Heuristic Algorithms for Design of Integrated Monitoring of Geologic Carbon Storage Sites

Alexander C. Hanna[a], Jonathan Whiting[a], Brian Huang[a], Delphine Appriou[a], Xianjin Yang[b], Julia de Toledo Camargo[a], Seunghwan Baek[a], Diana Bacon[a], Catherine Yonkofski[a]

[a] *Pacific Northwest National Laboratory, Richland, WA 99354*
[b] *Lawrence Livermore National Laboratory, Livermore, CA 94550*

Abstract

The Designs for Risk Evaluation and Management (DREAM) tool was developed as part of the effort to quantify the risk of lack of containment for geologic storage of carbon dioxide (CO_2) under the U.S. Department of Energy's National Risk Assessment Partnership (NRAP). DREAM is an optimization tool created to identify optimal monitoring schemes that minimize the time to first detection of CO_2 leakage from a subsurface storage formation. DREAM acts as a post-processor on user-provided output from subsurface leakage simulations. While DREAM was developed for CO_2 leakage scenarios, it is applicable to any subsurface leakage simulation of the same output format. The DREAM tool is comprised of three main components: (1) a Java wizard used to configure and execute the simulations, (2) a visualization tool to view the domain space and optimization results, and (3) a plotting tool used to analyze the results. A secondary Java application is provided to aid users in converting common American Standard Code for Information Interchange (ASCII) output data to the standard DREAM hierarchical data format (HDF5). DREAM employs a simulated annealing approach that searches the solution space by iteratively mutating potential monitoring schemes built of various configurations of monitoring locations and leak detection parameters. This approach has proven to be orders of magnitude faster than an exhaustive search of the entire solution space. The user's manual illustrates the program graphical user interface (GUI), describes the tool inputs, and includes an example application.

The latest version of the software, DREAMv3.0, introduces the ability to optimize surface geophysical surveys taken over various areas of the field site at irregular intervals. This optimization can be run using surface geophysics only, wellbore instruments only, or using a combination of surface geophysical surveys and wellbore instruments. These two approaches generally have broadly overlapping sensitivities, as a surface geophysical survey detects over a wide area of the field but only detects signals that are detectable along the surface at the time the survey was conducted, whereas a wellbore sensor can generally detect signals in the subsurface continuously the entire time it is in place, but only at one point in space. Therefore a pairing of wellbore sensors limited in their spatial coverage with geophysical surveys limited in their temporal coverage can produce a much improved monitoring plan. This software assumes that a set of hypothetical simulated CO_2 leaks are available, and that those leaks are representative of the underlying risk profile for a site. This allows the software to remain lightweight and functional on a typical laptop, and reduces the barrier to entry for users who lack the expertise of budget for high performance computing. This also improves compatibility across various application domains, allowing the tool to remain independent from any software dependencies related to for-

ward modelling or model calibration. Leak scenarios can be provided in the HDF5 file format, as NRAP-Open-IAM output files, or as STOMP, NUFT, Tecplot or Tough2 simulation files.

Keywords: carbon sequestration, subsurface monitoring design, gravity surveys, optimization, genetic algorithms

1. Introduction

Geological carbon capture and storage is one proposed means of mitigating human-induced climate change, by diverting and concentrating anthropogenic point sources of carbon dioxide into a supercritical fluid and then injecting it into a permeable underground rock formation for long-term storage. One application of this technique is in enhanced oil recovery, where the physicochemical properties of supercritical CO2 are used as a solvent to strip oil from mineral grain surfaces and mobilize it towards an oil production well. Another financial incentive for industry to pursue this technology is 45Q tax credits, which allow private industry to receive tax credits for capturing and storing carbon dioxide rather than releasing it into the atmosphere. This tax credit is currently priced at $10-20 per metric ton with some proposals to gradually increase the price to $35-50 per metric ton, and current CO_2 emissions are approximately 5 billion metric tons per year in the US. Therefore if monitoring costs and operational risks are sufficiently minimized, geologic carbon storage can potentially become an industry on the order of $100 billion/year in the near future.

One of the main challenges of this endeavor is verifying that the stored carbon dioxide remains stable and immobilized, as drilling enough monitoring wells near enough to the storage reservoir to detect every possible combination of CO_2 leaks can be prohibitively expensive. The DREAM (Designs for Risk Evaluation and Management) software package was therefore developed in order to design optimal arrays of monitoring wells. DREAM accepts an ensemble of potential leak scenarios as input, and then develops a Pareto-optimal set of monitoring plans representing the range of possible tradeoffs between cost and risk. DREAM is designed to maintain a low technical barrier to entry, and can be run on a typical laptop or workstation rather than a supercomputer.

The latest version of DREAM also incorporates a mix of direct monitoring through wellbores as well as remote sensing through gravity surveys conducted along the surface. While incorporating gravity surveys can improve the cost and performance of the overall monitoring plan, it can also make the optimization problem more challenging. Heuristic algorithms were therefore developed to efficiently solve these mixed continuous-combinatorial optimization problems.

2. Background

The algorithmic design of optimized sensor networks for environmental monitoring is a well-studied area of research [8, 10, 21], with many studies [17, 18, 19] focusing on simulated annealing [See Section 3.2.3] while others [5, 6, 7, 12, 14, 20, 24] use genetic algorithms [See Section

*Corresponding author. Tel.: +1 (509) 375-5945
Email address: alexander.hanna@pnnl.gov (Alexander C. Hanna)

3.2.4]. The goal of the DREAM software [2, 25, 26, 27] has been to integrate these methods into a single, user-friendly interface that uses a modular approach allowing it to be integrated into any arbitrary forward modeling and history matching workflow.

3. Methods

3.1. Software Application

This software is written primarily in Java, and has a graphical user interface with a series of windows allowing the user to configure the knowns, unknowns and constraints of their monitoring design problem. The software accepts a set of hypothetical simulated CO_2 leakage files as inputs, and treats those leakage files as representative of the underlying risk profile for a site. Leak scenarios are composed of physical parameters such as pore pressure, CO2 concentration, or salinity as a function of x,y,z and time, and can be provided in the HDF5 file format, as NRAP-Open-IAM output files, or as STOMP, NUFT, Tecplot or Tough2 simulation files.

Once the leak ensemble is loaded into DREAM, the user defines leakage thresholds indicating the value at which each subsurface parameter constitutes an impact or degradation to the aquifer. For some values such as pore pressure, there may not be a value that harms the aquifer, whereas other values such as salinity might be based on an MCL (Maximum Contaminant Level) defined by a state or federal agency. The user then defines detection thresholds for each parameter, which may be based off of the sensitivity of the instrument or the background noise levels particular to the site. A distinction is made between leakage thresholds and detection thresholds, as the amount of a given contaminant needed to impact an aquifer can be different from the amount needed for its presence to become detectable.

The user can then define a set of weighting coefficients which guide the software in prioritizing sensor coverage for each leak. By default these weights are set equal, however they can be re-defined based on the volume of aquifer degraded according to the leakage thresholds, or by a user-supplied estimate of the relative probability of each leak occurring or the relative magnitude or impact of each leak. These values can be defined manually within the GUI, or by reference to an external csv file. There is also a collection of Python scripts available which parse the csv files along with the input HDF5 files to generate visualizations of the optimization results as static images files.

3.2. Optimization Algorithms

For each optimization algorithm, we programmatically develop a series of monitoring plans each including a variety of wellbores, sensors and gravity surveys as well as their locations and timings, and we then assess how each plan performs in terms of our monitoring objectives. The five objectives available to consider are

1. Minimizing the purchasing, labor, installation and operational costs of the monitoring plan.
2. Maximizing the number of leak scenarios that would be detectable by the monitoring plan.
3. Maximizing the total mass of CO_2 of the leak scenarios that would be detectable by the monitoring plan.
4. Minimizing the time to detect our leak scenarios.

5. Minimizing the volume of aquifer degraded at the time of detection.

These objectives broadly overlap in some cases, however they can also conflict with one another to various degrees. For example, objective 1 conflicts strongly with objectives 2-5, as improving any of those objectives almost always requires more sensors and boreholes drilled. Objectives 2 and 3 also tend to conflict with objective 4, as detecting leak scenarios as early as possible requires locating sensors very near their points of origin which tends to preclude also detecting a large number of different potential leaks. Objectives 4 and 5 tend to correlate strongly with one another, however they can also conflict at times. For example a small, high-concentration but slow-moving leak may exceed the lower detection limits of our sensors and become detectable very quickly, but may take a long time for the plume to expand and contaminate a large volume of the aquifer. By contrast, a fast-moving but low concentration plume may contaminate a large area of the aquifer before finally reaching a high enough concentration to be detectable by our sensors.

A tradeoff surface therefore exists, representing the range of compromises that can be made between these five objectives. Some possible monitoring plans would provide a lot of broad coverage and would eventually detect most or all of the potential leaks, while others may provide more targeted coverage to preferentially detect the largest, most impactful (in terms of volume of aquifer degraded) and the most likely leaks as quickly as possible, while others make some compromises on coverage to minimize operating costs. Understanding the structure of this five-dimensional Pareto-optimal surface is critical to making good site monitoring decisions.

3.2.1. Grid Search

An exhaustive grid search requires that we employ some combinatorics to evaluate every possible combination of sensor type and sensor location. We consider installing ${}_nC_r$ combinations of sensors, where n is the number of sensor types we install and r is the number of sensors we propose to install. This can be represented by the equation

$$ {}_nC_r = \frac{n!}{r!(n-r)!}, \tag{1} $$

which can itself quickly become computationally intractable if a large array of sensors are needed. For example, if we consider employing three distinct sensor types (pressure, saturation, gravity) and intend to install 5 sensors in total, we get 21 possible combinations

$$ \text{Sensor Types} = \left\{ \begin{array}{ccccc} \text{p} & \text{p} & \text{p} & \text{p} & \text{p} \\ \text{p} & \text{p} & \text{p} & \text{p} & \text{s} \\ \text{p} & \text{p} & \text{p} & \text{p} & \text{g} \\ \text{p} & \text{p} & \text{p} & \text{s} & \text{s} \\ \multicolumn{5}{c}{\text{...14 more rows...}} \\ \text{s} & \text{s} & \text{g} & \text{g} & \text{g} \\ \text{s} & \text{g} & \text{g} & \text{g} & \text{g} \\ \text{g} & \text{g} & \text{g} & \text{g} & \text{g} \end{array} \right\}, \tag{2} $$

however if we consider installing 15 sensors, that leads to 136 possible sensor combinations. We then consider the list of possible sensor locations, as discretized by the *xyz* grid used in our ensemble of leak scenario files. Many of these grid nodes can be neglected, as they never see

a signal that exceeds the threshold and would never detect any of the leaks from the ensemble. However, there may still be thousands to millions of feasible sensor location combinations.

$$\text{Node Locations} = \left\{ \begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 \\ \multicolumn{5}{c}{...65774 \text{ more rows...}} \\ 20 & 20 & 21 & 21 & 21 \\ 20 & 21 & 21 & 21 & 21 \\ 21 & 21 & 21 & 21 & 21 \end{array} \right\}, \tag{3}$$

performing spatial downsampling on the aquifer simulation files can manage this to some degree, but not without sacrificing some spatial resolution. We then compute the Cartesian product of the list of possible sensor locations with the list of possible sensor location combinations. This can readily lead to millions to trillions of possible monitoring plans, making it an impractical approach for most full-reservoir models.

3.2.2. Monte Carlo

The Monte Carlo algorithm is a much simpler approach, randomly selecting the number and locations of sensors at each iteration. While this is less thorough than the full grid search, it tends to search the decision space fairly uniformly with a reasonable amount of computational effort. In many cases this may quickly find the same ideal solution that would have eventually been found by grid search.

3.2.3. Simulated Annealing

The Simulated Annealing algorithm [11, 16] begins by generating a single random monitoring plan, considering how well this plan would perform in terms of cost and leak detection, then making a series of small, iterative modifications to this plan. At each iteration, it considers whether the most recent change improved or degraded the monitoring plan performance, and accepts the change if it was deemed an improvement. If the change degrades performance, a random number generator is used to probabilistically accept or reject the change, with the acceptance probability depending on how much the performance was degraded. This encourages the algorithm to occasionally "climb out" of a local minima and explore new areas of the decision space. This algorithm can find "good" solutions very quickly, but often fails to find the best possible solution. It should generally be repeated multiple times to ensure that the decision space has been adequately explored.

3.2.4. Non-dominated Sorting Genetic Algorithm II

The DREAMv3 NSGA-II option is based on the Non-dominated Sorting Genetic Algorithm [4], which uses Pareto ranking [Figure 1] to iteratively develop a set of optimal monitoring plans. Pareto ranking assigns each monitoring plan a rank based on how well it manages tradeoffs between the various objectives, with rank 1 monitoring plans rated the best. In this example, monitoring plans are being evaluated based on minimizing two objectives f_1 and f_2. Monitoring plans 1 and 2 [Figure 1a] are superior to monitoring plan 3 in terms of both objectives f_1 and f_2, and therefore they are both a lower (better) rank than plan 3. However, plan 1 and 2 are

ambiguous, as plan 1 is better in terms of objective f_1 while plan 2 is better in terms of objective f_2. These two must therefore be assigned the same rank.

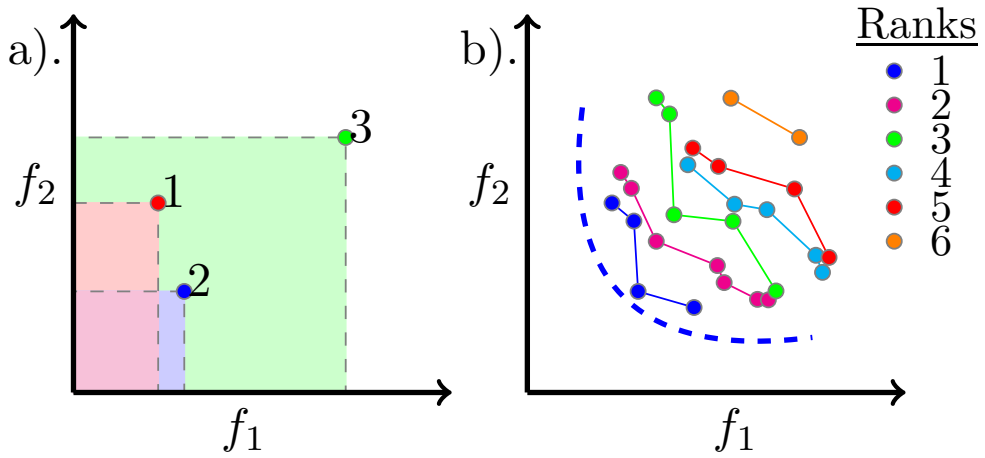

Figure 1: Illustration of Pareto optimality. Blue dashed line represents the 'true' Pareto front. Rank 1 solutions (blue dots) represent the best currently-available approximation of the Pareto front.

The NSGA-II algorithm uses a Monte Carlo approach to generate an initial population of monitoring plans, then semi-randomly selects pairs of monitoring plans from that ensemble. Monitoring plans with lower Pareto ranks are preferential selected, and a de-clustering method is used to preferentially select monitoring plans that are more unique in terms of their objective values. Sensors from these pairs of monitoring plans are then randomly selected and used to construct new monitoring plans that often combine the best features of the "parent" pair. Over enough iterations, this generates a set of monitoring plans that evenly explore the Pareto front.

3.2.5. Heuristic Sampling

The DREAM heuristic algorithm divides the monitoring design problem into a sequence of sensor placement decisions. The first sensor is placed by evaluating the full set of available sensor placement locations in terms of the added drilling and installation costs, the number of hypothetical leaks that would be detectable at that location [Figure 2], and the average time to detection of those leaks. The algorithm then considers the tradeoff relationship between these objectives, and randomly selects one of the resulting Pareto-optimal locations as the first sensor placement. An example of the tradeoff relationship between two objectives is visualized in 2d in Figure 3, and between three objectives in 3d in Figure 4.

Once the first sensor location has been determined, the set of hypothetical leaks that would be detected by that sensor are neglected and the objectives are re-computed considering only the set of leaks that would not be detectable by the first sensor. This process is repeated many times with various numbers and types of sensors as indicated by the users specifications. Where surface geophysical survey methods are being considered, several (1-4, randomly selected) points along the surface are sequentially chosen in the manner described above, and a rectangular survey geometry is constructed so as to be just large and dense enough to include those points.

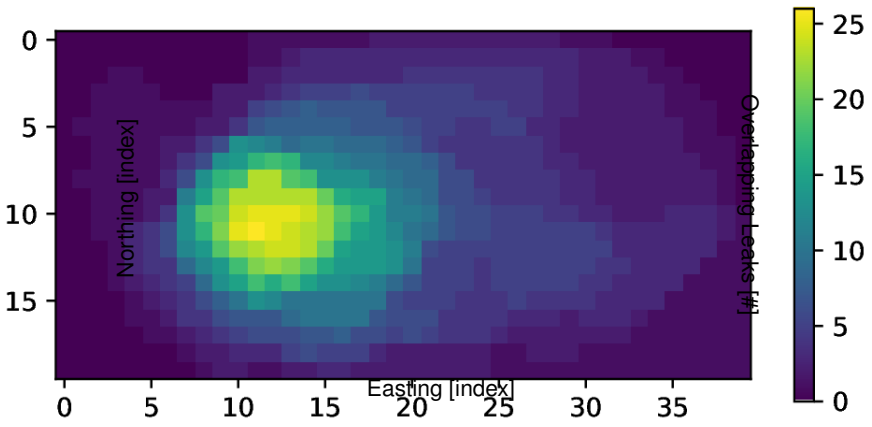

Figure 2: Color indicates the number of leaks that overlap at each point (x,y), considering only the leaks that as yet would be undetected by the current monitoring plan.

4. Results

4.1. NRAP-Open-IAM FutureGen2 Aquifer ROM

To create a set of leakage scenarios, a model of a hypothetical GCS site was created using the NRAP-Open-IAM [22], and integrated assessment model for quantifying the risk of unintended migration of CO2 and brine out of a storage reservoir. The hypothetical GCS project considered in this study treats the Vedder Sandstone, a sedimentary formation in the Southern San Joaquin Valley of CA, as a GCS system. The injection site is located approximately 20 miles northwest of Bakersfield, CA in Kern County. The region was previously characterized as part of the Westcarb Kimberlina Pilot Project [23].

Building on a previous study [13], a 50-year basin-scale injection of CO_2 and a 50-year post-injection period were simulated. CO_2 and brine leakage rates into a USDW were calculated along the 1000 wellbores in the study area, the volume of impacted aquifer for each leaky wellbore location determined, and compiled into a time-to-detection map for the entire site. We used the Monte Carlo framework of the NRAP-Open-IAM to run stochastic simulations that captured the uncertainty associated with the effective permeability of the wellbores in the model. The NRAP-Open-IAM represents discrete elements of a GCS system using individual surrogate models. Our integrated assessment model of the site consisted of a single reservoir lookup table joined with 1000 individual multisegmented wellbore models connected to FutureGen2 aquifer models in a one-way forward coupling.

In this preliminary example only a few wellbores are found to have detectable leaks, meaning only a very few locations are feasible for consideration as monitoring locations. This means that a full grid search is computationally possible, requiring only a few minutes to consider the 35,442 possible combinations (Figure 5, gray dots) of 5 sensor locations. We then use a Monte Carlo

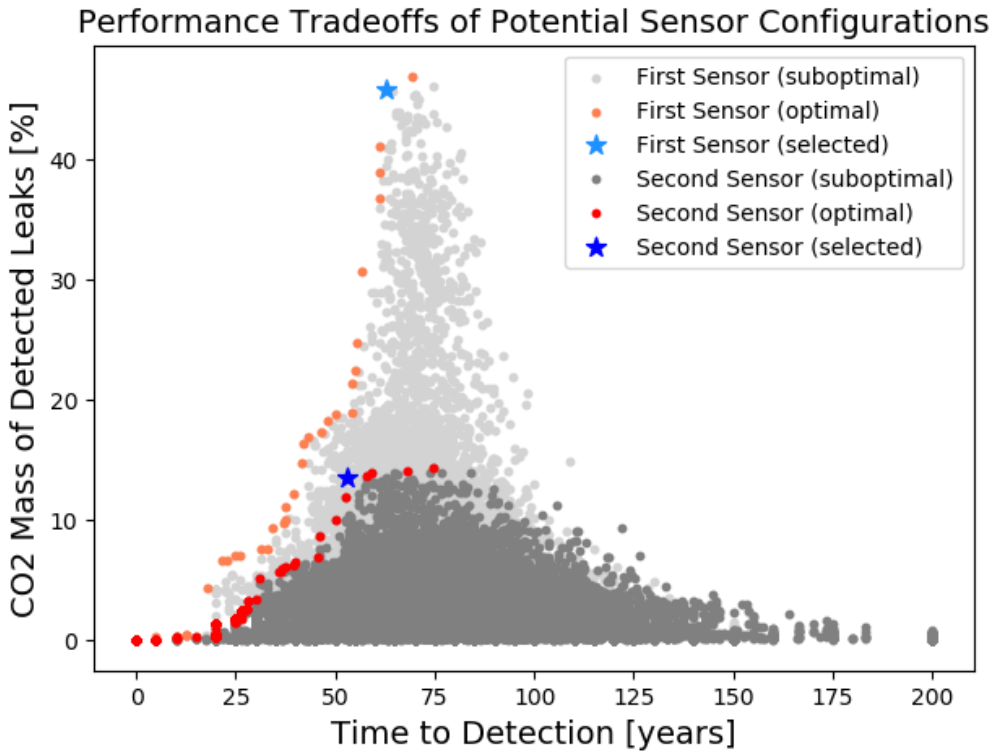

Figure 3: Each possible sensor placement is evaluated in terms of average time to detection and the percentage of leaks detected (weighted by mass of leak). Suboptimal options for the first sensor placement (light gray) are discounted because they each have worse performance than at least one other sensor location in terms of time to detection, poorer coverage or both. A single sensor placement is selected (light blue) from along the optimal tradeoff curve (orange). Suboptimal options (red) for the second sensor are then discounted in the same manner, and the second sensor placement is selected (blue) from along the new tradeoff curve (red).

approach to generate 10,000 random monitoring plans, and find almost the exact same monitoring performance (Figure 5a). By reducing the Monte Carlo search another order of magnitude (Figure 5b), we see that a fair number of possible monitoring outcomes are missed, however the general trend remains the same. The Simulated Annealing and Heuristic Algorithms are also compared to the grid search method (Figure 5c,d), and produce similar results.

This demonstrates that in simple cases at least, a relatively modest Monte Carlo sampling can nearly fully explore the decision space. For more complex monitoring design problems, however, a very large sample size may be required for Monte Carlo to adequately explore the decision space. This may become computationally impractical, and therefore the remaining algorithms will simply be evaluated by comparison to a ”very large” Monte Carlo sampling rather than a full grid search.

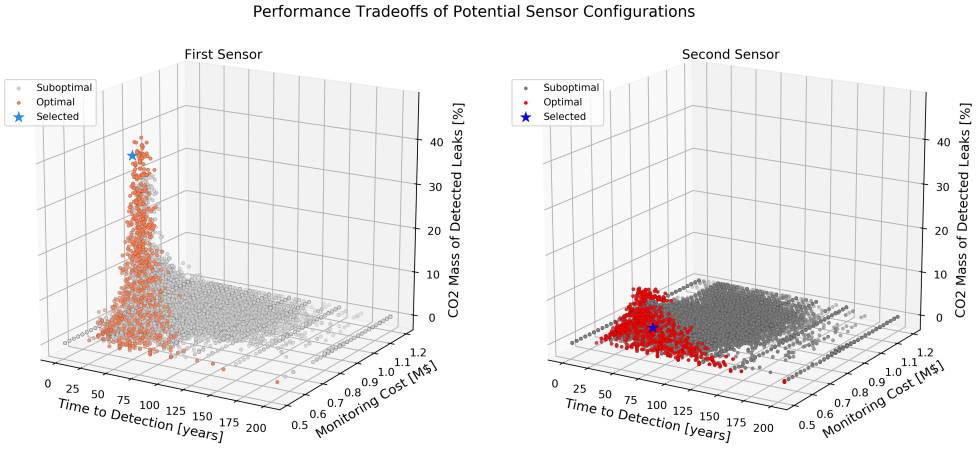

Figure 4: Each possible sensor placement is evaluated in terms of monitoring cost, average time to detection and the percentage of leaks detected (weighted by mass of leak). Suboptimal options for the first sensor placement (light gray) are discounted because they each have worse performance than at least one other sensor location in terms of time to detection, poorer coverage or both. A single sensor placement is selected (light blue) from along the optimal tradeoff curve (orange). Suboptimal options (red) for the second sensor are then discounted in the same manner, and the second sensor placement is selected (blue) from along the new tradeoff curve (red).

4.2. NRAP-Open-IAM Generic Aquifer ROM

The Reservoir component model provides spatial and temporal information on pressure and saturation in the storage reservoir. A reservoir simulation look-up table in United States (US) Department of Energy (DOE)'s National Risk Assessment Partnership Open-source Integrated Assessment Model (NRAP-Open-IAM) [22] is used for this study. The table is based on a set of numerical simulation models with varying rock properties for the Kimberlina reservoir in southern San Joaquin basin in California, USA. [28] The homogeneous reservoir model No. 14 (reservoir porosity: 0.276; reservoir permeability: 1.585×10^{-13} m^2; caprock permeability: 1.995×10^{-18} m^2) for the first 100 years (i.e., 50 years of CO_2 injection at a rate of 5 million metric tons/year and 50 years of post-injection) was used.

Coupled to the reservoir model, a wellbore model predicts leakage amount of the injected CO_2 and formation brine to an overlying aquifer through leaky wells based on input pressure and CO_2 saturation. A caprock segment wellbore model in NRAP-Open-IAM [22] was adopted [3]. This model assumes no leakage into caprock layers, and thus it provides a conservative estimate for risk evaluation. 74 wells within 5 km from the injection well were analyzed.

The output leakage rates of CO_2 and brine from the wellbore model is fed into a genetic aquifer impact reduced-order-model [1] as input to forecast the impacts of the well leakage on aquifers. The model generates 3D temporal and spatial datasets of mass fractions of the dissolved CO2 and salts throughout the aquifer. The model performs individual calculation for each of wells, and once dissolved CO_2 or salts plumes from different wells are overlapped, maximum value in the cell were used.

For a stochastic analysis, 200 realizations of input parameters of the wellbore and aquifer models were sampled using Latin-hypercube sampling (LHS) [9]. Table 1 shows the paramater ranges used. Of these stochastically generated models, 50 represent very small leaks which never

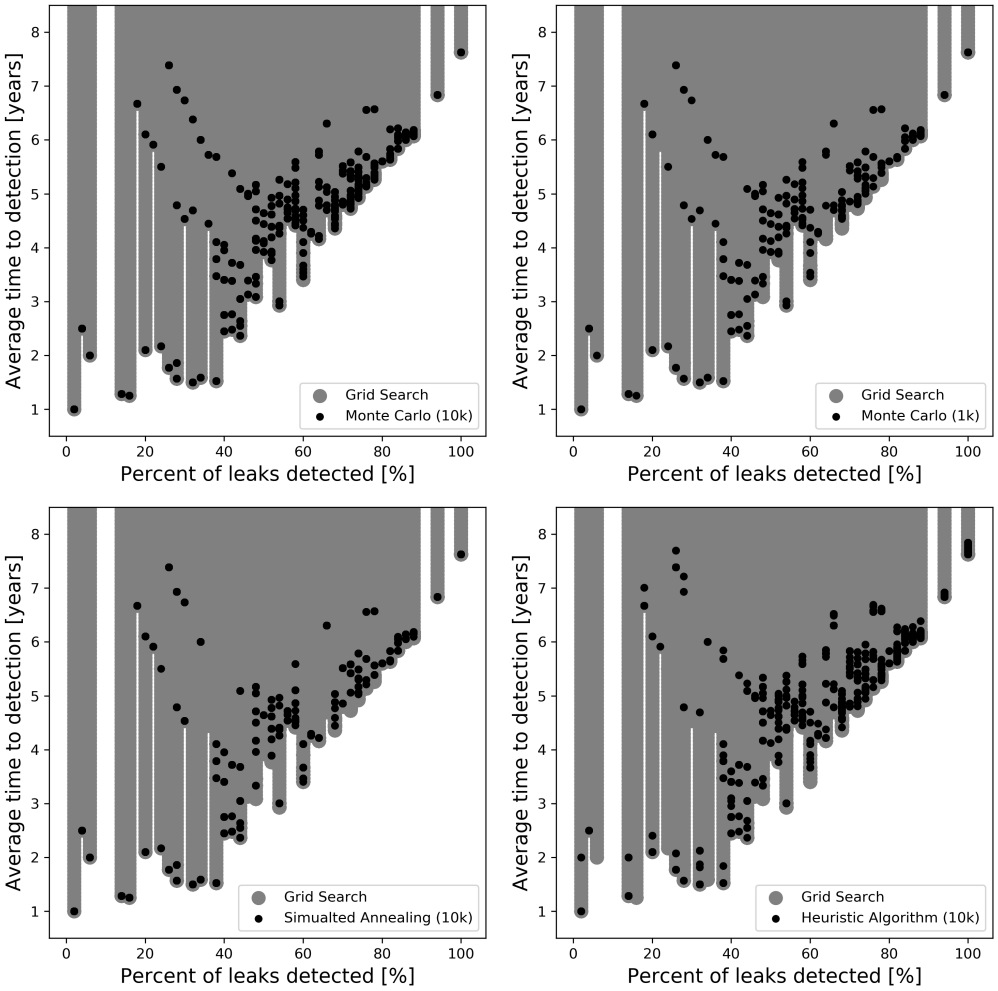

Figure 5: Comparison between Monte Carlo searches (for n=10,000 and n=1,000) and an exhaustive grid search of the decision space using the NRAP-Open-IAM Open Wellbore ROM.

exceed 0.01% CO_2 concentration [Figure 6] and are therefore neglected from the analysis.

This input dataset was run using DREAM's Monte Carlo, Simulated Annealing and Heuristic algorithms described above [Section 3.2], however due to the large size of the input files the grid search and genetic algorithm were too computationally intensive to be practical. Each of the 150 input files were 800-900MB, however DREAM had a peak memory usage of 2.1GB of the machine's 16GB RAM. As working with this dataset was time-intensive, only 100 iterations were run for each algorithm. For this example [Figure 7], the Monte Carlo algorithm was able to identify one monitoring plan capable of detecting 134/150 leak scenarios using 8 monitoring wells, while the heuristic algorithm was able to detect 138/150 leak scenarios using 2 monitoring wells.

Table 1: Parameter ranges for OpenIAM Generic Aquifer model

Parameter	Unit	Minimum	Maximum
Reservoir salinity	%	0.03	0.05
Well permeability	m^2	1.0×10^{-14}	1.778×10^{-12}
Aquifer thickness	m	25	100
Aquifer top depth	m	2,000	2,500
Aquifer porosity	%	0.02	0.10
Aquifer horizontal permeability	m^2	1.0×10^{-13}	1.0×10^{-10}
Aquifer permeability anisotropy	-	0.01	1.0

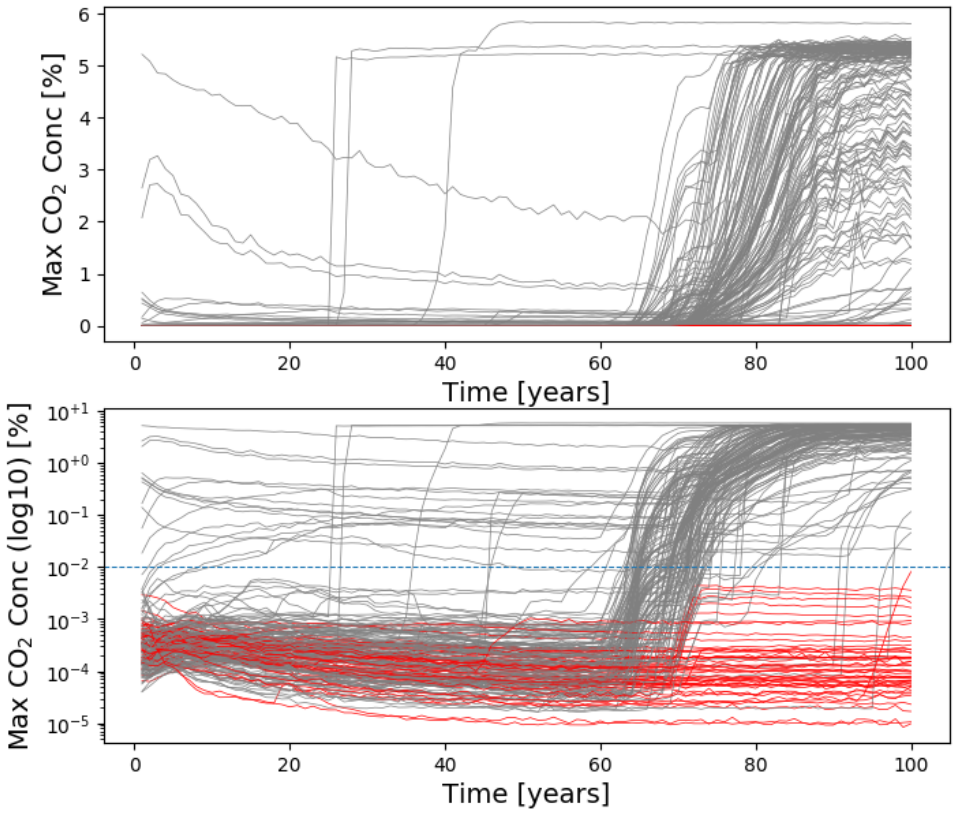

Figure 6: The maximum CO_2 concentration (taken over the entire model space) is shown for each leak in both linear-linear and linear-log space. Leaks which never exceed 0.01% CO_2 are highlighted in red and neglected from the analysis.

4.3. Kimberlina 1.2

We then ran monitoring plan optimizations based on the Kimberlina 1.2 dataset [15], supplemented by a gravity model later applied by Xianjin Yang. These optimizations considered combinations of pressure and saturation sensors installed within wellbores, as well as gravity

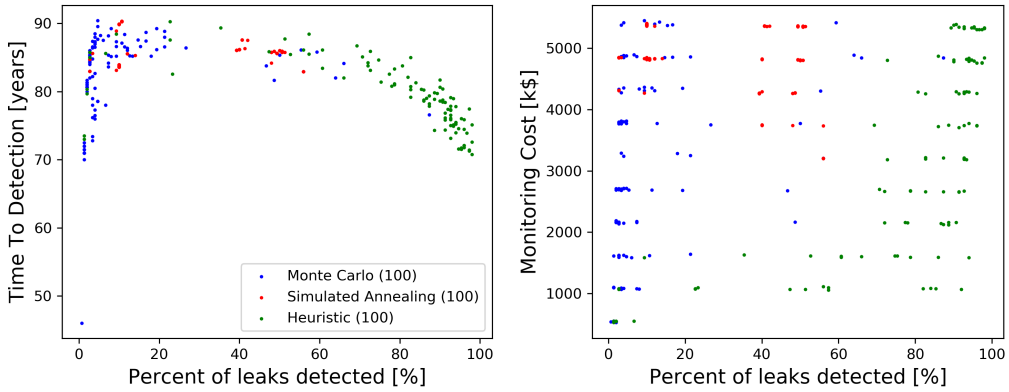

Figure 7: Comparison between the performance of Monte Carlo, Simulated Annealing and Heuristic algorithm using the OpenIAM Aquifer simulation dataset.

surveys. The spatial extent and resolution of this case allows us to consider many more (+3500) possible sensor locations, and combinations of as many as 5 sensors were considered. As the grid search method would therefore require an astronomical amount of computational effort, a large (100,000) set of Monte Carlo iterations were compared to 1,000 iterations of our heuristic algorithm. The heuristic algorithm produced much better results than either the Monte Carlo or Simulated Annealing algorithms, with some plans detecting more than 99.7% of the total CO2 leakage potential, while maintaining a relatively low cost and time to detection.

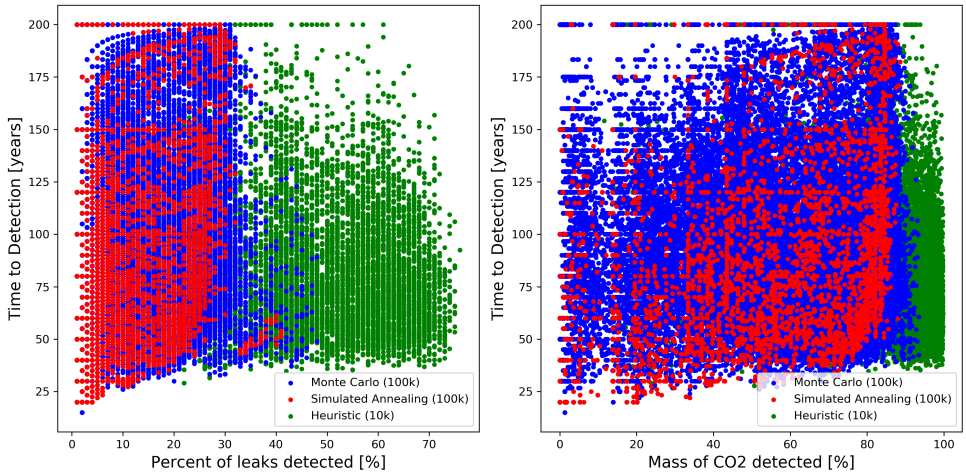

Figure 8: Comparison between the performance of Monte Carlo, Simulated Annealing and the DREAMv3 heuristic algorithm using the Kimberlina 1.2 leakage simulation dataset.

The best-performing monitoring plans were then identified [blue dots, Figure 10b], having greater than 98% detection of CO2 leakage by mass, and minimal monitoring costs and time to

detection. This pool of candidates can then be viewed one at a time and examined with an eye for practical field deployment. If the optimal monitoring plans are deemed viable for real-world application, they can be implemented, otherwise the algorithm can be run for additional iterations or using a different algorithm.

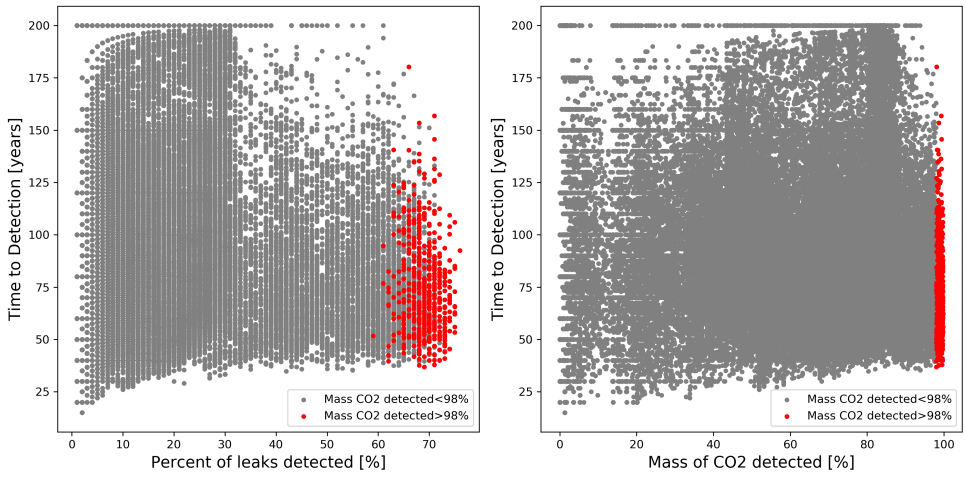

Figure 9: Monitoring plans with >98% mass CO2 detection are highlighted in red

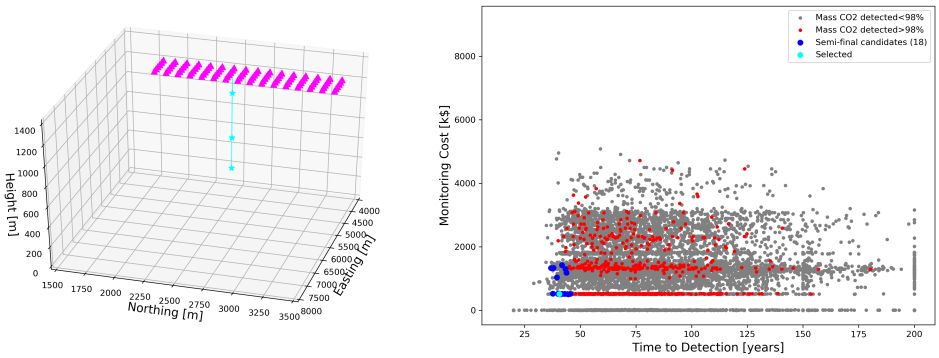

Figure 10: Set of optimal monitoring plan candidates

5. Conclusion

The DREAM software tool takes as input a set of (unlikely) CO_2 leakage scenarios and designs a monitoring plan tailored to the risk profile of the site. The current generation (version 3.0) of the software includes a heuristic algorithm which produces much cheaper and more effective monitoring plans while still providing a range of possible tradeoffs for the user to decide from. This version also includes passive surface geophysics data types such as gravity surveys, optimizing the survey size and station density.

References

[1] Diana H Bacon. Nrap-open-iam: Generic aquifer component development and testing. Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2022.

[2] Diana H Bacon, Catherine MR Yonkofski, Christopher F Brown, Deniz I Demirkanli, and Jonathan M Whiting. Risk-based post injection site care and monitoring for commercial-scale carbon storage: Reevaluation of the futuregen 2.0 site using nrap-open-iam and dream. *International Journal of Greenhouse Gas Control*, 90:102784, 2019.

[3] S. Baek, D.H. Bacon, and N.J. Huerta. Deep learning-based multisegmented wellbore model development for many wells site study. In *NRAP Annual Technical Meeting Conference*. National Risk Assessment Partnership, 2022.

[4] Kalyanmoy Deb. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[5] Anirban Dhar and Bithin Datta. Multiobjective design of dynamic monitoring networks for detection of groundwater pollution. *Journal of water resources planning and management*, 133(4):329–338, 2007.

[6] Anirban Dhar and Bithin Datta. Logic-based design of groundwater monitoring network for redundancy reduction. *Journal of water resources planning and management*, 136(1):88–94, 2010.

[7] Anirban Dhar and Rajvardhan S Patil. Multiobjective design of groundwater monitoring network under epistemic uncertainty. *Water resources management*, 26(7):1809–1825, 2012.

[8] Paul F Hudak and Hugo A Loaiciga. An optimization method for monitoring network design in multilayered groundwater flow systems. *water resources research*, 29(8):2835–2845, 1993.

[9] Ronald L Iman, Jon C Helton, and James E Campbell. An approach to sensitivity analysis of computer models: Part i—introduction, input variable selection and preliminary variable assessment. *Journal of quality technology*, 13(3):174–183, 1981.

[10] Abdelhaleem I Khader and Mac McKee. Use of a relevance vector machine for groundwater quality monitoring network design under uncertainty. *Environmental modelling & software*, 57:115–126, 2014.

[11] Scott Kirkpatrick, C Daniel Gelatt Jr, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220 (4598):671–680, 1983.

[12] Joshua B Kollat, Patrick M Reed, and Joseph R Kasprzyk. A new epsilon-dominance hierarchical bayesian optimization algorithm for large multiobjective monitoring network design problems. *Advances in Water Resources*, 31(5):828–845, 2008.

[13] Greg Lackey, Veronika S. Vasylkivska, Nicolas J. Huerta, Seth King, and Robert M. Dilmore. Managing well leakage risks at a geologic carbon storage site with many wells. *International Journal of Greenhouse Gas Control*, 88:182–194, 2019. ISSN 17505836. doi: 10.1016/j.ijggc.2019.06.011.

[14] Qiankun Luo, Jianfeng Wu, Yun Yang, Jiazhong Qian, and Jichun Wu. Multi-objective optimization of long-term groundwater monitoring network design using a probabilistic pareto genetic algorithm under uncertainty. *Journal of Hydrology*, 534:352–363, 2016.

[15] Kayyum Mansoor, Thomas A Buscheck, Xianjin Yang, Susan A Carroll, and Xiao Chen. Llnl kimberlina 1.2 nuft simulations june 2018 (v2). Technical report, National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV . . . , 2020.

[16] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[17] LM Nunes, MC Cunha, and L Ribeiro. Groundwater monitoring network optimization with redundancy reduction. *Journal of Water Resources Planning and Management*, 130(1):33–43, 2004.

[18] LM Nunes, MC Cunha, and L Ribeiro. Optimal space-time coverage and exploration costs in groundwater monitoring networks. *Environmental monitoring and assessment*, 93(1-3):103–124, 2004.

[19] Om Prakash and Bithin Datta. Sequential optimal monitoring network design and iterative spatial estimation of pollutant concentration for identification of unknown groundwater pollution source locations. *Environmental monitoring and assessment*, 185(7):5611–5626, 2013.

[20] Patrick Reed, Joshua B Kollat, and VK Devireddy. Using interactive archives in evolutionary multiobjective optimization: A case study for long-term groundwater monitoring design. *Environmental Modelling & Software*, 22 (5):683–692, 2007.

[21] J Sreekanth and Bithin Datta. Simulation-optimization models for the management and monitoring of coastal aquifers. *Hydrogeology Journal*, 23(6):1155–1166, 2015.

[22] Veronika Vasylkivska, Robert Dilmore, Greg Lackey, Yingqi Zhang, Seth King, Diana Bacon, Bailian Chen, Kayyum Mansoor, and Dylan Harp. Nrap-open-iam: A flexible open-source integrated-assessment-model for geologic carbon storage risk assessment and management. *Environmental Modelling & Software*, 143:105114, 2021.

[23] Haruko M. Wainwright, Stefan Finsterle, Quanlin Zhou, and Jens T. Birkholzer. Modeling the performance of

large-scale co2 storage systems: A comparison of different sensitivity analysis methods. *International Journal of Greenhouse Gas Control*, 17:189–205, 2013. ISSN 17505836. doi: 10.1016/j.ijggc.2013.05.007.

[24] Jianfeng Wu, Chunmiao Zheng, Calvin C Chien, and Li Zheng. A comparative study of monte carlo simple genetic algorithm and noisy genetic algorithm for cost-effective sampling network design under uncertainty. *Advances in Water Resources*, 29(6):899–911, 2006.

[25] Catherine Yonkofski, G Tartakovsky, N Huerta, and A Wentworth. Risk-based monitoring designs for detecting co2 leakage through abandoned wellbores: An application of nrap's wlat and dream tools. *International Journal of Greenhouse Gas Control*, 91:102807, 2019.

[26] Catherine MR Yonkofski, Jason A Gastelum, Ellen A Porter, Luke R Rodriguez, Diana H Bacon, and Christopher F Brown. An optimization approach to design monitoring schemes for co2 leakage detection. *International Journal of Greenhouse Gas Control*, 47:233–239, 2016.

[27] Catherine MR Yonkofski, Casie L Davidson, Luke R Rodriguez, Ellen A Porter, Sadie R Bender, and Christopher F Brown. Optimized, budget-constrained monitoring well placement using dream. *Energy Procedia*, 114:3649–3655, 2017.

[28] Quanlin Zhou and Jens T Birkholzer. On scale and magnitude of pressure build-up induced by large-scale geologic storage of co2. *Greenhouse Gases: Science and Technology*, 1(1):11–20, 2011.