S = 0$ to 100. Each combination was run for 1000 realizations.

The first step in the simulation process was to generate a 1000-by- n_L matrix, $[\mathbf{A}_{nL}]$, of leaky wells, which was populated according to deep well leakage potential scores of the random subsample of 100 wells. The second step in the simulation process was to generate a 1000-by- n_S matrix, $[\mathbf{B}_{nS}]$, of sampled wells, which was populated with random integers from 1 to n without replacement. Thus each row of matrix $[\mathbf{B}_{nS}]$ represents a unique simulation and contains the index of the randomly-sampled wells. The well leakage probability was calculated for each well as the individual deep well leakage potential score for that well divided by the sum of deep well leakage potential scores across all $n=100$ wells (Equation 3-5).

$$\text{Prob}[\text{Leakage}_i] = \frac{\text{DWLPS}_i}{\sum_{i=1}^n \text{DWLPS}_i} \quad (3-5)$$

Where:

- n = total number of wells ($n=100$);
 DWLPS_{*i*} = the deep well leakage potential score assigned to each well.

The Monte Carlo simulation of leaky wells for Scenario 3 was conducted using a generalization of the hypergeometric distribution where wells are selected with bias according to a matrix of weights (odds) that a particular well will be selected. The probability that a particular well is leaking in a particular draw depends not only on its initial Prob[Leakage_{*i*}], but also on the total sum of the competing wells that remain in the legacy oil and gas field at that moment. In other words, Prob[Leakage_{*i*}] must be recalculated after each draw because the leaky wells are selected without replacement. The solution for this type of simulation is the multivariate form of Wallenius' non-central hypergeometric distribution (Equations 3-6 through 3-10) (Wallenius, 1963; Fog, 2008; 2014):

$$f(\mathbf{x}|n_S, \mathbf{m}, \boldsymbol{\omega}) = \Lambda(\mathbf{x})I(\mathbf{x}), \text{ where} \quad (3-6)$$

$$\Lambda(\mathbf{x}) = \sum_{i=1}^n \binom{m_i}{x_i}, \quad (3-7)$$

$$I(\mathbf{x}) = \int_0^1 \prod_{i=1}^n (1 - t^{\omega_i/d})^{x_i} dt, \quad (3-8)$$

$$d = \boldsymbol{\omega} \cdot (\mathbf{m} - \mathbf{x}) = \sum_{i=1}^n \omega_i(m_i - x_i), \quad (3-9)$$

$$\mathbf{x} = (x_1, x_2, \dots, x_n), \mathbf{m} = (m_1, m_2, \dots, m_n), \boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_n), \quad (3-10)$$

Where:

$f(\mathbf{x}|n_S, \mathbf{m}, \boldsymbol{\omega})$ = the probability mass function of the multivariate Wallenius' non-central hypergeometric distribution;

ω_i = the individual “weights”, or Prob[Leakage_{*i*}] for each well;

n = total number of wells ($n=100$);

n_S = number of wells sampled ($\leq n$);

m_i = the number of each well available ($m_1=m_2=m_n=1$); and

x_i = the number of each well of a given type that is sampled.

The Monte Carlo simulations of the multivariate form of Wallenius' non-central hypergeometric distribution were implemented in R using the package “BiasedUrn” (Fog, 2008; 2014). Leaky wells are randomly selected, without replacement, from this distribution where the probability of leakage for each well was defined by Eq. 3-5. Leaky wells are simulated sequentially, whereby the leakage probabilities of the remaining wells are updated to reflect the ‘removal’ of the previous leaky well from the pool of potential monitoring wells. The probability that a particular well is sampled in a particular draw depends not only on its own weight, but also on the total weight of the competing wells that remain in the AoR at that moment (Wallenius, 1963). One thousand random draws were generated in R, and this output was then used as the 1000-by- n_L matrix, [\mathbf{A}_{nL}]. Appendix B provides the input file and R code that was used to generate the matrix [\mathbf{A}_{nL}] from BiasedUrn. The POLWS, PALWS, and EFLWS were computed using logical

functions to compare the indices of sampled wells and leaky wells as described for Scenario 2.

Scenario 4 Simulations

In Scenario 4, each well has a leakage probability that is proportional to their deep well leakage potential score, and wells are sampled based on these deep well leakage potential scores, from highest to lowest. The POLWS, PALWS, and EFLWS for Scenario 4 were computed through simulation for $n_L = 1, 2, 5,$ and 10 wells and $n_S = 0$ to 100 . Each combination was run for 1000 realizations. The 1000-by- n_L matrix, $[\mathbf{A}_{nL}]$, of leaky wells was populated according to deep well leakage potential scores of the random subsample of 100 wells using the multivariate form of Wallenius' non-central hypergeometric distribution and the same process as was described for Scenario 3. The 1000-by- n_S matrix, $[\mathbf{B}_{nS}]$, of sampled wells was populated according to deep well leakage potential scores of the random subsample of 100 wells using the same process as was described for Scenario 2. The POLWS, PALWS, and EFLWS were computed using logical functions to compare the indices of sampled wells and leaky wells as described for Scenario 2.

Scenario 5 Simulations

In Scenario 5, each well exhibits a relationship between their leakage probability and their deep well leakage potential score that is similar to Scenario 4, but with error introduced into the relationship. The sampling in Scenario 4 represents a form of judgmental sampling, where the well sampling order is informed by the deep well leakage potential scoring system. However, Scenario 4 assumes that both the “true”

increased probability of well leakage (i.e., $\text{Prob}[\text{Leakage}_i]$) and the information used to develop the well sampling plan are identical. In other words, Scenario 4 assumes that the sampler has perfect knowledge of the true increased probability of well leakage. In reality, that knowledge is likely to be imperfect and dependent upon the quality of the well integrity survey(s) conducted prior to and/or during CO₂ injection and other factors. Scenario 5 assesses the change in POLWS, PALWS, and EFLWS as a function of the accuracy in the relationship between the deep well leakage potential score and the leakage probability for each well.

The 1000-by- n_L matrix, $[\mathbf{A}_{nL}]$, of leaky wells was populated according to deep well leakage potential scores of the random subsample of 100 wells using the multivariate form of Wallenius' non-central hypergeometric distribution and the same process as was described for Scenario 4. The 1000-by- n_S matrix, $[\mathbf{B}_{nS}]$, of sampled wells was populated according to deep well leakage potential scores of the random subsample of 100 wells using the same process as was described for Scenario 4. However, the $[\mathbf{B}_{nS}]$ matrix of sampled wells was adjusted from Scenario 4 using a two-step process to simulate imperfect knowledge of the leakage probability. First, the "true" deep well leakage potential score for a specific well was multiplied by 0.8 and 1.2 to define the lower and upper bounds, respectively, of a range. Next, a pseudo-leakage potential score was selected randomly within this range using a uniform distribution. For example, if the "true" deep well leakage potential score was equal to 10, then the range would be 8 to 12 from which a pseudo-leakage potential score was randomly selected. The 0.8 and 1.2 factors were used to develop a pseudo-leakage potential score reflecting a small degradation in the accuracy in the sampler's beliefs, which was designated Scenario 5a.

In addition, factors of 0.2 and 5.0 were used to simulate the lower and upper bounds, respectively, of a range representing a large degradation in accuracy. In this case, if the “true” deep well leakage potential score was equal to 10, then the range would be 2 to 50 from which a pseudo-leakage potential score was randomly selected. The large degradation in accuracy was designated Scenario 5b. The pseudo-leakage potential scores from the small and large accuracy change cases were used to inform the sampling order. The POLWS, PALWS, and EFLWS were computed using logical functions to compare the indices of sampled wells and leaky wells as described for Scenario 2.

Scenario 6 Simulations

In Scenario 6, each well has a leakage probability that is proportional to their well leakage potential score, and wells are sampled based on these leakage potential score, from highest to lowest. However, Scenario 6 examines the cases where only two (Scenario 6a) or three (Scenario 6b) of the four well attributes are available; thus the sampler only has partial prior knowledge of the true well leakage potential, which is represented by the complete score using all four attributes.

There were six combinations of well treatments for Scenario 6a (two well attributes) and four combinations for Scenario 6b (three well attributes), for a total of 10 additional simulations (Table 3-3). Analogous to the model form used to fit the complete deep well leakage potential scores, the cdfs for each of the combinations were also fit to a 3-parameter gamma distribution (Figures 3-2 and 3-3). The cdfs of the two- and three-attribute partial deep well leakage potential scores were computed using numerical methods from commands available within Minitab.

Table 3-3. Combinations of well attributes for Scenarios 6a and 6b.

Scenario 6a	Scenario 6b
Fracture/Acid	Fracture/Acid/Abandonment
Fracture/Completions	Fracture/Acid/Completions
Fracture/Abandonment	Fracture/Abandonment/Completions
Acid/Abandonment	Acid/Abandonment/Completions
Acid/Completions	
Abandonment/Completions	

In Scenario 6, the 1000-by- n_L matrix, $[A_{nL}]$, of leaky wells was populated according to deep well leakage potential scores of the random subsample of 100 wells (the *complete* deep well leakage potential scores, i.e., all four well attributes) using the multivariate form of Wallenius' non-central hypergeometric distribution and the same process as was described for Scenario 4. However, the 1000-by- n_S matrix, $[B_{nS}]$, of sampled wells was populated according to *partial* deep well leakage potential scores of the random subsample of 100 wells (the *partial* deep well leakage potential scores, i.e., using only two well attributes [Scenario 6a] or three well attributes [Scenario 6b]). Thus the well sampling order is informed by the *partial* deep well leakage potential scores, which are not fully representative of the true leakage probability. The POLWS, PALWS, and EFLWS were computed using logical functions to compare the indices of sampled wells and leaky wells as described for Scenario 2.

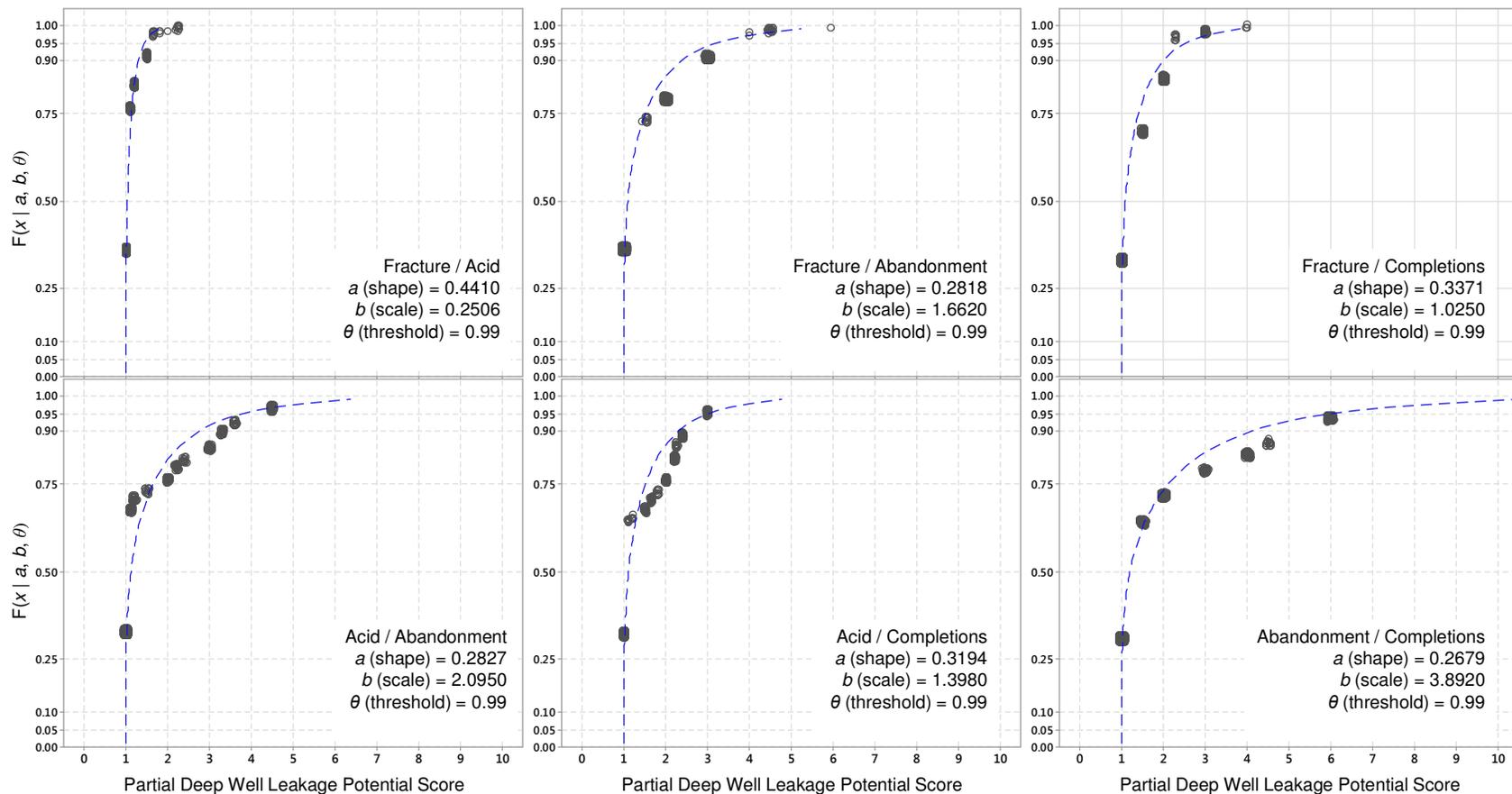


Figure 3-2. Fitted 3-parameter gamma cumulative distribution functions (cdfs, blue dashed line) to the two-attribute partial deep well leakage potential scores for 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada border (gray open circles) (Glazewski et al., 2013). A small jitter of 0.015 was added to the x - and y -directions of this plot to help visualize the number of points used to fit the cdfs. The combinations of two well attributes are shown for each panel.

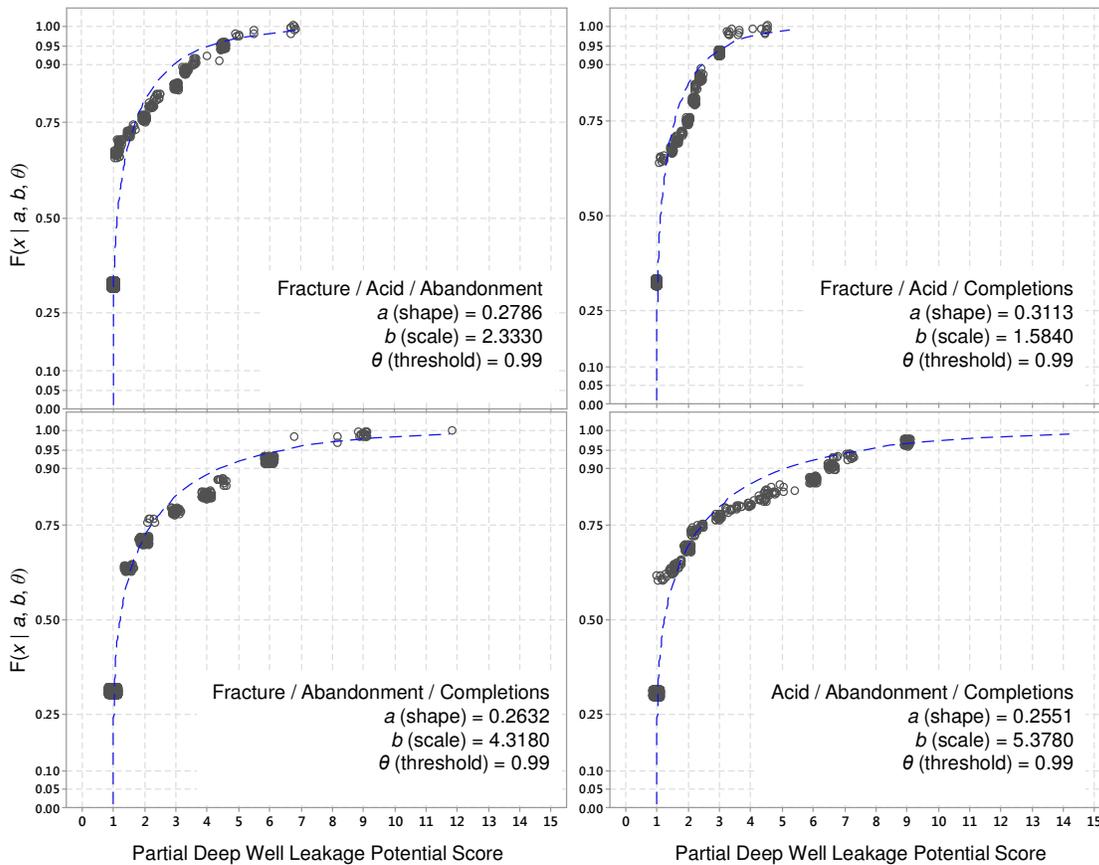


Figure 3-3. Fitted 3-parameter gamma cumulative distribution functions (cdfs, blue dashed line) to the three-attribute partial deep well leakage potential scores for 826 legacy oil and gas wells penetrating the basal Cambrian system on the U.S. side of the U.S./Canada border (gray open circles) (Glazewski et al., 2013). A small jitter of 0.015 was added to the x- and y-directions of this plot to help visualize the number of points used to fit the cdfs. The combinations of three well attributes are shown for each panel.

3.3 Results

3.3.1 Statistical Power Curves

The results for Scenario 1 (equiprobable well leakage and random well sampling), Scenario 2 (equiprobable well leakage and score-based well sampling), and Scenario 3 (score-based well leakage and random well sampling) were identical (Figures 3-4, 3-5, and 3-6); randomizing leakage, sampling, or both yields equivalent results. Therefore, all three scenarios produce a hypergeometric distribution as a function of n_s . Scenarios 2 and

3 are not discussed further and comparisons focus on the random, uninformed base case (Scenario 1) and Scenarios 4, 5 and 6.

Statistical power curves for Scenarios 1 and 4 for the outputs POLWS, PALWS, and EFLWS are shown in Figure 3-7. Under Scenario 1, the probability of sampling at least one leaky well (POLWS) is low when the number of leaky wells is small. For example, when $n_L=2$, it is not until $n_S \geq 69$ where POLWS is ≥ 0.9 . However, when $n_L=10$, POLWS is ≥ 0.9 when $n_S \geq 20$. The probability that all leaky wells are sampled (PALWS) is extremely low, regardless of the number of leaky wells. For example, it is not until $n_S \geq 95$ (nearly complete coverage of the $n=100$ well AoR) where POLWS is ≥ 0.9 when $n_L=2$. The lack of information provided by random sampling is further emphasized by the straight-line EFLWS for Scenario 1, which is equivalent to $n_S/100$, or “even odds” (Figure 3-7).

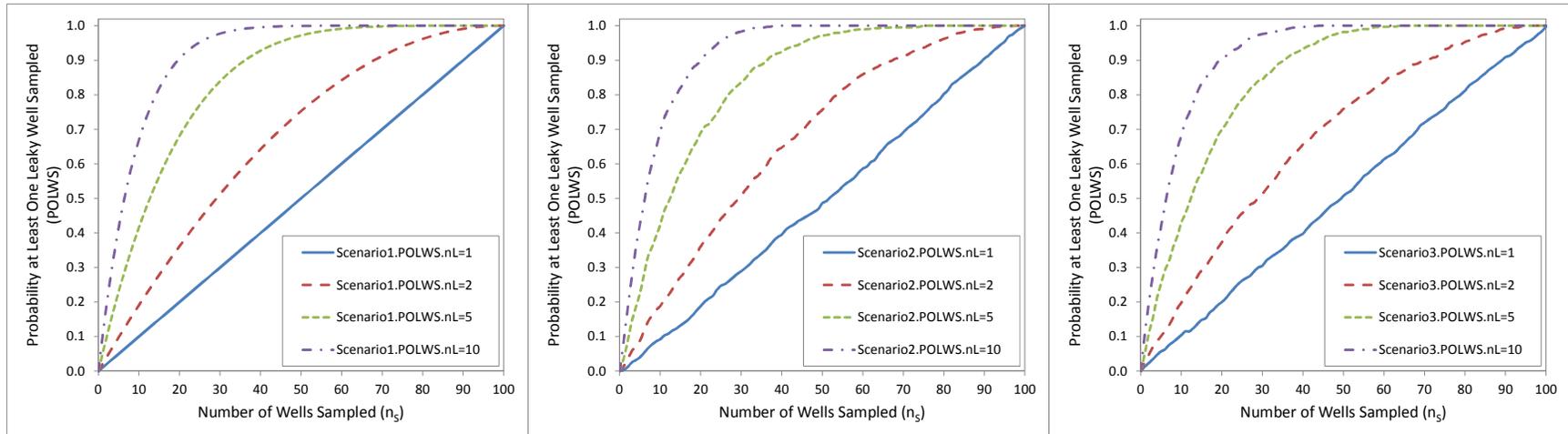


Figure 3-4. Probability of at least one leaky well sampled (POLWS) for Scenario 1 (equiprobable well leakage/random well sampling, [left]), Scenario 2 (equiprobable well leakage/score-based well sampling, [middle]), and Scenario 3 (score-based well leakage/random well sampling [right]).

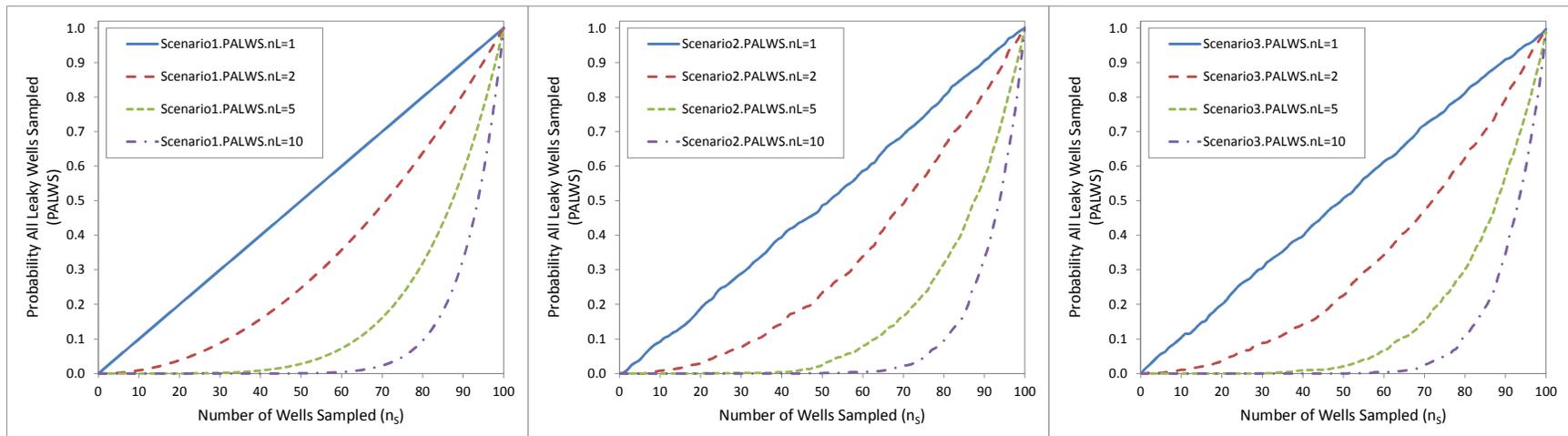


Figure 3-5. Probability that all leaky wells are sampled (PALWS) for Scenario 1 (equiprobable well leakage/random well sampling, [left]), Scenario 2 (equiprobable well leakage/score-based well sampling, [middle]), and Scenario 3 (score-based well leakage/random well sampling [right]).

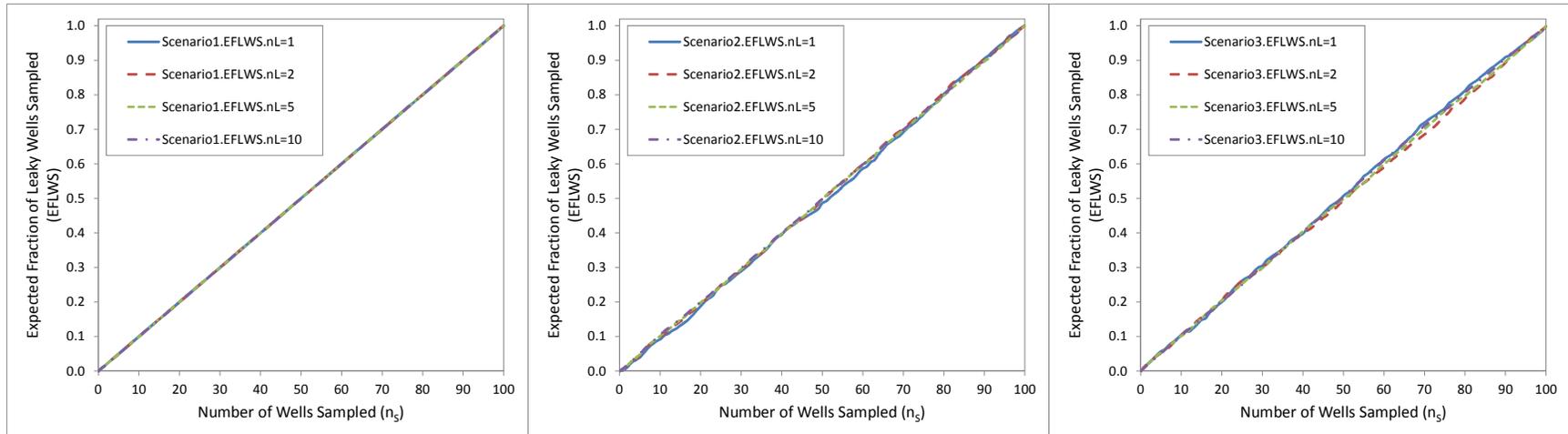


Figure 3-6. Expected fraction of leaky wells sampled (EFLWS) for Scenario 1 (equiprobable well leakage/random well sampling, [left]), Scenario 2 (equiprobable well leakage/score-based well sampling, [middle]), and Scenario 3 (score-based well leakage/random well sampling [right]).

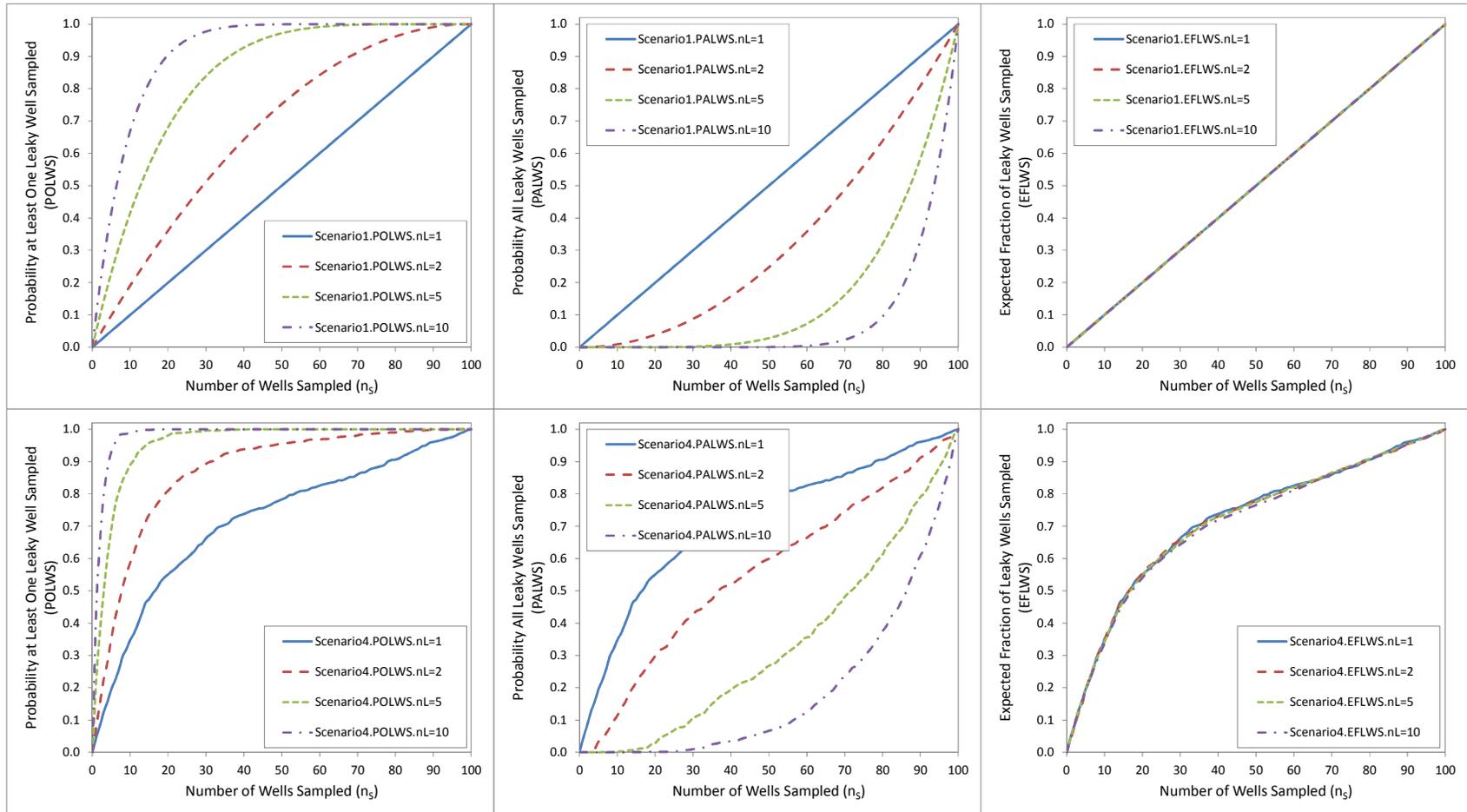


Figure 3-7. Probability of at least one leaky well sampled (POLWS, left column); probability that all leaky wells are sampled (PALWS, middle column); and expected fraction of leaky wells sampled (EFLWS, right column) for Scenario 1 (equiprobable well leakage/random well sampling, [top]) and Scenario 4 (score-based well leakage/score-based well sampling, [bottom]).

Information-based sampling using the deep well leakage potential scores in Scenario 4 produces significantly greater statistical power. While this result is intuitive, comparison of the statistical power curves quantifies the information value of the fully-informative deep well leakage potential scores for helping to select the leaky well(s). For example, when $n_L=2$, POLWS is ≥ 0.9 when $n_S \geq 32$, which is 37 fewer wells (54% fewer wells) than needed to obtain a similar detection power in Scenario 1. When $n_L=10$, POLWS is ≥ 0.9 when $n_S \geq 5$, which is 15 fewer wells (75% fewer wells) than needed under Scenario 1. Therefore comparable statistical power for detecting one or more leaky wells is achieved by sampling significantly fewer wells in Scenario 4. The results are similar for PALWS and EFLWS, though not as dramatic. For example, when $n_L=2$, PALWS is ≥ 0.9 when $n_S \geq 90$, which is 5 fewer wells (5% fewer wells) than needed to obtain a similar detection power in Scenario 1. To achieve an EFLWS of ≥ 0.9 (i.e., 90 percent of leaky wells sampled) when $n_L=2$ requires $n_S \geq 79$ wells for Scenario 4, as opposed to 90 wells in Scenario 1 (12% fewer wells).

Ratios of (EFLWS for Scenario 4 / EFLWS for Scenario 1) as a function of n_S and n_L are shown in Figure 3-8. Scenario 4 provides clear value over Scenario 1, especially when n_S is small. For example, when n_S is between 1 and 10 wells, the improvement in the power for achieving higher EFLWS provided by Scenario 4 is between 4.5 and 3.5 times, respectively. Therefore, in a situation where only a few wells are leaking CO₂ and limited resources are available to sample from n available wells, full prior knowledge of the deep well leakage potential adds significant value to the ability of the monitoring plan to detect the leaking well(s). This situation may arise quite often at geologic CO₂ storage sites in legacy oil and gas fields. The CO₂ injection phase is likely to span decades and

the post-injection site care would extend beyond that timeline. Current recommendations suggest long-term monitoring for CO₂ plume and pressure stability for 50 years post-injection (U.S. DOE, 2009). Thus, sampling all wells within the AoR at fixed frequency (e.g., annually) for 50 years may be cost-prohibitive (i.e., n_S would be limited by cost).

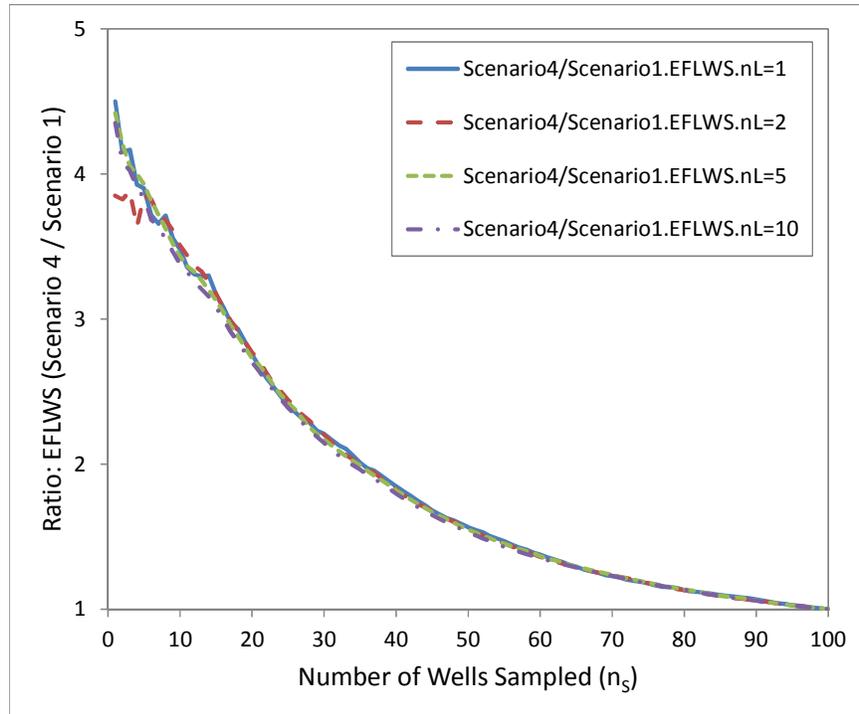


Figure 3-8. Ratio of the expected fraction of leaky wells sampled (EFLWS) for Scenario 4 (score-based well leakage/score-based sampling) / Scenario 1 (equiprobable well leakage/random well sampling) as a function of the number of wells sampled (n_S).

3.3.2 Sensitivity to Wellbore Leakage Potential Accuracy

Scenario 1 represents the worst-case assumptions considered; performance under this scenario is no better than chance (even odds) at selecting a leaking well. Worse performance could result from mis-informative leakage potential scores that are negatively correlated with leakage probability, though that case is not considered here. Scenario 4 represents the best-case, where the wells leak according to their deep well

leakage potential score and the wells are sampled from highest-to-lowest leakage potential score. Barring additional knowledge about the order in which wells might leak (beyond the scope of this study) the detection performance would not exceed the Scenario 4 results. Scenario 5 evaluates how Scenario 4 degrades toward Scenario 1 as a function of the change in accuracy of the *believed* well leakage risk score relative to the “true” well leakage probability.

Ratios of EFLWS for (Scenario 4 / Scenario 5a or 5b) as a function of n_S and n_L are shown in Figure 3-9 (Note: when $n_L = 1$, POLWS=PALWS=EFLWS). This comparison shows that a small change in accuracy does not substantively affect the statistical power, as the ratio for Scenario 5a is close to 1 (Figure 3-9, blue line). This is due to the right-skewed distribution of deep well leakage potential scores, which causes the first few wells to continue to be ranked highly despite the error introduced in the case with a small loss of accuracy. In other words, the top 10 highest-ranked wells generally remain in the top 10 when there is only a small loss in the accuracy of predicted leakage probabilities from deep well leakage potential scores in Scenario 5a. A large decrease in accuracy (Scenario 5b) does lead to a noticeable decrease in statistical power. For example, for $n_L=2$ and $n_S=20$, the ratio of EFLWS (Scenario 4 / Scenario 5b) is approximately 1.1, or a 10 percent decrease in statistical power for Scenario 5b as compared to the fully-informed Scenario 4. This effect decreases as the number of leaky wells increases and the number of sampled wells increases, as is seen in Figure 3 moving from $n_L=1$ to $n_L=10$.

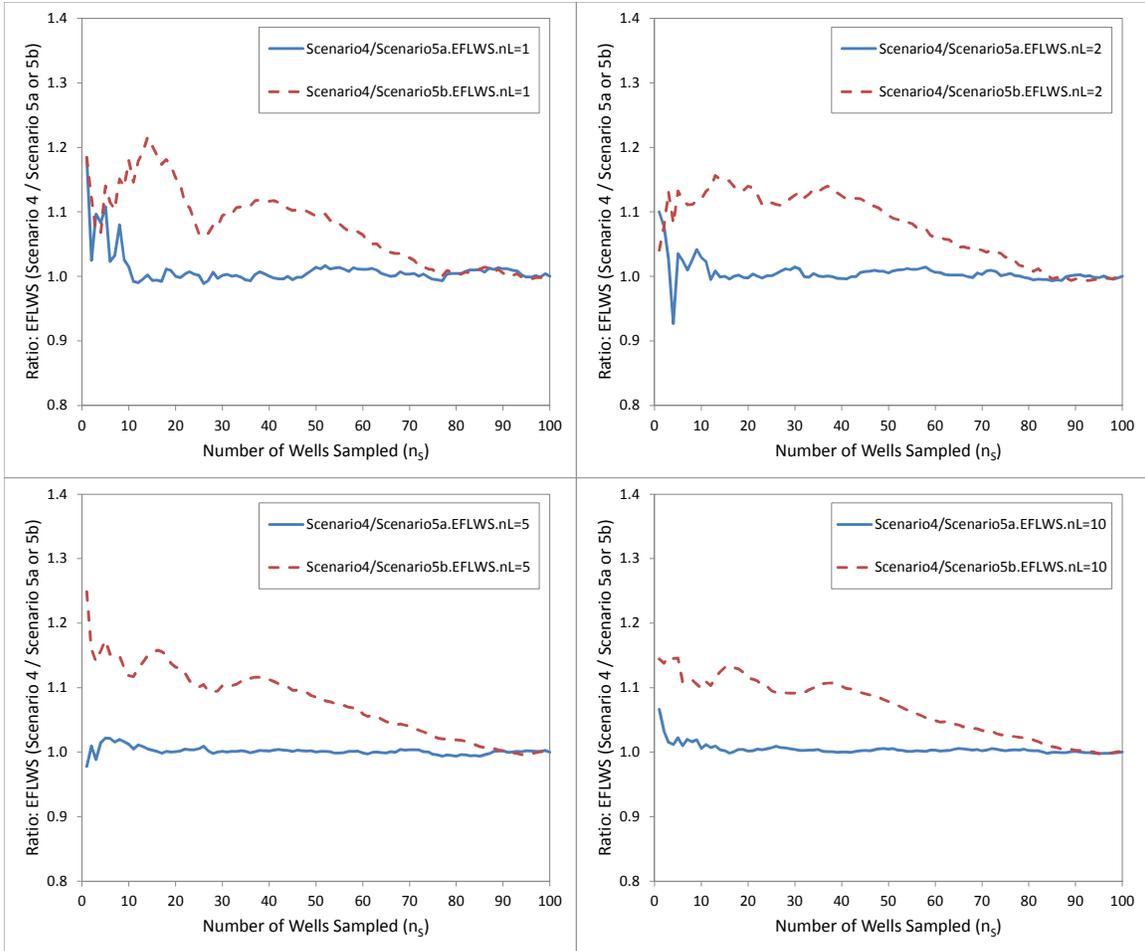


Figure 3-9. Ratios of the expected fraction of leaky wells sampled (EFLWS) for Scenario 4 / Scenario 5a (small accuracy loss, blue line) or Scenario 4 / Scenario 5b (large accuracy loss, dashed red line) as a function of n_s for $n_L = 1$ (top left), 2 (top right), 5 (bottom left), and 10 (bottom right) wells. Note: when $n_L = 1$, POLWS (probability at least one leaky well sampled) = PALWS (probability all leaky wells sampled) = EFLWS.

3.3.3 Sensitivity to Partial Wellbore Leakage Potential Scores

Figure 3-10 shows the ratio of EFLWS for (Scenario 4 / Scenario 6a) as a function of n_s and n_L . As would be expected, using two-attribute partial scores to sequence the sampling order results in a loss of statistical power as compared to the fully-informed Scenario 4. Due to the discrete nature of the scores (Table 3-3), the partial scores result in several “groups” of wells with tied scores. For example, most of the

partial scores yield sampling matrices that consist of between 6 to 10 groups of wells within which the sampling order is random. The complete deep well leakage potential scores using all four attributes also result in tied scores, with 19 groups of wells within which the sampling order is random. As a result, Figure 3-10 shows little difference (ratios near 1) between Scenarios 4 and 6a for some combinations when n_S is less than about five wells. However, the overarching trend is clear – there is a loss of statistical power with partial scores for each of the leaky well scenarios and it is most pronounced when n_S is small (generally less than 30 wells). Figure 3-11 shows the ratio of EFLWS for (Scenario 4 / Scenario 6b) as a function of n_S and n_L . There is less effect than the two-attribute case in Scenario 6a; however, there is a loss of statistical power with the three-attribute partial scores when n_S is small (generally less than 15 wells).

Figures 3-10 and 3-11 show that statistical power is *most* sensitive to omitting the number of fracture treatments, acid treatments, or well abandonment type from the deep well potential score, and *least* sensitive to omitting the number of completions. For example, the three two-attribute combinations in Scenario 6a which result in the smallest loss of statistical power are: (1) fracture/acid, (2) fracture/abandonment, and (3) acid/abandonment (Figure 3-10). Among the three-attribute scores in Scenario 6a, the combination fracture/acid/abandonment provides nearly identical performance to Scenario 4 (Figure 3-11). Analogous to Scenarios 5a and 5b, the closer the sampling knowledge to the true leakage probability, the quantifiably better the leaky well detection becomes. While these results seem to state the obvious, the methods used in this work illustrate statistical approaches that can be used to quantify different sampling plans and

thus justify one sampling plan over another at a site where $\text{Prob}[\text{Leakage}_i]$ can be quantified.

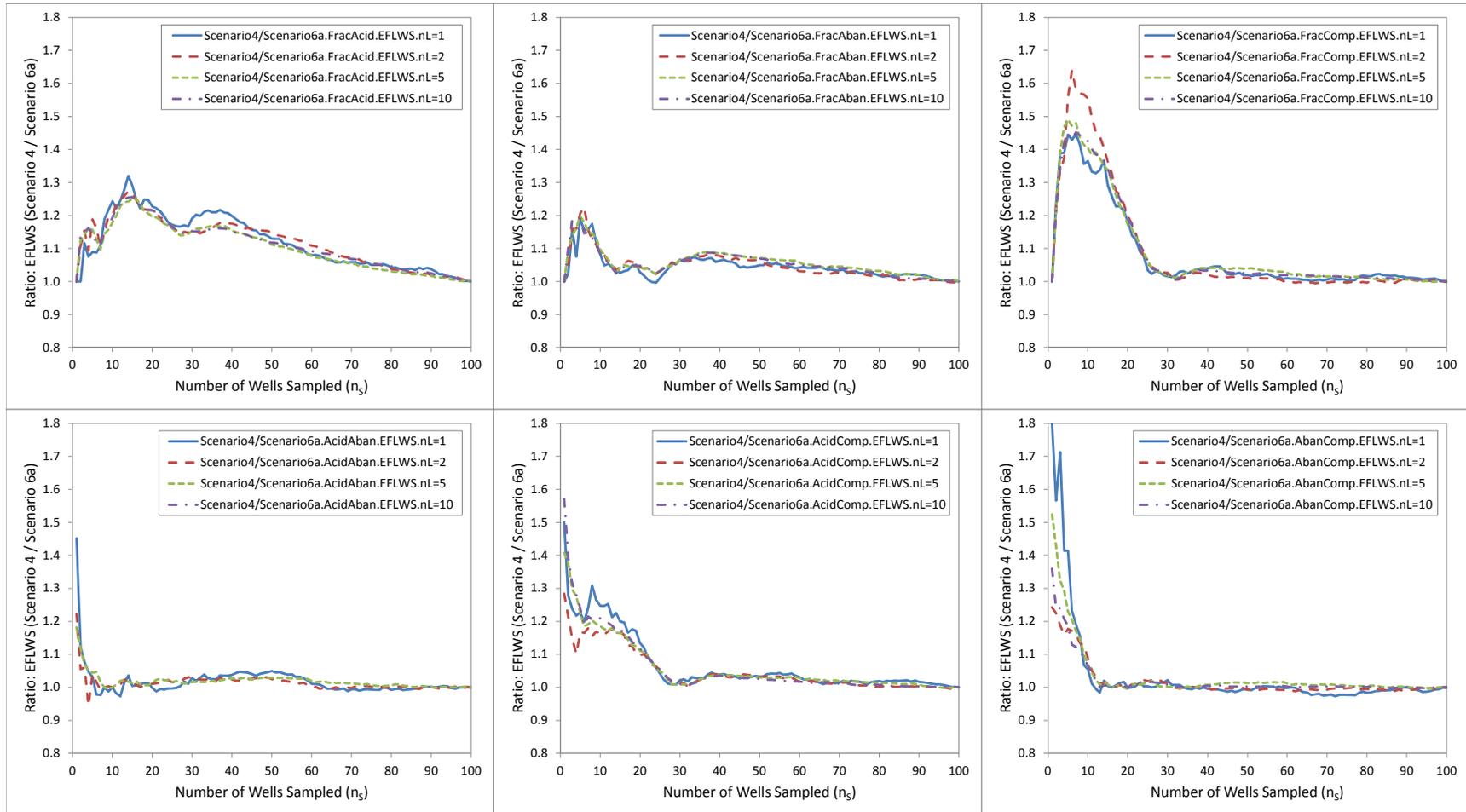


Figure 3-10. Ratio of EFLWS (Scenario 4) / EFLWS (Scenario 6a), where Scenario 6a utilizes only two of the four well attributes to develop the well sampling. The six combinations of two attributes include: (1) no. of fracture and acid treatments (upper left); (2) no. of fracture treatments and abandonment type (upper middle); (3) no. of fracture treatments and completions (upper right); (4) no. of acid treatments and abandonment type (lower left); (5) no. of acid treatments and completions (lower middle); and (6) no. of completions and abandonment type (lower right).

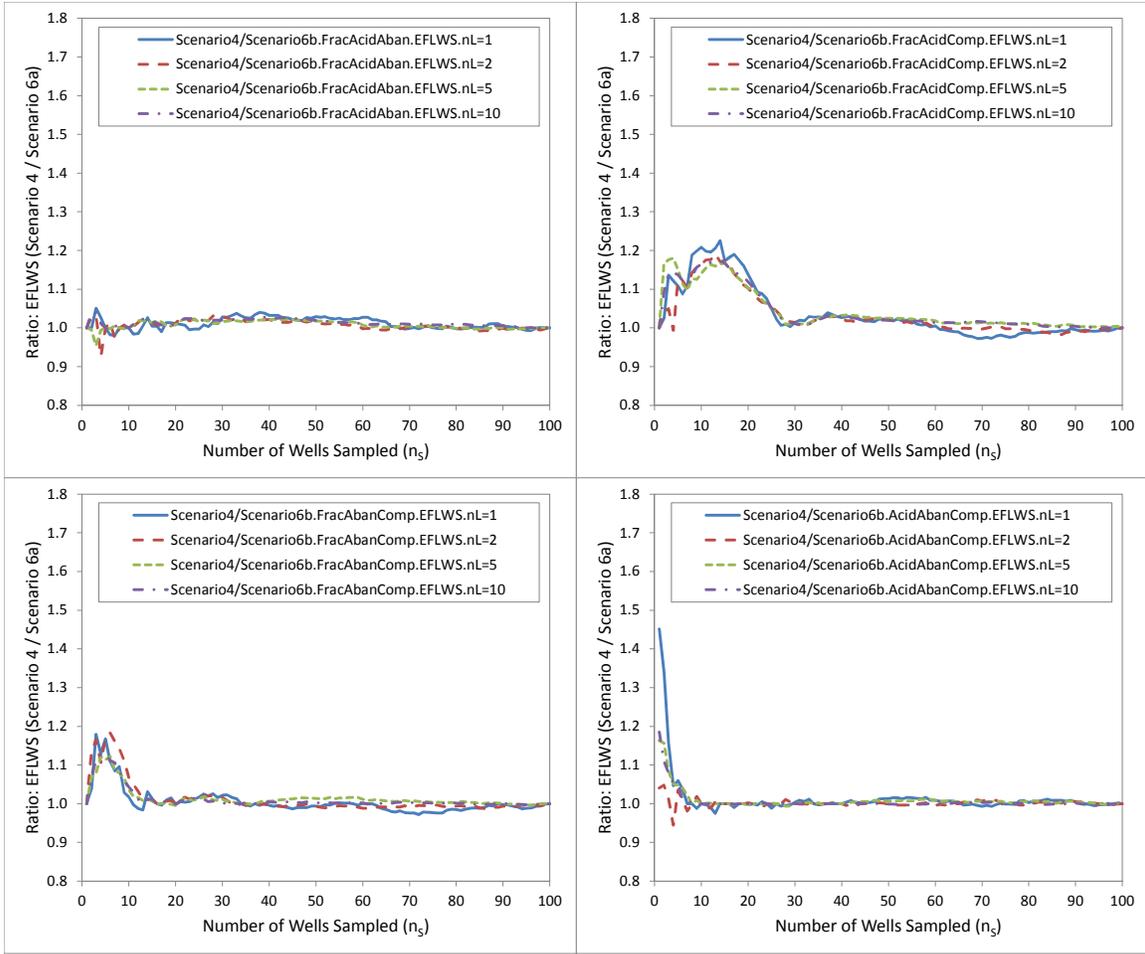


Figure 3-11. Ratio of EFLWS (Scenario 4) / EFLWS (Scenario 6b), where Scenario 6b utilizes only three of the four well attributes to develop the well sampling. The four combinations of three attributes include: (1) no. of fracture treatments, no. of acid treatments, and abandonment type (upper left); (2) no. of fracture treatments, acid treatments, and completions (upper right); (3) no. of fracture treatments, no. of completions, and abandonment type (lower left); and (4) no. of acid treatments, no. of completions, and abandonment type (lower right).

3.4 Implications for Long-Term Monitoring at CO₂ Storage Sites

Prior knowledge about the probability of well leakage can add measurable value to the ability to detect a leaking well during the monitoring program. For example, for the case of two leaking wells ($n_L=2$) and a budgetary limit of sampling 20 wells ($n_S=20$), the statistical power of detecting one leaking well (POLWS) for Scenario 1 (equiprobable well leakage/random well sampling) is only 36%. In contrast, the POLWS for Scenario 4,

where the deep well leakage potential scores are fully informative and used to develop the sampling plan, is 81%. Detecting one leaking well is important, as this will drive additional investigation work in the region surrounding the leaking well to confirm or deny the presumptive leak and to characterize the nature and extent of the leak should it turn out to be present.

The Scenario 5a and 5b results underscore the value of quality well integrity testing during the site feasibility study phase to estimate Prob[Leakage_i] for all wells within the AoR. A baseline survey which more closely reflects the true increased probability of well leakage provides better information for sampling plan development, which in turn increases the statistical power of locating a leaking well should one exist. The greater the error of the deep well leakage potential score for inferring the true well leakage probability, the more the detection performance of Scenario 4 degrades to that of Scenario 1 (even odds). Similarly, Scenarios 6a and 6b illustrate the loss of statistical power when only a subset of the well attributes used to define the leakage probability is known. Thus, Scenarios 5a, 5b, 6a, and 6b illustrate methods for quantifying the loss in detection ability due to imperfect knowledge, which may provide a means for justifying the economics associated with additional wellbore characterization during the pre-injection characterization phase for a geologic CO₂ storage project.

A host of atmospheric, near-surface, and subsurface monitoring techniques are available for CO₂ leak detection. The spatial and temporal scales of measurement for the different methods vary greatly, from those covering broad, field-wide scales such as Interferometric Synthetic Aperture Radar (e.g., InSAR) and three-dimensional (3D) seismic imaging, to small-scale samples collected from individual wells (points within a

field), such as down-hole geophysical logs or discrete well fluid samples (U.S. DOE, 2009). While the former provide field-wide information with respect to CO₂ plume behavior and migration, the latter would likely be needed as supplementary measurement tools and would almost certainly be required for confirmatory sampling in the event of a presumptive CO₂ leak. For geologic CO₂ storage in legacy oil and gas fields with large numbers of wells, the well sampling design must be considered along with the additional monitoring techniques. Ultimately, the long-term monitoring component of an MVA plan would be based on the combined knowledge of reservoir simulation, geologic characterization data, well integrity surveys, and observations during CO₂ injection, allowing the monitoring network to integrate combinations of technologies to provide the appropriate balance between risk reduction and cost effectiveness (Seto and McRae, 2011). This work underscores the value of well integrity surveys and estimates of deep well leakage potential to the well sampling design, which may be incorporated into the overall monitoring design to provide sufficient detection sensitivity of CO₂ leakage throughout the AoR.

3.5 Implicit Assumptions of the Statistical Simulation Approach

Implicit in this statistical simulation approach are five fundamental assumptions: (1) leakage may be reduced into a discrete (binary) outcome; (2) there is no time-dependence in the problem; (3) buoyant CO₂ is the emphasis; (4) no geospatial information is relevant; and (5) the scoring methodology of Watson and Bachu (2008) provides a valid estimate of well leakage probability. Discussions for each of these assumptions are provided below.

3.5.1 Monitoring as a Discrete Outcome

Monitoring is often reduced to a discrete (binary) outcome. For example, control charts (or Shewhart charts) are often used in manufacturing and business to assess whether a process is in statistical control. These charts often flag a measurement when it is above or below three standard deviations of the mean, or within 99.7% of the normal distribution probability mass (Shewhard, 1931). Multivariate extensions of control charts have also been developed (Rohlf, 1975; Beckman, 1983). Other authors investigating monitoring CO₂ leakage at geologic CO₂ storage sites have also cast monitoring as a dichotomous response of either (a) within the 99% prediction interval of baseline or (b) above this threshold concentration (Yang et al., 2011a; 2011b; 2012). The reduction of the sampling problem in this work into a simple “yes/no” for “did select/did not select” a leaky well simply extends the idea of a monitoring threshold to the well, i.e., if the well is leaking CO₂, and we conduct measurements at that well, then we are assuming that we will detect the leakage. While one might argue that the size of the leak is important, one could also argue that if you sample the wrong well, then regardless of how low your method detection limit you will still not observe a leak. Therefore selecting the correct wells is paramount to using the right technology or the size of the leak. As has been shown by other authors, detection limits are important and must be included in the overall MVA program (Yang et al., 2011a; 2011b; 2012). However, this work assumes that if you select a leaking well as part of the sampling plan, then the CO₂ leak will be detected by one or more monitoring technologies used at that discrete location.

3.5.2 *Time Independence*

Time is obviously an important variable in the fate and transport of CO₂ in the subsurface at a geologic CO₂ storage site. Annual injection rates of one million tonnes CO₂ per year add significant CO₂ mass to the subsurface and greater CO₂ mass injected over time means a greater potential for CO₂ leakage. Leakage risk profiles for CO₂ storage sites have been conjectured to track with CO₂ injection, peaking at the end of site operations and decaying over time as the CO₂ plume, pressure front, and reservoir equilibrate (Benson, 2007). However, for buoyant CO₂ leakage of the type we are exploring in this work, the free-phase CO₂ plume not only extends radially away from the CO₂ injector(s) but grows in a conic shape outward from the CO₂ injector(s), i.e., gains thickness over time. As a result, the rate of change in the radial extent of the free-phase CO₂ plume does not increase as rapidly in later time as it does earlier in the injection program.

For example, we used the analytical solution of Nordbotten et al. (2005) and the semi-analytical solutions of Vilarrasa et al. (2013) to model the CO₂-brine interface at the top of the injection reservoir (i.e., the base of the seal formation) over a 50-year injection period. The analytical solution of Nordbotten et al. (2005) was one of the first to derive analytical solutions to two-phase flow (i.e., CO₂ and brine) in porous media, and the results were shown to have reasonable agreement with numerical models such as ECLIPSE. However, Vilarrasa et al. (2013) showed that because the Nordbotten et al. (2005) analytical solution ignores CO₂ compressibility and buoyancy effects in the injection well (i.e., they inject CO₂ uniformly along the whole thickness of the aquifer), the accurate prediction of the CO₂ plume position is sensitive to the gravity number, or

the ratio of gravity to viscous forces. In contrast, the Vilarrasa et al. (2013) semi-analytical approach was shown to provide accurate estimates of CO₂ plume position and fluid pressure regardless of gravity number (as compared to numerical simulations using the finite element numerical code CODE_BRIGHT). Therefore we modeled the CO₂ plume position (CO₂-brine interface) in this work using both the analytical solution of Nordbotten et al. (2005) and the semi-analytical solutions of Vilarrasa et al. (2005). The modeling assumed a constant annual CO₂ injection rate of one million tonnes per year for 1, 5, 10, 20, 30, 40, and 50 years. The brine and CO₂ density and viscosity were estimated according to Bachu (2002) based on the estimated reservoir temperature, pressure, and salinity the basal Cambrian system on the U.S. side of the U.S./Canada border. The input parameters for the aquifer, brine, CO₂, and additional data into the models are summarized in Table 3-4. At one year into the injection program, the CO₂ plume radius was 1.4 km, and at five years it was out to 3.2 km, or an average rate of change in area of approximately 100% per year. However, at 30, 40, and 50 years into the injection program, the CO₂ plume radius was 4.0, 4.3, and 4.5 km, respectively, or an average rate of change in area of about 1% per year (Figure 3-12).

This work is aimed at the late-stage operational monitoring phase and into the post-injection site care phase, so time has less of an effect on the spatial distribution of the CO₂ plume in relation to the wells within the AoR as it would in earlier time. While CO₂ plume anisotropy and the rates of CO₂ fate and transport processes in the subsurface are obviously site-specific and therefore sensitive to the input parameters, these simulations illustrate that in late-stage operational monitoring, time dependency is less of a concern than it would be early in the injection program. As a result, this work focuses

on the steady-state condition and removes time from the solution, recognizing that a transient (i.e., time-dependent) CO₂ leakage model would have significantly greater relevance early in the injection program.

Table 3-4. Input parameters used in the analytical model of Nordbotten et al. (2005) and the semi-analytical model of Vilarrasa et al. (2013) to estimate the radial extent of the CO₂ plume during a 50-years CO₂ injection period.

Variable	Units	Symbol	Value
AQUIFER INPUT DATA			
Aquifer thickness	m	D	152
Aquifer porosity		ϕ	0.15
Aquifer permeability	m ²	k	1E-12
Aquifer temperature	K	T	380
Depth of the top of the aquifer	m	z_0	2440
Aquifer salinity	mg L ⁻¹	TDS	175,000
Hydrostatic pressure	MPa	p_0	24.4
BRINE INPUT DATA			
Brine density	kg m ⁻³	ρ_{brine}	1082
Brine viscosity	Pa·s	μ_{brine}	3.9E-04
CO₂ INPUT DATA			
CO ₂ density	kg m ⁻³	ρ_{CO_2}	550
CO ₂ viscosity	Pa·s	μ_{CO_2}	4.3E-05
ADDITIONAL INPUT DATA			
Injection well radius	m	r_p	0.5
CO ₂ volumetric injection rate	m ³ s ⁻¹	Q_0	0.06
Prescribed CO ₂ mass flow rate	kg s ⁻¹	Q_m	31.7

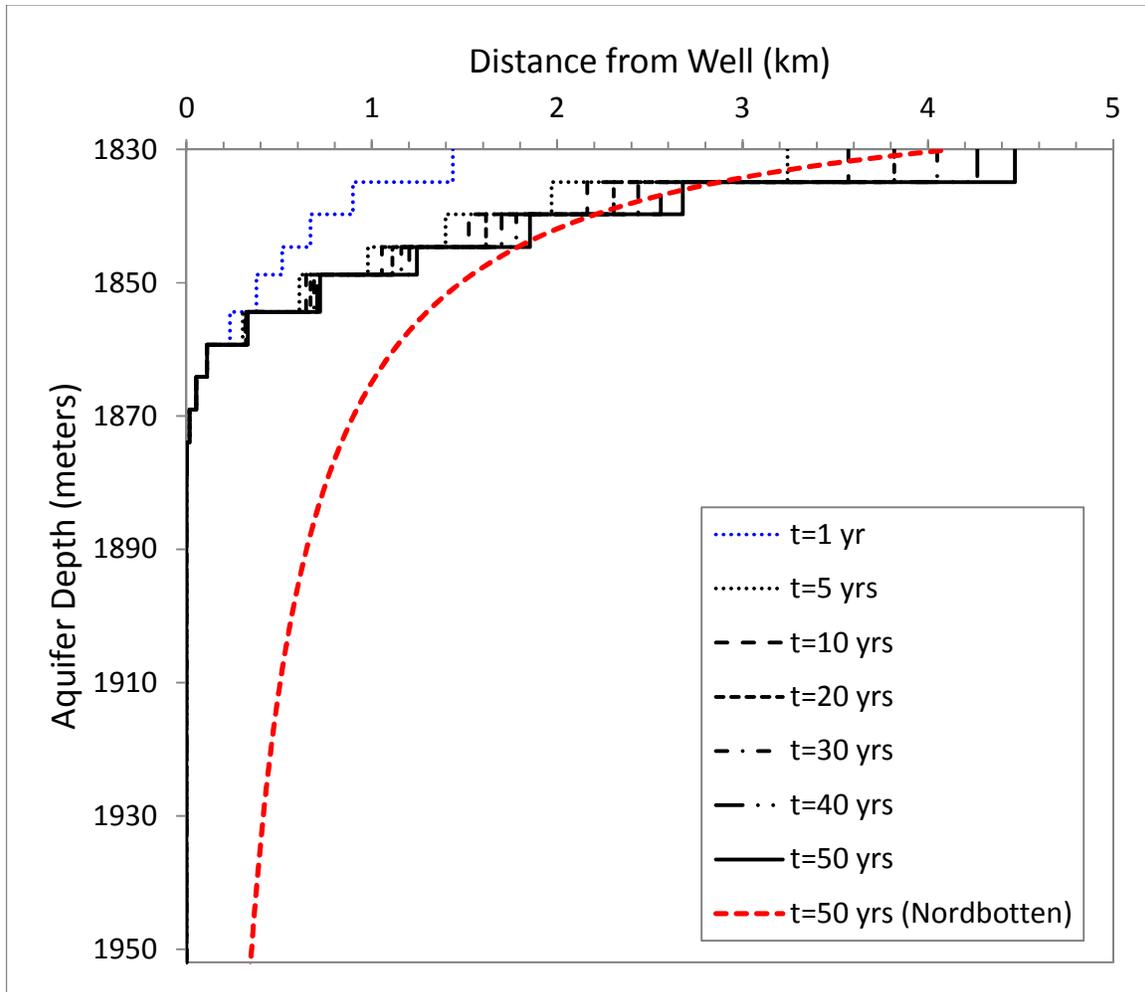


Figure 3-12. Profile from the injector ($x=0$) outward of a hypothetical CO_2 -brine interface at 1, 5, 10, 20, 30, 40, and 50 years into the CO_2 injection period using the input parameters for the aquifer, brine, and CO_2 shown in Table 3-4 and the semi-analytical model of Vilarrasa et al. (2013). The solution at 50 years for the same input parameters using the analytical model of Nordbotten et al. (2005) is shown for comparison (red dashed line).

3.5.3 Emphasis on Buoyant CO_2

This work focuses on buoyant CO_2 leakage and detection at a wellbore. A number of processes have been described in the literature for trapping CO_2 during geologic storage. These include physical trapping (stratigraphic or structural trapping of CO_2 beneath low-permeability seals), hydrodynamic trapping (CO_2 dissolution into slowly-migrating groundwater), and geochemical trapping (mineralization of CO_2 through rock-

water interactions) (IPCC, 2005). However, hydrodynamic and geochemical trapping mechanisms are relatively slow, perhaps taking centuries or millennia. In the time period that is the focus of this work (i.e., late-stage operational monitoring phase and into the post-injection site care phase – perhaps 50 years following the end of CO₂ injection), migration of buoyant CO₂ is the dominant leakage risk. For example, at the pressure and temperature conditions estimated for our analytical and semi-analytical modeling (24.4 MPa and 107 degrees Celsius), CO₂ is expected to have a density of approximately 550 kg/m³, or roughly half the density of the formation water (1082 kg/m³). This density difference drives the buoyant migration of CO₂ upwards towards the top of the reservoir layer (Eke et al., 2011). As noted by Watson and Bachu (2007) for a leak to occur, there must be (1) a leak source; (2) a driving force such as buoyancy, and (3) a leakage pathway. This work appropriately assumes that during the late-stage injection and into the post-injection site care phases, the first two conditions are met – buoyant CO₂ is present in the subsurface and buoyancy is the dominant driving force, and that any leakage pathway along the wellbore will permit CO₂ to migrate from the storage site.

3.5.4 Geospatial Information

This work uses a vector of deep well leakage potential scores (i.e., no X- or Y-coordinates), thus geospatial phenomena are not included in the calculations. Site-specific geologic characteristics such as preferential flow paths could lead to portions of the AoR having a higher risk of CO₂ leakage than others. The methods used in this work would need to be coupled with a broader reservoir simulation model to properly address geospatial correlation and other site-specific geospatial features.

3.5.5 Accuracy of the Deep Well Leakage Potential Score

This work assumes that the leakage probability is directly proportional to the deep well leakage potential score. While the scoring methods of Watson and Bachu (2007; 2008) were derived from real-world data on SCVF/GM well leaks against available well completion and abandonment factors, the weighting factors and scoring approach are unlikely to be exact predictors of the increased probability of CO₂ leakage. The scoring has not been validated against field data of wellbore permeability estimates, which are the actual drivers of wellbore leakage, not scores. For instance, it could be that all wells with a score higher than some threshold value are equally leaky in terms of their effective hydraulic parameters, and therefore the distribution of leakage probability would look quite different than the one used in this work. However, the methods used in this work are relevant regardless of the method used to assign the leakage probability. For example, other approaches for quantifying wellbore leakage probability estimates such as direct measurements of wellbore permeability could be used to define Prob(Leakage_{*i*}). The subsequent modeling steps would then be identical to those used in this work, i.e., use the multivariate form of Wallenius' non-central hypergeometric distribution to simulate the probability that a particular well is leaking and compare sampling designs to quantify the statistical power of selecting one or more leaky wells.

3.6 Conclusions

Despite the few limitations discussed above, the current work shows the value of high-quality wellbore characterization data for informing MVA sampling plans. Site-specific scoring systems using methods like the one provided in this work or empirical

measurements of wellbore integrity/permeability may be used to derive the probability of leakage, which may then be modeled to assess the number and location of wellbore samples needed to satisfy the site-specific MVA objective.

Acknowledgements

This study was funded by the U.S. Department of Energy National Energy Technology Laboratory.

3.7 References

- Azzolina, N.A.; Small, M.J.; Nakles, D.V.; and Bromhal, G.S. (2014) Effectiveness of subsurface pressure monitoring for brine leakage detection in an uncertain CO₂ sequestration system. *Stochastic Environmental Research and Risk Assessment*, 28:895-909; DOI 10.1007/s00477-013-0788-9.
- Bachu, S. (2002) Sequestration of CO₂ in geological media in response to climate change: Road map for site selection using the transform of the geologic space into the CO₂ phase space. *Energy Conversion and Management*, 43:87-102.
- Bachu, S.; Faltinson, J.; Hauck, T.; Perkins, E.; Peterson, J.; Talman, S.; and Jensen, G. (2012) *The basal aquifer in the prairie region of Canada—characterization for CO₂ storage: Final report for Stage I (Phases 1 and 2)*, Alberta Innovates Technology Futures.
- Beckman, R.J.; and Cook, R.D. (1983) Outlier.....s, *Technometrics*, 25:119-149.
- Benson, S.A. (2007) *Geological Storage of CO₂: Analogues and Risk Management*, presentation to Carbon Sequestration Leadership Forum, May 7, Pittsburgh, PA.
- Berkopec, A. (2007) HyperQuick algorithm for discrete hypergeometric distribution. *Journal of Discrete Algorithms*, 5(2):341; DOI 10.1016/j.jda.2006.01.001.
- Celia, M.A.; Bachu, S.; Nordbotten, J.M.; Gasda, S.E.; and Dahle, H.K. (2007) *Quantitative estimation of CO₂ leakage from geological storage: analytical models, numerical models and data needs*. Proc. 7th Int. Conf. Greenhouse Gas Control Technology, Vancouver, Canada.
- Celia, M.A.; Nordbotten, J.M.; Court, B.; Dobossy, M.; and Bachu, S. (2011) Field-scale application of a semi-analytical model for estimation of CO₂ and brine leakage along old wells. *International Journal of Greenhouse Gas Control*, 5:257-269; DOI 10.1016/j.ijggc.2010.10.005.
- Eke, P.E.; Naylor, M.; Haszeldine, S.; and Curtis, A. (2011) *CO₂ leakage prevention technologies*, paper SPE 145263 presented at the SPE Offshore Europe Oil and Gas Conference and Exhibition, Aberdeen, UK, 6-8 September.
- Fog, A. (2008) Sampling methods for Wallenius' and Fisher's noncentral hypergeometric distributions. *Communications in Statistics – Simulation and Computation*, 37(2): 241-257; DOI 10.1080/03610910701790236.
- Fog, A. (2014) *BiasedUrn: Biased Urn model distributions*. R package version 3.02. <http://CRAN.R-project.org/package=BiasedUrn> (accessed May 7, 2014).

Gilbert, R.O. (1987) *Statistical Methods for Environmental Pollution Monitoring*; John Wiley and Sons, Inc.: New York.

Glazewski, K.; Hamling, J.; Peck, W.; Doll, T.; Laumb, J.; Gorecki, C.; Steadman, E.; and Harju, J. (2013) *Wellbore evaluation of the basal Cambrian system*; Plains CO₂ Reduction Partnership Phase III Task 16 Deliverable D90 for U.S. Department of Energy National Energy Technology Laboratory Cooperative Agreement No. DE-FC26-05NT42592; Energy & Environmental Research Center: Grand Forks, North Dakota, September.

Hitchon, B., Ed. (2012) *Best Practices for Validating CO₂ Geological Storage: Observations and Guidance from the IEAGHG Weyburn-Midale CO₂ Monitoring and Storage Project*; Geoscience Publishing: Alberta, Canada.

Intergovernmental Panel on Climate Change (IPCC). (2005). *Special Report on Carbon Dioxide Capture and Storage; Prepared by Working Group III of the Intergovernmental Panel on Climate Change* [Metz, B., O. Davidson, H. C. de Coninck, M. Loos, and L. A. Meyer (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jaynes, E.T. (2003) *Probability Theory: The Logic of Science*; Cambridge University Press.

Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com)

NIST/SEMATECH. (2014) *e-Handbook of Statistical Methods*, 1.3.6.6.11. Gamma Distribution; <http://www.itl.nist.gov/div898/handbook/eda/section3/eda366b.htm>.

Nogues, J.P.; Nordbotten, J.M.; and Celia, M.A. (2011) Detecting leakage of brine or CO₂ through abandoned wells in a geological sequestration operation using pressure monitoring wells. *Energy Procedia*, 4:3620-3627; DOI 10.1016/j.egypro.2011.02.292.

Nordbotten, J.M.; Celia, M.A.; and Bachu, S. (2005) Injection and storage of CO₂ in deep saline aquifers: analytical solution for CO₂ plume evolution during injection. *Transport in Porous Media*, 58: 339-360.

R Core Team. (2014) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/> (accessed May 7, 2014).

Rohlf, F.J. (1975) Generalization of the gap test for the detection of multivariate outliers, *Biometrics*, 31:93-102.

Seto, C.J.; and McRae, G.J. (2011) Reducing risk in basin scale CO₂ sequestration: A framework for integrated monitoring design. *Environmental Science and Technology*, 45:845–859; DOI 10.1021/es102240w.

Shewhart, W.A. (1931) *Economic Control of Quality of Manufactured Product*. New York: D. Van Nostrand Company. pp. 501.

U.S. Department of Energy (U.S. DOE). (2009) *Best Practices for: Monitoring, Verification, and Accounting of CO₂ Stored in Deep Geologic Formations*; U.S. Department of Energy, National Energy Technology Laboratory, DOE/NETL-311/081508, March.

U.S. Department of Energy (U.S. DOE). (2011) *Best Practices for: Risk Analysis and Simulation for Geologic Storage of CO₂*; U.S. Department of Energy, National Energy Technology Laboratory, DOE/NETL-2011/1459, March.

U.S. Department of Energy (U.S. DOE). (2012) *The 2012 United States Carbon Utilization and Storage Atlas – Fourth Edition (Atlas IV)*; U.S. Department of Energy, National Energy Technology Laboratory.

U.S. Environmental Protection Agency (U.S. EPA). (2002) *Guidance on Choosing a Sampling Design for Environmental Data Collection*, Office of Environmental Information; U.S. Environmental Protection Agency, EPA-240-R-02-005: Washington, D.C., December.

U.S. Environmental Protection Agency (U.S. EPA). (2011) *Geologic Sequestration of Carbon Dioxide Draft Underground Injection Control (UIC) Program Class VI Well Area of Review Evaluation and Corrective Action Guidance for Owners and Operators*; U.S. Environmental Protection Agency, Office of Water (4606M), EPA-816-D-10-007, March.

Vilarrasa, V.; Carrera, J.; Bolster, D; and Dentz, M. (2013) Semianalytical solution for CO₂ plume shape and pressure evolution during CO₂ injection in deep saline formations. *Transport in Porous Media*, 97(1):43-65.

Wallenius, K.T. (1963) *Biased Sampling: The Non-central Hypergeometric Probability Distribution*. Ph.D. Dissertation, Stanford University, Department of Statistics.

Watson, T.L.; and Bachu, S. (2007) *Evaluation of the potential for gas and CO₂ leakage along wellbores*. SPE Paper 106817 presented at the Society of Petroleum Engineers (SPE) E&P Environmental and Safety Conference, Galveston, Texas, March 5-7.

Watson, T.L.; and Bachu, S. (2008) *Identification of wells with high CO₂ leakage potential in mature oil fields developed for CO₂ enhanced oil recovery*. SPE Paper 112924 presented at the Society of Petroleum Engineers (SPE) Improved Oil Recovery Symposium, Tulsa, Oklahoma, April 19-23.

Yang, Y.M.; Small, M.J.; Ogretim, E.O.; Gray, D.D.; Bromhal, G.S.; Strazisar, B.R.; and Wells, A.W. (2011a) Probabilistic design of a near-surface CO₂ leak detection system. *Environmental Science and Technology*, 45:6380–6387; DOI 10.1021/es104379m.

Yang, Y.M.; Small, M.J.; Junker, B.; Bromhal, G.S.; Strazisar, B.R.; and Wells, A.W. (2011b) Bayesian hierarchical models for soil CO₂ flux and leak detection at geologic sequestration sites. *Environmental Earth Science*, 64(3):787-798; DOI 10.1007/s12665-011-0903-5.

Yang, Y.M.; Small, M.J.; Ogretim, E.O.; Gray, D.D.; Wells, A.W.; Bromhal, G.S.; and Strazisar, B.R. (2012) A Bayesian belief network (BBN) for combining evidence from multiple CO₂ leak detection technologies. *Greenhouse Gases: Science and Technology*, 2(3):185-199; DOI 10.1002/ghg.1284.

Chapter 4: CO₂ storage associated with CO₂ enhanced oil recovery: A statistical analysis of historical operations⁶

Abstract

This work analyzes a database of 31 existing CO₂ enhanced oil recovery (EOR) projects that was compiled for the estimation of oil reserves to better understand the CO₂ retention, incremental oil recovery, and net CO₂ utilization for these oil fields. The measured data begin at the start date of the CO₂ flood and extend through the year 2007. Cumulative CO₂ retention (in the formation), incremental oil recovery factors, and net CO₂ utilization factors were calculated for each of the sites. To express all site data on a common dimensionless scale, the data were extrapolated to 300% cumulative hydrocarbon pore volume (HCPV) by fitting nonlinear functions. Summary statistics were then calculated from 0% to 300% HCPV. Across all 31 sites, the 10th, 50th (median), and 90th percentile values for the three factors at 300% HCPV were: CO₂ retention: 23.1%, 48.3%, and 61.8% retained; incremental oil recovery: 5.3%, 12.2%, and 21.5% of OOIP (original oil in place); and net CO₂ utilization: 4.8, 8.7, and 10.5 Mscf/STB (stock-tank barrel). This work employs a novel approach that incorporates nonlinear functions to quantify uncertainty in the estimated values as a function of HCPV and to describe the shape of the CO₂ retention or incremental oil recovery response with a handful of parameters, providing insight into the behavior of the reservoir across the entire timeline of the CO₂ flood. These nonlinear curve fits are focused on statistical

⁶ Chapter 4 was the basis for the peer-reviewed publication. In the publication, some of the text and tables were shortened to conform to the journal format. Azzolina, N.A.; Nakles, D.V.; Gorecki, C.D.; Peck, W.D.; Ayash, S.C.; Melzer, L.S.; and Chatterjee, S. (2015). CO₂ storage associated with CO₂ enhanced oil recovery: A statistical analysis of historical operations. *International Journal of Greenhouse Gas Control*, 37:384-397.

inference – i.e., what is the likely outcome and uncertainty ranges for CO₂ retention, incremental oil recovery, and net CO₂ utilization given the historical data from the 31 CO₂ EOR sites? However, the approach described in this work also provides useful information for prediction – i.e., given a set of inputs from another site with similar geology, what are plausible ranges in outcomes for each of these factors? Consequently, the results of this work can be used to estimate the potential range of expected performance for similar candidate oil fields that are not currently undergoing CO₂ injection, including estimates of the associated CO₂ storage potential of these candidate fields. The results of this work allow estimation of CO₂ storage capacity in CO₂-EOR operations with various degrees of confidence. The sites in the dataset reflect water–alternating gas CO₂ floods – all within the continental United States and heavily dominated by the West Texas carbonate floods. Other floods outside of this region, where the data were available, are also included in this study (i.e., the Rocky Mountain region and the State of Oklahoma).

4.1 Introduction

The need to simultaneously reduce atmospheric concentrations of carbon dioxide (CO₂) and provide energy to satisfy an ever-growing worldwide energy demand presents a seemingly intractable societal challenge. One proven technology which both produces oil and permanently stores CO₂ in the subsurface is CO₂ enhanced oil recovery (CO₂ EOR). CO₂ EOR refers to the process whereby CO₂ is injected into the subsurface at an oil field, after which it mixes with the oil to swell it and reduce the oil viscosity, making it lighter and detaching it from the rock surfaces. These subsurface alterations cause the

oil to flow more freely within the reservoir so that it will flow from the injection well to producer wells. During this process, nearly all of the purchased CO₂ delivered to the oil field remains securely trapped within the deep geologic formation (Melzer, 2012). Consequently, CO₂ EOR provides one viable means for offsetting carbon emissions from oil production and combustion via geologic storage of the carbon. The CO₂ EOR technique was first tested at large scale in the 1970s in the Permian Basin of West Texas and southeastern New Mexico, thus it has been used in the United States for over 40 years. As of 2014, a total of 136 active CO₂ EOR projects have been identified in the U.S. (Kuuskraa and Wallace, 2014).

Historically, most of the CO₂ supply to CO₂ EOR projects has come from larger *natural* sources, which obviously do not contribute toward a reduction in *anthropogenic* CO₂ emissions (Kuuskraa and Wallace, 2014). However, in an effort to reduce concentrations of CO₂ in the atmosphere that are caused by anthropogenic sources, the U.S. Department of Energy (DOE) is pursuing carbon capture, utilization, and storage (CCUS) as one approach in a portfolio of greenhouse gas reduction strategies. CCUS involves (1) separating CO₂ from an industrial process; (2) transporting the CO₂ to a geologic storage location; and (3) injecting and storing the CO₂ in a geologic reservoir for long-term isolation from the atmosphere (IPCC, 2005). In the case of CO₂ EOR, the CO₂ is “utilized” in the sense that it is serving an additional value-added purpose of producing incremental oil prior to geologic storage (in contrast to CO₂ injection into saline reservoirs which only serve as CO₂ storage units and do not provide incremental oil recovery). As part of its Carbon Storage Program, DOE is working toward the demonstration and commercialization of CCUS through its Infrastructure, Core Research

and Development (R&D), and Global Collaborations areas (U.S. DOE, 2013), which was also the source of funding for this study.

CO₂ EOR is most commonly a “tertiary production phase” process, which is used after the primary and secondary production phases have been completed. During the life-cycle of the CO₂ EOR process (commonly referred to as the “CO₂ flood”), CO₂ is injected into the reservoir at designated injection wells and a mixture of oil, CO₂, and water is produced at production wells. These produced fluids are separated at the surface. Oil is transported off-site to refineries and other end users. The produced CO₂ is separated from the other fluids (processed), dried, re-compressed and re-injected into the subsurface. This “recycling” of the produced CO₂ prevents it from being released to the atmosphere and provides substantial savings to the oil field operator that would otherwise have had to purchase replacement CO₂ volumes (U.S. DOE, 2010; Melzer, 2012). Therefore, the CO₂ EOR process is essentially a “closed loop” for CO₂. However, not all of the injected CO₂ is produced, as a significant fraction of the CO₂ is retained in the reservoir. To maintain a specified injection ratio of water and CO₂, the recycled CO₂ is supplemented with the purchased CO₂. Therefore, as the CO₂ flood continues through time, the total volume of injected CO₂ is comprised of both recycled CO₂ and purchased CO₂. Often, CO₂ floods proceed by alternating the injection of CO₂ and water, which are designated as “water-alternating-gas”, or WAG floods. The WAG approach helps to reduce the tendency for the lower viscosity CO₂ to finger its way ahead of the displaced oil and therefore improves the sweep efficiency of the flood (U.S. DOE, 2010). Sweep efficiency refers to the effectiveness of the injected CO₂ to contact the volume of the reservoir where the oil resides (Green and Willhite, 1998). The produced water may also

be re-injected into the reservoir as part of the WAG process. Figure 4-1 illustrates the fluid flows (oil, CO₂, and water) in a typical CO₂ EOR project.

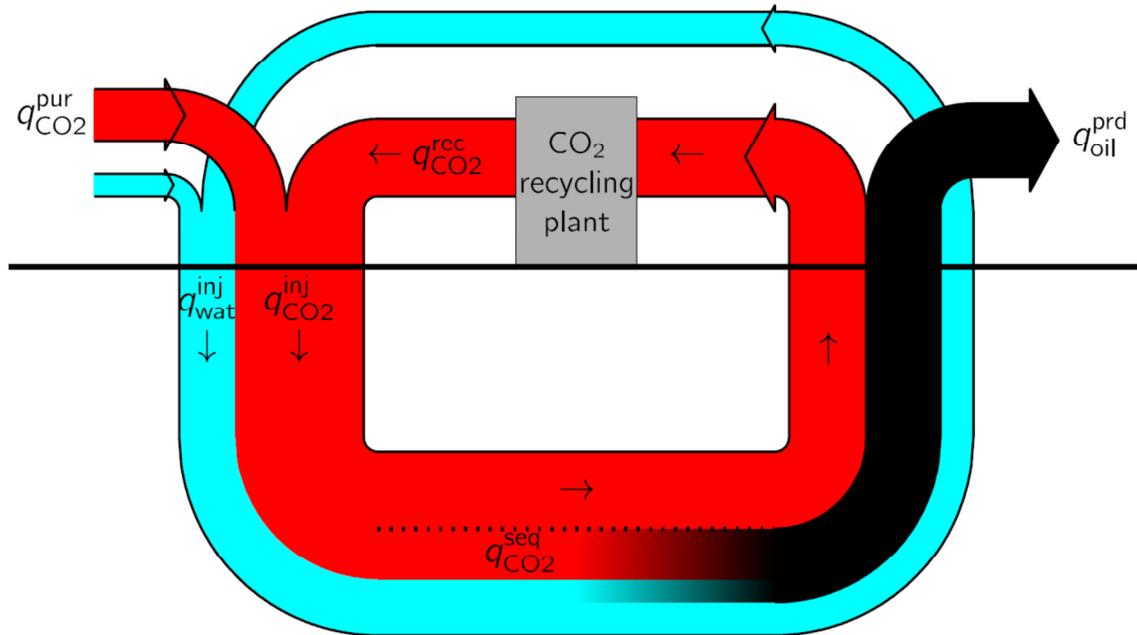


Figure 4-1. Schematic of fluid flows in a CO₂ EOR project. $q_{CO_2}^{pur}$ refers to purchased CO₂; $q_{CO_2}^{rec}$ refers to recycled CO₂; $q_{CO_2}^{inj}$ refers to total injected CO₂, which is comprised of both purchased CO₂ and recycled CO₂; $q_{CO_2}^{seq}$ refers to the sequestered CO₂ that is permanently stored in the subsurface and not available for recycle; q_{wat}^{inj} refers to injected water; and q_{oil}^{prd} refers to incremental oil production (Source: van 't Veld et al., 2013).

Industry experience and simulation studies suggest that the amount of CO₂ stored is a function of both geologic and operational controls. Geologic controls are specific to the reservoir, and include lithology, permeability, heterogeneity, and other physical features unique to the reservoir. Operational controls are specific to the design and operation of the CO₂ flood, including injection pattern (the geometrical arrangement of injector and producer wells), pattern spacing (the distance between injectors and producers), the volume of CO₂ injected, and the ratio of injected water to injected CO₂ (the “WAG ratio”). The interplay of geologic and operational controls leads to

uncertainty in the amount of CO₂ storage that occurs at an individual site. Analogously, these same controls affect the quantity of incremental oil that is produced (ARI and Melzer, 2010; Hill et al., 2013; van 't Veld et al., 2013; 2014; Ettehadtavakkol, 2014).

The primary objective of this work is to improve the overall quantitative understanding of CO₂ retention, incremental oil recovery, and net CO₂ utilization, which are factors that significantly influence the long-term performance and economic viability of CO₂ EOR projects. CO₂ retention is a metric that expresses the fraction of total injected CO₂ that is not recycled but remains in the subsurface. CO₂ storage is a metric that expresses the fraction of total purchased CO₂ that remains in the subsurface. Incremental oil recovery is the amount of oil that is produced solely from the CO₂ flood (i.e., not from the previous primary and secondary production phases). Lastly, net CO₂ utilization tracks the amount of purchased CO₂ that is required to produce a barrel of oil. This work provides a statistical analysis of an historical industry dataset, which allows modeling results to be compared against real-world, observed injection and production data. Other authors have summarized multi-site CO₂ EOR field data in the past (e.g., Brock and Bryan, 1989; Hadlow, 1992; Manrique et al. 2007), or have extended predictions of CO₂ EOR performance into the future using simulation models (e.g., Merchant, 2010). However, this work differs significantly from past publications in three major ways. First, empirical models in this study were fit across 31 sites with real-world injection and production data, which represents one of the largest CO₂ EOR datasets ever to be analyzed and summarized in the peer-reviewed literature. Second, by incorporating a statistical model that fits the entire data record, as opposed to extending long-term trends via decline-curve analysis or other forecasting tools, this approach allows

inference about the overall shape of the CO₂ retention, incremental oil recovery, and net CO₂ utilization responses through time. Finally, this approach and data set significantly extend the understanding of CO₂ EOR behavior and uncertainty quantification further into the future (i.e., at higher total volumes of fluid injected into the reservoir) than has been presented in previous work.

4.2 Methods

4.2.1 Industry Data Set and Parameters

Monthly reservoir performance data from 31 CO₂ EOR sites (hereafter referred to as the “Industry Data Set”) were provided by Melzer Consulting, Inc. and a worldwide reservoir appraisal company (confidential). These data have been used to develop petroleum reserve estimates for the operators of the field and to prepare annual petroleum reserves certifications for filers with the U.S. Securities and Exchange Commission. They include the following five parameters for each of the 31 CO₂ EOR sites:

- 1) **Time (years):** This is the time in years since the start of the CO₂ flood. All 31 sites are expressed on a common time scale such that time zero (t_0) represents the beginning of the CO₂ flood across all projects, with data continuing up to and including the year 2007;
- 2) **CO₂ injected (%HCPV):** This is the total cumulative volume of CO₂ injected (purchased plus recycled CO₂) expressed as %HCPV;
- 3) **CO₂ produced (%HCPV):** This is the total cumulative volume of CO₂ produced expressed as %HCPV. It is equivalent to the recycled volume of CO₂;

- 4) **CO₂ + H₂O injected (%HCPV):** This is the total cumulative volume of CO₂ and water injected expressed as %HCPV. Most, but not all, CO₂ floods utilize the WAG injection method wherein stages of gas injection (predominantly CO₂) are followed by injection of liquid only (predominantly water). The normalizing factor used herein tracks both the CO₂ and the water volumes in contrast to just the CO₂ volumes; and
- 5) **Cumulative incremental oil recovery (% original oil in place [OOIP]):** The cumulative incremental oil recovery data are expressed as %OOIP, where the baseline waterflood and any infill drilling contributions have been subtracted from the total oil production to determine the incremental oil production resulting from CO₂ injection.

4.2.1 Dimensionless Variables

In comparing reservoir performance data across multiple sites, it is useful to express the data using dimensionless variables, which are independent of any particular system. These calculations were conducted by the worldwide reservoir appraisal company who compiled the database. The first step is to define OOIP, which is the amount of oil in the reservoir prior to any production and is a function of the volume of the reservoir (area and thickness), porosity, and the original oil saturation, as shown in Eq. 1 (Green and Willhite, 1998).

$$N = Ah\phi S_{oi} \frac{1}{B_{oi}} \left(\frac{1}{5.615 \text{ ft}^3/\text{bbl}} \right) \quad (4-1)$$

Where:

N	= Original oil in place (stock tank barrels [STB]);
A	= Reservoir area (ft ²);
h	= Reservoir thickness (ft);
ϕ	= Reservoir porosity (fraction);
S _{oi}	= Original oil saturation (volume fraction);
B _{oi}	= Initial oil formation volume factor (reservoir barrels [RB]/STB);

Reservoir barrel (RB) refers to the volume of oil at reservoir pressure and temperature conditions. Stock tank barrel (STB) refers to the volume of oil after production, at surface pressure and temperature, which has been established as 15.6°C and 101.325 kPa [60°F and 14.696 psia], respectively. OOIP is commonly reported in units of STB.

The volumes of CO₂ injected, CO₂ produced, and CO₂ + H₂O injected may be expressed as dimensionless variables in units of HCPV, as opposed to a volume or mass measure. HCPV represents the pore volume of the reservoir that is occupied by hydrocarbons. One hundred percent HCPV (which is equivalently reported in the literature as 1.0 HCPV) is equivalent to the OOIP. However, the amount of CO₂ or CO₂ + H₂O injected is not bounded between zero and 100%, as more than 100% HCPV may be injected into the reservoir over the life-cycle of the CO₂ flood due to the fact that there is not perfect displacement, i.e., sweep efficiency, of the OOIP, which requires the recycling of produced CO₂ plus additionally purchased CO₂ beyond 100% HCPV to achieve the desired oil recoveries (Figure 1). The formula for determining %HCPV of CO₂ injected or produced is as follows (Green and Willhite, 1998):

$$HCPV = \frac{RB \text{ CO}_2}{OOIP \times B_{oi}} \quad (4-2)$$

Where:

HCPV = Hydrocarbon pore volume (%);

RB CO₂ = Reservoir barrels (RB) of CO₂;

OOIP = Original oil in place [stock tank barrels (STB)]

B_{oi} = Initial oil formation volume factor (RB/STB)

Incremental oil recovery represents oil that is produced following the primary or secondary oil production, i.e., during tertiary recovery from the CO₂ flood, and hence “incrementally” adds to the entire oil production from the field. Incremental oil recovery may also be expressed as a dimensionless variable, the incremental oil recovery factor, which is expressed in units of %OOIP, as opposed to a volumetric measure, as shown in Eq. 3 (Green and Willhite, 1998).

$$RF = \frac{N_p}{N} \quad (4-3)$$

Where:

RF = Incremental oil recovery factor [unitless];

N_p = Cumulative incremental oil production [STB]; and

N = Original oil in place [STB].

4.2.2 *CO₂ Retention*

The formula for CO₂ retention that is used throughout this paper (Eq. 4) is widely used within the oil and gas industry and expresses CO₂ retention as a fraction between 0 and 1 (Melzer, 2012):

$$\text{CO}_2 \text{ retention} = \frac{(\text{total CO}_2 \text{ injected} - \text{CO}_2 \text{ produced})}{\text{total CO}_2 \text{ injected}} \quad (4-4)$$

Where:

CO₂ retention = percent of injected CO₂ retained in the reservoir (%);

total CO₂ injected = total injected volumes of CO₂ [purchased plus recycled CO₂] (%HCPV); and

CO₂ produced = total produced volumes of CO₂ [recycled CO₂] (%HCPV).

The CO₂ retention computed using Eq. 4 is a *cumulative* CO₂ retention, and represents the fraction of total injected CO₂ that is not recycled.

4.2.3 *Net CO₂ Utilization Factors*

This paper calculates and discusses *net* CO₂ utilization expressed as Mscf of CO₂ per STB of oil produced. Net CO₂ utilization does not include the recycled CO₂ component and, therefore, only incorporates the purchased CO₂ volumes into the calculation. In contrast, *gross* CO₂ utilization includes the total amount of CO₂ injected, which incorporates both purchased and recycled CO₂ volumes into the calculation as discussed further below.

The net CO₂ utilization factor is commonly expressed as the amount of purchased CO₂ used to recover a barrel of oil, in units of Mscf/STB (Wiggins and Hughes, 2005). The units needed to derive the CO₂ utilization factor are different from those provided in the Industry Data Set, which expresses injection and production data using dimensionless variables – injected CO₂ volumes are in units of %HCPV (not Mscf) and incremental oil production is in units of %OOIP (not STB).

For the purposes of calculating net CO₂ utilization, this work assumed that the volume of purchased CO₂ is equal to the total CO₂ injected less the CO₂ recycled from the previous time step. This assumption was based on the fact that CO₂ losses are *de minimus*, and therefore, contribute a relatively small amount of error into the calculation (Fox, 2009; DOE, 2010a; Melzer, 2012). We then used OOIP, initial oil formation volume factor (B_{oi}), and CO₂ formation volume factor (B_{CO2}) to derive the purchased CO₂ volume in units of Mscf, as per Eq. 5 (Green and Willhite, 1998):

$$V_{CO2,purchased,Mscf} = \frac{V_{CO2,purchased,HCPV} \times N \times B_{oi}}{B_{CO2}} \quad (4-5)$$

Where:

- V_{CO2,purchased,Mscf} = Cumulative volume of purchased CO₂ injected (Mscf);
- V_{CO2,purchased,HCPV} = Cumulative volume of purchased CO₂ injected (HCPV);
- N = Original oil in place [STB];
- B_{oi} = Initial oil formation volume factor (RB/STB); and
- B_{CO2} = CO₂ formation volume factor (RB/Mscf).

The B_{oi} , B_{CO_2} , and N values for each site that were used in Eq. 5 to convert the units of the purchased CO_2 from %HCPV to Mscf were obtained from the literature; however, literature values for all of these parameters were not available for all 31 sites. As a result, net CO_2 utilization factors could only be computed for 16 of the 31 sites.

Cumulative incremental oil recovery in the Industry Data Set is expressed as %OOIP. We converted %OOIP into units of STB prior to computing the net CO_2 utilization factor by rearranging Eq. 3, as shown in Eq. 6:

$$N_p = RF \times N \quad (4-6)$$

Where:

N_p = Cumulative incremental oil production [STB];

N = Original oil in place [STB]; and

RF = Incremental oil recovery factor [unitless].

For the 16 sites for which we could accurately express both the cumulative volume of purchased CO_2 injected (Mscf) and the cumulative incremental oil production (STB) in the appropriate units, the net CO_2 utilization factor for each site was calculated as follows (Eq. 7):

$$UF_{CO_2,net} = \frac{V_{CO_2,purchased,Mscf}}{N_p} \quad (4-7)$$

Where:

- $UF_{CO_2,net}$ = Net CO₂ utilization factor (Mscf/STB);
 $V_{CO_2,purchased,Mscf}$ = Cumulative volume of purchased CO₂ injected (Mscf);
 and
 N_p = Cumulative incremental oil production (STB).

The net CO₂ utilization factors computed in this paper are *cumulative* net CO₂ utilization factors for each site, where the cumulative CO₂ injected at the site is divided by the cumulative oil produced at the site up to the %HCPV that was selected for the calculation.

4.2.4 Nonlinear Regression

As previously noted the Industry Data Set includes measured data that begins at the start date of the CO₂ flood and extends up to and including the year 2007. It is difficult to assess CO₂ retention, incremental oil recovery, and net CO₂ utilization across multiple sites when there is no common basis for comparison. For example, at 10 years into the CO₂ flood, Sites A and B had injected 50% and 65% HCPV of CO₂ + H₂O , respectively; therefore, simply using the time since the start of the CO₂ flood is not a sufficient parameter with which to compare multiple sites. Therefore, to express all site data on a common dimensionless scale, the reservoir performance data were evaluated as a function of total CO₂ + H₂O injected in units of %HCPV, and were extrapolated to 300% HCPV for each of the sites to allow for multi-site comparisons of the above factors from 0% to 300% HCPV. This extrapolation was done by fitting a nonlinear function to the data using nonlinear regression. There were two steps to the nonlinear regression

modeling: 1) selecting the nonlinear regression model form and 2) fitting the nonlinear regression model parameters.

The curves produced by plotting CO₂ retention and incremental oil recovery versus CO₂ + H₂O injected all exhibited a sigmoidal shape (Figure 4-2 and Figure 4-3, respectively). CO₂ retention and incremental oil recovery factors are both expressed as a value between 0 and 1 (i.e., a percentage). Sigmoidal-shaped curves bounded by the interval from 0 to 1 are commonly modeled using a logistic function (Gelman and Hill, 2007; Carlin and Louis, 2009). However, the logistic function is symmetrical around its inflection point. Since the CO₂ retention and incremental oil recovery curves show a different type of response, i.e., the slopes are steeper at the early part of the CO₂ flood, it was necessary to fit these sigmoidal curves using a four-parameter log-logistic function. This function defines the response, y , at any value of x by an equation with four parameters: minimum response, maximum response, the natural log(inflection point), and the slope factor (Minitab Statistical Software, 2014):

$$y = A + (B - A) / \{1 + \exp[D \times \ln(x / C)]\}, \quad (4-8)$$

Where:

y = CO₂ retention (%) or incremental oil recovery (%OOIP);

x = total cumulative volume of CO₂ + H₂O injected (%HCPV);

\ln = natural log to the base e (approximately equal to 2.71828);

- A = fitted model parameter equal to the lower asymptote at infinite CO₂ + H₂O injected (for CO₂ retention) or the upper asymptote at infinite CO₂ + H₂O injected (for incremental oil recovery);
- B = fixed (constant) model parameter equal to the upper asymptote of 100% (for CO₂ retention) or the lower asymptote of 0% (for incremental oil recovery);
- C = fitted model parameter equal to the inflection point in the sigmoidal shape where the curve changes from concave to convex (for CO₂ retention) or from convex to concave (for incremental oil recovery); and
- D = fitted model parameter for the slope factor of the curve between the upper and lower asymptotes.

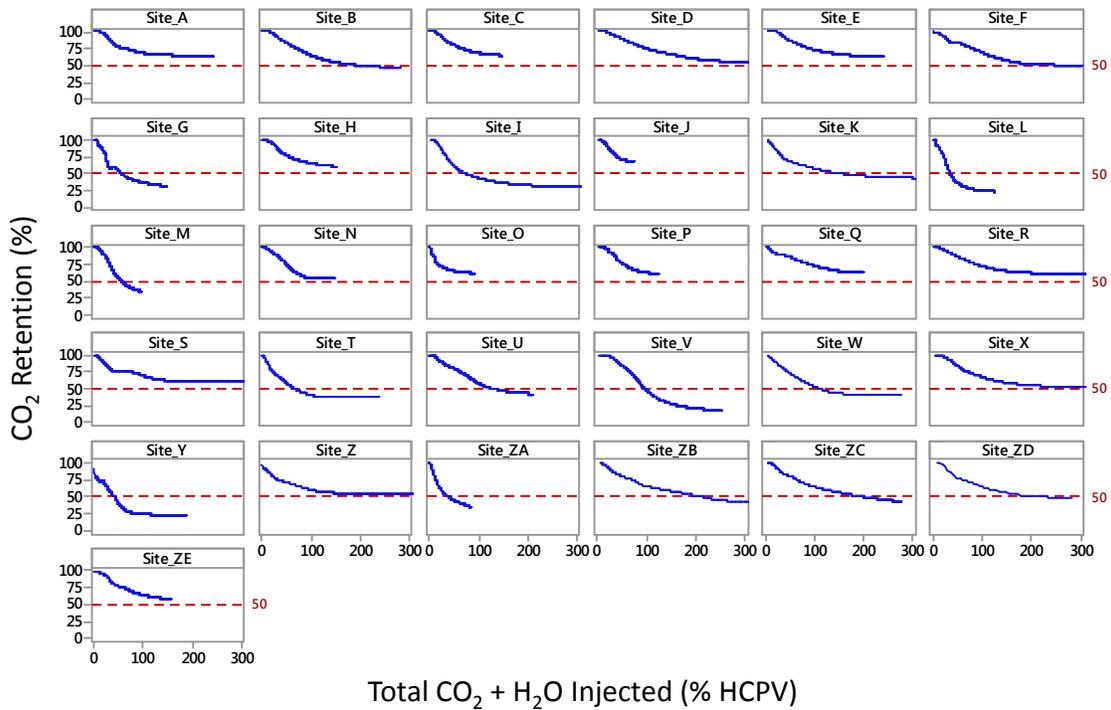


Figure 4-2. Scatterplots of x = cumulative total CO₂ + H₂O injected and y = CO₂ retention for each of the 31 CO₂ EOR sites included in the Industry Data Set. The red horizontal line across each panel represents 50% CO₂ retention.

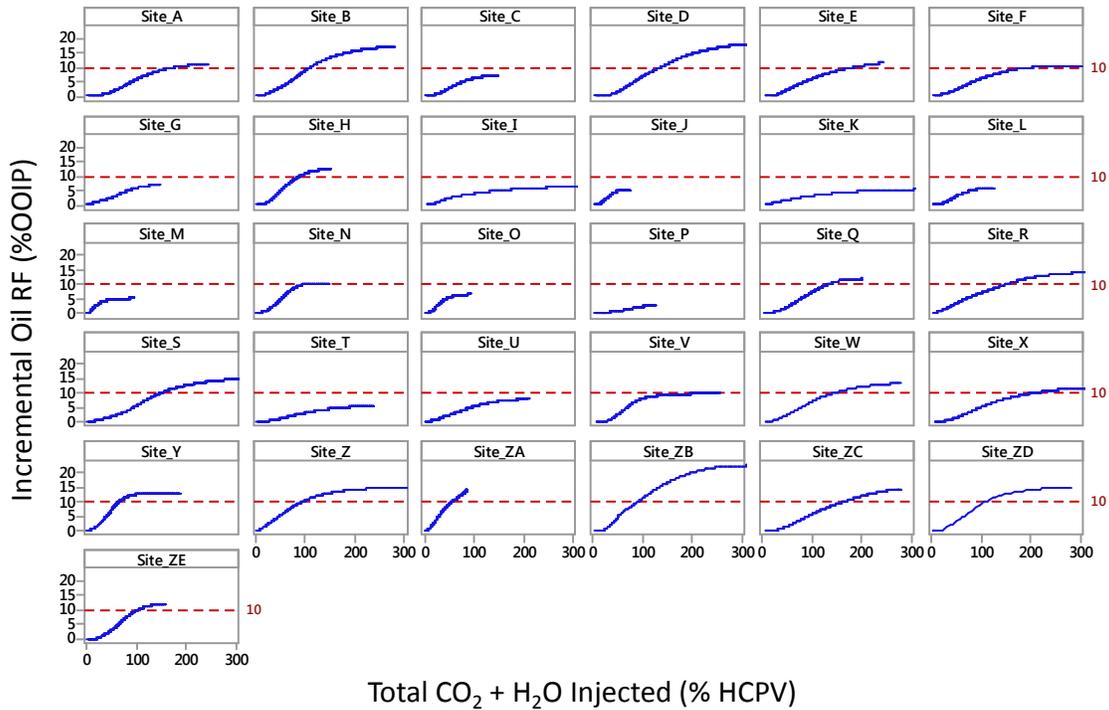


Figure 4-3. Scatterplots of x = cumulative total CO₂ + H₂O injected and y = incremental oil recovery for each of the 31 CO₂ EOR sites included in the Industry Data Set. The red horizontal line across each panel represents 10% OOIP.

As for the net CO₂ utilization versus CO₂ + H₂O injected curves, they all exhibited an asymptotic shape, with high values at low injection volumes and rapid decline to an asymptotic value of approximately 10 Mscf/STB (Fig. 4-4). Several different asymptotic model forms were explored, and ultimately these curves were fit using a two-parameter asymptotic model of the form:

$$\ln(y) = A + (B/x), \tag{4-9}$$

Where:

$\ln(y)$ = natural log of CO₂ net utilization (Mscf/STB);

- x = total cumulative volume of CO₂ + H₂O injected (%HCPV);
- A = fitted model parameter equal to the lower asymptote at infinite CO₂ + H₂O injected; and
- B = fitted model parameter related to the slope of the curve from the maximum value to the asymptote.

Other model forms were evaluated for extrapolation to 300% HCPV, such as the exponential, harmonic, power, and hyperbolic forms, which are commonly used in decline-curve analysis (e.g., Arps, 1994; Guo et al., 2007; Poston and Poe, 2008). However, these decline-curve models are used to extend trends into the future, and are insufficient for modeling the entire shape of the sigmoidal curve. This study was interested in not only extrapolating the data record to 300% HCPV, but also in fitting a function to the entire data record to allow inference about the overall shape of the curve and how that shape might relate to reservoir conditions and CO₂ flood design.

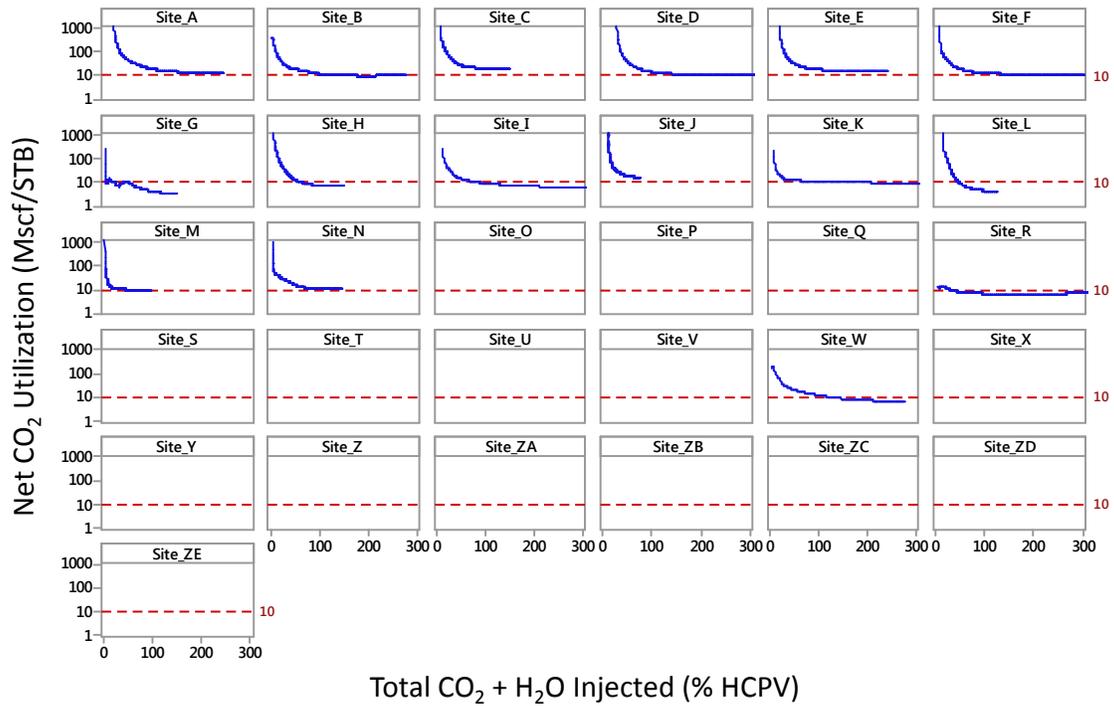


Figure 4-4. Scatterplots of x = cumulative total $\text{CO}_2 + \text{H}_2\text{O}$ injected and y = net CO_2 utilization factor for 16 of the 31 CO_2 EOR sites included in the Industry Data Set. The red horizontal line across each panel represents 10 Mscf/STB. Note: y -axis is presented on a log-scale.

Model fits were evaluated using the root mean squared error (RMSE) as a gross measure of model goodness-of-fit across the data record. In addition, the residuals ($y_{\text{observed}} - y_{\text{predicted}}$) were plotted against cumulative total $\text{CO}_2 + \text{H}_2\text{O}$ injected (x) to assess model fit as a function of %HCPV (Motulsky and Ransnas, 1987). The RMSE values and graphical assessment of residual fits were used to justify selection of a single model form used in the multi-site analysis, with priority given to models with better fits at higher %HCPV and to models that could reproduce the full shape of the curve. While there is no absolute criterion for a “good” value of RMSE, lower RMSE values across model forms represent better-fitting models. In addition, models with smaller residuals in the region of the curve where predictions are needed most (i.e., between 200% and 300% HCPV)

represent better models for the goals of this study. Approximately 87% of the sites in the Industry Data Set have observed data through 100% HCPV; 60% of the sites have observed data through 200% HCPV; and less than 30% of the sites have observed data through 300% HCPV. Thus a target area for prediction in our work is the region between 200% and 300% HCPV. The reason for selecting a single model was so that the fitted parameters could be assessed across all sites. Using this logic, the four-parameter log-logistic model was determined to provide the best fit for CO₂ retention and incremental oil recovery, while the asymptotic model was determined to provide the best fit for net CO₂ utilization.

Parameters for each model were fit using the nonlinear regression module in Minitab® Statistical Software, Version 17 (Minitab Statistical Software, 2014). Initial starting values for each parameter were selected based on a graphical assessment and the geometric interpretation of each parameter (i.e., asymptotes, inflection point, and slope factor). The Gauss-Newton algorithm was then used to minimize the sum of squared errors (SSE). Step-by-step examples of fitting these curves for Site A are provided in Appendix C. Model fits were evaluated using a graphical assessment in addition to inspection of residuals and SSE.

4.3 Results

4.3.1 Patterns in CO₂ Retention, Incremental Oil Recovery Factor, and Net CO₂ Utilization

The 31 sites had similar CO₂ retention patterns over time (Figure 4-2) and show that a larger proportion of the injected CO₂ is retained in the reservoir (i.e., CO₂ recycle

volumes are low) during the early years of a CO₂ EOR project (at low %HCPV), until reaching a near-horizontal slope (i.e., asymptote) in later years (at high %HCPV).

The incremental oil recovery factor trends were the inverse of CO₂ retention; the 31 sites showed low oil recovery at low %HCPV and an increased until reaching a stable, near-horizontal slope at higher %HCPV. These patterns show that during the early years of a CO₂ EOR project, a larger proportion of the injected CO₂ is retained in the reservoir with no commensurate increase in the incremental oil production, resulting in relatively low oil recovery factors (Figure 4-3).

The 16 sites with corresponding data also showed similar net CO₂ utilization factor patterns over time, with higher net CO₂ utilization factors at low %HCPV and an exponential-type decline until reaching a stable, near-horizontal slope at higher %HCPV (Figure 4-4). These patterns show that during the early years of a CO₂ EOR project (at low %HCPV), a larger proportion of the injected CO₂ is retained in the reservoir while little oil is produced, resulting in significantly higher net CO₂ utilization factors. At many of the sites, the CO₂ utilization factor exceeded 1000 Mscf/STB in the first several years before converging toward a more stable asymptotic value near 10 Mscf/STB.

4.3.2 Nonlinear Curve Fits of CO₂ Retention

The four-parameter log-logistic function was able to accurately describe the overall shape of the CO₂ retention curve across the 31 sites. Representative fits for six of the 31 sites are shown in Figure 4-5. These sites reflect the variety of slopes, asymptotic values, and overall shapes observed in the data set and demonstrate the robustness of the four-parameter log-logistic function to accurately describe the patterns observed in the

data. The median RMSE value across the CO₂ retention model fits from 50% to 300% HCPV for the 31 sites was 0.496, indicating a good overall model fit with average differences between observed and predicted values of about 0.5% CO₂ retention. A list of the parameter estimates and standard errors for A, C and D for each of the 31 sites is compiled in Table 4-1 (parameter B is a constant at 100%).

The lower asymptotic values (parameter A) across the 31 sites, which represent CO₂ retention at infinite CO₂ + H₂O injected, ranged from 13.0% to 62.2%, with a median value of 47.7%. As noted in the introduction, uncertainty in CO₂ retention reflects the various site-specific geologic and operational factors that contribute to incidental CO₂ storage over the lifetime of the CO₂ flood.

From a CO₂ cost perspective, the ideal site would have low CO₂ retention and more rapidly transition from high to low CO₂ retention such that the CO₂ utilization would be less intensive over the lifetime of the CO₂ flood (i.e., less CO₂ would be required to produce an incremental barrel of oil). In terms of Equation 4-8, this would translate into a low parameter A value (low CO₂ retention), low parameter C value (early inflection point in the CO₂ flood), and high parameter D value (steep slope).

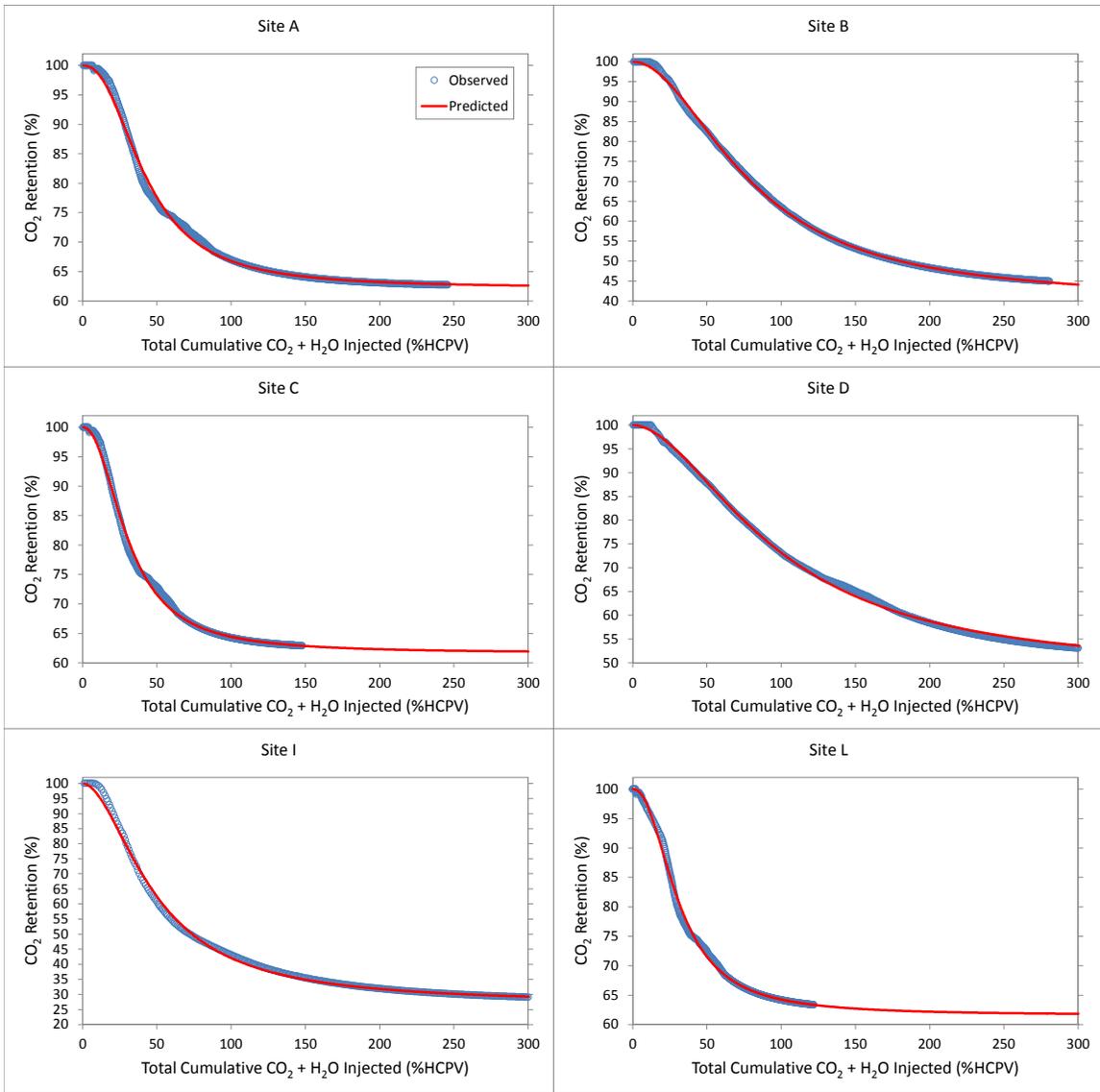


Figure 4-5. Fits of CO₂ retention four-parameter log-logistic functions to six representative sites from the Industry Data Set. Blue circles represent observed data and the red line represents the fitted response from the four-parameter log-logistic model, Equation 4-8.

Table 4-1. Fitted parameters A, C, and D and their standard errors for the four-parameter log-logistic model used to fit CO₂ retention. The column showing the maximum value for %HCPV is the length of the measured data record from the Industry Data Set. Parameter B was constant in the four-parameter log-logistic model (100 for CO₂ retention) and is therefore not shown in the table. Cells with an “---” indicate that insufficient data were available to calculate estimates.

Site Letter Code	$x = \text{CO}_2 + \text{H}_2\text{O}$ Injected (%HCPVI)	$y = \text{CO}_2$ Retention (%)					
	Maximum Value	A	std. error	C	std. error	D	std. error
Site_A	245.8	62.21	0.06	42.57	0.14	2.32	0.02
Site_B	280.6	40.00	0.08	79.31	0.16	1.96	0.01
Site_C	147.8	61.65	0.08	30.80	0.12	2.17	0.02
Site_D	449.9	47.69	0.08	97.08	0.26	1.83	0.01
Site_E	242.1	59.06	0.10	67.67	0.23	2.18	0.01
Site_F	301.8	44.92	0.16	80.87	0.26	2.27	0.02
Site_G	148.3	25.01	0.47	34.11	0.37	1.65	0.03
Site_H	151.7	57.96	0.16	46.53	0.23	2.17	0.02
Site_I	478.1	26.88	0.07	48.64	0.15	1.86	0.01
Site_J	70.4	56.72	0.51	21.60	0.47	1.45	0.03
Site_K	649.8	38.99	0.10	46.90	0.26	1.30	0.01
Site_L	121.9	61.59	0.11	31.08	0.12	2.22	0.02
Site_M	94.9	26.41	0.26	39.21	0.15	2.48	0.02
Site_N	148.2	47.94	0.24	47.58	0.27	2.65	0.03
Site_O	92.2	57.13	0.15	12.02	0.09	1.21	0.01
Site_P	122.0	56.54	0.10	43.18	0.11	2.56	0.02
Site_Q	198.8	56.09	0.16	74.63	0.18	2.19	0.02
Site_R	378.0	57.56	0.07	66.29	0.21	2.08	0.01
Site_S	310.7	60.61	0.07	37.26	0.29	1.82	0.02
Site_T	238.1	38.16	0.11	26.35	0.20	2.04	0.03
Site_U	208.6	31.72	0.31	81.16	0.48	1.89	0.02
Site_V	253.6	15.30	0.09	85.34	0.10	3.12	0.01
Site_W	277.5	38.69	0.20	45.66	0.63	1.91	0.03
Site_X	434.7	50.59	0.06	73.00	0.18	2.11	0.01
Site_Y	188.3	12.96	0.85	25.79	0.55	1.31	0.04
Site_Z	446.5	50.30	0.20	32.19	0.40	1.25	0.02
Site_ZA	83.1	17.71	0.35	24.70	0.23	1.15	0.01
Site_ZB	341.1	13.14	1.56	151.03	5.55	1.07	0.02
Site_ZC	276.7	21.99	0.56	116.21	1.61	1.21	0.01
Site_ZD	281.0	40.62	1.07	73.75	2.19	1.58	0.06
Site_ZE	158.3	53.79	0.29	55.96	0.47	1.90	0.02
minimum =	70.4	12.96	0.06	12.02	0.09	1.07	0.01
median =	242.1	47.69	0.16	46.90	0.23	1.91	0.02
mean =	252.3	42.90	0.28	56.08	0.54	1.90	0.02
std. deviation =	136.4	16.11	0.33	30.42	1.03	0.50	0.01
maximum =	649.8	62.21	1.56	151.03	5.55	3.12	0.06

4.3.3 Nonlinear Curve Fits of Incremental Oil Recovery

The four-parameter log-logistic function was also able to accurately describe the overall shape of the incremental oil recovery curve across the 31 sites. Representative fits for six of the 31 sites are shown in Figure 4-6. These sites reflect the variety of slopes, asymptotic values, and overall shapes observed in the data set, and demonstrate the robustness of the four-parameter log-logistic function to accurately describe the data. The median RMSE value across the incremental oil recovery model fits from 50% to 300% HCPV for the 31 sites was 0.075, indicating a good overall model fit with average differences between observed and predicted values of less than 0.1% OOIP. A list of the parameter estimates for A, C and D for each of the 31 sites is compiled in Table 4-2 (parameter B is a constant at 0% OOIP).

The upper asymptotic values (parameter A) across the 31 sites, which represent incremental oil recovery at infinite CO₂ + H₂O injected, ranged from 4.8% to 26.1%, with a median value of 13.5%. Analogous to CO₂ retention, the uncertainty in incremental oil recovery reflects the various site-specific geologic and operational factors.

From an oil production perspective, the ideal site would have a high incremental oil recovery and would more rapidly transition from low to high recovery such that the oil production would be more rapidly achieved in the early phases of the CO₂ flood. In terms of Equation 4-8, this would translate to a high parameter A value (high incremental oil recovery), low parameter C value (early inflection point in the CO₂ flood), and high parameter D value (steep slope).

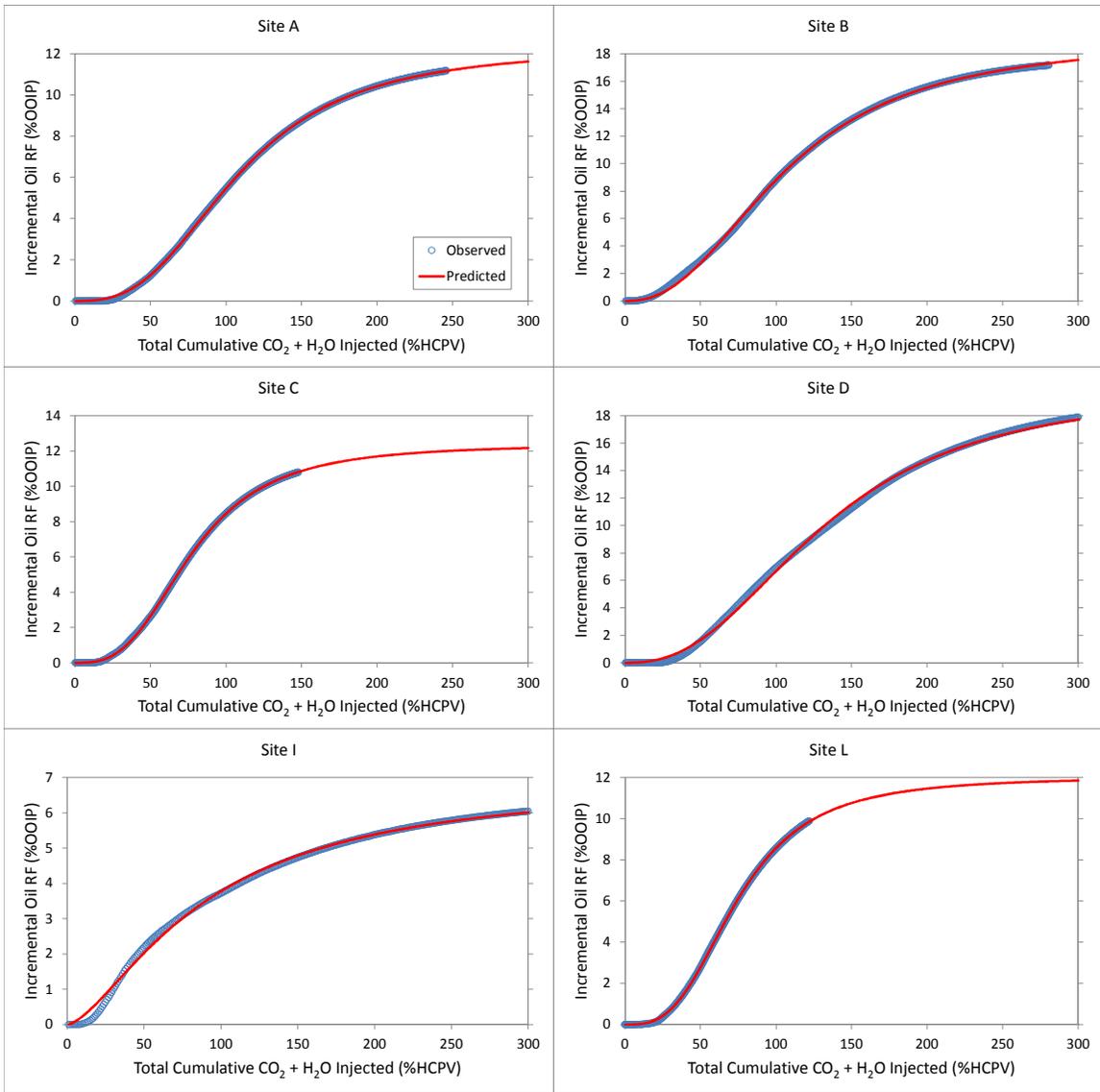


Figure 4-6. Fits of incremental oil recovery four-parameter log-logistic functions to six representative sites from the Industry Data Set. Blue circles represent observed data and the red line represents the fitted response from the four-parameter log-logistic model, Equation 4-8.

Table 4-2. Fitted parameters A, C, and D and their standard errors for the four-parameter log-logistic model used to fit incremental oil recovery. The column showing the maximum value for %HCPV is the length of the measured data record from the Industry Data Set. Parameter B was constant in the four-parameter log-logistic model (0 for incremental oil recovery) and is therefore not shown in the table. Cells with an “---” indicate that insufficient data were available to calculate estimates.

Site Letter Code	$x = \text{CO}_2 + \text{H}_2\text{O}$ Injected (%HCPVI)	$y = \text{Incremental Oil Recovery Factor}$ (%OOIP)					
	Maximum Value	A	std. error	C	std. error	D	std. error
Site_A	245.8	12.29	0.01	108.37	0.08	2.81	0.00
Site_B	280.6	19.11	0.03	106.85	0.17	2.34	0.01
Site_C	147.8	12.38	0.01	77.38	0.06	2.98	0.00
Site_D	449.9	20.08	0.02	132.69	0.21	2.46	0.01
Site_E	242.1	13.76	0.06	114.57	0.57	2.05	0.01
Site_F	301.8	11.28	0.03	92.52	0.27	2.29	0.01
Site_G	148.3	9.27	0.09	83.12	0.89	2.02	0.02
Site_H	151.7	13.48	0.02	61.98	0.08	2.82	0.01
Site_I	478.1	6.95	0.01	89.46	0.35	1.53	0.01
Site_J	70.4	24.89	1.23	78.12	2.55	2.67	0.05
Site_K	649.8	5.70	0.01	92.71	0.23	1.68	0.01
Site_L	121.9	12.01	0.02	74.13	0.08	3.05	0.00
Site_M	94.9	5.16	0.01	17.76	0.08	2.08	0.02
Site_N	148.2	11.24	0.03	52.54	0.16	3.12	0.03
Site_O	92.2	6.73	0.01	27.81	0.06	2.76	0.01
Site_P	122.0	4.76	0.08	101.78	1.60	2.25	0.02
Site_Q	198.8	13.89	0.05	92.06	0.31	2.65	0.02
Site_R	378.0	16.99	0.07	114.41	0.78	1.63	0.01
Site_S	310.7	16.71	0.05	127.88	0.38	2.54	0.01
Site_T	238.1	6.45	0.01	103.48	0.16	2.16	0.00
Site_U	208.6	9.81	0.02	89.86	0.28	1.87	0.01
Site_V	253.6	10.07	0.00	64.34	0.04	3.47	0.01
Site_W	277.5	15.30	0.03	104.35	0.25	2.10	0.01
Site_X	434.7	13.10	0.01	118.16	0.14	2.27	0.01
Site_Y	188.3	13.54	0.02	45.85	0.11	3.14	0.02
Site_Z	446.5	15.95	0.02	68.77	0.19	1.94	0.01
Site_ZA	83.1	25.34	0.54	71.02	2.21	1.42	0.02
Site_ZB	341.1	26.14	0.16	108.05	0.92	1.95	0.02
Site_ZC	276.7	18.54	0.06	146.21	0.56	2.07	0.01
Site_ZD	281.0	16.16	0.06	86.20	0.38	2.38	0.02
Site_ZE	158.3	13.79	0.05	72.96	0.23	3.09	0.02
minimum =	70.4	4.76	0.00	17.76	0.04	1.42	0.00
median =	242.1	13.48	0.03	89.86	0.23	2.29	0.01
mean =	252.3	13.58	0.09	87.92	0.46	2.37	0.01
std. deviation =	136.4	5.69	0.23	29.08	0.61	0.52	0.01
maximum =	649.8	26.14	1.23	146.21	2.55	3.47	0.05

4.3.4 Nonlinear Curve Fits of Net CO₂ Utilization

As previously mentioned, net CO₂ utilization could only be calculated for 16 of the 31 sites because the B_{oi}, B_{CO₂}, or N values needed to convert from dimensionless variables to volumetric units were not available for all 31 sites. For the 16 sites for which we could accurately express both the cumulative volume of purchased CO₂ injected and the cumulative incremental oil production, the two-parameter asymptotic function was able to accurately describe the overall shape of the net CO₂ utilization curves. Unlike CO₂ retention and incremental oil recovery, which had very low residuals across the entire data record for nearly all 31 sites, the net CO₂ utilization residuals were slightly higher. This was particularly true near the focus of the hyperbolic shape where the steep, near-vertical decline in net CO₂ utilization transitions to the asymptote. However, despite this phenomenon, the two-parameter asymptotic function was able to accurately describe the overall shape of the net CO₂ utilization curve and had small residuals at higher %HCPV. Representative fits for six of the 16 sites are shown in Figure 4-7. The median RMSE value across the net CO₂ utilization model fits from 50% to 300% HCPV for the 16 sites was 0.593, indicating a good overall model fit with average differences between observed and predicted values of about 0.6 Mscf/STB. A list of the parameter estimates for A and B for each of the 16 sites is compiled in Table 4-3. Analogous to CO₂ retention and incremental oil recovery, the uncertainty in net CO₂ utilization reflects the various site-specific geologic and operational factors.

From a profitability perspective, the ideal site would have a steep slope, reflecting a rapid transition from high to low net CO₂ utilization, and asymptote to a low net CO₂ utilization. This combination of observations would yield a more rapid decline in the

amount of CO₂ required to produce a barrel of oil over the lifetime of the CO₂ flood. In terms of Equation 4-9, this translates to a low parameter A value (low net CO₂ utilization asymptote) and a high parameter B value (steep slope).

The lower asymptote (exponentiation of parameter A) represents net CO₂ utilization at infinite (CO₂ + H₂O) injected. However, it should be noted that it is not correct to interpret an infinitely asymptotic behavior to net CO₂ utilization, as the true model form may be slightly curvilinear at higher %HCPV. Stated differently, net CO₂ utilization may increase beyond 300% HCPV and therefore may not be a true asymptote. The residuals within the modeled region used in this study (between 50% to 300% HCPVI) were generally less than 0.2 Mscf/STB, and therefore, the model form was appropriate for this intended use. However, interpretations should not be extended to outside this region of injection.

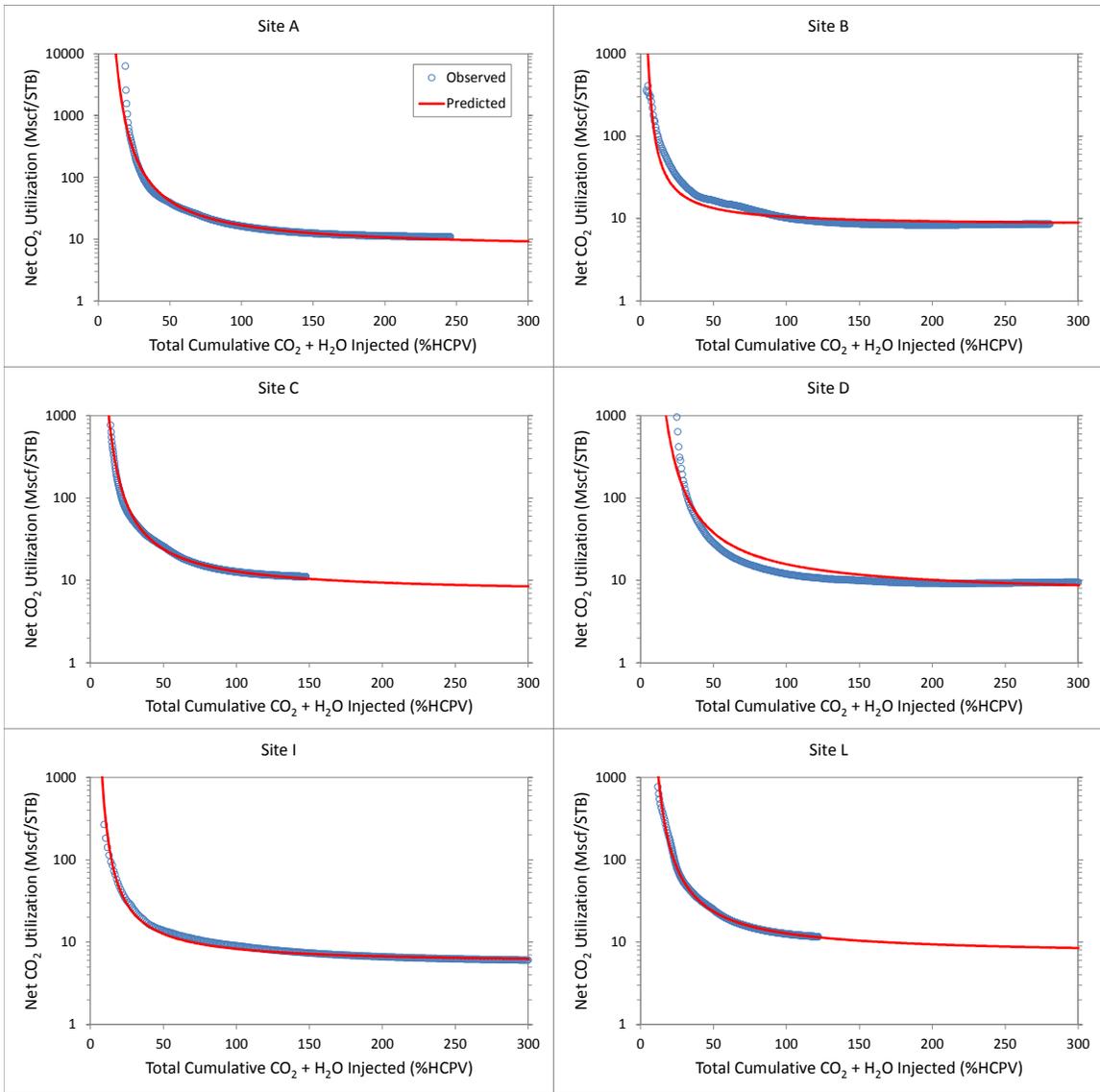


Figure 4-7. Fits of net CO₂ utilization two-parameter asymptotic functions to six representative sites from the Industry Data Set. Blue circles represent observed data and the red line represents the fitted response from the two-parameter asymptotic model, Equation 4-9.

Table 4-3. Fitted parameters A and B and their standard errors for the two-parameter asymptotic function used to model net CO₂ utilization. The column showing the maximum value for %HCPV is the length of the measured data record from the Industry Data Set. Cells with an “---” indicate that insufficient data were available to calculate estimates.

Site Letter Code	$x = \text{CO}_2 + \text{H}_2\text{O}$ Injected (%HCPV)	$y = \text{Net CO}_2$ Utilization (Mscf/STB)			
	Maximum Value	A	std. error	B	std. error
Site_A	245.8	1.92	0.01	90.48	0.61
Site_B	280.6	2.10	0.01	24.58	0.34
Site_C	147.8	1.93	0.01	62.65	0.39
Site_D	449.9	1.87	0.01	88.07	1.37
Site_E	242.1	2.37	0.02	43.11	1.38
Site_F	301.8	2.20	0.01	19.00	0.21
Site_G	148.3	0.65	0.01	77.87	0.42
Site_H	151.7	1.61	0.01	48.09	0.34
Site_I	478.1	1.70	0.00	41.93	0.23
Site_J	70.4	2.18	0.01	28.06	0.32
Site_K	649.8	2.17	0.00	16.48	0.08
Site_L	121.9	1.94	0.01	61.22	0.21
Site_M	94.9	2.13	0.00	5.14	0.03
Site_N	148.2	2.10	0.01	27.79	0.29
Site_O	92.2	---	---	---	---
Site_P	122.0	---	---	---	---
Site_Q	198.8	---	---	---	---
Site_R	378.0	1.81	0.00	9.75	0.16
Site_S	310.7	---	---	---	---
Site_T	238.1	---	---	---	---
Site_U	208.6	---	---	---	---
Site_V	253.6	---	---	---	---
Site_W	277.5	1.70	0.01	62.65	1.33
Site_X	434.7	---	---	---	---
Site_Y	188.3	---	---	---	---
Site_Z	446.5	---	---	---	---
Site_ZA	83.1	---	---	---	---
Site_ZB	341.1	---	---	---	---
Site_ZC	276.7	---	---	---	---
Site_ZD	281.0	---	---	---	---
Site_ZE	158.3	---	---	---	---
minimum =	70.4	0.65	0.00	5.14	0.03
median =	242.1	1.93	0.01	42.52	0.33
mean =	252.3	1.90	0.01	44.18	0.48
std. deviation =	136.4	0.40	0.01	27.35	0.46
maximum =	649.8	2.37	0.02	90.48	1.38

4.3.5 Uncertainty Quantification – Percentile Estimates from 0% to 300% HCPVI

The modeled responses from 0% to 300% HCPVI from each of the sites (31 sites for CO₂ retention and incremental oil recovery and 16 sites for net CO₂ utilization) were combined into a single data set to generate empirical percentile estimates across the sites. The model fits for CO₂ retention, incremental oil recovery, and net CO₂ utilization are shown in Figure 4-8. The 10th, 25th (first quartile), 50th (median), 75th (third quartile), and 90th percentile values (P₁₀, P₂₅, P₅₀, P₇₅, and P₉₀, respectively) were used to describe the uncertainty in the response as a function of %HCPV. The P₅₀ estimate is the central value of the distribution. The P₂₅ to P₇₅ is the interquartile range (IQR) and the central 50 percent of the data would be found within this interval (dark gray shaded regions in Figure 4-8). Lastly, the P₁₀ to P₉₀ is the range within which 80 percent of the data would be found (light gray shaded regions in Figure 4-8).

The minimum, maximum, average, standard deviation, and percentile estimates are summarized in Table 4-4. The P₁₀, P₅₀, and P₉₀ estimates at 300% HCPV for the three responses modeled in this work were: CO₂ retention – 23.1%, 48.3%, and 61.8%; incremental oil recovery – 5.3%, 12.2%, and 21.5% OOIP; and net CO₂ utilization – 4.8, 8.7, and 10.5 Mscf/STB.

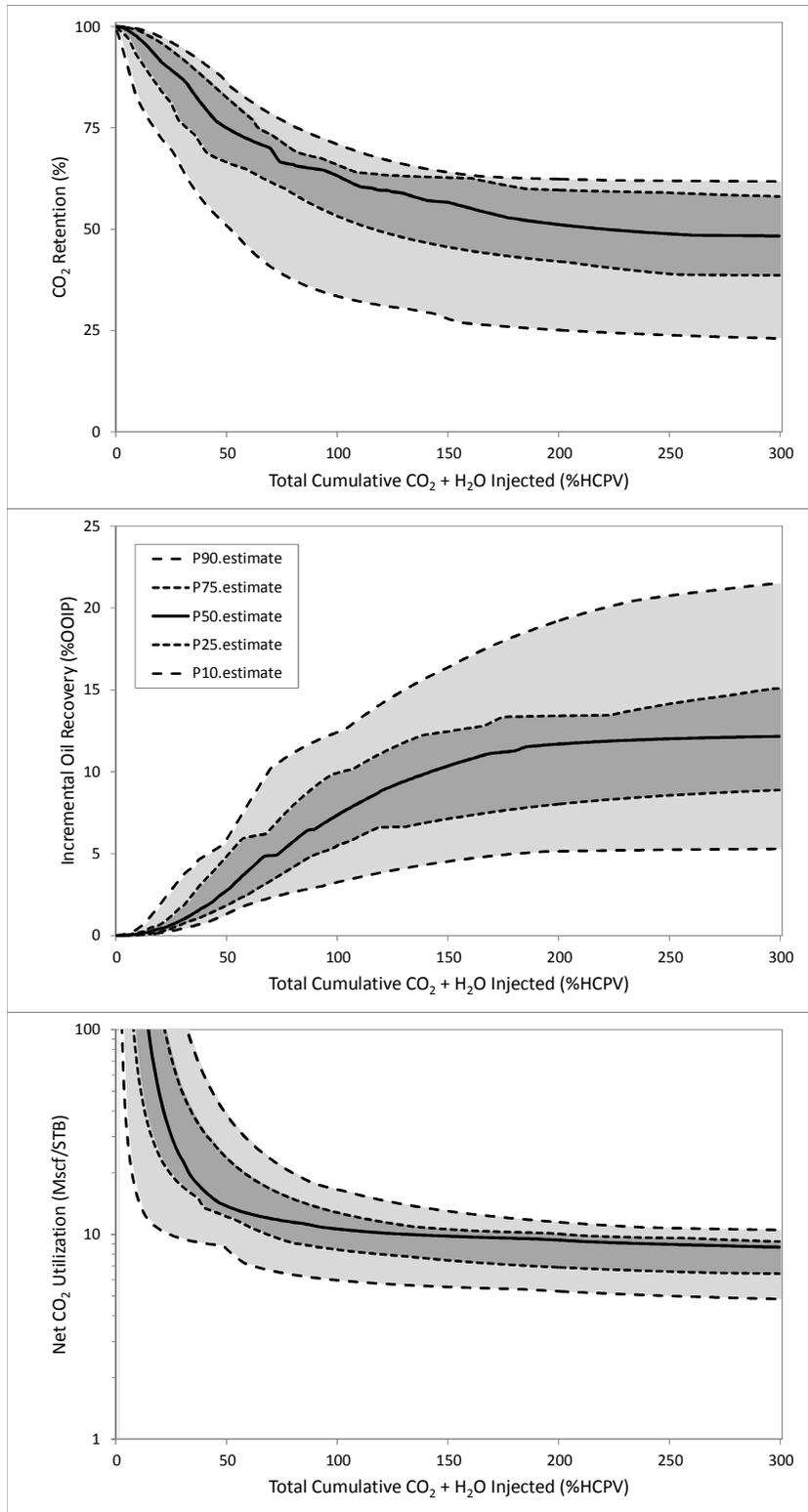


Figure 4-8. Percentile estimates generated from multi-site nonlinear regression fits for CO₂ retention (top), incremental oil recovery (middle), and net CO₂ utilization (bottom). The light gray shaded regions bounded by dashed lines represent the 10th to 90th percentiles; the dark gray shaded regions bounded by dotted lines represent the 25th to 75th percentiles (interquartile range); and the P₅₀ estimate (solid black line) represents the central value of the distribution.

Table 4-4. Multi-site empirical percentile estimates for CO₂ retention, incremental oil recovery factor, and net CO₂ utilization as a function of total CO₂ + H₂O injected (50%, 100%, 150%, 200%, 250%, and 300% HCPV). The modeled responses from each of the sites (31 sites for CO₂ retention and incremental oil recovery and 16 sites for net CO₂ utilization) were combined into a single data set to generate empirical percentile estimates across the sites.

$x = \text{CO}_2 + \text{H}_2\text{O}$ Injected (%HCPV)	$y = \text{CO}_2$ Retention %								
	Min	P ₁₀	P ₂₅	Average	Std Dev	P ₅₀	P ₇₅	P ₉₀	Max
50	38.8	51.1	66.6	72.3	13.3	75.1	82.7	86.5	88.0
100	25.6	33.6	53.3	57.7	12.8	63.3	66.0	71.0	73.1
150	20.9	28.0	45.6	51.8	13.2	56.6	62.7	64.1	65.2
200	18.6	25.1	42.1	48.9	13.5	51.2	59.7	62.4	63.2
250	17.2	23.9	39.0	47.4	13.8	48.9	59.0	61.9	62.8
300	16.4	23.1	38.6	46.4	14.0	48.3	58.1	61.8	62.6
$x = \text{CO}_2 + \text{H}_2\text{O}$ Injected (%HCPV)	$y = \text{Incremental Oil Recovery Factor}$ (%OOIP)								
	Min	P ₁₀	P ₂₅	Average	Std Dev	P ₅₀	P ₇₅	P ₉₀	Max
50	0.8	1.3	1.8	3.3	2.0	2.7	4.7	5.8	9.6
100	3.2	5.5	7.7	3.4	7.3	9.9	12.4	16.4	
150	3.4	4.5	7.1	10.2	4.1	10.3	12.4	16.3	21.2
200	3.9	5.1	8.0	11.4	4.6	11.7	13.4	19.2	23.0
250	4.2	5.2	8.6	12.1	4.9	12.0	14.1	20.7	23.8
300	4.4	5.3	8.9	12.5	5.1	12.2	15.1	21.5	24.2
$x = \text{CO}_2 + \text{H}_2\text{O}$ Injected (%HCPV)	$y = \text{Net CO}_2$ Utilization (Mscf/STB)								
	Min	P ₁₀	P ₂₅	Average	Std Dev	P ₅₀	P ₇₅	P ₉₀	Max
50	7.4	8.6	12.3	18.2	10.0	13.8	23.9	39.1	41.7
100	4.2	6.0	8.4	11.0	3.5	10.6	12.8	16.6	16.9
150	3.2	5.5	7.5	9.4	2.6	9.8	10.6	13.0	14.3
200	2.8	5.3	6.9	8.7	2.4	9.4	10.1	11.5	13.3
250	2.6	5.0	6.6	8.4	2.3	8.9	9.6	10.8	12.7
300	2.5	4.8	6.4	8.1	2.2	8.7	9.2	10.5	12.4

4.4 Discussion

4.4.1 Comparison to Other Published Studies

The results of this work at 300% HCPV compare well with other published studies that projected end-of-project estimates. For example, the patterns shown by Hadlow (1992) for simulation studies of incremental oil recovery as a function of time

and different CO₂ injection methods (continuous CO₂; 1:1 conventional WAG; 1:1 hybrid WAG) demonstrate the sigmoidal shape used in the modeling work in this study. In addition, Hadlow (1992) also described the estimated end-of-project net CO₂ utilization for five major CO₂ EOR projects (Denver, Means, Rangely, Seminole, and E. Vacuum). The ranges in net CO₂ utilization were 6 to 11 Mscf/STB. These ranges agree well with the percentile estimates developed in the current work, which ranged from a P₁₀ value of 4.8 Mscf/STB to a P₉₀ value of 10.5 Mscf/STB across the 16 sites for which net CO₂ utilization values could be calculated.

Merchant (2010) evaluated oil recovery for WAG floods beyond 80% HCPV using numerical models. Data from that work, which included several of the sites that were also part of the Industry Data Set of this work, suggest tertiary oil recovery beyond 80% HCPV in the range of 5% to 25% OOIP, depending on the specific reservoir formation, field, or lease. These ranges agree well with the percentile estimates developed in the current work, which ranged from a P₁₀ value of 5.3% to a P₉₀ value of 21.5% across the 31 sites. Similar results were also observed for net CO₂ utilization. The ranges for net CO₂ utilization reported by Merchant (2010) were approximately 1 to 10 Mscf/STB, again dependent upon the specific reservoir formation, field, or lease, as compared to the P₁₀ and P₉₀ estimates developed in the current work, as noted above. Thus the model predictions of Merchant (2010) are verified by the observations of this study, which are based on empirical data from actual field results.

Lastly, Zhou et al. (2012) discuss incremental oil recovery in the range from 8% to 25% OOIP for Permian Basin CO₂ WAG injection, which agrees well with the percentile estimates developed in the current work. In addition, Zhou et al. (2012)

estimated the net CO₂ utilization at continuous injection and WAGs with different WAG ratios. For continuous CO₂ injection up to WAG ratios of 4:1, predicted net CO₂ utilization ranged from 13 to 5 Mscf/STB, respectively. These ranges also agree well with the P₁₀ and P₉₀ values discussed in this current work.

4.4.2 Predictions of CO₂ Storage

The nonlinear curve fits and percentile estimates are focused on statistical *inference* – i.e., what is the likely outcome, and uncertainty ranges, for CO₂ retention, incremental oil recovery, and net CO₂ utilization given the historical data from 31 CO₂ EOR sites? However, the approach described in this work also provides useful information for *prediction* – i.e., given a set of inputs from another site with similar geology and operational factors, what are plausible ranges in outcomes? Generic predictions of CO₂ storage are possible using the results from this work. The volume of CO₂ stored is a function of OOIP, incremental oil recovery factor, and net CO₂ utilization factor, is shown below in Equation 4-10. Generic predictions.

$$\text{CO}_2 \text{ stored [Mscf]} = (\text{OOIP} \times \text{RF}) \times \text{UF}_{\text{CO}_2, \text{net}}, \quad (4-10)$$

Where:

OOIP = original oil in place (STB);

RF = incremental oil recovery factor (%); and

UF_{CO₂,net} = net CO₂ utilization factor (Mscf/STB).

For example, assuming an OOIP of 500 million bbl, a point estimate of the volume of CO₂ stored using the median value for incremental oil recovery factor and UF_{CO₂,netnet} at 300% HCPV from Table 4-2 is:

$$\begin{aligned}\text{CO}_2 \text{ stored [Mscf]} &= (500,000,000 \text{ bbl} \times 0.122) \times 8.7 \text{ Mscf/STB} \\ &= 530,700,000 \text{ Mscf} \\ &= 27.6 \text{ Mt}\end{aligned}$$

The conversion from Mscf to tonnes assumes 1 tonne of CO₂ per 19.25 Mscf at standard conditions of 101.4 kPa (14.7 psi) and 21.1 degrees Celsius (70 degrees Fahrenheit) (DOE, 2010b).

For generic screening, however, it is recognized that the incremental oil recovery factor and UF_{CO₂,netnet} at each %HCPV are both uncertain quantities, with summary statistics from the Industry Data Set provided in Table 4-2. Therefore, estimates of these factors which better quantify the inherent uncertainty can be achieved through Monte Carlo simulation that takes into account this uncertainty. For example, using the same OOIP as above but accounting for the uncertainty in the estimates for incremental oil recovery factor and UF_{CO₂,netnet} at 300% HCPV, Monte Carlo simulation produces an interval estimate for the amount of CO₂ stored of 13.1, 25.3, and 42.4 Mt for the P₁₀, P₅₀, and P₉₀, respectively. A histogram of CO₂ stored using this Monte Carlo simulation approach is shown in Figure 4-9. While this generic screening approach is rapid and straightforward to implement, the uncertainty in the inputs produces uncertainty in the estimates of CO₂ stored. This generic screening approach would be suitable for broad

regional assessments, and could perhaps provide useful information to make site-specific decisions about incremental oil recovery and associated CO₂ storage.

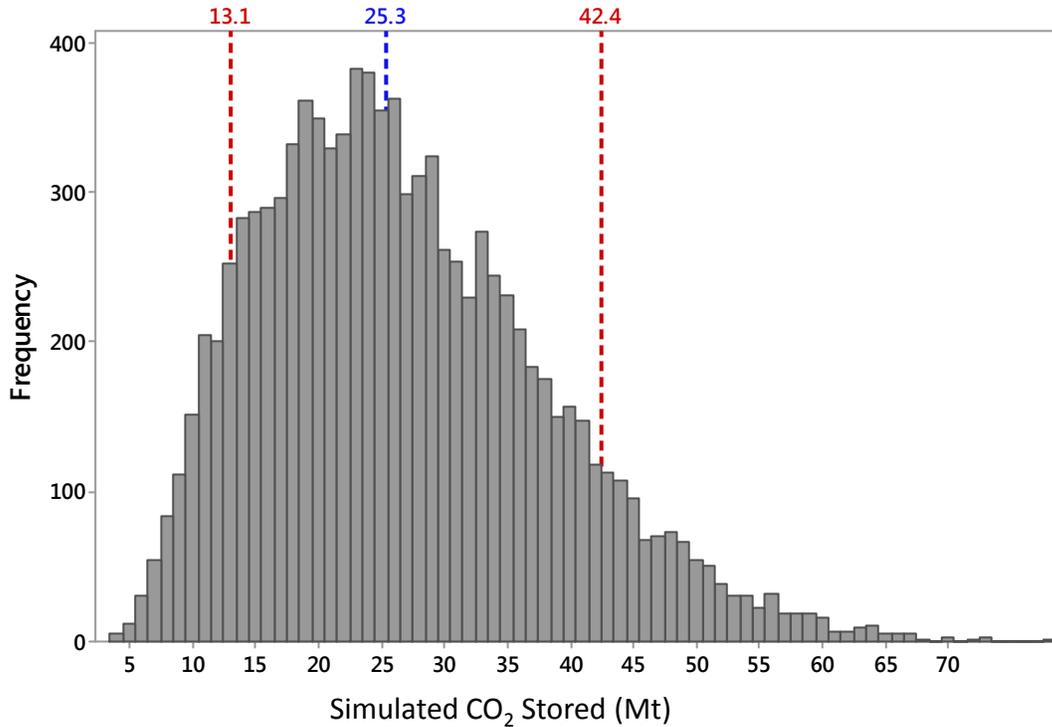


Figure 4-9. Monte Carlo simulation results of the volume of CO₂ stored as a function of OOIP, incremental oil recovery factor, and net CO₂ utilization factor using Equation 4-10 and an OOIP of 500,000,000 bbl. The y-axis shows the frequency out of 10,000 simulations. The red vertical lines at 13.1 and 42.4 Mt show the P₁₀ and P₉₀ estimates, respectively. The blue vertical line at 25.3 shows the P₅₀ (median).

4.4.3 CO₂ Retention versus CO₂ Storage

It is important to emphasize that Equation 4-4 quantifies CO₂ retention, not CO₂ storage, as the latter requires knowledge of purchased CO₂ volumes. For example, the formula for calculating CO₂ storage would be as follows (Equation 4-11) (Melzer, 2012):

$$\text{CO}_2 \text{ storage} = \frac{(\text{total CO}_2 \text{ injected} - \text{CO}_2 \text{ produced} - \text{CO}_2 \text{ losses})}{\text{purchased CO}_2 \text{ injected}} \quad (4-11)$$

Where:

- CO₂ storage = percentage of purchased CO₂ stored (%);
- total CO₂ injected = total injected volumes of CO₂ [purchased plus recycled CO₂];
- CO₂ produced = total produced volumes of CO₂ [recycled CO₂];
- CO₂ losses = CO₂ losses from the injector/producer system; and
- purchased CO₂ injected = purchased CO₂ volumes.

Purchased CO₂ volumes are protected by confidentiality agreements between the buyers and sellers of CO₂, and were not available as part of the Industry Data Set. As a result of the protected nature of these data, there is widespread use of the non-confidential quantity of total injected CO₂ volumes (which includes the recycled CO₂ volumes), and hence the use of Equation 4-4 rather than Equation 4-11 (Melzer, 2012).

CO₂ losses from the injector/producer system can occur during an EOR project. For example, power outages or equipment repair can cause periods where CO₂ (and other gaseous substances) must be released to the atmosphere during the flaring of the gas. This is an example of surface CO₂ losses. In addition, lateral migration within the target injection zone and the inability to assure perfect lateral containment within the flooded area can also lead to CO₂ losses in the reservoir, i.e., CO₂ migrates outside of the producer wells and is therefore not captured in the recycling loop but remains in the subsurface (Melzer, 2012). These down-hole losses do not represent a loss of CO₂ storage; the CO₂ is still retained in the subsurface, but it is not captured in the recycle

loop. These surface and down-hole losses would be subtracted from the numerator of Equation 4-11, as they represent a component of the CO₂ stream that leaves the injection/production system and is not recycled. CO₂ losses (either surface or down-hole [i.e., in the reservoir]) were not quantified in the Industry Data Set; however, industry experience suggests that total CO₂ losses from these are *de minimus* (<5%), with half allocated to surface losses and half allocated to down-hole losses (which do not contribute to a loss of storage) (Fox, 2009; DOE, 2010a; Melzer, 2012). This leads to CO₂ storage values greater than 95%.

As a result of the inability to quantify purchased CO₂ volumes and CO₂ losses from the Industry Data Set, this paper defines CO₂ retention as presented in Equation 4-4, recognizing that the computed value represents the fraction of total CO₂ injected that is retained in the subsurface (i.e., not recycled), and that this quantity is not directly comparable to the percentage of stored CO₂ volumes as a fraction of the purchased CO₂ volumes. For example, the median CO₂ retention value of 48.5% determined as part of this work (i.e., the P₅₀ value) does not imply that the remaining 51.5% of the CO₂ was emitted to the atmosphere – the recycled CO₂ is captured for use at another CO₂ EOR site. The scarcity and cost of the CO₂ drives the operator to recapture and conserve as much of the CO₂ as possible. The stored volumes of CO₂ in the reservoir are the purchased volumes of CO₂ minus negligible losses (Hill et al., 2013). Using this basis, life-cycle assessments have shown that more than 97% of purchased CO₂ volumes are retained in the subsurface (Fox, 2009). Thus, losses of purchased CO₂ to the atmosphere are on the order of a few percent.

4.5 Conclusions

CO₂ flooding is a proven method for extending oilfield life through EOR. Until recently, most of the CO₂ used for these operations was from natural geologic sources; however, with concerns over human-induced climate change, a lack of CO₂ supply for EOR, and hundreds of fields that are amenable to EOR, several anthropogenic CO₂ sources are now supplying EOR operations of the industry and resulting in the storage of CO₂ that would otherwise be emitted to the atmosphere. This work provides a statistical analysis of an historical industry data set, which allows modeled responses to be compared against real-world, observed injection and production data. The results of this study provide quantitative ranges for the potential EOR market for CO₂ as well as the potential incremental oil recovery from, and storage of CO₂ in, these oil fields. This work showed that a four-parameter log-logistic function was able to accurately describe the overall shape of the CO₂ retention and incremental oil recovery curves across the 31 sites, and that a two-parameter simple asymptotic function was able to accurately describe the overall shape of the net CO₂ utilization curves across the 16 sites for which net CO₂ utilization values could be calculated. These functions provide a practical approach for operators to summarize key performance metrics of their CO₂ EOR field data and forecast the performance into the future. In addition, these functions also provide useful screening tools for broad regional assessments of the CO₂ storage resource associated with candidate CO₂ EOR fields. The results of this work allow estimation of CO₂ storage capacity in CO₂ EOR operations with various degrees of confidence. Additional investigation into the geologic and operational factors that control the parameters of the four-parameter log-logistic or two-parameter asymptotic functions may help to refine

these screening tools and their estimates of associated CO₂ storage. The results of this work provide useful statistical information to estimate CO₂ storage in CO₂ EOR operations.

Acknowledgements

The authors are grateful for the financial support provided by the U.S. Department of Energy National Energy Technology Laboratory under Award No. DE-FE0009114. The authors would also like to thank project partners Schlumberger Carbon Services and Computer Modelling Group for their support.

4.6 References

Advanced Resources International, Inc. [ARI] and Melzer Consulting [Melzer]. (2010) *Optimization of CO₂ Storage in CO₂ Enhanced Oil Recovery Projects*. Department of Energy & Climate Change (DECC), Office of Carbon Capture & Storage, November 30, 2010.

Arps, J.J. (1994) Analysis of Decline Curves. *Trans. AIME* 160: 228–247.

Bachu, S. (2003) Screening and ranking of sedimentary basins for sequestration of CO₂ in geological media in response to climate change. *Environmental Geology*, 44, 277-289.

Brock, W.R., Bryan, L.A. (1989) *Summary Results of CO₂ EOR Field Tests, 1972–1987*. Paper SPE 18977 presented at the SPE Joint Rocky Mountain Regional/Low Permeability Reservoirs Symposium and Exhibition, Denver, Colorado, 6–8 March.

Carlin, B.P., Louis, T.A. (2009) *Bayesian Methods for Data Analysis*, Third Edition. CRC Press, Boca Raton.

DOE (United States Department of Energy). (2010a) *An Assessment of Gate-to-Gate Environmental Life Cycle Performance of Water-Alternating-Gas CO₂-Enhanced Oil Recovery in the Permian Basin*, National Energy Technology Laboratory, DOE/NETL-2010/1433, September 30, 2010.

DOE. (2010b) *Carbon Dioxide Enhanced Oil Recovery: Untapped Domestic Energy Supply and Long Term Carbon Storage Solution*, Strategic Center for Natural Gas and Oil (SCNGO), National Energy Technology Laboratory, March, 2010.

DOE. (2013) *Carbon Storage Research and Development*, www.netl.doe.gov/research/coal/carbon-storage/research-and-development/core-r-d (accessed 8 April 2014).

Ettehadtavakkol, A., Lake, L.W., and Bryant, S.L. (2014) CO₂-EOR and storage design optimization. *International Journal of Greenhouse Gas Control*, 25:79-92.

Fox, C. (2009) *CO₂-EOR Carbon Balance 7th Annual EOR Carbon Management Workshop*, December 2009.

Gelman, A., Hill, J. (2006) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York.

Green, D.W., Willhite, G.P. (1998) *Enhanced Oil Recovery*, Vol. 6. Dallas, Texas: SPE Textbook Series, SPE.

Guo, B., Lyons, W.C., Ghalambor, A. (2007). *Petroleum Production Engineering: A Computer-Assisted Approach*. Elsevier, Inc., 288 p.

Hadlow, R.E. (1992) *Update of Industry Experience with CO₂ Injection*. Paper SPE 24928 presented at the 67th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers held in Washington, D.C., 4–7 October.

Hill, B., Hovorka, S., Melzer, S. (2013) Geologic carbon storage through enhanced oil recovery. *Energy Procedia*, 37:6808-6830.

IEA Greenhouse Gas R&D Programme (IEA GHG). (2009) *CCS Site Characterisation Criteria*, 2009/10 July.

Intergovernmental Panel on Climate (IPCC). (2005) *IPCC Special Report on Carbon Dioxide Capture and Storage*. Prepared by Working Group III of the Intergovernmental Panel on Climate Change, ed. B. Metz, O. Davidson, H.C. de Coninck, M. Loos, and L.A. Meyer. Cambridge, UK, and New York City: Cambridge University Press, 442 pp.

Jarrell, P.M., Fox, C., Stein, M., Webb, S. (2002) *Practical Aspects of CO₂ Flooding*, SPE Monograph Series Vol. 22, ISBN:978-1-55563-096-6, Society of Petroleum Engineers.

Kuuskräa, V., and Wallace, M. (2014) CO₂-EOR set for growth as new CO₂ supplies emerge. *Oil & Gas Journal*, May, pp.92-105.

Manrique, E.J., Muci, V.E., Gurfinkel, M.E. (2007) *EOR Field Experiences in Carbonate Reservoirs in the United States*. SPE Reservoir Evaluation & Engineering, December.

Melzer, L.S. (2012) *Carbon Dioxide Enhanced Oil Recovery (CO₂ EOR): Factors Involved in Adding Carbon Capture, Utilization and Storage (CCUS) to Enhanced Oil Recovery*. Midland Texas: Melzer Consulting. National Enhanced Oil Recovery Initiative Resource: http://neori.org/Melzer_CO2EOR_CCUS_Feb2012.pdf (accessed 9 April 2014).

Merchant, D.H. (2010) *Life Beyond 80: A Look at Conventional WAG Recovery Beyond 80% HCPV Injection in CO₂ Tertiary Floods*. Paper SPE 139516 presented at SPE International Conference on CO₂ Capture, Storage, and Utilization, New Orleans, Louisiana, 10–12 November.

Minitab Statistical Software. (2014) [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com).

Motulsky, H.A., Ransnas, L.A. (1987) Fitting Curves to Data Using Nonlinear Regression: A Practical and Nonmathematical Review. *FASEB J.* 1: 365–374.

Poston, S.W., Poe, B.D. Jr. (2008). *Analysis of Production Decline Curves*. SPE.

van 't Veld, K., Mason, C.F., and Leach, A. (2013). The economics of CO₂ sequestration through enhanced oil recovery. *Energy Procedia*, 37:6909-6919.

van 't Veld, K., Wang, X., and Alvarado, V. (2014). *Economic Co-optimization of Oil Recovery and CO₂ Sequestration*. Paper SPE 170880-MS presented at the SPE Annual Technical Conference and Exhibition, Amsterdam, The Netherlands, 27-29 October.

Wiggins, M.L., Hughes, R.G. (2005). *Evaluation and Enhancement of Carbon Dioxide Flooding Through Sweep Improvement*. Annual Technical Report: Project Period October 1, 2004 through September 30, 2005, DE-FC26-04NT15536, December.

Zhou, D., Yan, M., Calvin, W.M. (2012) *Optimization of a Mature CO₂ Flood*. Paper SPE 154181 presented at the 18th SPE Improved Oil Recovery Symposium, Tulsa, Oklahoma, 14–18 April.

Chapter 5: Conclusions and Future Research Directions

5.1 Conclusions

The following conclusions were developed in three separate papers which were presented in Chapters 2, 3 and 4, and are summarized below in Sections 5.1.1, 5.1.2, and 5.1.3, respectively. These conclusions address each of the three specific research objectives that were discussed in Chapter 1.

5.1.1 Chapter 2: Effectiveness of subsurface pressure monitoring for brine leakage detection in an uncertain CO₂ sequestration system

The first objective (Objective 1) of this dissertation was to estimate the statistical power of a pressure monitoring system to detect leakage of brine through the cap rock as a function of (1) spatial orientation of the pressure monitoring (i.e., the lateral distance from the injection wells and vertical orientation within the reservoir); (2) time (i.e., years since CO₂ injection); and (3) the specified leakage rate of brine through the cap rock.

The work in Chapter 2 highlights the importance of uncertainty in the design of a reliable pressure monitoring network for geologic CO₂ storage sites. The results suggest that measurements of the absolute change in pressure within the target injection aquifer would not be able to distinguish small leakage rates from baseline conditions, and that only large leakage rates would be discriminated with sufficient statistical power. Combining measurements, for example by taking the ratio of formation pressure in the aquifer above the cap rock to the target injection aquifer, provides better statistical power for distinguishing smaller leakage rates at earlier times in the injection program. Thus an optimized monitoring system which integrates multiple measurements could provide improved leakage detection sensitivity.

In addition, the work in Chapter 2 shows that detection sensitivity for pressure is a function of lateral distance from the injection wells, vertical position within the system, and time, or $\sim f(x, z, t)$. Therefore, design of an adequate monitoring network for subsurface pressure should account for this time-space variability to ensure that the monitoring system performs to the necessary design criteria. There is unlikely to be a “one-size fits all” monitoring program for geologic CO₂ storage sites. However, a critical insight from the work in Chapter 2 is that the final monitoring program must be adaptive and dynamic, changing over the project life-cycle to adjust to changing subsurface conditions in response to CO₂ injection.

5.1.2 Chapter 3: Quantifying the benefit of wellbore leakage potential estimates for prioritizing long-term MVA well sampling at a CO₂ storage site

The second objective (Objective 2) of this dissertation was to estimate the statistical power of selecting one or more leaky wells from among a population of legacy wells as a function of (1) the number of leaking wells within the Area of Review; (2) the sampling design (random or judgmental); (3) the number of wells included in the monitoring sampling plan; and (4) the relationship between a well’s leakage potential score and its relative probability of leakage.

The work in Chapter 3 demonstrates that accurate prior knowledge about the probability of well leakage adds measurable value to the ability to detect a leaking well during the monitoring program, and that the loss in detection ability due to imperfect knowledge of the leakage probability can be quantified. Statistical power curves quantifying the probability that a leaky well was selected were developed as a function of (1) the number of leaking wells within the Area of Review; (2) the sampling design

(random or judgmental, choosing first the wells with the highest deep leakage potential scores); (3) the number of wells included in the monitoring sampling plan; and (4) the relationship between a well's leakage potential score and its relative probability of leakage. Cases where the deep well leakage potential scores are fully or partially informative of the relative leakage probability are compared to a non-informative base case in which leakage is equiprobable across all wells in the Area of Review. These statistical power curves underscore the value of well integrity surveys and estimates of deep well leakage potential to the well sampling design.

5.1.3 Chapter 4: CO₂ storage associated with CO₂ enhanced oil recovery: A statistical analysis of historical operations

The third objective (Objective 3) of this dissertation was to quantify the uncertainty associated with the CO₂ retention, incremental oil recovery, and net CO₂ utilization for CO₂ enhanced oil recovery (EOR) sites using historical reservoir performance data from 31 CO₂ EOR projects located within the continental United States and heavily dominated by West Texas carbonate floods.

The results of this study showed that across all 31 sites, the 10th, 50th (median), and 90th percentile values for the three factors at 300% hydrocarbon pore volume (HCPV) were: CO₂ retention: 23.1%, 48.3%, and 61.8% retained; incremental oil recovery: 5.3%, 12.2%, and 21.5% OOIP (original oil in place); and net CO₂ utilization: 4.8, 8.7, and 10.5 Mscf/STB (stock-tank barrel). These results can be used to estimate the potential range of expected performance for similar candidate oil fields that are not currently undergoing CO₂ injection, including estimates of the associated CO₂ storage potential of these candidate fields.

5.2 New Contributions

The focus of this dissertation is to contribute statistical methods and/or approaches for quantifying uncertainty with respect to both monitoring and performance at geologic CO₂ storage sites. These methods are intended for future use by NRAP or other geologic CO₂ storage practitioners and may be incorporated into broader modeling approaches. The results help to advance the risk assessment tools that are needed for safe, permanent geologic CO₂ storage. However, the results and contributions from this work extend beyond geologic CO₂ storage and apply to other subsurface engineered-natural systems. Several new contributions to the state-of-the-science were made in this dissertation, as summarized below.

5.2.1 Integration of Analytical Model and First-Order Uncertainty Analysis to Answer Fundamental Questions

The work in Chapter 2 provides an example of how to integrate an analytical model with first-order uncertainty analysis to answer fundamental questions about pressure monitoring. The NRAP technical approach for developing the risk assessment tools for geologic CO₂ storage entails an iterative modeling approach that integrates component models into a system model which may then be used to provide quantitative assessments of potential risks and to design monitoring protocols. A key part of this “system modeling” approach is the integration of reduced-order models (ROMs) as simplified and efficient computational tools for reproducing predictions from the detailed 3D numerical models over the range of conditions of interest (DOE, 2015). Subsurface pressure is a potentially useful monitoring tool because the pressure pulse from CO₂

injection will cover a significantly broader portion of the system than the CO₂ plume. The analytical ROM that was used in Chapter 2 (Lawrence Berkeley National Laboratory [LBNL] Model) is one available tool for rapidly estimating pressure buildup within a geologic storage system in response to CO₂ injection. The approaches that were used in Chapter 2 illustrate how to adapt the LBNL Model, which is fundamentally a deterministic calculation tool, to account for uncertainty using first-order uncertainty analysis. Therefore, Chapter 2 provides one approach for integrating the LBNL Model (or similar analytical solutions) into broader system modeling efforts.

5.2.2 Simulation Approach for Integrating Wellbore Leakage Risk and Monitoring

The simulation approach in Chapter 3 allows an operator to combine well leakage risk estimates with different sampling designs to help select a sampling plan from among several alternatives. Legacy wells at geologic CO₂ storage sites (or other types of subsurface storage sites) are potential preferential pathways for leakage. For monitoring these sites, a site-specific sampling plan would need to specify not only the measurements to be made at each well, but the specific wells at which the measurements would be conducted. A key research question is how to incorporate estimates of well leakage risk into a quantitative simulation approach that allows an operator to justify one sampling plan among several alternatives by using risk-based, data-driven decisions. The solution approach in Chapter 3 uses a stepwise procedure that was implemented using open-source (free) software; therefore, the process is reproducible and easily transferable to other modeling and simulation platforms. The probabilistic simulation incorporates sampling from the multivariate form of Wallenius' non-central hypergeometric distribution, which is the first time that this discrete distribution has been applied to this

particular problem. The approach can accommodate any type of well leakage risk scoring system into the process – the operator need only assign a leakage probability score to each well. The simulation approach therefore provides a practical example of how a field operator might incorporate prior knowledge about the probability of well leakage into their sampling plan design and thus quantitatively evaluate different plans to satisfy the site-specific monitoring objectives.

5.2.3 Novel Approach for Fitting CO₂ EOR Performance Metrics

The work in Chapter 4 employs a novel approach that incorporates nonlinear functions to describe the shape of the CO₂ retention, incremental oil recovery, and net CO₂ utilization responses with a handful of parameters. These three metrics are factors that significantly influence the long-term performance and economic viability of CO₂ EOR projects. This work showed that a four-parameter log-logistic function was able to accurately describe the overall shape of the CO₂ retention and incremental oil recovery curves across the 31 sites, and that a two-parameter simple asymptotic function was able to accurately describe the overall shape of the net CO₂ utilization curves across the 16 sites for which net CO₂ utilization values could be calculated. These functions provide a practical approach for operators to summarize key performance metrics of their CO₂ EOR field data and forecast the performance into the future. In addition, these functions also provide useful screening tools for broad regional assessments of the CO₂ storage resource associated with candidate CO₂ EOR fields. Lastly, these analytical approaches lend themselves to Monte Carlo simulation and therefore rapid uncertainty quantification as compared to complex three-dimensional numerical simulation models.

5.3 Recommendations for Future Research

Over the course of developing each paper, several challenges emerged and limitations in some of the tools and approaches were apparent. Based on these experiences, we suggest the following additional research areas which are worth investigating in the future.

- **CO₂ injection simulators which allow multiple realizations.** A deficit of the LBNL model that was used in Chapter 2 is the computational burden in post-processing associated with the first-order uncertainty analysis (FOUA) process – the total simulation time was less than one minute per run; however, the subsequent data handling and statistical computations took hours per run. The FOUA process demonstrated in Chapter 2 required 10 nominal model runs and $180 \pm \Delta x$ model runs, for a total of 190 model runs to generate the first-order uncertainty assessment. Altogether, the labor-intensive process took tens of hours and was an inefficient procedure consisting of (1) modifying the input text file; (2) executing the FORTRAN model; (3) importing the output data into Excel; and (4) performing the matrix calculations needed to conduct the FOUA assessment. There is a need for CO₂ injection simulation programs capable of running multiple realizations with user-defined sets of parameters analogous to those which are currently available within the LBNL model. Moreover, the FOUA process that was used in Chapter 2 assumes a local linear approximation for the model output to variations in the input parameters. A CO₂ storage system may exhibit nonlinear behavior, in which case the FOUA estimates would be inadequate. Having a simulation program that is capable of running multiple

realizations would allow for a Monte Carlo analysis of uncertainty, which would be better-suited for estimating nonlinear behavior, and for providing estimates of the full probability distribution function for the model outputs (Ramaswami et al., 2005). The CO₂ injection simulation program should be user-friendly and capable of being run by individuals who are not necessarily “expert-level” programmers such that practitioners with diverse backgrounds beyond programming may evaluate different modeling scenarios and ask fundamental questions.

- **Improved programming for post-processing computations.** Analogous to the computational intensity of the pressure monitoring work in Chapter 2, one deficit of the solution approach used in the wellbore leakage work in Chapter 3 is the calculations using logical statements for determining the probability of detecting one or more leaky wells. The solution approach compares a matrix of sampled wells to a matrix of leaky wells, and then uses logical statements (i.e., IF, OR, AND) to count “hits” and to generate the different performance metrics. The process that was used in the Chapter 3 work involved exporting files from R into Excel and then performing the computations. This process is time-consuming and potentially error-prone. Additional programming should be developed to seamlessly integrate the simulations of leaky wells in R using BiasedUrn, simulations of sampled wells using random or judgmental sampling, and the matrix comparisons. Improving this process would allow a practitioner to explore additional well leakage and well sampling scenarios without incurring the lengthy data handling and computational times associated with the current process.

- **Coupling of geospatial data to wellbore leakage.** The work in Chapter 3 uses a vector of well leakage potential scores (i.e., no *X*- or *Y*-coordinates), thus geospatial phenomena are not included in the calculations. However, it is likely that site-specific geologic characteristics such as preferential flow paths could lead to portions of the AoR having a higher risk of CO₂ leakage than others regardless of their well-specific leakage potential score. As a result, wells located closer in space could potentially suffer from “common cause failure”, or increased probability of leakage due to geologic conditions which place a greater risk onto specific regions of the site. In other words, the wellbore leakage potential would not be independent of the site geology, or stated differently, leakage would be dependent upon geospatial consideration beyond the well leakage potential score. The methods used in the Chapter 3 work would need to be coupled with a broader reservoir simulation model to properly address geospatial correlation and other site-specific geospatial features.
- **Accurate measurements related to wellbore leakage potential.** The work in Chapter 3 assumes that the leakage probability is directly proportional to the deep well leakage potential score. While the scoring methods of Watson and Bachu (2007; 2008) were derived from real-world data on surface-casing vent flow (SCVF) through wellbore annuli, soil gas migration (GM) outside casing, and available well completion and abandonment factors, the weighting factors and scoring approach are unlikely to be exact predictors of the increased probability of CO₂ leakage. The scoring has not been validated against field data of wellbore permeability estimates, which are the actual drivers of wellbore leakage, not

scores. For instance, it could be that all wells with a score higher than some threshold value are equally leaky in terms of their effective hydraulic parameters, and therefore the distribution of leakage probability would look quite different than the one used in the Chapter 3 work. Additional field data are needed to validate the scoring approach (or alternatively to devise a new and improved scoring approach) and accurately quantify the relative probability of increased well leakage potential for wells within an Area of Review.

- **Additional CO₂ EOR injection and production data from other regions.** The 31 CO₂ EOR sites in the database that was used in Chapter 4 reflect water-alternating gas CO₂ floods – all within the continental United States and heavily dominated by the West Texas carbonate floods (25 of 31 sites). A handful of other floods outside of this region are also included in this study (one from Colorado, three from New Mexico, one from Oklahoma, and one from Utah). While the West Texas carbonate floods reflect the region of the U.S. where CO₂ EOR has been most active historically, other areas within the U.S. are also anticipated to be targets for future CO₂ EOR. These include regions such as the Gulf Coast (Louisiana, Mississippi, Texas, and off-shore in the Gulf of Mexico), Midcontinent (primarily Oklahoma), and the Rocky Mountain Region (Colorado and Wyoming) (Kuuskra and Wallace, 2014). To get a broader sample of CO₂ EOR sites, additional injection and production data from sites located within these alternate regions should be aggregated into a similar database and analyzed using the methods described in Chapter 4. The results would provide a more representative set of parameters which could then be used to develop a more

accurate picture of the CO₂ storage capacity throughout the U.S. at CO₂ EOR sites.

- **Quantifying factors related to CO₂ EOR prediction.** The nonlinear curve fits that were used to evaluate the CO₂ EOR sites in Chapter 4 are focused on statistical inference – i.e., what are the likely outcomes and uncertainty ranges for CO₂ retention, incremental oil recovery, and net CO₂ utilization given the historical data from the 31 CO₂ EOR sites? However, the approach described in this work also provides useful information for prediction – i.e., given a set of inputs from another site with similar geology, what are plausible ranges in outcomes for each of these factors? While the work in Chapter 4 provides a generic approach for making broad predictions of CO₂ storage from estimates of original oil in place, future work is needed to provide more accurate predictions. This is especially true when trying to make predictions early in the CO₂ flood, as the curve-fitting techniques employed in Chapter 4 are sensitive to the length of the data record used to fit each function. Additional investigation into the geologic and operational factors that control the parameters of the four-parameter log-logistic or two-parameter asymptotic functions may help to refine these screening tools and their estimates of associated CO₂ storage.

5.4 References

DOE (2015) *National Risk Assessment Partnership*, Retrieved from <https://edx.netl.doe.gov/nrap/>

Fog, A. (2008) Sampling methods for Wallenius' and Fisher's noncentral hypergeometric distributions. *Communications in Statistics – Simulation and Computation*, 37(2): 241-257; DOI 10.1080/03610910701790236.

Fog, A. (2014) *BiasedUrn: Biased Urn model distributions*. R package version 3.02. <http://CRAN.R-project.org/package=BiasedUrn> (accessed May 7, 2014).

Kuuskra, V.; Wallace, M. (2014) CO₂-EOR set for growth as new CO₂ supplies emerge. *Oil & Gas Journal*, May:92-105.

Leach, A.; C.F. Mason; and K. van 't Veld. (2011) Co-optimization of Enhanced Oil Recovery and Carbon Sequestration. *Resource & Energy Economics*, 33(4): pp 893-912.

Peridas, G. (2008) *Written testimony submitted to the Natural Resources Committee, Subcommittee on Energy and Mineral Resources*, Unites States House of Representatives hearing on "Spinning Straw Into Black Gold: Enhanced Oil Recovery Using Carbon Dioxide," June 12, 2008.

Ramaswami, A.; Milford, J.B.; and Small, M.J. (2005) *Integrated Environmental Modeling: Pollutant Transport, Fate, and Risk in the Environment*. John Wiley & Sons, Inc. New Jersey. Chapter 14: Tools for Evaluation, Analysis, and Optimization of Environmental Models.

U.S. Department of Energy (U.S. DOE). (2009) *Best Practices for: Monitoring, Verification, and Accounting of CO₂ Stored in Deep Geologic Formations*; U.S. Department of Energy, National Energy Technology Laboratory, DOE/NETL-311/081508, March.

Watson, T.L.; and Bachu, S. (2007) *Evaluation of the potential for gas and CO₂ leakage along wellbores*. SPE Paper 106817 presented at the Society of Petroleum Engineers (SPE) E&P Environmental and Safety Conference, Galveston, Texas, March 5-7.

Watson, T.L.; and Bachu, S. (2008) *Identification of wells with high CO₂ leakage potential in mature oil fields developed for CO₂ enhanced oil recovery*. SPE Paper 112924 presented at the Society of Petroleum Engineers (SPE) Improved Oil Recovery Symposium, Tulsa, Oklahoma, April 19-23.

Appendix A: LBNL Model Input File for the Baseline Model

The Lawrence Berkley National Lab Model (LBNL Model) uses a text file as the input, which must be saved into the same directory as the LBNL Model executable file. The text below represents the input file for the baseline model that was used in Chapter 2. This text may be copied and pasted directly into a text file to reproduce the baseline model results. Figure 2-2 in Chapter 2 provides the values that were used to derive the input parameters for each aquifer and aquitard. Table 2-1 in Chapter 2 lists the uncertain parameter inputs and their statistical distributions.

```
*TITLE: DIFFUSE LEAKAGE IN EIGHTH-AQUIFER-SEVEN-AQUITARD SYSTEM (BIRKHOFF ET AL., 2009)
*-----*
* Model Selection
* Model Type=1: Focused Leakage Only
* Model Type=2: Diffuse Leakage Only
* Model Type=3: Coupled Focused and Diffuse Leakage
* Enter Model Type
  2

*-----*
* 1- RESERVOIR DESCRIPTION
* 1.1. Layer types at bottom and top
* AQUIFER = 1, AQUITARD = 0
* BOTTOM OF THE DOMAIN, BL
  1

* TOP OF THE DOMAIN; TL
  1

* 1.2: Number of Aquifers, NAQ
  8

* 1.3: Aquifer Properties
* BAQ HCONX ANSR  SS
60    0.101452974    1.    1.69284E-06
60    0.103743737    1.    1.70399E-06
60    0.103814996    1.    1.72306E-06
60    0.100094923    1.    1.74163E-06
60    0.095347086    1.    1.76346E-06
60    0.088516151    1.    1.7757E-06
60    0.081488598    1.    1.79016E-06
60    0.075849035    1.    1.80335E-06

* 1.4. Number of Aquitards
* NAQT=NAQ+1-BL-TL
  7

* 1.5. Aquitard Properties
* BAQT HCONP  SSP
100    1.02834E-06    1.98821E-06
100    1.04022E-06    1.97796E-06
100    1.021E-06 1.97604E-06
100    9.80039E-07    1.98082E-06
100    9.19177E-07    1.98942E-06
100    8.49152E-07    1.99954E-06
100    7.88179E-07    2.00954E-06

* 1.6. Domain Boundary Conditions. ENTER ONLY IF MODEL TYPE IS EQUAL TO 2 OR 3
* IF BOTB OR TOPB EQUAL TO Zero, ZERO DRAWDOWN BOUNDARY
* IF BOTB OR TOPB DIFFERENT THAN ZERO, NO FLOW BOUNDARY USED
* BOTB, TOPB
  1 1
```

* 2 - ACTIVE WELLS

* 2.1.Number of Injection/Pumping Wells and Number of Periods

* NIW NP

1 100

* 2.2. ENTER END TIME OF NP PERIODS

365
730
1095
1460
1825
2190
2555
2920
3285
3650
4015
4380
4745
5110
5475
5840
6205
6570
6935
7300
7665
8030
8395
8760
9125
9490
9855
10220
10585
10950
11315
11680
12045
12410
12775
13140
13505
13870
14235
14600
14965
15330
15695
16060
16425
16790
17155
17520
17885
18250
18615
18980
19345
19710
20075
20440
20805
21170
21535
21900
22265
22630
22995

0
0
0
0
0
0.

* 3 - LEAKY WELLS

* 4 - Solution Control and Time Steps

* 4.1. NINT: Number of terms for the Stehfest Laplace Inversion Formula (Series)
* NINT must be EVEN, and suggested to assign $8 \leq \text{NINT} \leq 16$ for efficient and accurate results
* NT: (Number of time steps to write output)
* NINT NT
16 100

* 4.2. Times to write output
* DT3 (I=1,...,NT); MUST BE > 0

365
730
1095
1460
1825
2190
2555
2920
3285
3650
4015
4380
4745
5110
5475
5840
6205
6570
6935
7300
7665
8030
8395
8760
9125
9490
9855
10220
10585
10950
11315
11680
12045
12410
12775
13140
13505
13870
14235
14600
14965
15330
15695
16060
16425
16790
17155
17520
17885

18250
18615
18980
19345
19710
20075
20440
20805
21170
21535
21900
22265
22630
22995
23360
23725
24090
24455
24820
25185
25550
25915
26280
26645
27010
27375
27740
28105
28470
28835
29200
29565
29930
30295
30660
31025
31390
31755
32120
32485
32850
33215
33580
33945
34310
34675
35040
35405
35770
36135
36500.

* 4.3. Variables written to output

* S:(Drawdown or buildup)
* P:(Pressure) (not available in this version)
* H:(Head) (not available in this version)
* OUT(I),I=1,3
1 0 0

* 4.4. MESH Option

* FLAG=1, Create Mesh from external file, (A FILE CALLED 'MESH_CONTOUR' SHOULD BE PROVIDED)
* FLAG= otherwise, provide manually
1

* IF FLAG DIFFERENT THAN ZERO, PROVIDE THE FOLLOWING

* Number of grid blocks IN X AND Y DIRECTIONS (for plotting)

* NI NJ

* 1 1

* Distance to nodes in x-direction. (m)

* IF FLAG=1, LEAVE EMPTY!, THE PROGRAM READS X(I) FROM MESH FILE

```

* X(i),i=1,NI
* Distance to nodes in Y-direction. (m)
* IF FLAG=1, LEAVE EMPTY!, THE PROGRAM READS Y(J) FROM MESH FILE
* Y(j),j=1,NJ

* 4.5: ENTER NUMBER OF AQUIFERS FOR CONTOUR PLOT
* NAQP
1

* 4.6: AQUIFER LAYER NUMBERS WHOSE CONTOUR PLOTS WILL BE PREPARED
* AQL(I); I=1,....,NAQP
1

* 4.7: ENTER NUMBER OF LOCATIONS FOR CALCULATION OF BUILDUP IN AQUIFERS
* NB
20

* 4.8: ENTER X AND Y FOR PLOTTING BUILDUP AT AQUIFERS AS FUNCTION OF TIME
* X Y
1000. 0.
2000. 0.
3000. 0.
4000. 0.
5000. 0.
6000. 0.
7000. 0.
8000. 0.
9000. 0.
10000. 0.
11000. 0.
12000. 0.
13000. 0.
14000. 0.
15000. 0.
16000. 0.
17000. 0.
18000. 0.
19000. 0.
20000. 0.

* 4.9: ENTER NUMBER OF LOCATIONS FOR Vertical Profile Plot
* NAB
20

* ENTER X AND Y FOR PLOTTING BUILDUP IN AQUITARDS
1000. 0.
2000. 0.
3000. 0.
4000. 0.
5000. 0.
6000. 0.
7000. 0.
8000. 0.
9000. 0.
10000. 0.
11000. 0.
12000. 0.
13000. 0.
14000. 0.
15000. 0.
16000. 0.
17000. 0.
18000. 0.
19000. 0.
20000. 0.

* end of input file
*-----*

```

Appendix B: Biased Urn R Code

In Chapter 3, one thousand random draws were generated in R using the package “BiasedUrn”, and this output was then used as the 1000-by- n_L matrix, $[A_{nL}]$. The original CRAN version of BiasedUrn (available for download at: <http://cran.r-project.org/web/packages/BiasedUrn/index.html>) is only compiled to run a maximum of 32 “colors” (MAXCOLORS=32) or 32 different odds that can be used to weight each random draw. Since the work in Chapter 3 used 100 wells each with their own leakage potential score, additional colors were needed. The work in Chapter 3 therefore used a precompiled Windows version of BiasedUrn that was modified by Emory University to allow a maximum of 10,000 colors (MAXCOLORS=10000). Both 32- and 64-bit versions compiled for R-3.0.1 are available for download at the following website:

<http://genetics.emory.edu/labs/epstein/software/BiasedUrn/burn-install-windows.html>

Once the BiasedUrn package is installed, to simulate 1000 random draws in R for the case with 10 leaky wells would use the following code:

```
mymatrix = rMWNCHypergeo(1000, m=m, n=10, odds=odds) (10)
```

Where:

rMWNCHypergeo	= specifies the random variate generation function of the multivariate form of Wallenius’ non-central hypergeometric distribution;
1000	= specifies the number of random variables to generate;
m	= refers to a vector of $n=100$ wells (all equal to 1);
n	= refers to n_L , the number of leaky wells (10 in this example); and
odds	= refers to a vector of weights (Prob[Leak Risk _{<i>i</i>}]) for each well.

The output, “mymatrix”, may then be exported from R and used to derive the 1000-by- n_L matrix, $[A_{nL}]$.