



This is a repository copy of *Stochastic injection-strategy optimization for the preliminary assessment of candidate geological storage sites.*

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/89694/>

Version: Accepted Version

Article:

Cody, B.M., Baù, D. and González-Nicolás, A. (2015) Stochastic injection-strategy optimization for the preliminary assessment of candidate geological storage sites. *Hydrogeology Journal*, 23 (6). 1229 -1245. ISSN 1431-2174

<https://doi.org/10.1007/s10040-015-1250-5>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Stochastic injection-strategy optimization for the preliminary assessment of candidate geological storage sites

Brent M. Cody^{1*}, Domenico Baù², Ana González-Nicolás³
1. Natural Resources Consulting Engineers, bcody@nrce.com
2. University of Sheffield, Sheffield, UK
3. Colorado State University, Fort Collins, CO, USA
*corresponding author

Keywords: Climate Change, Analytical Models, Multi-objective Optimization, Carbon Sequestration

Hydrogeology Journal, 23 (6). 1229 -1245. ISSN 1431-2174
Published online 25 April 2015: DOI 10.1007/s10040-015-1250-54.

Abstract

Geological carbon sequestration (GCS) has been identified as having the potential to reduce increasing atmospheric concentrations of carbon dioxide (CO₂). However, a global impact will only be achieved if GCS is cost-effectively and safely implemented on a massive scale. This work presents a computationally efficient methodology for identifying optimal injection strategies at candidate GCS sites having uncertainty associated with caprock permeability, effective compressibility, and aquifer permeability. A multi-objective evolutionary optimization algorithm is used to heuristically determine non-dominated solutions between the following two competing objectives: 1) maximize mass of CO₂ sequestered and 2) minimize project cost. A semi-analytical algorithm is used to estimate CO₂ leakage mass rather than a numerical model, enabling the study of GCS sites having vastly different domain characteristics. The stochastic optimization framework presented herein is applied to a feasibility study of GCS in a brine aquifer in the Michigan Basin (MB), USA. Eight optimization test cases are performed to investigate the impact of decision-maker (DM) preferences on Pareto-optimal objective-function values and carbon-injection strategies. This analysis shows that the feasibility of GCS at the MB test site is highly dependent upon the DM's risk-adversity preference and degree of uncertainty associated with caprock integrity. Finally, large gains in computational efficiency achieved using parallel processing and archiving are discussed.

1 Introduction

Geological carbon sequestration (GCS) has been identified as a prominent technology to manage increasing atmospheric concentrations of carbon dioxide (CO₂) (Nordbotten and Celia 2012, Pacala and Socolow 2004), however, the effective application of GCS will require a global implementation of large numbers of carbon injection projects. While an individual large coal-fired power plant may emit up to 5-10 megatonnes (Mt) of CO₂ per year (Birkholzer et al. 2012), total annual global anthropogenic carbon emissions measured in mass of CO₂ are approximately 30,000 Mt (Nordbotten and Celia 2012). Results from Eccles et al. (2012) suggest that specific regions of a small number of candidate aquifers will provide the majority of low cost geological CO₂ storage. As the selection of the appropriate reservoir is crucial to the success of a GCS project (Bossie-Codreanu and Le Gallo 2004), many potential injection sites will need to be assessed world-wide for GCS suitability. Therefore, the efficient preliminary characterization of candidate GCS injection sites has the potential for massive resource savings. In addition, a comprehensive pre-screening effort will increase GCS storage reliability by eliminating “bad” and identifying “good” GCS reservoirs.

Large-scale conjunctive preliminary project planning will involve the characterization, optimization, and risk assessment of several potential GCS sites. There are, however, several difficulties associated with these tasks. The first is that the large-scale, multiphase numerical modeling of several potential injection sites for the purpose of initial assessment is infeasible due to the effort involved in model construction and calibration. Data characterizing the subsurface domain are typically scarce, which introduces parameter uncertainty and adds to the complexity of modeling GCS. Also, because of their propensity to be computationally expensive (Espinete and Shoemaker 2013), the direct use of large-scale, multiphase numerical models would be unrealistic in simulating the high volume of realizations needed for risk assessment and optimization.

The high computational cost associated with numerical models may be reduced by the use of data-based response surface methods (e.g. Box and Draper 2007). However, it is the authors' intent for the resulting framework to ultimately be used to optimize and compare large numbers of potential injection sites having vastly different domain characteristics. Creating and calibrating each potential injection site's numerical model, as well as training the resulting response surface would require user expertise and large investments of computational time. This work has therefore chosen to use a semi-analytical multiphase flow model presented by Nordbotten et al. (2009) for multiphase subsurface flow simulation. The semi-analytical leakage algorithm is very general and can be applied to simplified computational models of the vast majority of potential injection sites.

An additional difficulty associated with preliminary GCS project planning is that potential storage reservoirs typically exhibit a high degree of uncertainty associated with physical parameters. Celia et al. (2011) identified abandoned (herein referred to as "passive") well permeabilities as the most dominant uncertainty parameter when estimating fluid leakage due to GCS. In North America, significant numbers of passive wells may perforate the caprock in formations suitable for GCS (Celia et al. 2009, Nogues and Dobossy 2012, Nordbotten and Celia 2012). Most likely, very little information exists on the location and/or sealing properties of these wells. However, several efforts have been made to investigate and account for the uncertainty associated with passive well permeability. Watson and Bachu (2008; 2009) developed a passive well integrity scoring index based upon typically available information (e.g. completion date, regulatory requirements, etc.). Crow et al. (2009) physically sampled and analyzed segments of a 30 yr. old passive well that had been continuously exposed to 96% CO₂ finding that cement interfaces are more important than the cement matrix when quantifying migration pathways.

Multiphase subsurface optimization problems are typically highly non-linear due to the irregular spatial location of preferential flow pathways and the multiphase flow (i.e. CO₂ and brine) equations governing pressure response and CO₂ plume migration. Therefore, a robust global optimization tool is needed to

find best performing injection strategies that maximize the mass of CO₂ sequestered while minimizing project cost by selecting optimal injection well locations and injection rates. In multi-objective problems, a Pareto-optimal, or non-dominated, solution outperforms all other solutions with respect to all objectives (Reed 2013). Multi-objective evolutionary algorithms (MOEAs) have been shown to be effective in providing Pareto-optimal solutions for a large number of subsurface flow applications possessing several decision variables (Alzraiee et al. 2013, Baú 2012, Chen et al. 2007, Hansen et al. 2013, Kumphon 2013, Mantoglou and Kourakos 2006, Nicklow et al. 2010, Peralta et al. 2014, Reed and Kollat 2013, Singh 2013, Singh 2014, Singh and Chakrabarty 2010, Tabari and Soltani 2012, Zheng, F., Zecchin 2014). In particular, (Reed 2013) presents a comprehensive review of state-of-the-art MOEAs highlighting key algorithm advances which may be used to identify critical tradeoffs in water resources problems. A non-dominated sorting genetic algorithm (NSGA-II; Deb et al. 2002) with ϵ -dominance (Laumanns et al. 2002) has been selected as the computational optimization tool because it is among the most reliable and best performing multi-objective optimization evolutionary algorithms (Cheng et al. 2009).

If computationally feasible, stochastic methods should be applied in cases where parameter uncertainty is of significant concern. A popular approach for accomplishing this is to apply a Monte Carlo (MC) method where simulation is performed for an ensemble of uncertain parameter sets to estimate the statistics of optimization objectives and constraints. There are several examples in the literature where MOEAs are coupled with MC techniques to optimize groundwater problems having parameter uncertainty. A multi-objective groundwater flow optimization problem with aquifer hydraulic conductivity uncertainty is solved by Baú (2012) using an NPGA. Alzraiee et al. (2013) used a MC-based Bayesian update scheme to approximate posterior uncertainty in hydraulic conductivity and head using an NSGA-II when performing multi-objective design of aquifer monitoring networks. A MC approach was also used by Mantoglou and Kourakos (2006) when determining optimal remediation methods for groundwater aquifers having hydraulic conductivity uncertainty. MC techniques are also used to investigate parameter uncertainty associated with GCS (Goda and Sato 2011, Oladyshkin et al.

2011, Viswanathan et al. 2008, Wagner and Gorelick 1987, Wriedt et al. 2014). In particular, Celia et al. (2011) applied a stochastic Monte Carlo approach to estimate leakage risk associated with passive well permeability uncertainty. Nogues and Dobossy (2012) used a large-scale Monte Carlo method to explore the effects of caprock permeability uncertainty on fluid leakage estimation, finding that the amount of CO₂ leakage from GCS is typically acceptable for climate change mitigation.

Herein, several computational tools are integrated into a stochastic multi-objective optimization framework for the purpose of performing large-scale GCS site feasibility studies. These tools include 1) a semi-analytical leakage algorithm to rank the performance of trial injection strategies; 2) a Monte Carlo procedure to quantify risk resulting from parameter uncertainty; and 3) a NSGA-II with ϵ -dominance (Deb et al. 2002) to heuristically determine Pareto-optimal solutions between competing objectives. The paper begins with a detailed description of the stochastic optimization methodology including an overview of both the semi-analytical CO₂ leakage algorithm and the NSGA-II with ϵ -dominance. Next, a general description and GCS-specific characterization of a saline aquifer situated within the Michigan Basin (MB) is presented. Following this is the formulation of the optimization case study at a MB test site. The framework is then used to investigate the following three goals regarding the stochastic optimization of the MB site: 1) quantify the impact of decision maker (DM) preferences on heuristically determined Pareto-optimal objective values (i.e. mass sequestered and project cost); 2) quantify the impact of DM preferences on carbon injection strategy (i.e. the selection of injection well flow rate and location), and; 3) assess the suitability of the MB test site for GCS. Finally, gains in computational efficiency using parallel processing and simulation archiving are discussed.

2 Methodology

2.1 Semi-Analytical CO₂ Leakage Estimation

While other semi-analytical algorithms provide insight regarding specific processes, this work has chosen to use a modified version of the ELSA (Estimating Leakage Semi-analytically) multiphase subsurface flow model (Nordbotten et al. 2009, Celia et al. 2011). This model has been used in several applications (Nogues and Dobossy 2012, Bielicki et al. 2013 & 2014) because it is the only semi-analytical model able to simulate multiphase flow in domains having multiple injection wells, geological layers, and passive wells (i.e. weak caprock areas).

Figure 1 shows a schematic of the semi-analytical leakage model's computational domain. This domain is structured as a stack of $L+1$ caprock layers separated by L aquifer layers, perforated by M carbon injection wells and N passive wells. Aquifers are assumed to be horizontally level, homogenous, and isotropic. Caprock layers are assumed to be impermeable, except where perforated by passive wells. Injection wells are able to inject into any layer. Initially, fluid is not flowing through any of the passive wells because the entire domain is assumed to be saturated with brine at hydrostatic pressure. Additional assumptions made by this model include: a) aquifers exhibit horizontal flow; b) capillary pressure is negligible, resulting in a sharp CO₂-brine interface; c) CO₂ plume thickness at any given location is assumed to be the maximum plume thickness from all sources and sinks in the aquifer; d) pressure response from sources and sinks can be superimposed in each aquifer; and e) the injectivity of the formation remains constant. Several of these processes are important (Cihan et al. 2013, Court et al. 2012, Doster et al. 2013, Gasda et al. 2011, Gasda et al. 2012, Heße et al. 2013, Juanes et al. 2009) and should be included (Buscheck et al. 2012, Gasda et al. 2009, Huang et al. 2014, Nordbotten and Flemisch 2012) when accuracy is more important than efficiency (e.g. final project design).

At the start of injection, aquifer fluid pressures throughout the domain begin to change resulting in pressure differentials across caprock layers and fluid flux through passive wells. It is therefore very important to understand aquifer fluid pressure response resulting from changes in the mass storage of CO₂ and brine. A pressure response function for the injection of CO₂ into a brine-filled confined aquifer was derived by Nordbotten and Celia (2005). Celia et al. (2011) expresses this radial pressure response p [ML⁻¹T⁻²] at the bottom of the aquifer where a single well injects CO₂ as:

$$p = p_0 + (\rho_b - \rho_c) \cdot g \cdot H \cdot \Delta p' \quad (1)$$

where p_0 is the initial fluid pressures at the bottom of the aquifer, ρ_α is fluid density [ML⁻³] (α denotes the phase type, b for brine and c for CO₂), g is gravitational acceleration [LT⁻²] and H is aquifer thickness [L]. In Eqn (1), $\Delta p' (/)$ is defined by:

$$\Delta p'(\chi) = \begin{cases} 0 & \text{for } \chi \geq \psi \\ -\frac{1}{2\Gamma} \ln\left(\frac{\chi}{\psi}\right) + \Delta p'(\psi) & \text{for } \psi > \chi \geq 2\lambda \\ \frac{1}{\Gamma} - \frac{\sqrt{\chi}}{\Gamma\sqrt{2\lambda}} + \Delta p'(2\lambda) + F(h') & \text{for } 2\lambda > \chi \geq \frac{2}{\lambda} \\ -\frac{1}{2\lambda\Gamma} \ln\left(\frac{\chi\lambda}{2}\right) + \Delta p'\left(\frac{2}{\lambda}\right) & \text{for } \frac{2}{\lambda} > \chi \end{cases} \quad (2)$$

where,

$$\chi = \frac{2\pi H\phi(1 - S_b^{\text{res}})r^2}{Q \cdot t} \quad (3)$$

$$\Gamma = \frac{2\pi(\rho_b - \rho_c)gkH^2}{\mu_b Q} \quad (4)$$

$$\psi = \frac{4.5\pi H\phi k(1 - S_b^{\text{res}})}{\mu_b c_{\text{eff}} Q} \quad (5)$$

$$h' = \frac{h(\chi)}{H} = \frac{1}{\lambda - 1} \left(\frac{\sqrt{2\lambda}}{\sqrt{\chi}} - 1 \right) \quad (6)$$

$$F(h') = \frac{-\lambda}{\lambda - 1} \left[h' - \frac{\ln[(\lambda - 1)h' + 1]}{\lambda - 1} \right] \quad (7)$$

In Equations (2-7): B is aquitard thickness [L]; h is CO₂ plume thickness [L]; h' [/] is the CO₂ plume thickness relative to the aquifer thickness H ; S_b^{res} is the residual saturation of the brine [/]; t is time [T]; k is the aquifer permeability [L²]; μ is the dynamic viscosity [ML⁻¹T⁻¹]; ϕ is the aquifer porosity [/]; Q is the total volumetric well flux [L³T⁻¹]; c_{eff} is the effective compressibility of the fluid and solid matrix [M⁻¹L²T²]; and r is the radial distance [L]. Also, $F(h')$ is an offset term related to the vertical pressure distribution (Celia et al. 2011) and the mobility ratio is defined as $\lambda = \lambda_c/\lambda_b$, where $\lambda_\alpha = k_{r,\alpha}/\mu_\alpha$ and $k_{r,\alpha}$ is the relative permeability of phase α ($\alpha = b$ for brine or $\alpha = c$ for CO₂).

To determine the fluid overpressure at any given time throughout the aquifer system, Nogues and Dobossy (2012) applied superposition of effects derived from the application of Eqn (1) for all the volumetric sources and sinks corresponding to CO₂ injection wells and passive wells. Consequently, the fluid pressure at any given time t , at the bottom of the generic aquifer l ($l=1,2,\dots,L$) and for each passive well j ($j=1,2,\dots,N$) can be expressed as:

$$p_{j,l} = p_{0l} + (\rho_b - \rho_c) \cdot g \cdot H_l \cdot \left[\sum_{i_w=1}^M \Delta p'(\chi_{i_w,j,l}) + \sum_{i=1}^N \Delta p'(\chi_{i,j,l}) \right] \quad (8)$$

where

$$\chi_{i_w,j,l} = \frac{2\pi H_l \phi_l (1 - S_{b,l}^{\text{res}}) r_{i_w,j}^2}{(Q_{i_w,l} \cdot t)} \quad (9)$$

and

$$\chi_{i,j,l} = 2\pi H_l \phi_l (1 - S_{b,l}^{\text{res}}) r_{i,j}^2 / \int_0^t (Q_{j,l} - Q_{j,l+1}) \cdot d\tau \quad (10)$$

By following this approach, the fluid pressures at the bottom of each aquifer and at each passive well can be grouped into the $N \times L$ vector \mathbf{p} :

$$\mathbf{p}(t) = \mathbf{p}[\mathbf{P}_1, \mathbf{M}(t)] \quad (11)$$

Eqn (11) shows that $\mathbf{p}(t)$ is a function of the array

$$\mathbf{P}_1 \equiv [\mathbf{H}, \boldsymbol{\phi}, \mathbf{k}, \mathbf{S}_b^{\text{res}}, \mathbf{r}, \mathbf{Q}_{i_w}, \rho_c, \rho_b, g, \lambda, \pi, \mu_b, c_{\text{eff}}] \quad (12)$$

In Eqn (12): the $L \times 1$ vectors \mathbf{H} , $\boldsymbol{\phi}$, $\mathbf{S}_b^{\text{res}}$ and \mathbf{k} include the thicknesses, porosities, brine residual saturations and permeabilities of all aquifers; \mathbf{Q}_{i_w} is $M \times L$ vector including the CO_2 inflow rates for each

aquifer l ($l=1,2,\dots,L$) for injection well iw ($iw=1,2,\dots,M$); and the $(M+N)\times(M+N)$ \mathbf{r} matrix includes the relative distances between injection and passive wells.

In addition, $\mathbf{p}(t)$ (Equation 9) is a function of the $N\times L$ vector $\mathbf{M}(t)$, whose generic component $M_{j,l}(t)$ represents the net cumulative fluid mass transferred into aquifer l through passive well j . This mass is calculated as:

$$M_{j,l}(t) = \int_0^t \rho_{\text{eff},j,l}(\tau) \cdot [Q_{j,l}(\tau) - Q_{j,l+1}(\tau)] \cdot d\tau \quad (13)$$

where $\rho_{\text{eff},j,l}(\tau)$ is the effective fluid density in aquifer l , at passive well j which is time-dependent since the leaking fluid composition is a function of the phase saturations of CO_2 and brine, which vary based on the CO_2 plume location. The effective fluid density is estimated as $\rho_{\text{eff},j,l} = \rho_c \cdot S_{c,j,l} + \rho_b \cdot (1 - S_{c,j,l})$.

Since application of Eqn (11) requires knowing the temporal evolution of leakage rates through passive wells $Q_{j,l}$, Nogues and Dobossy (2012) propose to use the sum of the flow rates, $Q_{\alpha,j,l}$, for each phase α , calculated using the multiphase version of Darcy's law across each confining layer l for each passive well j :

$$Q_{j,l} = \sum_{\alpha=b,c} \left[\pi r_{\text{pw},j,l}^2 \frac{k_{r,\alpha,j,l} k_{\text{pw},j,l}}{\mu_{\alpha} B_l} (p_{j,l-1} - \rho_{\alpha} g B_l - g \rho_{\alpha} H_{l-1} - p_{j,l}) \right] \quad (14)$$

In Eqn (14), $r_{\text{pw},j,l}$ is the passive well radius and $k_{\text{pw},j,l}$ is the single phase passive well permeability for passive well j ($j=1,2,\dots,N$) and aquitard layer l . Note that in order to apply Eqn (14), fluid pressures $p_{j,l}$ as

well as CO₂ relative thicknesses in passive well pathways, must be known to estimate pressure gradients, fluid saturations and relative permeability values. Given Eqn (14), the flow rates across each aquitard l ($l=1,2,\dots,L$) for each passive well j ($j=1,2,\dots,N$) can be grouped into the $N \times L$ vector \mathbf{Q} :

$$\mathbf{Q}(t) = \mathbf{Q}[\mathbf{P}_2, \mathbf{p}(t)] \quad (15)$$

where the array \mathbf{P}_2 is given by:

$$\mathbf{P}_2 \equiv [\mathbf{B}, \mathbf{H}, \mathbf{r}_{pw}, \mathbf{k}_{pw}, \mathbf{k}_{r,c}, \mathbf{k}_{r,b}, \rho_c, \rho_b, \mu_c, \mu_b, g] \quad (16)$$

In Eqn (16), the $(L+1) \times 1$ vector \mathbf{B} includes the aquitard thicknesses, the $N \times (L+1)$ matrices \mathbf{r}_{pw} and \mathbf{k}_{pw} contain the passive well radii and permeabilities, and the $N \times (L+1)$ matrices $\mathbf{k}_{r,c}$ and $\mathbf{k}_{r,b}$ include the relative permeabilities of CO₂ and brine at passive wells.

By combining Equations (9) and (13), a set of $N \cdot L$ non-linear equations in $N \cdot L$ unknowns is obtained. These unknowns are the fluid pressures at the bottom of each aquifer and at each passive well (Equation 8), and the flow rates (Equation 12) across each aquitard for each passive well.

In an effort to increase computational efficiency, this work has chosen to neglect effects from upconing and apply an Iterative Global Pressure Solution (IGPS) modification where a fixed-point iteration scheme (Takahashi 2000) is applied to solve the system of non-linear equations introduced above at a generic time t given the solution at time $t - \Delta t$ obtained at the previous time step. The resulting CO₂ leakage simulator is highly efficient, capable of simulating 50 years of injection into a domain with 100 passive wells and 4 aquifer/aquitard layers in less than 1 second on a standard laptop.

2.2 Stochastic Multi-objective GCS Problem

For any given potential GCS reservoir, a set of Pareto-optimal injection scenarios exist which (a) maximize the deterministic mass of CO₂ injected, Mass_{inj} , and (b) minimize the stochastic project cost, C :

$$\text{objective (a):} \quad \max\{\text{Mass}_{\text{inj}}(\mathbf{Q})\} \quad (17)$$

$$\text{objective (b):} \quad \min\{C(\mathbf{x}, \mathbf{Q}, \mathfrak{S})\} \quad (18)$$

In Eqns (17-18), spatial position, \mathbf{x} , and flow rate, \mathbf{Q} , are the two decision variables comprising each injection strategy and \mathfrak{S} is the set of stochastic state variables (i.e. aquifer pressure distribution and CO₂ plume locations and thicknesses) dependent on each simulation's injection strategy and stochastic parameter set. Mass_{inj} is simply the total mass of CO₂ injected into the domain during simulation, and is calculated by multiplying the sum of injection well flow rates by the injection duration, t_{inj} :

$$\text{Mass}_{\text{inj}} = \sum_{iw=1}^M Q_{c_{iw}} * t_{\text{inj}} \quad (19)$$

where M is the number of injection wells, and $Q_{c_{iw}}$ is the CO₂ mass injection rate for injection well iw .

Note that the project cost, C , is a stochastic variable and is here estimated using a Monte Carlo (MC) simulation based on a stochastically generated parameter set. In order to quantify the risk of CO₂ leakage associated with uncertain aquifer permeability, effective compressibility, and passive-well segment permeability, an ensemble of N_{MC} parameter set realizations $[(\mathbf{k}^{(1)}, \mathbf{k}^{(2)}, \dots, \mathbf{k}^{(N_{\text{MC}})}), (\mathbf{c}_{\text{eff}}^{(1)}, \mathbf{c}_{\text{eff}}^{(2)}, \dots, \mathbf{c}_{\text{eff}}^{(N_{\text{MC}})}), (\mathbf{k}_{\text{pw}}^{(1)}, \mathbf{k}_{\text{pw}}^{(2)}, \dots, \mathbf{k}_{\text{pw}}^{(N_{\text{MC}})})]$, with each having an equal probability of occurrence, is generated for

each trial injection strategy. For each parameter realization, aquifer permeabilities and effective compressibilities are assigned a value using log-normal distributions. Each passive-well segment is assigned either an “intact” (low) or “degraded” (high) (Celia et al. 2011, Nogues and Dobossy 2012) permeability value. This is accomplished using a simple bi-value probability distribution function (PDF). Passive-well segments have a P_{intact} chance of being assigned an “intact” permeability value and a $(1 - P_{\text{intact}})$ chance of being assigned a “degraded” permeability value. The cost, Cost_i , associated with each MC realization, i ($i = 1, 2, \dots, N_{\text{MC}}$), consists of the summation of each injection well’s (iw) capital cost, Cap, operational cost, OP, surface maintenance cost, SurM, subsurface maintenance cost, SubM, and variable cost, Var, added to the cost associated with CO₂ leakage, LC:

$$\begin{aligned} \text{Cost}_i(\mathbf{Q}, \text{Mass}_{\text{leak}}) &= \sum_{iw=1}^M [\text{Cap}_{iw} + \text{OP}_{iw}(t_{\text{inj}}) + \text{SurM}_{iw}(t_{\text{inj}}) + \text{SubM}_{iw}(t_{\text{inj}}) \\ &+ \text{Var}_{iw}(Q_{iw}, t_{\text{inj}})] + \text{LC}[\text{Mass}_{\text{leak}}(\mathbf{Q}, \mathbf{s}_i)] \end{aligned} \quad (20)$$

where \mathbf{s}_i is the set of state variables for each MC realization including the CO₂ saturation and total flow rate in each passive-well segment. LC is estimated as:

$$\text{LC}[\text{Mass}_{\text{leak}}(\mathbf{Q}, \mathbf{s}_i)] = c_L \cdot \text{Mass}_{\text{leak}}(\mathbf{Q}, \mathbf{s}_i)^{r_A} \quad (21)$$

In Eqn (21), c_L is the coefficient representing penalty cost per unit of CO₂ leakage (\$/kg) and r_A is a risk adversity factor reflecting the preferences of the DM by exponentially increasing LC in relation to the mass of CO₂ leakage. The mass of CO₂ leakage (kg), $\text{Mass}_{\text{leak}}$, is quantified numerically using the semi-analytical leakage algorithm described in Section ‘*Semi-analytical CO₂ leakage estimation*’ above as:

$$\text{Mass}_{\text{leak}}(\mathbf{Q}, \mathbf{s}_i) = \sum_{j=1}^N \int_0^{t_{\text{inj}}} \rho_c \cdot S_{c_{j,L}}^{(i)}(t) \cdot Q_{j,L}^{(i)}(t) \cdot dt \quad (22)$$

Empirical cumulative distribution functions (CDFs) of each injection strategy's MC **Cost** ($\text{Cost}_1, \text{Cost}_2, \dots, \text{Cost}_{N_{\text{MC}}}$) are then compiled by first sorting the **Cost** vector in ascending order from $i = 1 \dots N_{\text{MC}}$, then assigning the non-exceedance probability of Cost_i as $(i - 0.5)/N_{\text{MC}}$ (Hahn and Shapiro 1967). The stochastic cost objective, C , is finally assigned as the z^{th} percentile value, P_z , of the sorted **Cost** vector:

$$C(\mathbf{x}, \mathbf{Q}, \tilde{\mathbf{s}}) = P_z(\mathbf{Cost}) \quad (23)$$

As discussed later in Section '*NSGA-II parameter calibration*', larger N_{MC} values are needed to accurately quantify C as z becomes farther from 50%. Several constraints are included in this problem. First, the number of injection wells is limited to a maximum integer value by constraining M to a value between 0 and M_{max} :

$$0 \leq M \leq M_{\text{max}} \quad (24)$$

All candidate injection wells must be located horizontally within prescribed minimum and maximum spatial bounds representing the areal extent selected for the construction of GCS facilities:

$$x_{\text{min}} \leq x_{\text{iw}} \leq x_{\text{max}} ; y_{\text{min}} \leq y_{\text{iw}} \leq y_{\text{max}} ; \text{iw} = 1, 2, \dots, M \quad (25)$$

Also, each injection well's flow rate must be between prescribed minimum and maximum flow rates Q_{min} and Q_{max} :

$$Q_{\min} \leq Q_{iw} \leq Q_{\max} ; iw = 1, 2, \dots, M \quad (26)$$

In Eqn (26), Q_{\min} and Q_{\max} are physical constraints related to the technical capacity of injection pumps and wells. Finally, down-hole fluid pressures at each injection well in each layer, $p_{iw,l}$, must not exceed each layer's fracture fluid pressure, p_{frac_l} .

$$p_{iw,l}(\mathbf{Q}, \mathbf{s}) < p_{frac_l} ; iw = 1, 2, \dots, M ; l = 1, 2, \dots, L \quad (27)$$

where p_{frac_l} is calculated by multiplying a specified fracture gradient (Celia et al. 2011) by the layer depth and $p_{iw,l}$ is estimated for each injection well iw at the end of the injection duration at a small effective distance from the injection location. The value for effective distance must be carefully selected as Eqn (1) is very sensitive to radial distance when estimating pressure changes nearby the injection location. A conservative estimate for this value would be the injection well's casing radius. Note that because the left-hand side of Eqn (27) contains stochastic state variables, these stochastic constraints have to be transformed into deterministic. This is carried out using a chance-constraint approach (Wagner and Gorelick 1987) such that:

$$\text{Prob}\{p_{iw,l}(\mathbf{Q}, \mathbf{s}) < p_{frac_l}\} \geq S_p ; iw = 1, 2, \dots, M ; l = 1, 2, \dots, L \quad (28)$$

Eqn (28) requires the probability of the fluid pressure at injection wells not exceeding the layer fracture pressure to be larger than or equal to a prescribed fracture safety probability, S_p .

2.3 Multi-objective GCS Optimization Algorithm

In order to solve the stochastic multi-objective optimization problem presented in Section '*Stochastic Multi-objective GCS Problem*', the semi-analytical CO₂ leakage algorithm (Section '*Semi-Analytical CO₂*

Leakage Estimation') and the Monte Carlo procedure (Section '*Stochastic Multi-objective GCS Problem*') are integrated into a NSGA-II with ε -dominance. Each trial injection strategy is encoded into a "chromosome" (i.e. a unique sequence of binary numbers representing the values of each decision variable). The first segment of the chromosome contains spatial location information for M_{\max} injection wells represented by N_{cl} predefined candidate injection-well location indices, thus intrinsically satisfying the spatial location constraint defined by Eqn (25). The number of digits required to represent each injection well's location is a function of the number of candidate locations, N_{cl} . For example, if $N_{cl} = 16$ for a potential GCS site, the locational information of each injection well for each injection strategy would require four binary digits (i.e. '0000' = 1, '0001' = 2, ..., '1111'=16). In this case, the spatial data component of the chromosome has a length of $M \cdot 4$. The last segment of the chromosome represents the injection rate index of each candidate well. Injection rate indices represent N_{ir} prescribed discrete flow rate values within the constraints of Eqn (26). For example, injection rate index values of 0, 1, 2, 3, and 4 may represent injection rates of 0, 20, 30, and 40 kg/s, respectively ($N_{ir} = 4$). In this case, two binary digits would be used to encode each well's injection index ('00' = 0, '01' = 1, ..., '11'=3) bringing the total chromosome length to $M \cdot 4 + M \cdot 2$.

The general procedure followed by the NSGA-II with ε -dominance is presented in Figure 2. Once input files containing domain characteristics and algorithm parameters are read, an initial population of N_{pop} trial injection scenarios, including the spatial location and flux rate for each injection well, is randomly generated. In step (b), objective function values are calculated for each population member and simulation archiving keeps a record of simulation results, preventing the NSGA-II from recalculating objective values for identical scenarios. As described in Section '*Stochastic Multi-objective GCS Problem*', an ensemble of parameter set realizations are generated for each trial injection strategy in step (c). After this, fitness values (i.e. an assessment of performance related to objective function values) are calculated for each injection strategy. The semi-analytical leakage algorithm is used to estimate $Mass_{leak}$

for each of these realizations in step (d), then objective values $Mass_{inj}$ and \tilde{C} are calculated using Equations (16-20). Also in step (d), if Eqn (28) is violated the injection strategy is deemed infeasible.

Next, the NSGA-II is used to generate new injection strategy populations of size N_{pop} using selection, crossover, and mutation operators. In step (e), population members are first ranked (i_{rank}) as the number of solutions dominating population member i using the fast non-dominated sorting procedure (Deb et al. 2002), and then assigned a crowding distance ($i_{distance}$) as the largest cuboid in objective space enclosing the point i without including any other point in the population (Deb et al. 2002). A partial order is established using the crowding comparison operator, \succeq_n (Deb et al. 2002). Population member i outperforms member j if the following conditions are met:

$$i \succeq_n j \quad \text{IF } \{(i_{rank} < j_{rank}) \text{ OR } ((i_{rank} = j_{rank}) \text{ AND } (i_{distance} > j_{distance}))\} \quad (29)$$

The NSGA-II's ranking process is further improved with the concept of ε -dominance (Laumanns et al. 2002). Suppose population members p_1 and p_2 have fitness values of f_1 and f_2 , respectively. Using ε -dominance, p_1 is allowed to dominate p_2 if $(1+\varepsilon)f_1$ is greater than or equal to f_2 . The ε -dominance allows for the inclusion of additional, well-performing population members to each rank's Pareto-optimal front.

Next, in step (f), an iterative tournament-style selection process is used to select parent injection strategies for the next generation. A subset of T_{size} population members is selected randomly and the population member with highest partial order according to Eqn (29) is chosen as a parent. Once parents are selected, a crossover operator is used to create the new generation of population members in step (g). During each crossover operation, components of chromosomes from two randomly selected parents are used to build a new trial injection strategy. Crossover is repeated until the new population is filled. Finally, there is a

chance, quantified as the mutation rate, M_{rate} , in which chromosome elements of this new population will be randomly altered during step (h). Steps (b-h) are repeated until the maximum prescribed number of generations, N_{gens} , is reached.

2.4 Efficient Computational Implementation

This framework utilizes parallel computing and simulation archiving to improve the computational efficiency. Due to the iterative nature of both evolutionary search and Monte Carlo analyses, large numbers of model simulations are needed for each stochastic optimization run. Without using simulation archiving, the total number of model calls required for each stochastic optimization run, N_{ct} , is equal to the product of N_{MC} , N_{pop} , and N_{gens} . CO₂ leakage estimation calculations for each trial injection strategy are independent and therefore may be processed in parallel (i.e. simultaneously) rather than sequentially by distributing processes on different computer cores.

A custom optimization program is created to apply both parallel processing and simulation archiving. An efficient, multi-tier archive lookup method is implemented. As results from new trial injection scenarios are received, the archive dataset continually remains sorted from smallest to largest based upon the sum of each archived injection strategy's decision variable values. When the code is searching through the archive database to determine if an identical trial injection strategy had previously been evaluated, the code first looks for a match between the summations of decision variable values. If a match is found, the code checks to see if each decision variable is identical. The archive search process is stopped if either an identical set of decision variables is found or the remainder of archived injection scenarios have a larger decision variable value sum.

3 Characterization of the Michigan Basin test site

The Michigan Basin (MB) is near the town of Thompsonville in northwest Michigan. Michigan Technological University's data library provides detailed subsurface data in this region. A cross-sectional schematic of the MB is provided in Figure 3 along with a map of the study area, showing a nearly depleted hydrocarbon reservoir between depths of approximately 4660 and 5000 feet (1420-1520 meters) overlain by multiple confining and saline aquifer layers. This reservoir's only production well, Merit 1-20A, was originally drilled by the Shell Oil Company and is located between two exploration boreholes, Burch 1-20B and Stech 1-21A. This work explores the simulation and optimization of GCS into the saline Grey Niagaran formation immediately below the hydrocarbon reservoir.

The Burch 1-20B (Burch) and Stech 1-21A (Stech) boreholes have provided a wide variety of high resolution well logs that may be used to characterize subsurface domain properties such as density, porosity, electrical resistivity, and compressional and shear wave velocities. For this analysis, neutron porosity hydrogen index (NPHI) data were used, gathered from the boreholes to estimate aquifer and caprock locations, thicknesses, and permeabilities. High NPHI values indicate relatively high permeability while regions exhibiting low NPHI values indicate low permeability caprock layers. Figure 4 shows NPHI versus depth for both the Burch and Stech boreholes. Aquifer (highlighted light blue) and caprock (highlighted brown) layers are then defined using this data. Derived formation thicknesses and depths shown in Figure 4 correspond with aggregated layer sets displayed in Figure 3.

Five derived aquitards and four interlaying aquifer layers (including the Grey Niagaran formation) are used in the computational model ($L = 4$). Caprock thicknesses are estimated to be 16.8, 18.3, and 109.1 m for layers 2, 3, and 4, respectively, while layers 1 and 5 are assumed to be completely impermeable even

at passive well locations. Average aquifer permeability, k , in milliDarcys (mD; $1 \text{ mD} \cong 10^{-15} \text{ m}^2$) is estimated for all aquifer layers using the following from Trebin (1945):

$$k = \begin{cases} 2 \cdot e^{31.6 \cdot \phi} & \text{if } \phi < 0.124 \\ 4.94 \cdot 10^4 \cdot \phi^2 - 763 & \text{if } \phi \geq 0.124 \end{cases} \quad (30)$$

where, ϕ is assumed to equal NPHI. Table 1 shows the permeability, thickness, and porosity value for each aquifer layer obtained using the preceding methodology.

Table 1 Aquifer permeability, thickness, and porosity values used in this study

Parameter	Symbol	Value for Aquifer Layer l				Units
		$l = 1$	$l = 2$	$l = 3$	$l = 4$	
Permeability	k	24	4.5	6.7	655	mD
Thickness	H	119.5	35.1	36.6	75.3	m
Porosity	ϕ	0.079	0.026	0.038	0.169	/

The Michigan Department of Environmental Quality maintains a database of producing and inactive oil and gas wells in the state. Although 65,560 well records containing permit number, depth, and spatial coordinates are retrieved from the department's website (Michigan Department of Environmental Quality Oil and Gas Database 2014), only 131 wells are located within 4 km of the reservoir's production well, Merit 1-20A, and intersect the four aquifer layers defined above.

A GCS test site is envisioned herein using MB data characterized above. Optimal injection rates and well locations are determined for a maximum of 3 injection wells ($M_{max} = 3$). Injection wells are allowed to operate at a constant injection rate of either $Q_1 = Q_{min} = 0$, $Q_2 = 10$, $Q_3 = 20$, $Q_4 = 30$, or $Q_5 = Q_{max} = 40$ kg/s ($N_{ir} = 5$) over a 50-year duration ($t_{inj} = 50$ yrs.), thereby satisfying Eqn (26). Injection well locations

may be selected from 16 candidate locations uniformly distributed over a 2.25 km² square grid. Figure 5 provides a plan view of the MB site with candidate injection well locations ($N_{ci} = 16$) shown as orange circles and existing passive well locations ($N = 131$) shown as blue pluses. Horizontal positions are relative to the Merit 1-20 production well with $x_{min} = y_{min} = -750$ m and $x_{max} = y_{max} = 750$ m), thereby satisfying constraint Eqn (25).

In Eqn (21), $Mass_{leak}$ is defined as the cumulative mass of CO₂ that escapes into the top layer ($l = 4$) at the end of the 50-year injection duration. The radius, r_{pw} , used in Eqn (14) is assumed to be 0.2 m for all 131 passive wells. A fracture safety probability of $S_p = 95\%$ is assumed for the pressure constraint (Eqn (28)). Also a fracture gradient of 14.2 kPa/m (Oldenburg et al. 2011) is used for this analysis. Table 2 lists additional deterministic hydrogeological parameter values used herein.

Table 2 Deterministic hydrogeological parameter symbols and values

Parameter	Symbol	Value	Units
Brine density	ρ_b	1,045	kg/m ³
CO ₂ density	ρ_c	479	kg/m ³
Brine viscosity	μ_b	2.94×10^{-4}	Pa·s
CO ₂ viscosity	μ_c	3.95×10^{-5}	Pa·s
Brine residual saturation	S_{res}	0.3	/

The costs associated with installation, operation, and maintenance (Jim Lepinski, Headwaters Clean Carbon Services LLC, personal communication, 2012) used for each candidate well location in this study are shown in Table 3. This preliminary analysis does not include costs involved with site characterization, permitting, lease/purchase of land/pore space, financing, insurance, monitoring, verification, Environmental Protection Agency (EPA) financial bond requirement, post-injection site care, and long-term liabilities although these may be significant in the final financial assessment (Jim Lepinski,

Headwaters Clean Carbon Services LLC, personal communication, 2012). The leakage cost parameter c_L in Eqn (21) is assigned a values of 0.6 \$/kg.

Table 3 Cost parameter values for each candidate well location in Eqn (20)

Capital Cost, Cap (\$/well)	Fixed O&M Cost, OP (\$/day/well)	Surface Maintenance, SurM (\$/yr/well)	Subsurface Maintenance, SubM (\$/yr/well)	Variable Cost, Var (\$/kg of CO ₂)
3,537,104	11,566	120,608	37,612	0.009

As stated previously, aquifer permeability, effective compressibility, c_{eff} , and passive-well segment permeability are assumed to be uncertain variables. For each MC simulation, aquifer permeabilities and effective compressibilities are assigned a random value using log-normal distributions with a standard deviation of 0.4, mean aquifer permeabilities from Table 1, and mean c_{eff} of 4.6×10^{-10} m²/N (Celia et al. 2011). Passive-well segment permeabilities are randomly assigned as either 0.01 or 1000 mD, representing “intact” and “degraded” cement, respectively (Celia et al. 2011, Court et al. 2012, Nogues and Dobossy 2012). Table 4 shows the probability of each passive-well segment being assigned an "intact" or “degraded” permeability value during the MC ensemble generation for each uncertainty scenario. Herein, two uncertainty scenarios are tested and compared including 1) data supporting an abundance of intact passive-well segments (U1) and 2) data supporting an abundance of degraded passive-well segments (U2). Passive-well segment permeabilities are assumed to be fully uncorrelated (i.e. passive wells may have differing permeabilities at different depths).

Table 4 Probability of passive-well segments being assigned either an "intact" or “degraded” permeability value for each uncertainty scenario.

Uncertainty Scenario	Description	Probability of being assigned an "intact" permeability value	Probability of being assigned a "degraded" permeability value
U1	Intact passive-well data	90%	10%
U2	Degraded passive-well data	10%	90%

4 Results and Discussion

4.1 NSGA-II Parameter Calibration

The number of Monte Carlo simulations, N_{MC} , generated and simulated for each trial injection strategy is determined from an analysis of the uncertainty scenario U2. U2 is assumed to be the worst case (i.e. most difficult to solve) optimization problem because it has the largest probability of degraded passive wells and, therefore, the greatest potential for CO₂ leakage. Both $P_{50\%}(\mathbf{Cost})$ and $P_{95\%}(\mathbf{Cost})$ (Eqn (23)) are estimated for ten different injection strategies ($N_{sims} = 10$). For each of these simulations, the value of N_{MC} is ranged between 50 and 1000 in intervals of 50. An inspection of the results of this analysis shows that a N_{MC} value of 400 produces convergence for both $P_{50\%}(\mathbf{Cost})$ and $P_{95\%}(\mathbf{Cost})$.

Optimal NSGA-II parameter values of N_{pop} , M_{rate} , T_{size} , and ϵ are selected from a series of preliminary tests using a simplified deterministic optimization problem having 50% degraded and 50% intact passive-well segments. A deterministic problem is used because it requires much less computational time for each simulation as there is no need for MC analysis. First, a single measure is established to quantify the performance of each set of NSGA-II parameter values. The average project cost per unit mass sequestered (\$/kg) over all non-dominated solutions is assumed to represent the fitness of any given Pareto set. A lower value of this measure indicates less overall project cost and therefore a more “fit” solution. Next, the true optimal Pareto set is found for the test deterministic optimization problem by performing three exhaustive evolutionary searches, each with 5,000 generations. Each of these converged to the same average project cost per unit mass sequestered. Finally, the set of parameter values requiring the least number of simulation calls to reach within 0.1% of the minimal average project cost per unit mass sequestered found from exhaustive search are selected to be used by the stochastic optimization problem. Table 5 shows the NSGA-II parameter values and maximum number of optimization generations found using this analysis.

A performance comparison between the NSGA-II and a completely random search of the decision space is also performed to further validate the effectiveness of the NSGA-II used in this study. One hundred deterministic optimization trials having differing random number seeds are processed. The genetic algorithm is found to greatly outperform the random search algorithm. The percent of trials reaching 0.1% of the minimal project cost per unit mass sequestered in 200 generations is 100% for the NSGA-II compared to 0% for the random search algorithm. Results from this convergence test provide strong evidence that, for this problem, optimal or close-to-optimal Pareto sets are found using the algorithmic parameter set shown in Table 5.

Table 5 NSGA-II parameter values and maximum number of optimization generations used for the stochastic case study. All parameters are dimensionless.

N_{pop}	M_{rate}	T_{size}	ϵ	N_{gens}
25	1.6%	2	0.001	200

4.2 Stochastic Optimization Analysis

Three goals regarding the stochastic optimization of the MB site are investigated and discussed within this section: 1) quantify the impact of DM preferences on heuristically determined Pareto-optimal objective values (i.e. mass sequestered and project cost); 2) quantify the impact of DM preferences on carbon injection strategy (i.e. the selection of injection-well flow rate and location), and; 3) assess the suitability of the MB test site for GCS. To accomplish this, results from several optimization test cases having differing DM preference values, including the selection of the risk adversity factor, r_A , (Eqn (21)), stochastic non-exceedance cost probability, z , (Eqn (23)), and passive-well uncertainty scenario (Table 4), are compared and discussed. Eight stochastic optimization cases are performed for the MB test site data described in Section ‘*Characterization of the Michigan Basin test site*’. The risk adversity factor, r_A , is a sensitive exponential term and is set to either 1.0 or 1.1 representing the preference of either a “risk-neutral” or “risk-adverse” decision maker, respectively. Also, the stochastic non-exceedance cost

probability, z , is set to either 50% or 95% for each of the two possible passive-well uncertainty scenarios (U1, U2), hence obtaining a total of $2 \cdot 2 \cdot 2 = 8$ combinations of DM preference selections. Table 6 shows DM preferences for each of these stochastic optimization cases (C1-C8).

Table 6 Decision maker (DM) preferences for each of the eight stochastic optimization runs (C1-C8)

Stochastic Optimization Case	Risk Adversity Factor, r_A	Passive-Well Uncertainty Scenario	Stochastic non-exceedance cost probability, z (%)
C1	1.0	U1	50
C2	1.0		95
C3	1.0	U2	50
C4	1.0		95
C5	1.1	U1	50
C6	1.1		95
C7	1.1	U2	50
C8	1.1		95%

4.2.1 Impact of DM Preferences on Objective Function Values

The first step in investigating relationships between DM preferences and objective function values is to determine Pareto-optimal (or close-to-optimal) tradeoff sets for each stochastic optimization case described in Table 6 using the methodology presented in Section ‘*Methodology*’. The quantity of total mass sequestered (Eqn (19)) is a function of the injection duration, t_{inj} , the maximum number of injection wells, M_{max} , as well as the number, N_{ir} , and flow rates, \mathbf{Q} , of prescribed injection rates. Because values for these three variables are prescribed before the start of optimization, all possible values of the total mass sequestered objective function are known *a priori*. Recall from Section ‘*Characterization of the Michigan Basin test site*’ that each injection well may only assume an injection rate of either 0, 10, 20, 30, or 40 kg/s. Because of this, there are only 12 possible non-zero, discrete values of total mass sequestered when $M_{max} = 3$, as the total injection rate for all three wells may range between 10 and 120 kg/s in increments of 10 kg/s.

A visual representation of resulting Pareto sets is provided in Figure 6, where project cost versus mass of CO₂ sequestered is displayed for each of the 8 optimization cases described in Table 6. Figures 6a-d plot the objective function tradeoff curve associated with each uncertainty scenario (i.e. U1 and U2) for given values of r_A and z . Capital, operation and maintenance (CO&M) costs, defined herein as all project costs other than the penalties incurred from CO₂ leakage, are also plotted. Policies found to violate the fracture constraints (Eqn (28)) are infeasible and are therefore not shown in Figure 6. Note that, due to higher project costs, a larger scale is used on the ordinate axis when $r_A = 1.1$.

A general trend exhibited by these tradeoff profiles shows that project cost increases with the mass of CO₂ sequestered in all optimization runs. A sharper increase in capital project cost is observed in CO&M costs when $r_A = 1.0$ (Figure 6a-b) because, due to the maximum prescribed injection rate (Eqn (26)), an additional injection well is needed when sequestering more than 63.1 Mt of CO₂. While increases in injection well capital costs are also incurred when $r_A = 1.1$ (Figure 6c-d), they are overshadowed by much larger total project costs, as increasing r_A drastically increase leakage penalty costs.

In this work, DM risk adversity preference, r_A , is quantified through an exponential term when estimating leakage cost (Eqn (21)) and therefore does not affect the estimated mass of CO₂ leakage. However, as seen in optimization cases C7 and C8, when the estimated mass of CO₂ leakage is high, r_A is found to have a profound effect on project cost. While all optimization cases under uncertainty scenario U1 (90% probability of any passive well being assigned as intact) exhibit very little CO₂ leakage, substantial CO₂ leakage masses are estimated for test cases assigned a U2 (10% probability of any passive well being assigned as intact) uncertainty scenario, resulting in large leakage costs when $r_A = 1.1$.

The selection of uncertainty scenario (Table 4) represents a DM's knowledge of the GCS site's caprock integrity. Uncertainty scenario selection directly impacts the estimated mass of CO₂ leakage and, therefore, indirectly affects the estimated cost associated with CO₂ leakage. Cases assigned uncertainty scenarios having greater percentages of intact well segments are found to exhibit less CO₂ leakage. All optimization cases assigned a U1 uncertainty scenario exhibit fairly negligible CO₂ leakage costs as their minimized total project costs are very similar to CO&M costs. There is, however, one exception to this observation: the fracture constraint (Eqn (28)) does not allow for any well to operate at 40 kg/s and therefore forces the installation of additional injection wells when sequestering more than 47.3 Mt of CO₂

Finally, the choice of stochastic non-exceedance cost probability, z , affects the estimated cost associated with CO₂ leakage because greater project costs are required to increase non-exceedance probability. Data shown in Figure 6 suggest that the value chosen for z has a minor impact on resulting Pareto-optimal objective function values, thus indicating that spread of each **Cost** CDF is relatively contained. Estimated project costs increase when a greater value of z is used. However, the project cost variability associated with stochastic non-exceedance cost probability is much smaller than the project cost variability associated with risk adversity or uncertainty scenario selection.

4.2.2 Impact of DM Preferences on Injection Strategy Selection

The impact of DM preferences upon the heuristic selection of optimal injection locations and flow rates is also investigated. Figure 7 displays a close-up, plan view of the MB test site showing the 16 candidate injection-well location indices.

As explained in Section '*Impact of DM Preferences on Objective Function Values*', there are 12 possible mass sequestered values for each of the 8 optimization cases. Hence, for this problem there is a maximum of $12 \cdot 8 = 96$ Pareto optimal injection strategies which may be found using the NSGA-II.

However several of these solutions are rendered infeasible by the fracture constraints (Eqn (28)) and therefore not displayed in the following discussion. As a general presentation of decision variable results, Table 7 shows both candidate well location indices (see Figure 7) and injection rates in kg/s for each injection strategy. Note that injection strategies may use multiple injection wells.

Table 7 Candidate well location indices from Figure 7 and injection rates for each injection, displayed as Location(Injection Rate). Policies indicated as "-" are infeasible due to a violation of the fracture constraints (Eqn (28)).

Stochastic Optimization Run Index	Mass Sequestered (MT)						
	15.8	31.5	47.3	63.1	78.8	94.6	110.4
C1	1(10)	1(20)	1(30)	9(20), 13(20)	1(30), 13(20)	-	-
C2	4(10)	1(20)	5(30)	1(20), 5(20)	1(20), 13(30)	-	-
C3	1(10)	1(20)	1(30)	1(20), 2(20)	1(30), 2(20)	1(30), 5(30)	1(10), 5(30), 13(30)
C4	1(10)	1(20)	1(30)	1(20), 2(20)	1(30), 2(20)	1(30), 5(30)	1(20), 2(30), 13(20)
C5	1(10)	1(20)	1(30)	1(20), 2(20)	3(20), 9(30)	-	-
C6	4(10)	1(20)	5(30)	1(20), 5(20)	1(20), 13(30)	-	-
C7	1(10)	1(10), 5(10)	1(10), 2(10), 5(10)	1(20), 2(10), 5(10)	1(20), 2(20), 5(10)	1(30), 2(20), 5(10)	2(30), 3(30), 13(10)
C8	1(10)	1(10), 5(10)	1(10), 2(10), 5(10)	1(20), 2(10), 5(10)	1(20), 2(10), 5(20)	1(30), 3(10), 6(20)	2(30), 5(20), 13(20)

Using the data presented in Table 7, two quantitative analyses are performed to study how DM preferences ultimately influence the heuristic selection of carbon injection strategies. First, the relative sensitivity of carbon-injection-strategy selection in relation to each DM parameter (i.e. r_A , z , and uncertainty scenario) is quantified as the percentage of injection-well rate/location combinations that change when varying each DM preference. These percentages are calculated as the sum of differing

injection-well rate/location combinations divided by the total number of injection-well rate/location combinations in each comparison set. Optimization cases having all but one identical DM preferences are individually compared. When contrasting r_A ($r_A=1.0$ with $r_A =1.1$) a set of four injection strategy comparisons are made; stochastic optimization cases C1-C4 are thus compared with cases C5-C8, respectively. Four injection strategy comparisons are also made when contrasting z ($z=50\%$ with $z=95\%$); case C1 with case C2, case C3 with case C4, case C5 with case C6, and case C7 with case C8. Finally, a set of four injection strategy comparisons are made when contrasting uncertainty scenario (U1 with U2); cases C1-C2 are compared with cases C3-C4, respectively, and cases C5-C6 are compared with cases C7-C8, respectively. The percentage of injection strategies that remain constant when varying values of r_A , uncertainty scenario, and values of z is quantified as 49.3%, 37.5%, and 58.3%, respectively. These findings are used to augment the following discussion.

Secondly, a categorical distribution analysis is used to identify general injection strategy trends associated with DM preferences. The number of times each candidate location is selected for injection-well placement is counted for all cases having each given DM preference value. For example, candidate location 1 is found to be selected for well placement 21 times when $r_A=1.0$ (i.e. for cases C1-C4), 14 times when assuming uncertainty scenario U1 (i.e. for cases C1, C2, C5, and C6) and 21 times when $z=50\%$ (i.e. for cases C1, C3, C5, C7). Table 8 provides the number of selections of each candidate injection location for each DM preference value.

Table 8 Number selections for each candidate well location index. Indices having zero selections are not shown.

Candidate Location	Risk Adversity Factor, r_A		Passive-Well Uncertainty Scenario		Stochastic non-exceedance cost probability, z	
	1.0	1.1	U1	U2	50%	95%
1	21	19	14	26	21	19
2	5	10	1	14	8	7
3	0	3	1	2	2	1
4	1	1	2	0	0	2

5	5	12	4	13	7	10
6	0	1	0	1	0	1
9	1	1	2	0	2	0
13	5	3	4	4	4	4
15	0	0	0	0	0	0
16	0	0	0	0	0	0
Total	38	50	28	60	44	44

From the results shown in Table 8, the southwest corner of the candidate injection-well field is heavily favored by the optimization algorithm, regardless of parameter choice, with 81.8% of all injection-well placements being made at either candidate location 1, 2, or 5. This is due to the presence of a passive-well cluster approximately 1000 meters northeast of the candidate injection-well field (see Figure 5). Also, the furthest southwest candidate injection-well location is found to have a substantially greater number of selections than all other individual locations. Approximately 45.5% of all injection-well placements are being made at candidate location 1, compared with 17.0% and 19.3% for candidate locations 2 and 5, respectively.

DM risk adversity is found to have a great effect on injection strategy selection as only 49.3% of injection strategy selections remain constant when varying values of r_A (see second and third columns of Table 8). Also, increasing r_A from 1.0 to 1.1 is found to increase the total number of candidate well location selections from 38 to 50. This increased number in total candidate well location selections is caused by the optimization algorithm attempting to reduce CO₂ leakage cost by using additional injection wells to spread out and reduce the injection induced pressure distribution. For example, a total injection rate of 20 kg/s may be achieved by one well injecting at 20 kg/s or by two each injecting at 10 kg/s. This finding suggests that the leakage penalty savings from diversifying the injection-well field are greater in certain cases than the additional CO&M costs incurred from installing, operating, and maintaining more injection wells.

Passive-well uncertainty scenario selection is found to have the lowest percentage (37.5%) of injection strategy selections remaining constant, and therefore the greatest effect upon injection strategy selection (see fourth and fifth columns of Table 8). The total number of candidate well location selections is also found to increase from 28 for uncertainty scenario U1 to 60 for the more expensive uncertainty scenario U2, further validating the trend found when studying risk adversity. Also, greater estimated CO₂ leakage, as in the case of uncertainty scenario U2, is clearly observed to drive candidate injection-well location selections further southwest. The likelihood of selecting the three furthest southwest candidate locations (i.e. locational indices 1, 2, and 5) increases from 67.9% in cases assigned uncertainty scenario U1 to 88.3% in cases uncertainty scenario U2.

Finally, the stochastic non-exceedance cost probability, z , is found to have the highest percentage (58.3%) of injection strategy selections remaining constant and, therefore, has the least effect on injection strategy selection. In addition, there are very similar results shown in the sixth and seventh columns of Table 8 with the total number of candidate location selections being identical for both columns.

4.2.3 GCS Suitability Assessment for the MB test site

The final decision of whether or not to proceed with GCS project planning is made by the DM based on the assessment of a large number of political and financial indicators. However, the preliminary stochastic cost assessment presented above suggests that GCS feasibility at the MB test site would be highly dependent upon the DM's risk adversity preference. Figure 6 shows that each uncertainty scenario produces feasible project cost results if the DM selects $r_A = 1.0$. If the DM decides to select $r_A = 1.1$, both uncertainty scenarios provide much greater project cost results. In practice, the selection of a high risk adversity value effectively eliminates scenarios exhibiting a moderate to high mass of CO₂ leakage due to high CO₂ leakage costs resulting from the exponential leakage cost term, r_A . Economic drivers (i.e. carbon tax) will largely influence the decision of the quantity of CO₂ injected versus leakage cost incurred.

4.3 Computational Efficiency

It is interesting to note that the complete enumeration of this problem would require 34,800,000 simulation calls:

$$N_{MC} \cdot N_{ir}^{M_{max}} \cdot \sum_{i=1}^{M_{max}} \binom{N_{cl}}{i} = 400 \cdot 5^3 \cdot \left[\frac{16!}{1! \cdot 15!} + \frac{16!}{2! \cdot 14!} + \frac{16!}{3! \cdot 13!} \right] = 34,800,000 \quad (31)$$

Assuming that a numerical model would require one hour per simulation, the CPU time required to sequentially process 34,800,000 CO₂ leakage evaluations without archiving would be approximately 3,973 years. Therefore, a semi-analytical algorithm and an NSGA-II optimization approach is used to make this problem computationally feasible. For the NSGA-II optimization parameters provided in Table 5, each optimization run requires 2,000,000 simulation model calls, N_{ct} , to estimate CO₂ leakage without archiving (e.g. $N_{ct} = N_{MC} \cdot N_{pop} \cdot N_{gens} = 400 \cdot 25 \cdot 200 = 2,000,000$). Assuming that each semi-analytical simulation requires 1 second, the CPU time required to sequentially process 2 million CO₂ leakage evaluations without archiving would be 23.2 days. The computational time required for this problem may also be reduced using parallel processing and archiving. For example, if 25 computer processor cores are available, setting the number of parallel processes to N_{pop} will reduce this theoretical simulation evaluation time by 96% to 0.93 days. The actual CPU time required for a single optimization run with $N_{MC} = 400$, $N_{pop} = 25$, and $N_{gens} = 200$ using 12 processor cores and employing both parallel processing and simulation archiving is approximately 1.04 days, or about six orders of magnitude less than the theoretical time required for the complete enumeration of this problem using a numerical model. These findings show that the use of the optimization algorithm is essential in providing the computational efficiency required when deriving the Pareto optimal set, particularly if the number of candidate locations, M , is increased.

5 Conclusions

A stochastic methodology for determining optimal GCS injection strategies has been presented, where a semi-analytical CO₂ leakage algorithm and a Monte Carlo procedure were integrated into a NSGA-II with ϵ -dominance. In an effort to show the applicability of this method to real world potential injection sites, the stochastic optimization framework has been used to assess a hypothetical GCS project at a Michigan Basin (MB) test site in northern Michigan, USA. Eight MB test site stochastic optimization cases having differing DM preferences were evaluated.

DM risk adversity preference was found to have a profound effect on project cost when the estimated mass of CO₂ leakage was high. While all optimization cases that were assigned an “intact” uncertainty scenario exhibited very little CO₂ leakage, substantial CO₂ leakage masses were estimated for test cases assigned a “degraded” uncertainty scenario, resulting in large leakage costs with high risk adversity. The choice of stochastic non-exceedance cost probability had only a minor impact on resulting Pareto-optimal objective function values.

The southwest corner of the candidate injection-well field was heavily favored by the optimization algorithm, regardless of parameter choice, with 81.8% of all injection-well placements being made at the three furthest southwest candidate locations. Passive-well uncertainty scenario selection was found to have the greatest effect on injection strategy selection while stochastic non-exceedance cost probability was shown to have the least effect on injection strategy selection. In addition, DM risk adversity was found to significantly affect injection strategy selection. This finding suggests that the leakage penalty savings from diversifying the injection-well field were greater in certain cases than the additional project costs incurred from installing, operating, and maintaining more injection wells.

This work also discussed large gains in computational efficiency using semi-analytical modeling, NSGA-II optimization, parallel computing, and simulation archiving. The actual CPU time required for a typical optimization run using 12 processor cores and employing both parallel processing and simulation archiving was approximately 1.04 days, or about six orders of magnitude less than the theoretical time required for the complete enumeration of this problem using a numerical model.

Because of the large set of assumptions made by the semi-analytical CO₂ leakage algorithm, this framework may only be used for initial site planning and characterization. After ‘coarse scale’ project planning has been completed using this stochastic optimization framework, more rigorous, although slower, numerical models should be used for final project development of individual potential injection sites. However, this tool has potential for initial carbon-sequestration project planning as well as initial screening and ranking of large sets of potential carbon sequestration sites.

Looking forward, there are several possible variations of the optimization framework which may be explored. For example, the injection duration may be included as a third decision variable in addition to location and flow rate of each injection well. In addition, the minimization of risk of cost exceedance may be included as a third objective function in addition to maximizing the mass of CO₂ sequestered and minimizing the project cost. A multi-criteria decision analysis (MCDA) may also be performed upon the resulting Pareto-optimal sets of equations to quantify the importance of conflicting objectives and aid in the final selection of injection strategy selection. Finally, substantial gains in computational efficiency may be obtained by additional parallelization using large processor core clusters. Increased computational efficiency will lead to the ability to increase the number of model calls per optimization run or to process greater quantities of potential injection sites.

References

- Alzraiee, A.H., Bau, D.A., Garcia, L.A. (2013) Multiobjective design of aquifer monitoring networks for optimal spatial prediction and geostatistical parameter estimation. *Water Resour. Res.* 49, 3670–3684
- Baú, D.A. (2012) Planning of groundwater supply systems subject to uncertainty using stochastic flow reduced models and multi-objective evolutionary optimization. *Water Resour. Manag.* 26, 2513–2536
- Bielicki, J.M., Pollak, M.F., Fitts, J.P., Peters, C.A. (2014) Causes and financial consequences of geological CO₂ storage reservoir leakage and interference with other subsurface resources. *Int. J. Greenh. Gas Control.* 20, 272–284
- Bielicki, J.M., Pollak, M.F., Wilson, E.J., Fitts, J.P., Peters, C.A. (2013) A Method for monetizing basin-scale leakage risk and stakeholder impacts. *Energy Procedia.* 37, 4665–4672
- Birkholzer, J.T., Cihan, A., Zhou, Q. (2012) Impact-driven pressure management via targeted brine extraction - Conceptual studies of CO₂ storage in saline formations. *Int. J. Greenh. Gas Control.* 7, 168–180
- Bossie-Codreanu, D., Le Gallo, Y. (2004) A simulation method for the rapid screening of potential depleted oil reservoirs for CO₂ sequestration. *Energy.* 29, 1347–1359
- Box, G., Draper, N. (2007) Response surfaces, mixtures, and ridge analyses, *Second Edition*. Wiley.
- Buscheck, T.A., Sun, Y., Chen, M., Hao, Y., Wolery, T.J., Bourcier, W.L., Court, B., Celia, M.A., Julio Friedmann, S., Aines, R.D. (2012) Active CO₂ reservoir management for carbon storage: Analysis of operational strategies to relieve pressure buildup and improve injectivity. *Int. J. Greenh. Gas Control.* 6, 230–245
- Celia, M.A., Nordbotten, J.M., Bachu, S., Dobossy, M., Court, B. (2009) Risk of leakage versus depth of injection in geological Storage. *Energy Procedia.* 1, 2573–2580
- Celia, M.A., Nordbotten, J.M., Court, B., Dobossy, M., Bachu, S. (2011) Field-scale application of a semi-analytical model for estimation of CO₂ and brine leakage along old wells. *Int. J. Greenh. Gas Control.* 5, 257–269
- Chen, L., McPhee, J., Yeh, W.W.-G. (2007) A diversified multiobjective GA for optimizing reservoir rule curves. *Adv. Water Resour.* 30, 1082–1093
- Cheng, Y., Jin, Y., Hu, J. (2009) Adaptive epsilon non-dominated sorting multi-objective evolutionary optimization and its application in shortest path problem. *ICCAS-SICE, 2009.* 2545–2549
- Cihan, A., Birkholzer, J.T., Zhou, Q. (2013) Pressure buildup and brine migration during CO₂ storage in multilayered aquifers. *Ground Water.* 51, 252–67
- Court, B., Bandilla, K.W., Celia, M.A., Janzen, A., Dobossy, M., Nordbotten, J.M. (2012) Applicability of vertical-equilibrium and sharp-interface assumptions in CO₂ sequestration modeling. *Int. J. Greenh. Gas Control.* 10, 134–147
- Crow, W., Brian Williams, D., William Carey, J., Celia, M., Gasda, S. (2009) Wellbore integrity analysis of a natural CO₂ producer. *Energy Procedia.* 1, 3561–3569
- Deb, K., Agrawal, S., Meyarivan T. (2002) A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *Evolutionary Computation* 6(2): 182-197.

- Doster, F., Nordbotten, J.M., Celia, M.A. (2013) Impact of capillary hysteresis and trapping on vertically integrated models for CO₂ storage. *Adv. Water Resour.* 62, 465–474
- Eccles, J.K., Pratson, L., Newell, R.G., Jackson, R.B. (2012) The impact of geologic variability on capacity and cost estimates for storing CO₂ in deep-saline aquifers. *Energy Econ.* 34, 1569–1579
- Espinete, A.J., Shoemaker, C.A. (2013) Comparison of optimization algorithms for parameter estimation of multi-phase flow models with application to geological carbon sequestration. *Adv. Water Resour.* 54, 133–148
- Gasda, S.E., Nordbotten, J.M., Celia, M.A. (2009) Vertical equilibrium with sub-scale analytical methods for geological CO₂ sequestration. *Comput. Geosci.* 13, 469–481
- Gasda, S.E., Nordbotten, J.M., Celia, M.A. (2011) The impact of local-scale processes on large-scale CO₂ migration and immobilization. *Energy Procedia.* 4, 3896–3903
- Gasda, S.E., Nordbotten, J.M., Celia, M.A. (2012) Application of simplified models to CO₂ migration and immobilization in large-scale geological systems. *Int. J. Greenh. Gas Control.* 9, 72–84
- Goda, T., Sato, K. (2011) Optimization of well placement for geological sequestration of carbon dioxide using adaptive evolutionary Monte Carlo algorithm. *Energy Procedia.* 4, 4275–4282
- Hahn, G.J., and Shapiro, S.S. (1967) *Statistical models in engineering.* John Wiley, New York
- Hansen, A.K., Hendricks Franssen, H.-J., Bauer-Gottwein, P., Madsen, H., Rosbjerg, D., Kaiser, H.P. (2013) Well field management using multi-objective optimization. *Water Resour. Manag.* 27, 629–648
- Heße, F., Prykhodko, V., Attinger, S. (2013) Assessing the validity of a lower-dimensional representation of fractures for numerical and analytical investigations. *Adv. Water Resour.* 56, 35–48
- Huang, X., Bandilla, K.W., Celia, M.A., Bachu, S. (2014) Basin-scale modeling of CO₂ storage using models of varying complexity. *Int. J. Greenh. Gas Control.* 20, 73–86
- Juanes, R., MacMinn, C.W., Szulczewski, M.L. (2009) The footprint of the CO₂ plume during carbon dioxide storage in saline aquifers: storage efficiency for capillary trapping at the basin scale. *Transp. Porous Media.* 82, 19–30
- Kumphon, B. (2013) Genetic Algorithms for Multi-objective Optimization: Application to a multi-reservoir system in the Chi River Basin, Thailand. *Water Resour. Manag.* 27, 4369–4378
- Laumanns, M., Thiele, L., Deb, K., Zitzler, E. (2002) Combining convergence and diversity in evolutionary multiobjective optimization. *Evol. Comput.* 10, 263–82
- Mantoglou, A., Kourakos, G. (2006) Optimal groundwater remediation under uncertainty using multi-objective optimization. *Water Resour. Manag.* 21, 835–847
- The MathWorks, Inc. (2012) *MATLAB and Statistics Toolbox Release 2012b.* Natick, Massachusetts, United States
- McPhee, J., Yeh, W.W.-G. (2004) Multiobjective optimization for sustainable groundwater management in semi-arid regions. *J. Water Resour. Plan. Manag.* 130, 490–497

Michigan Department of Environmental Quality Oil and Gas Database (2014) http://www.michigan.gov/deq/0,4561,7-135-6132_6828-98518--,00.html Accessed 08 March 2014

Nicklow, J., Reed, P., Savic, D. (2010) State of the art for genetic algorithms and beyond in water resources planning and management. *J. Water Resour. Plan. Manag.* 412–432

Nogues, J., Dobossy, M. (2012) A methodology to estimate maximum probable leakage along old wells in a geological sequestration operation. *Int. J. Greenh. Gas Control.* 7, 39–47

Nordbotten, J., Celia M. (2012) *Geological storage of CO₂*. Wiley

Nordbotten, J., Celia, M. (2005) Semi-analytical solution for CO₂ leakage through an abandoned well. *Environ. Sci. Technol.* 39, 602–611

Nordbotten, J., Celia, M., Bachu, S. (2004) Analytical solutions for leakage rates through abandoned wells. *Water Resour. Res.* 40, 1–10

Nordbotten, J., Flemisch, B. (2012) Uncertainties in practical simulation of CO₂ storage. *Int. J. Greenh. Gas Control.* 9, 234–242

Nordbotten, J., Kavetski, D., Celia, M., Bachu, S. (2009) Model for CO₂ leakage including multiple geological layers and multiple leaky wells. *Environ. Sci. Technol.* 43, 743–749

Oladyshkin, S., Class, H., Helmig, R., Nowak, W. (2011) An integrative approach to robust design and probabilistic risk assessment for CO₂ storage in geological formations. *Comput. Geosci.* 15, 565–577

Oldenburg, C.M., Jordan, P.D., Nicot, J.-P., Mazzoldi, A., Gupta, A.K., Bryant, S.L. (2011) Leakage risk assessment of the In Salah CO₂ storage project: Applying the certification framework in a dynamic context. *Energy Procedia.* 4, 4154–4161

Pacala, S., Socolow, R. (2004) Stabilization wedges: solving the climate problem for the next 50 years with current technologies. *Science.* 305, 968–972

Peralta, R.C., Forghani, A., Fayad, H. (2014) Multiobjective genetic algorithm conjunctive use optimization for production, cost, and energy with dynamic return flow. *J. Hydrol.*

Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B. (2013) Evolutionary multiobjective optimization in water resources: The past, present, and future. *Adv. Water Resour.* 51, 438–456

Reed, P.M., Kollat, J.B. (2013) Visual analytics clarify the scalability and effectiveness of massively parallel many-objective optimization: A groundwater monitoring design example. *Adv. Water Resour.* 56, 1–13

Singh, A. (2013) Simulation and optimization modeling for the management of groundwater resources. I: Distinct applications. *J. Irrig. Drain. Eng.* 1–10

Singh, A. (2014) Simulation and Optimization Modeling for the Management of Groundwater Resources. II: Combined Applications. *J. Irrig. Drain. Eng.* 1–9

Singh, T.S., Chakrabarty, D. (2010) Multi-objective optimization for optimal groundwater remediation design and management systems. *Geosci. J.* 14, 87–97

- Tabari, M.M.R., Soltani, J. (2012) Multi-objective optimal model for conjunctive use management using SGAs and NSGA-II models. *Water Resour. Manag.* 27, 37–53
- Takahashi W. (2000) *Nonlinear functional analysis: Fixed point theory and its applications*. Yokohama Publishers, Yokohama
- Trebin, F.A. (1945) *Oil permeability of sandstone reservoirs*. Gostoptekhizdat, Moscow
- Turpening, R.M., Toksöz, M.N., Born, A.E., et al. (1992) *Reservoir delineation consortium annual report*. Massachusetts Institute of Technology, Cambridge
- Viswanathan, H.S., Pawar, R.J., Stauffer, P.H., Kaszuba, J.P., Carey, J.W., Olsen, S.C., Keating, G.N., Kavetski, D., Guthrie, G.D. (2008) Development of a hybrid process and system model for the assessment of wellbore leakage at a geologic CO₂ sequestration site. *Environ. Sci. Technol.*, 2008, 42 (19), 7280-7286
- Wagner, B., Gorelick, S. (1987) Optimal groundwater quality management under parameter uncertainty. *Water Resour. Res.* 23, 1162–1174
- Walter, L., Binning, P., Oladyskin, S. (2012) Brine migration resulting from CO₂ injection into saline aquifers—An approach to risk estimation including various levels of uncertainty. *Int. J. Greenh. Gas Control.* 9, 495–506
- Watson, T., Bachu, S. (2008) Identification of wells with high CO₂ leakage potential in mature oil fields developed for CO₂-enhanced oil recovery. *SPE Symp. Improv. Oil Recover*
- Watson, T., Bachu, S. (2009) Evaluation of the potential for gas and CO₂ leakage along wellbores. *SPE Drill. Complet.*
- Wriedt, J., Deo, M., Han, W.S., Lepinski, J. (2014) A methodology for quantifying risk and likelihood of failure for carbon dioxide injection into deep saline reservoirs. *Int. J. Greenh. Gas Control.* 20, 196–211
- Zheng, F., Zecchin, A. (2014) An efficient decomposition and dual-stage multi-objective optimization method for water distribution systems with multiple supply sources. *Environ. Model. Softw.* 55, 143–155

FIGURES:

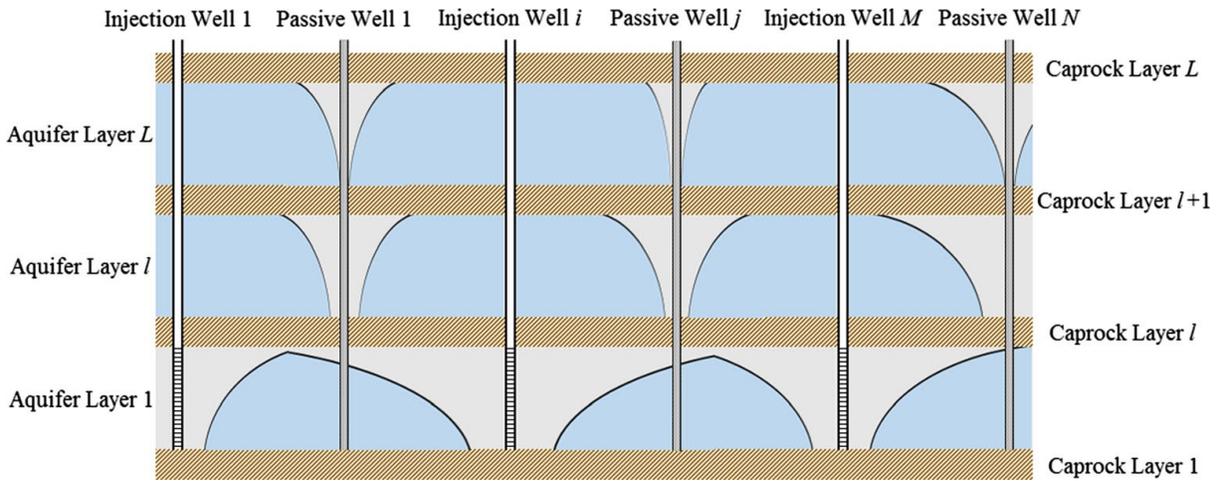


Fig. 1 Schematic of the semi-analytical leakage model's computational domain

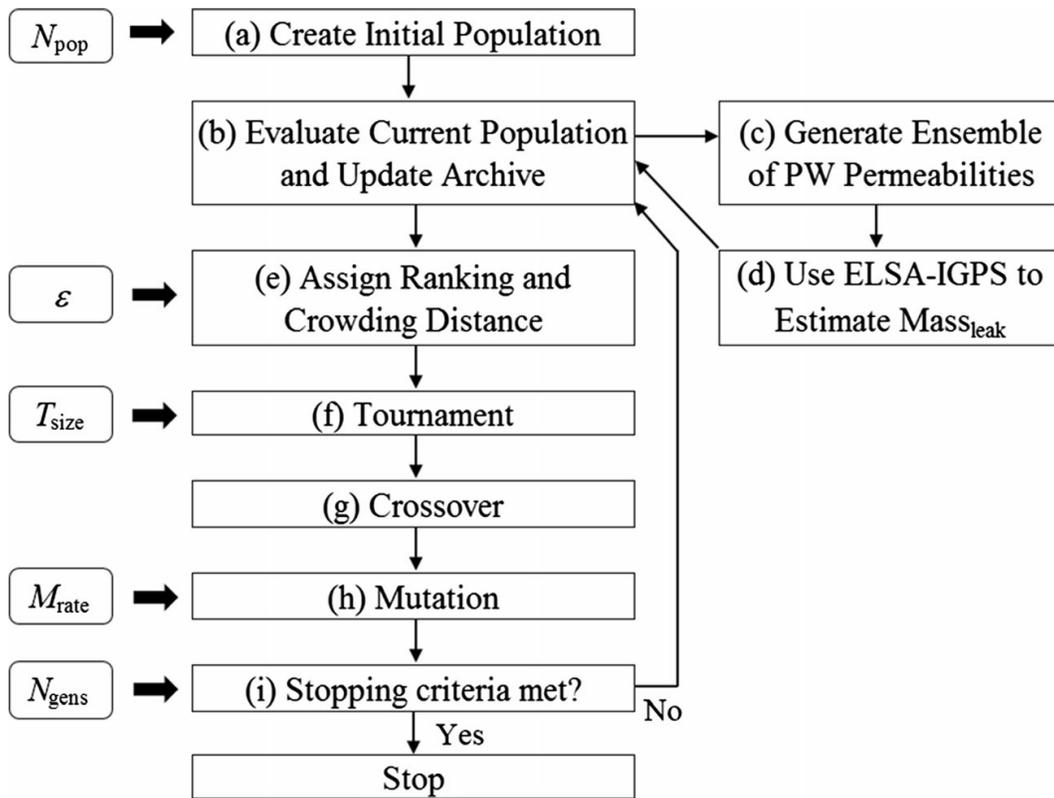


Fig. 2 Schematic of the stochastic optimization algorithm



(b)

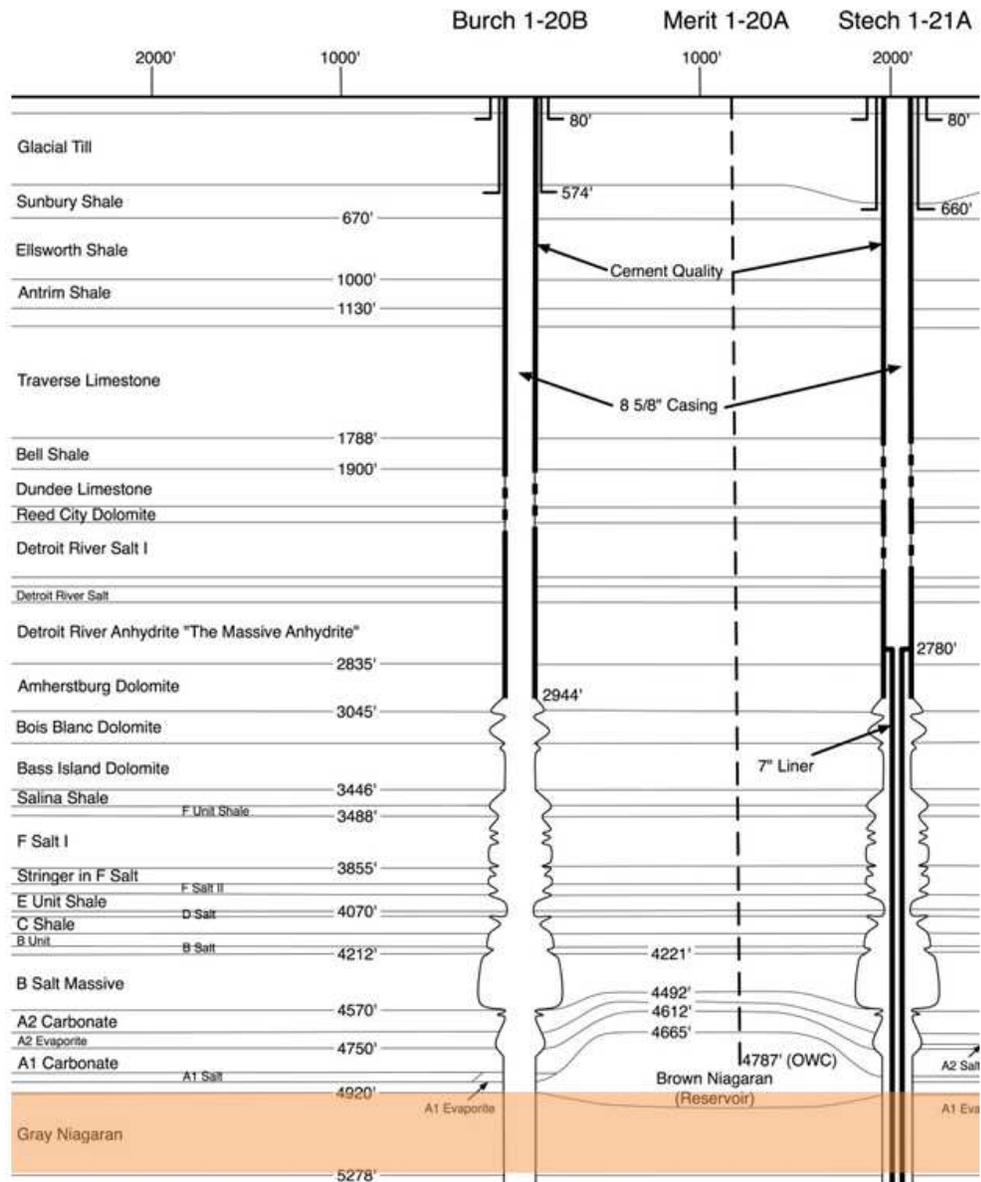


Fig. 3 a Map of study area and b a cross-sectional schematic of the Michigan Basin with the simulated injection layer highlighted orange (modified with permission from Turpening et al. (1992))

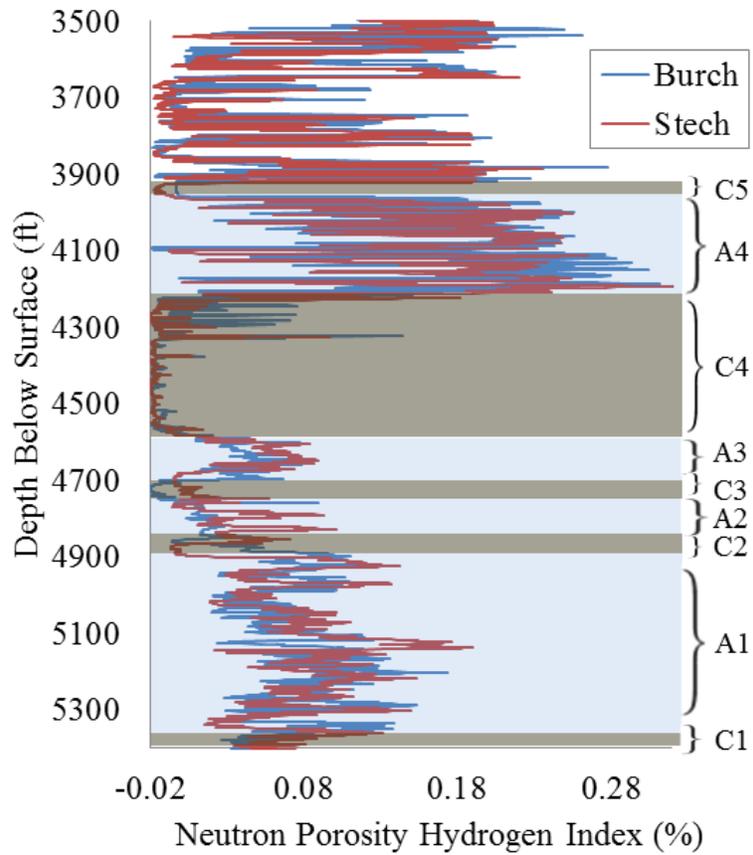


Fig. 4 Neutron porosity hydrogen index (NPHI) vs. depth for the Burch 1-20B (Burch) and Stech 1-21A (Stech) boreholes. Light blue areas are defined as aquifers (A1 – A4) while brown areas are defined as confining units (C1 – C5).

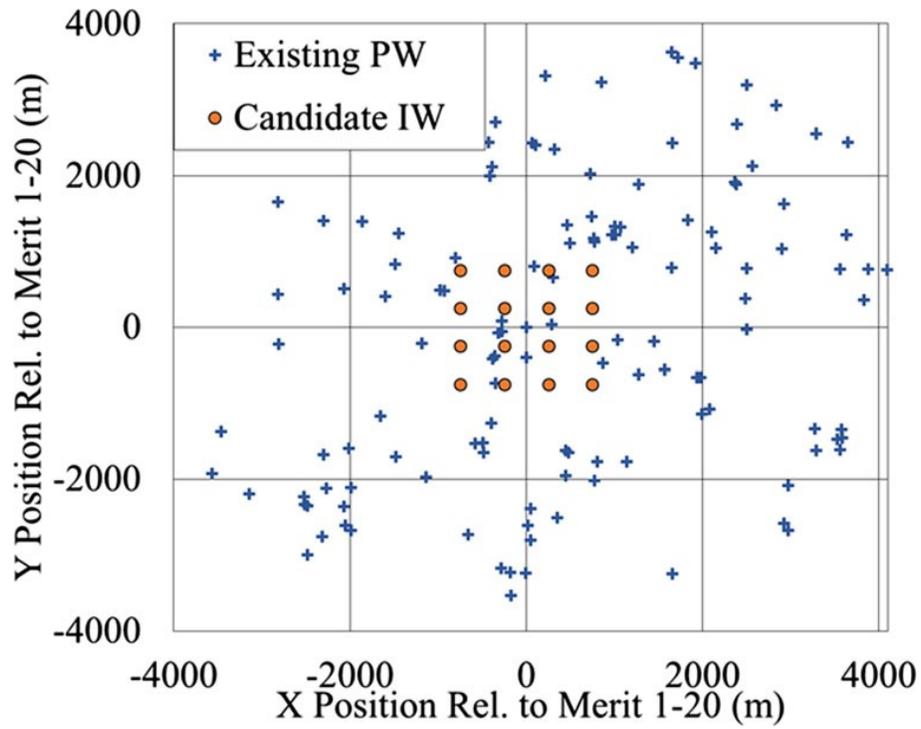


Fig. 5 Plan view of the Michigan Basin test site showing all pumping wells (PW) and candidate injection wells (IW) included in the study

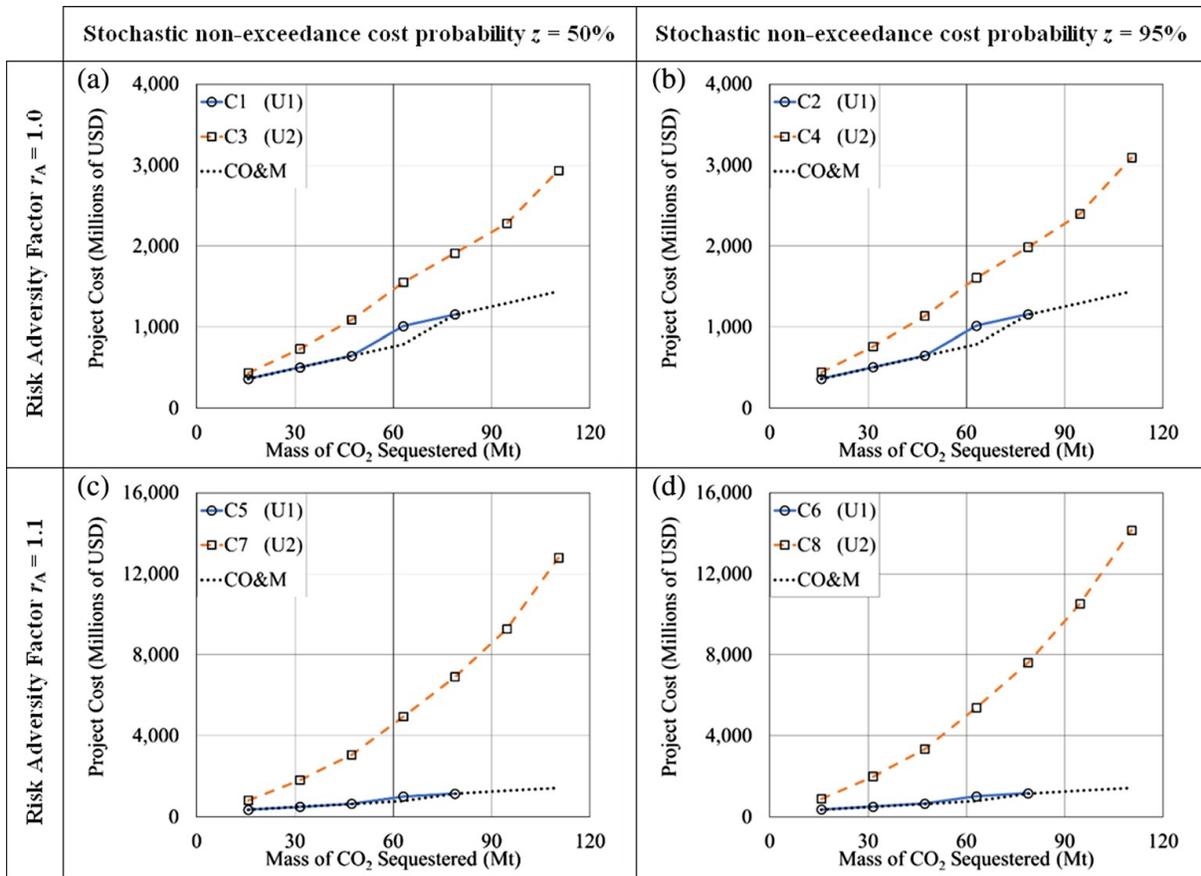


Fig. 6a-d Optimal project costs versus mass of CO₂ sequestered for the eight optimization runs described in Table 6. Policies found to violate the fracture constraints (Eqn (31)) are infeasible and are therefore not shown.

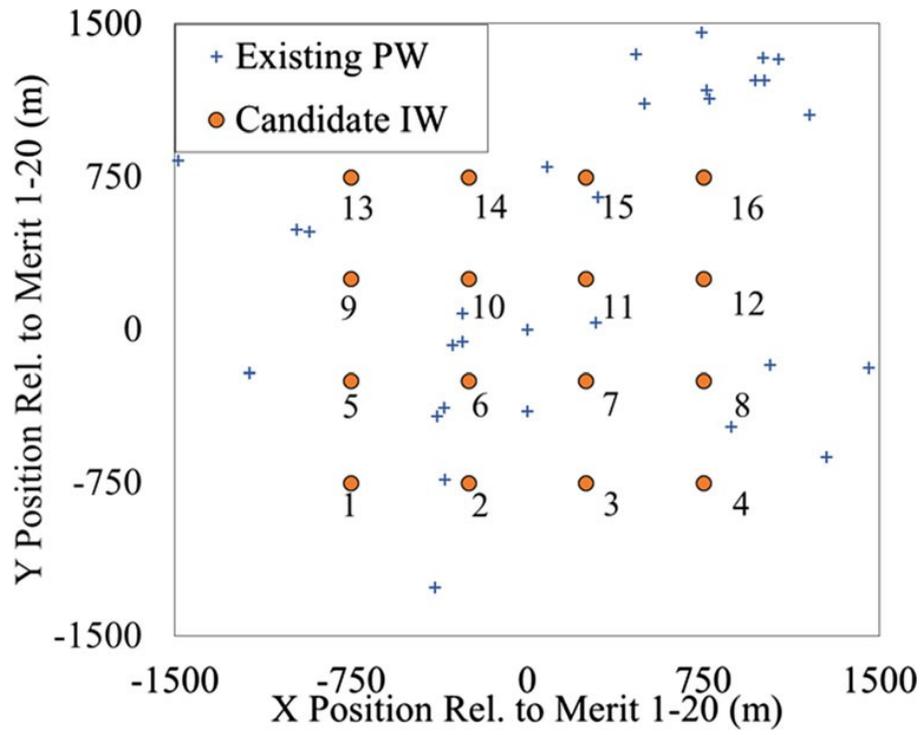


Fig. 7 Close-up plan view of the Michigan Basin test site highlighting candidate injection-well (IW) location indices.