233 (2024) 212473

Contents lists available at ScienceDirect

Geoenergy Science and Engineering

journal homepage: www.sciencedirect.com/journal/geoenergy-science-and-engineering

Passive-seismic sensor placement optimization for geologic carbon storage

Yuan Zi [a,*], Lei Fan [c], Xuqing Wu [d], Jiefu Chen [a], Zhu Han [a,b]

[a] Department of Electrical and Computer Engineering at the University of Houston, Houston, TX 77004, United States
[b] Department of Computer Science and Engineering, Kyung Hee University, Seoul, 446-701, South Korea
[c] Department of Engineering Technology, University of Houston, TX, 77004, United States
[d] Department of Information and Logistics Technology, University of Houston, TX, 77004, United States

ARTICLE INFO

Keywords:
Optimal sensor placement
Stochastic optimization
Passive-seismic
CO2 storage monitoring

ABSTRACT

The implementation of passive-seismic monitoring is essential for the geological carbon storage projects. For secure and continuous observing of induced passive-seismic events by CO2 injecting, a ground geophone network would be required. By determining the ideal number of seismic stations within a regular network, recent research has enhanced monitoring capabilities while also accommodating budgetary limitations. The primary objective of the optimal placement strategy for the surface geophone network is to create a cost-effective monitoring system. Given the cost limitations, a restricted quantity of sensors can be deployed at each location to maximize monitoring performance. To address this challenge, our approach involves the P-median stochastic programming formulation. The formulation aims to minimize the expected value of a monitoring target metric, which ultimately results in the optimal placement of detectors exhibiting superior expected behaviors. Our methodology is designed to choose the arrangement of detectors with best performance to improve both the early alarm detection time and localization capabilities of the sensor network. We utilize site-specific passive-seismic scenarios that capture the uncertainty in the characteristics of carbon leakage. The sensor grids provided by the optimization method invariably improve the capacity to detect passive-seismic events in comparison to sensors positioned in a regular grid configuration. We test the effectiveness of our approach on synthetic data based on a carbon storage site named Kimberlina. In conclusion, our approach provides a cost-effective solution for the optimal placement of sensors to achieve superior monitoring performance for the detection of passive-seismic events. The incorporation of site-specific scenarios with stochastic uncertainty coverage allows for more accurate and reliable results, leading to better decision-making for long-term geological carbon storage.

1. Introduction

Geological carbon dioxide (CO2) sequestration is a promising option for mitigating the increase in atmospheric CO2 and reducing the associated global climate change impacts. However, one of the principal issues of geological sequestration is the potential for CO2 leakage, which needs to be resolved. According to Ellsworth (2013), injecting CO2 into subsurface reservoirs during geological CO2 sequestration can cause geological strain and result in passive-seismic events around the intended carbon storage areas. Induced seismicity in CO2 geological storage is a primary concern, where CO2 injection and pressurization can trigger events ranging from minor tremors to potentially damaging earthquakes, posing risks to infrastructure and reservoir integrity. These seismic activities can induce ground motion, stress changes, and potential leakage pathways within the storage reservoir, emphasizing the importance of maintaining reservoir and caprock integrity. Even minor seismic events can displace the ground, affecting surface infrastructure such as wellheads and pipelines, necessitating stability measures. Additionally, seismic events may lead to pressure changes within the reservoir, impacting CO2 behavior, thus highlighting the need for proper pressure management. Regulatory oversight mandates thorough risk assessments and monitoring for CO2 storage projects to ensure compliance with safety standards. Furthermore, addressing public concerns and building trust through effective communication and community engagement is crucial for gaining acceptance of CO2 geological storage projects. Monitoring passive-seismic events can be critical in ensuring the secure and long period of geological sequestration of carbon. Monitoring passive-seismic events during carbon sequestration can aid in visualizing the CO2 plume's movement (Yin et al., 2021; Goertz-Allmann et al., 2017), examining stress changes within the reservoir, and identifying potential pathways for CO2 leakage, such as fractures and faults (Maxwell and Urbancic,

* Corresponding author.
E-mail address: ziyuan1501040205@gmail.com (Y. Zi).

https://doi.org/10.1016/j.geoen.2023.212473
Received 16 May 2023; Received in revised form 28 September 2023; Accepted 6 November 2023
Available online 20 November 2023

2001; Miyazawa et al., 2008; Verdon et al., 2012; Hu et al., 2023; Jin et al., 2021). In practice, a single downhole array of geophones with weak passive-seismic monitoring performance is commonly used due to cost limitations. Nevertheless, for monitoring CO2, a surface seismic station network with higher performance and regular spacing has been employed in the field (Boullenger et al., 2015; Kaven et al., 2015; Stork et al., 2018). The performance of locating passive-seismic events hinges on the arrangement of seismic sensors in geospatial. As a result, the optimal configuration of seismic sensor networks has garnered significant attention in the field of passive-seismic monitoring (Kijko, 1977; Rabinowitz and Steinberg, 1990; Steinberg et al., 1995). A straightforward approach that can examine the number of ground seismic detectors which give optimal performance for a grid network with regular spacing was recently introduced by Chen and Huang (2020). To achieve cost-effective passive-seismic monitoring, this method utilizes the potential passive-seismic events sources sites to select the optimal sensor placement strategy considering the balance between event localization accuracy and the total number of seismic sensors. In summary, a key challenge in monitoring carbon sequestration-induced passive-seismic events is to accurately observe large areas at high monitoring performance with a minimal sensor budget. Thus, the optimal placement of passive-seismic sensors is considered one of the crucial techniques for identifying potential leakage during carbon sequestration.

Optimizing sensor placement is a type of facility location problem (FLP), which has been extensively studied in various applications, such as methane sensor placement (Zi et al., 2022), water sensor placement (Berry et al., 2005), and wireless sensor placement (Ling et al., 2022; Younis and Akkaya, 2008; Liu et al., 2022). Optimizing sensor placement with respect to event impact is frequently modeled as a mixed-integer linear programming P-median formulation (Karatas et al., 2016). In the P-median formulation, the placement of sensors is optimized to minimize event impact measures, such as damage cost. The event cost is usually proportionate with the time taken to detect the event after it occurs, which is referred to as the first detection time. This formulation was originally proposed to address water sensor placement problems (Berry et al., 2005), and has since been used in the optimization of petrochemical facilities' gas sensor placement and in monitoring site-scale methane emissions in various studies, such as Legg et al. (2012), Benavides-Serrano et al. (2016), Klise et al. (2020), and Zi et al. (2022).

In this paper, our proposed approach involves using a ray-tracing simulation tool and an open-source, modern-MILP-solver-based, package called Chama to perform sensor placement optimization. The proposed sensor placement scheme is designed to address two key problems: optimizing the early detection time and accurately localizing the source. The optimization problem's solution has two general parts. In the first part, passive-seismic wave propagation simulation and data collection are performed based on the possible passive-seismic source location data and the history of the geophysical survey. In the second step, a mixed-integer linear programming (MILP) solver is utilized to solve the problem within a provided detector cost. To consider the impact of uncertainty, both the physical simulation and data-driven scheme employ multiple simulations, referred to as passive-seismic events, to represent the potential distribution of passive-seismic events under varying conditions. In terms of contributions. As far as the authors are aware, this is the first study to employ a MILP formulation in passive-seismic monitoring. Also, this paper provides a simple open-source passive-seismic sensor placement scheme based on readily available tools, allowing researchers and practitioners to easily use or modify it for their own projects or research. Finally, this paper presents an experiment that studies the effectiveness of the proposed stochastic optimal design of passive-seismic surveillance using synthetic data for the Kimberlina. According to Chen and Huang (2020), Kimberlina is a site under consideration as a possible demonstration field for the National Risk Assessment Partnership (NRAP) initiative project. The NRAP's goal is to build a risk assessment procedure for CCUS.

In addition to the current challenges in CO2 sequestration, there is growing interest in injecting CO2 into deep marine sediments to create stable CO2 hydrate for long-term storage. This approach introduces new challenges, including monitoring CO2 leakage in marine environments and detecting seismic events during injection. Addressing these issues in future research will make our methods more widely applicable (Liu et al., 2023). Additionally, researchers can adapt our algorithms to detect seismic events during methane (CH4) extraction from methane hydrate-bearing sediments. This can be valuable in regions like the Gulf of Mexico, Alaska, Japan, and China, where geological conditions vary significantly. These future research directions expand the potential applications of our approach Ren et al. (2022).

The remaining portion of this paper is organized in the subsequent manner. Section 2 provides a comprehensive explanation of the problem statement and formulation. Section 3 introduces the stochastic optimization approach for passive-seismic sensor placement. In Section 4, we present optimization results are presented using a synthetic dataset of passive-seismic events created from previous geophysical surveys. Finally, Section 5 summarizes the key findings and conclusions of this paper.

2. Method

2.1. Problem statement

The problem of placing passive-seismic sensors is a real-world example of location planning problems, which involves identifying suitable locations for facilities that offer necessary services. Location planning problems encompass various industries, such as healthcare, food service, retail, and emergency response, and may involve determining optimal locations for hospitals, restaurants, stores, and fire stations. This paper focuses on utilizing passive-seismic sensors as facilities to provide carbon-leakage-related event monitoring. The objective is to optimize the design of sensor networks that can quickly detect and accurately locate such events while minimizing the number of sensors used under various environmental conditions. The goal is to achieve this with an unknown leakage event using the least amount of sensor budget available. Few studies have been conducted to solve this problem, and the most recent study considered comparing different sensor numbers and regular grid sensor networks' performance to determine the optimal sensor number. The paper employs the P-median formulation in addressing the passive-seismic sensor placement problem. Specifically, its objective is to ascertain the ideal positioning of P facilities to diminish the average distance between a demand node and the facility's location.

2.2. Problem formulation

In this section, we first introduce a novel P-median MIP formulation for optimizing the placement of sensors in passive-seismic monitoring applications aiming at minimizing detection time. Table 1 provides the notations' summary employed in this problem. The proposed formulation can be expressed as follows:

$$\min_{y_l} \sum_{e\in\mathcal{E}_e} p_e \sum_{i\in\mathbb{L}_e} d'_{e,i} x_{e,i}, \tag{1}$$

subject to

$$y_l \in \{0,1\} \quad \forall\, l \in \mathbb{L}, \tag{2}$$

$$\sum_{l\in\mathbb{L}} c_i y_l \le c, \tag{3}$$

$$0 \le x_{e,i} \le 1 \quad \forall\, e \in \mathcal{E},\, i \in L_e. \tag{4}$$

$$\sum_{i\in\mathbb{L}_e} x_{e,i} = 1 \quad \forall\, e \in \mathcal{E}, \tag{5}$$

$$x_{e,i} \le y_i \quad \forall\, e \in \mathcal{E}, i \in L_e, \tag{6}$$

The optimization approach we propose is designed to enhance the detection time of passive-seismic scenarios by optimizing the positioning of sensors. The Eq. (1) is the objective function which represents the expected damage impact of all passive-seismic events, where the impact is assumed to be proportionated to the time at which the seismic event is first detected. In the equation, $x_{e,i}$, being a binary indicator, represents whether location i is the first to detect event e, and y_l is an additional binary variable that denotes whether a sensor is placed at location l. It should be noted that while multiple sensors might be activated by a passive-seismic event, only one is considered to be the first to detect it. Additionally, $d_{e,i}$ represents the damage coefficient associated with event e at location i, and p_e denotes the probability of event e occurring. The sets L, L_e, and $\mathcal{E}$ correspond to the collection of all possible sensor locations, the set of sensors triggered by event e, and the set of all events, respectively.

To ensure the decision variable y_l is binary, Eq. (2) is used as a constraint, and Eq. (3) provides a cost constraint of the highest number of sensors. In this equation, c_i represents the cost of sensor i, and c denotes the allocated budget for sensors. Constraints (6) and (5) serve to guarantee that every passive-seismic event is noticed by at no less than one sensor and that a place i is only able to be the sonnest one to catch a passive-seismic occurrence e if a sensor is installed at place i. To simplify computation, Eq. (4) relaxes the binary integer decision variable to real numbers, but it will converge to either 0 or 1 after solving the problem.

In addition to the detection time, passive-seismic source localization error is another commonly used performance metric for monitoring the sensor network (Chen and Huang, 2020). The present study introduces a bi-level mixed-integer programming formulation for the P-median problem that is aimed at minimizing the inversion errors associated with passive-seismic source localization. As in previous sections, Table 1 provides a summary of the notation employed in this problem.

$$\min_{x_i} \sum_{e\in\mathcal{E}_e} p_e MSE(s'_e, \hat{s}_e), \tag{7}$$

subject to

$$\hat{s}_e = \arg\min_{s_e} MSE(t_{obs}, t), \tag{8}$$

$$t = \mathcal{F}(y, v, s_e), \tag{9}$$

$$\sum_{i\in\mathbb{I}} c_i y_i \le c, \tag{10}$$

$$y_i \in \{0, 1\} \quad \forall\, i \in \mathbb{I}. \tag{11}$$

The outer objective function, as given by Eq. (7). The objective is to minimize the anticipated value of the consequence metric for the sensor network, which is the passive-seismic source inversion errors. Here, s_e represents the spatial location of the passive-seismic source for event e, $\hat{s}_e$ is the predicted seismic source spatial location for event e obtained through the inversion algorithm, and s'_e is the ground truth for the seismic source spatial location for event e. The inner objective function, given by Eq. (8), solves the source inversion problem and obtains the optimal estimated source localization that minimizes the mean square error between the observed seismic signal's first arrival time and the simulated seismic signal's first arrival time using the estimated source localization. The inversion problem aims to locate the source location with a ray tracing forward operator $\mathcal{F}$, subsurface velocity model v, and network trail design y. Eq. (9) provides a forward simulation operator that takes the sensor network, velocity, and source location and outputs simulated seismic event observation first arrival time. Constraint Eq. (10) offers a maximum threshold, c, on the permissible quantity of detectors. To obtain a minimum objective function in a standard MILP problem format, the objective function in Eq. (7) can be transformed similarly to the first problem's formulation presented by Eqs. (1) to (4). This can be achieved by solving the inversion optimization problem specified by Eqs. (8) and (9) using a trial sensor design.

Table 1
Tables of symbols.

Symbol	Meaning
$e \in \mathcal{E}$	The collection of all events.
L	The group of all potential sensors.
L_e	The set of sensors that can detect event e
p_e	The probability of event e taking place
$d_{e,i}$	The damage coefficient for passive-seismic event e at location i
$x_{e,i}$	Binary variable indicating if location i first detects event e
y_l	Binary variable indicating sensor presence at location l
c_i	The expense associated with sensor i
c	The allocation of funds for the sensors
s_e	Passive-seismic source spatial location for event e
$\hat{s}_e$	Predicted passive-seismic source spatial location for event e with heuristic optimization
s'_e	Passive-seismic source spatial location ground truth for the event e
t_{obs}	Observation of passive-seismic wave-arrival time.
t	Measurement of passive-seismic wave-arrival time.
$\mathcal{F}$	Ray tracing forward operator.
v	Subsurface velocity model.

Algorithm 1 Stochastic passive-seismic sensor placement detection time optimization algorithm

Input subsurface velocity model, events positions, number of sensor grid and strata layers: v_p, n, s_e and l_{depths}
Set the best placement to empty list: x_{best}
Initialize events impact to empty list: imp_{list}
Set min objective value to infinity: $obj_{min} = inf$
Create all sensor candidate matrix D of size $n \times n$
#Physical-based passive-seismic Wave Propagation Simulation
for *each s_e event position of events* **do**
    Simulate the impact (detection time of all sensor candidates) using the forward model: $t = \mathcal{F}(D, v, s_e)$
    Append simulated impact to events impact list
end
Initialize sensor cost list, event probability, and sensor budget: $sensor_{cost}$, p_e and c
# Optimize the Mix integer programming (MIP) problem with MIP Solver
for *each y_l (combination of c candidate locations)* **do**
    Simulate the earliest detection time using the current sensor placement: $d'_{e,i} = \text{argmin}(mask_{combination} \otimes d_e)$
    calculate objective value:
    $y_l \sim x_{e,i}$
    $obj = \sum_{e\in\mathcal{E}_e} p_e \sum_{i\in\mathbb{L}_e} d'_{e,i} x_{e,i}$
    if $obj < obj_{min}$ **then**
        Update min objective value and best placement
    end
end
Result: best placement

3. Solution

Following the math formulations mentioned in the last section, this solution section is going to introduce the components and details of the solution scheme. We leverage state-of-the-art optimization software such as Chama, heuristic optimization techniques, and open-source ray tracing physical numerical simulation models for sensor placement

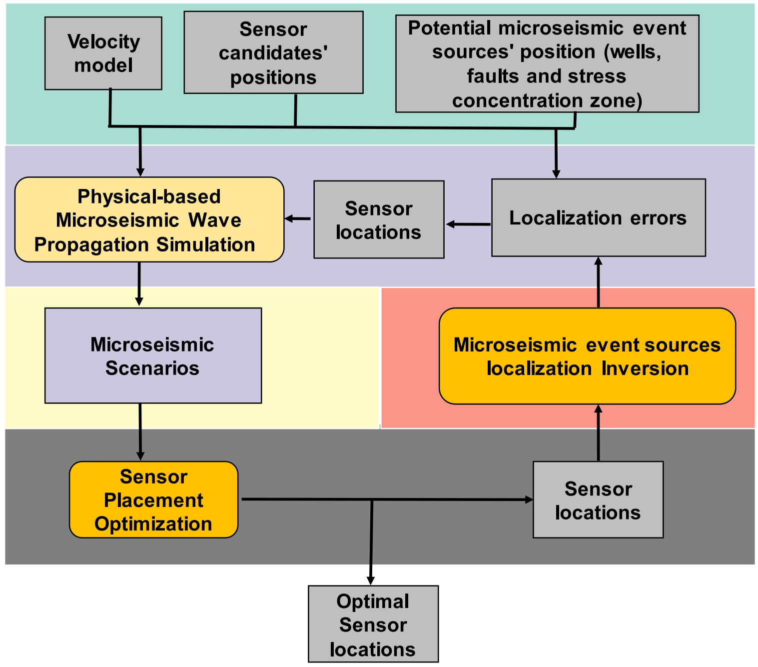

Fig. 1. Optimizing sensor placement for source localization in geologic carbon storage monitoring workflow.

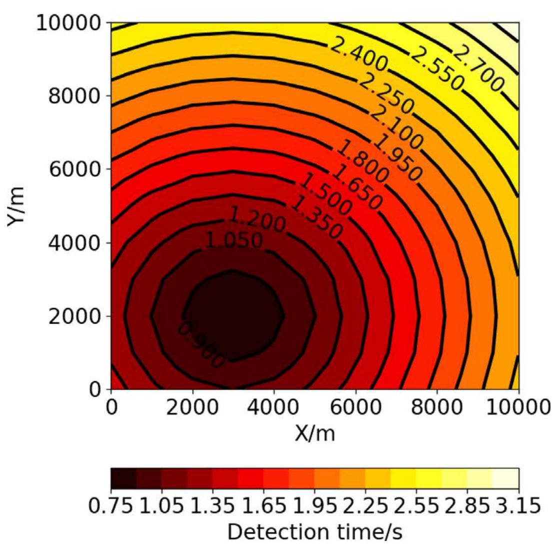

Fig. 2. Passive-seismic event simulation and detection time distribution of one source in the top-down view.

optimization, inversion, and forward problems. Building on the work of Chen and Huang (2020), this paper offers a fresh perspective on exploring the behavior of irregularly-spaced seismic sensors by focusing on the mathematical space of sensor location combinations. To this end, we consider all the discrete combinations of a given number of sensors and propose a scheme based on the diagram shown in Fig. 2, which contains three basic steps. The initial two stages, which involve preparing the geospatial attributes of passive-seismic events for optimization, are akin to the modeling procedures outlined in Klise et al. (2020). We present innovative data-driven techniques for optimizing the placement of passive-seismic sensors. Our approach involves utilizing a set of events generated with smoothed velocity model directly for mixed-integer linear programming.

Our proposed solution framework for optimizing the detection time problem consists of three core steps:

(1) *Input of historical data*: This stage involves acquiring and preprocessing the required data for the proposed solution. First, potential passive-seismic source locations, including wells, faults, and tectonic stress concentration zones, are identified as the region of interest using available exploration records. The velocity model is then defined using data collected from these records. Finally, potential sensor candidate positions are selected from a grid on the region's surface, which is available for sensor placement.
(2) *Passive-seismic event simulation*: This component of the framework involves a physics-based simulation that models passive-seismic events wave propagation in the subsurface across the CO_2 storage site using a wave propagation model based on ray-tracing. Multiple events are simulated to account for the uncertainty in passive-seismic event locations.
(3) *Mixed-integer linear programming*: The passive-seismic events simulated in this step follow the same data structure as the one used in the initial formulation of the P-media model for the problem of methane sensor placement (Klise et al., 2017). A standard MILP solver can solve this optimization problem. We utilize Chama, an open-source sensor placement optimization software implemented in Python (Klise et al., 2017). Chama formulates the optimization problem using Pyomo (Hart et al., 2017) and solves it using the open-source GNU linear programming kit (GLPK) (Makhorin, 2000).

Our proposed solution framework for the source localization optimization formula involves the same data preparation step as the previous formula. However, due to the bi-level nature of this formula, a standard optimization tool cannot be used directly to solve the problem. Therefore, se put forward a methodology to resolve both the localization inversion problem and the sensor placement optimization problem in an iterative scheme, as illustrated in Fig. 1. The two core steps of this solution scheme are:

(1) *Passive-seismic source localization inversion*: To evaluate the localization performance with a given trial sensor network design, we employ a heuristic inversion optimization method. Specifically, we use the particle swarm optimization (PSO) algorithm, an

Table 2
Sensors.

Sensors	Type	Sensitivity
Point detector	Ideal Seismometer	2000 V/m/s
Geophone-LGT4.5	Moving-coil	28.8 V/m/s
Broadband Seismometer-Geotech KS-1	Inertial	2400 V/m/s

open-source optimization method (de Rosa and Papa, 2019; Kennedy and Eberhart, 1995).

(2) *Sensor network design optimization*: We also utilize the open-source optimization method particle swarm optimization (PSO) algorithm (de Rosa and Papa, 2019; Kennedy and Eberhart, 1995) for this purpose. Each step's updated solution would be input to the inversion solver as the input and be assessed the localization performance of each proposed placement strategy. After the performance reach the convergency criteria, the iteration stops, and the system provides the final optimal solution.

Algorithm 2 Stochastic passive-seismic sensor placement source localization optimization algorithm

Input subsurface velocity model, events positions, and strata layers: v_p, s_e and l_{depths}
Set the best placement to empty list: x_{best}
Set min objective value to infinity: $obj_{min} = inf$
Initialize sensor cost and sensor budget: $sensor_{cost}$ and c
for *each* y_l *(combination of* c *candidate locations)* **do**
  Initialize events impact to empty list: imp_{list}
  # evaluation of placement strategy
  for *each* s_e *event position of events* **do**
    # Forward
    Simulate the ground truth observation signal using the forward model: $t = \mathcal{F}(D, v, s_e)$ (observation signal is first arrival detection time list of current sensor network)
    # Inversion
    Initial localization inversion solution space (boundary of possible source area)
    Set the inversion initial solution
    for *each inversion iteration* **do**
      Simulate data using the current sensor placement and current inversion solution: $t = \mathcal{F}(y, v, \hat{s}_e)$
      # evaluation inversion optimization convergence criteria
      if $MSE(t_{obs}, t) < \epsilon$ **then**
        $imp = MSE(s'_e, \hat{s}_e)$
        Append simulated impact to imp_{list}
      end
    end
  end
  Initialize event probability: p_e
  # calculate objective value: localization expected errors
  $obj = \sum_{e \in \mathcal{E}_e} p_e MSE(s'_e, \hat{s}_e)$
  if $obj < obj_{min}$ **then**
    Update min objective value and best placement
  end
end
Result: best placement

4. Experiment

The empirical analysis presented in the following simulation section is focused on the Kimberlina site. A site velocity model is constructed using geological strata and historical seismic surveys (Walter and Mooney, 1987; Birkholzer et al., 2011). This case study provides a representative passive-seismic detection scenario for the advancement and

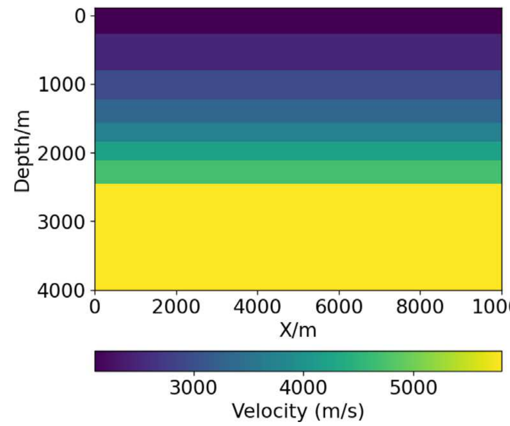

Fig. 3. Model of P-wave velocity at the Kimberlina site.

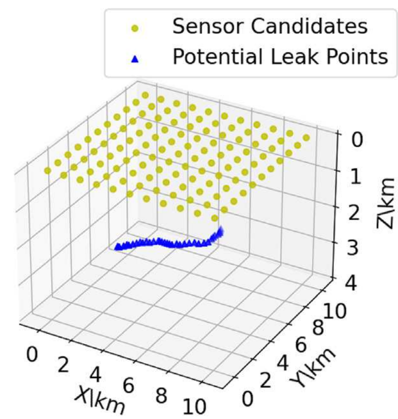

Fig. 4. The locations of sensor candidates (yellow dots) and passive-seismic events (blue triangles) used in the synthetic study.

assessment of diverse sensor placement optimization techniques. The simulation and optimization of this case study are conducted using the ray-tracing-based passive-seismic simulation model (Cervenỳ, 2001), the sensor placement optimization software called Chama, which is open-source (Klise et al., 2020), and the open-source heuristic solver Opytimizer (de Rosa and Papa, 2019)

4.1. Sensors discussion

The sensor utilized in our numerical simulation experiment is an ideal point sensor, and its characteristics are detailed in Table 2. We have also included two additional sensor types as practical references. However, for the purposes of this theoretical research, we exclusively consider standardized ideal sensors with exceptionally high sensitivity. Among the cost-effective sensors commonly employed in various environments is the moving-coil geophone (Zhang and Hu, 2010), exemplified by the LGT4.5 model, which boasts lower sensitivity. Alternatively, there exists a more advanced option in the form of the Broadband Seismometer (Ackerley et al., 2014), characterized by heightened sensitivity. An instance of such a high-sensitivity point sensor is the Geotech KS-1.

4.2. Simulations setting

Fig. 4 depicts a simulation volume in three dimensions (3D). The simulation domain is a cubic region with dimensions 10,000 m × 10,000 m × 4000 m. This area encompasses 35 potential source points (blue triangles), which correspond to locations where passive-seismic

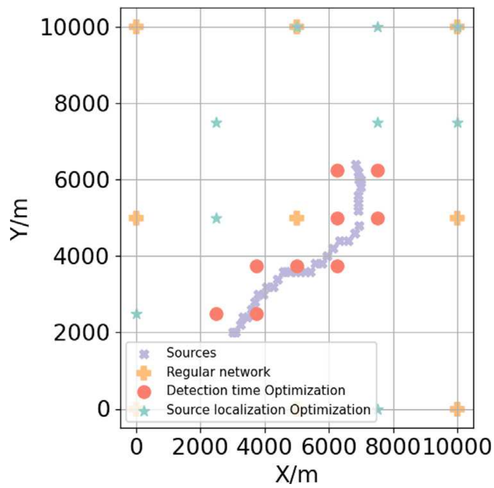

Fig. 5. The sensor placements obtained using two different approaches: stochastic optimization and regular grid methods.

Table 3
Localization testing.

Methods	Mean square error
Stochastic optimized network	95.257 m
Regular grid network	324.318 m

events could occur in the target carbon injection reservoir at an approximate depth of 2100 m. These points represent faults and fractures that may release carbon plumes. Candidate sensor locations at the surface (yellow circles) are represented as grid points where sensors can be installed for monitoring purposes.

In the physical simulation component, we utilized the Ray Tracing model (Cervenỳ, 2001) model for the passive-seismic waves' propagation. We performed the first arrival time picking processing using the modified energy ratio (MER) method (Han, 2010). We simulated a single passive-seismic event scenario (one event source point) for a 5-second record using a 2-millisecond time sampling step with a Ricker wavelet as the source wavelet, for each passive-seismic event. If an event failed to be detected, we set its missing impact to 50 s. To provide an example of a passive-seismic event, Fig. 2 shows the top view of the event measurement scenario and determines the initial arrival time of the passive-seismic wave in proximity to the source location. The time at which the candidate sensor detects the signal is influenced by the source positions, the geology model, the sensor's positions, and the velocity model, according to the physical model.

To capture the uncertainty of passive-seismic event source locations, we simulated multiple events as a set. By increasing the number of simulation scenarios and event diversity, we can capture the distribution of all uncertainties based on the law of large numbers. In this case study, we generated a passive-seismic event set with 35 sources based on the fixed Kimberlina velocity model and layered geological model (see Fig. 3). These source locations were sampled from the Pond-Poso Creek fault shape, one of the major faults in the area, at the target reservoir depth of 2100 m, as shown in Fig. 4.

In this case study, we defined candidate sensor locations by employing a uniform grid across the entire model domain. The grid had a spacing of 1000 m in both the x-axe and y-axe, 1 m along the z-axis, commencing from a below-ground elevation of -2 m. This setting results in 121 candidate sensors arranged in an 11×11 grid. The overall budget allocated to sensors is defined as 9000, with each sensor having an ideal active threshold cost of 1000.

4.3. Results discussion

In this section, we engage in a comprehensive discussion to assess the effectiveness of sensor placements derived through various methodologies in the context of passive seismic event monitoring. Our primary focus centers on the comparison between our innovative stochastic sensor placement network and the conventional regular grid sensor network. This comparative analysis encompasses two critical facets: the optimization of detection time and the enhancement of source localization accuracy. Our observations, as depicted in Fig. 6, shed light on the performance contrast as we progressively allocate a larger sensor budget. Notably, the stochastic optimal sensor placement consistently outshines its regular grid counterpart. This performance edge demonstrates the practical viability and efficiency of our approach. Furthermore, our deliberations, encapsulated in Table 3, emphasize the significance of source localization precision. Here, too, our stochastic optimization method emerges as the victor, clearly surpassing the performance metrics of the conventional regular grid sensor placement strategy. Drawing inspiration from the insights gleaned from state-of-the-art research, particularly the work of Chen and Huang (2020), we delve into the notion of the optimal trade-off point. In this context, our findings provide compelling evidence that our stochastic optimization method enhances this pivotal point, thereby offering a notably improved performance, as visually conveyed in Fig. 6. For instance, when we allocate a budget of nine sensors, our stochastic optimization method yields an expected minimal detection time of 0.8 s, a substantial 20% improvement compared to the regular grid method's 1.046 s. As we further scrutinize our results, we observe a consistent trend aligning with existing literature. The relationship between the number of seismic stations and location accuracy exhibits varying degrees of improvement, with the margin of performance enhancement diminishing as more sensors are incorporated into the network. In addition to our extensive evaluation of detection time optimization, we extend our analysis to assess the localization optimization performance of our proposed method. The results, as presented in Table 3, unequivocally affirm the superiority of our stochastic approach. With the same allocation of nine sensors, our network generates an impressive localization accuracy of 95.257 m, while the regular grid network trails behind with a substantially higher localization error of 324.318 m. To visually encapsulate our findings, Fig. 5 offers a compelling representation of the sensor networks, highlighting the efficacy of our stochastic method for both detection time optimization and source localization when juxtaposed with the conventional regular grid sensor network. The sensor placements obtained using two different approaches, stochastic optimization, and regular grid methods, are evaluated to satisfy the same detection time requirement of less than 0.81 s. Our analysis reveals that the stochastic optimization method while meeting the identical detection time criteria, leads to a reduced sensor budget with only 4 sensors required while 25 are required for the grid method as shown in Fig. 7. This result demonstrates the cost-saving potential of the stochastic optimization approach in comparison to the conventional regular grid method. These collective insights underscore the remarkable potential and practicality of our innovative sensor placement strategy in the domain of passive seismic event monitoring.

5. Conclusion

We devised a stochastic optimization approach to design an optimal seismic network for monitoring passive-seismic events during CO2 injection and storage. This is the first formal math formulation for a data-driven sensor placement solution for a passive-seismic sensor placement problem. Our methodology optimizes the expected earliest detection time or expected localization accuracy by designing optimized seismic stations surrounding the area of interest at the surface, where the injection and storage of CO2 may trigger passive-seismic events, potentially leading to severe consequences. The seismic sensor number

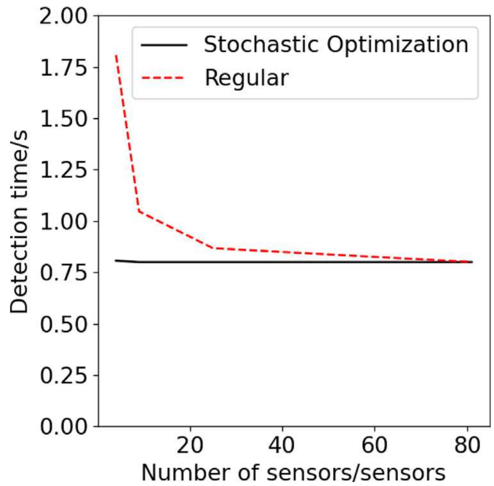

Fig. 6. A comparison between the total number of seismic stations and the expected detection time using two different methods. The black lines represent the outcomes obtained with the stochastic optimization method, while The red lines correspond to the results obtained using regular grid sensor networks, both using P-wave arrival times.

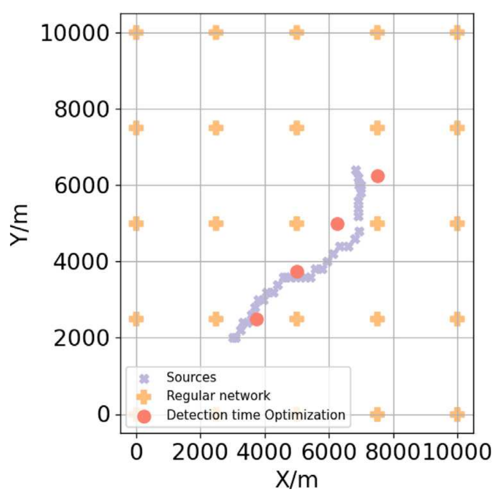

Fig. 7. The sensor placements obtained using two different approaches to satisfy the same detection time requirement: stochastic optimization and regular grid methods.

is altered to determine the earliest detection time via first arrival travel times. Our experiments demonstrate that the stochastic optimization placement strategy has a relatively robust performance compared to a regularly-spaced sensor network when varying the sensor numbers. In specific monitoring regions, adjusting the sensor location can further improve localization accuracy to ensure high monitoring performance with the most efficient cost. In this showcase, we make the assumption that all interesting seismic events can be accurately captured within the experimental region and that the first-arrival time picking is ideal. Our proposed methodology Could be utilized for different passive-seismic monitoring sites with more sophisticated passive-seismic event simulation tools.

6. Future work

Recent advances in carbon capture and storage (CCS) propose injecting CO2 into deep marine sediments, forming stable CO2 hydrates for long-term storage. This introduces the need to monitor CO2 leakage and seismic events during multiphase injection. Integrating this background into our introduction, citing relevant references, is planned (Liu et al., 2023; Ren et al., 2022).

Our future work will prioritize addressing geological uncertainty. Current research lacks real-world geological complexity. We aim to incorporate geological uncertainty, drawing from recent advanced drilling studies (Caers et al., 2022; Hall et al., 2022) and Bayesian approaches (Thibaut et al., 2022). It is also interesting to utilize advanced machine learning algorithms like reinforcement learning to solve more complex mixed integer programming when scaling up the problem and facing nonlinear high-dimension issues Sang et al. (2022), Tsai et al. (2022).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Zhu Han, Lei Fan reports financial support was provided by National Science Foundation.

Data availability

Data will be made available on request.

Acknowledgments

This work is supported by NSF CNS-2107216, NSF-EPCN-2045978, CNS-2128368, CMMI-2222810, ECCS-2302469, US Department of Transportation, Toyota and Amazon.

References

Ackerley, N., Beer, M., Kougioumtzoglou, I., Patelli, E., Au, S., 2014. Principles of broadband seismometry. In: Encyclopedia of Earthquake Engineering. Springer, Berlin.
Benavides-Serrano, A., Mannan, M., Laird, C., 2016. Optimal placement of gas detectors: AP-median formulation considering dynamic nonuniform unavailabilities. AIChE J. 62 (8), 2728–2739.
Berry, J.W., Fleischer, L., Hart, W.E., Phillips, C.A., Watson, J.-P., 2005. Sensor placement in municipal water networks. J. Water Resour. Plan. Manag. 131 (3), 237–243.
Birkholzer, J., Zhou, Q., Cortis, A., Finsterle, S., 2011. A sensitivity study on regional pressure buildup from large-scale CO2 storage projects. Energy Procedia 4, 4371–4378.
Boullenger, B., Verdel, A., Paap, B., Thorbecke, J., Draganov, D., 2015. Studying CO2 storage with ambient-noise seismic interferometry: A combined numerical feasibility study and field-data example for Ketzin, Germany. Geophysics 80 (1), Q1–Q13.
Caers, J., Scheidt, C., Yin, Z., Wang, L., Mukerji, T., House, K., 2022. Efficacy of information in mineral exploration drilling. Nat. Resour. Res. 31 (3), 1157–1173.
Cerveňỳ, V., 2001. Seismic Ray Theory, Vol. 110. Cambridge University Press.
Chen, T., Huang, L., 2020. Optimal design of microseismic monitoring network: Synthetic study for the Kimberlina CO2 storage demonstration site. Int. J. Greenhouse Gas Control 95, 102981.
Ellsworth, W.L., 2013. Injection-induced earthquakes. Science 341 (6142), 1225942.
Goertz-Allmann, B., Jordan, M., Bauer, R., Oye, V., Greenberg, S., 2017. Integrating active with passive seismic data to best constrain CO2 injection monitoring. In: EAGE/SEG Research Workshop 2017. European Association of Geoscientists & Engineers, pp. cp–522.
Hall, T., Scheidt, C., Wang, L., Yin, Z., Mukerji, T., Caers, J., 2022. Sequential value of information for subsurface exploration drilling. Nat. Resour. Res. 31 (5), 2413–2434.
Han, L., 2010. Microseismic Monitoring and Hypocenter Location (MS thesis). University of Calgary.
Hart, W.E., Laird, C.D., Watson, J.-P., Woodruff, D.L., Hackebeil, G.A., Nicholson, B.L., Siirola, J.D., et al., 2017. Pyomo-Optimization Modeling in Python, Vol. 67. Springer.
Hu, Y., Wei, X., Wu, X., Sun, J., Chen, J., Huang, Y., Chen, J., 2023. A deep learning-enhanced framework for multiphysics joint inversion. Geophysics 88 (1), K13–K26.
Jin, Y., Hu, W., Wang, S., Zi, Y., Wu, X., Chen, J., 2021. Efficient progressive transfer learning for full-waveform inversion with extrapolated low-frequency reflection seismic data. IEEE Trans. Geosci. Remote Sens. 60, 1–10.
Karatas, M., Razi, N., Tozan, H., 2016. A comparison of p-median and maximal coverage location models with Q–coverage requirement. Procedia Eng. 149, 169–176.
Kaven, J.O., Hickman, S.H., McGarr, A.F., Ellsworth, W.L., 2015. Surface monitoring of microseismicity at the Decatur, Illinois, CO2 sequestration demonstration site. Seismol. Res. Lett. 86 (4), 1096–1101.

Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proceedings of IEEE ICNN'95-International Conference on Neural Networks, Vol. 4. Perth, WA, Australia, pp. 1942–1948.

Kijko, A., 1977. An algorithm for the optimum distribution of a regional seismic network—I. Pure Appl. Geophys. 115, 999–1009.

Klise, K.A., Nicholson, B.L., Laird, C.D., 2017. Sensor Placement Optimization Using Chama. Tech. Rep. 146, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

Klise, K.A., Nicholson, B.L., Laird, C.D., Ravikumar, A.P., Brandt, A.R., 2020. Sensor placement optimization software applied to site-scale methane-emissions monitoring. J. Environ. Eng. 146 (7), 04020054.

Legg, S., Benavides-Serrano, A., Siirola, J.D., Watson, J.-P., Davis, S., Bratteteig, A., Laird, C.D., 2012. A stochastic programming approach for gas detector placement using CFD-based dispersion simulations. Comput. Chem. Eng. 47, 194–201.

Ling, Z., Hu, F., Zhang, H., Han, Z., 2022. Age of information minimization in healthcare IOT using distributionally robust optimization. IEEE Internet Things J. 1.

Liu, X., Li, Y., Chen, G.-J., Chen, D.-Y., Sun, B., Yin, Z., 2023. Coupling amino acid with THF for the synergistic promotion of CO2 hydrate micro kinetics: Implication for hydrate-based CO2 sequestration. ACS Sustain. Chem. Eng. 11 (15), 6057–6069.

Liu, Y., Zhang, H., Di, B., Wu, J., Han, Z., 2022. Deployment for high altitude platform systems with perturbation: Distributionally robust optimization approach. IEEE Commun. Lett. 26 (5), 1126–1130.

Makhorin, A., 2000. GLPK (GNU Linear Programming Kit), Available: https://www.gnu.org/software/glpk.

Maxwell, S.C., Urbancic, T.I., 2001. The role of passive microseismic monitoring in the instrumented oil field. Lead. Edge 20 (6), 636–639.

Miyazawa, M., Venkataraman, A., Snieder, R., Payne, M.A., 2008. Analysis of microearthquake data at Cold Lake and its applications to reservoir monitoring. Geophysics 73 (3), O15–O21.

Rabinowitz, N., Steinberg, D.M., 1990. Optimal configuration of a seismographic network: a statistical approach. Bull. Seismol. Soc. Am. 80 (1), 187–196.

Ren, J., Liu, X., Niu, M., Yin, Z., 2022. Effect of sodium montmorillonite clay on the kinetics of CH4 hydrate-implication for energy recovery. Chem. Eng. J. 437, 135368.

de Rosa, G.H., Papa, J.P., 2019. Opytimizer: A nature-inspired python optimizer, CoRR abs/1912.13002. Available: http://arxiv.org/abs/1912.13002.

Sang, J., Sun, H., Kou, L., 2022. Deep reinforcement learning microgrid optimization strategy considering priority flexible demand side. Sensors 22 (6), 2256.

Steinberg, D.M., Rabinowitz, N., Shimshoni, Y., Mizrachi, D., 1995. Configuring a seismographic network for optimal monitoring of fault lines and multiple sources. Bull. Seismol. Soc. Am. 85 (6), 1847–1857.

Stork, A., Nixon, C., Hawkes, C., Birnie, C., White, D., Schmitt, D., Roberts, B., 2018. Is CO2 injection at Aquistore aseismic? A combined seismological and geomechanical study of early injection operations. Int. J. Greenhouse Gas Control 75, 107–124.

Thibaut, R., Compaire, N., Lesparre, N., Ramgraber, M., Laloy, E., Hermans, T., 2022. Comparing well and geophysical data for temperature monitoring within a Bayesian experimental design framework. Water Resour. Res. 58 (11), e2022WR033045.

Tsai, K.-C., Yao, T.-J., Huang, P.-H., Huang, C.-S., Han, Z., Wang, L.-C., 2022. Deep reinforcement learning-based routing for space-terrestrial networks. In: 2022 IEEE 96th Vehicular Technology Conference. VTC2022-Fall, IEEE, pp. 1–5.

Verdon, J.P., Kendall, J.-M., White, D.J., 2012. Monitoring carbon dioxide storage using passive seismic techniques. Proc. Inst. Civ. Eng.-Energy 165 (2), 85–96.

Walter, A., Mooney, W., 1987. Interpretations of the SJ-6 seismic reflection/refraction profile, south-central California. USA: US Geol. Surv. Open-File Rep. 87 (73), 132.

Yin, Z., Louboutin, M., Herrmann, F.J., 2021. Compressive time-lapse seismic monitoring of carbon storage and sequestration with the joint recovery model. In: SEG International Exposition and Annual Meeting. SEG, D011S145R001.

Younis, M., Akkaya, K., 2008. Strategies and techniques for node placement in wireless sensor networks: A survey. Ad Hoc Netw. 6 (4), 621–655.

Zhang, G., Hu, S., 2010. Dynamic characteristics of moving-coil geophone with large damping. Int. J. Appl. Electromagn. Mech. 33 (1–2), 565–571.

Zi, Y., Fan, L., Wu, X., Chen, J., Han, Z., 2022. Distributionally robust optimal sensor placement method for site-scale methane-emission monitoring. IEEE Sens. J. 22 (23), 23403–23412.