processes

MDPI

Article

Optimizing Sensor Placement for Enhanced Source Term Estimation in Chemical Plants

Hao Tian [1], Ziqiang Lang [1,2], Chenxi Cao [1] and Bing Wang [1,*]

[1] Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China; y30210937@mail.ecust.edu.cn (H.T.); z.lang@sheffield.ac.uk (Z.L.); caocx@ecust.edu.cn (C.C.)
[2] Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK
* Correspondence: wangb07@ecust.edu.cn

Abstract: The leakage of hazardous chemical gases in chemical plants can lead to severe consequences. Source term estimation (STE) algorithms are effective in locating the leak source. The layout of the sensor network significantly affects the performance of the STE algorithm, yet the underlying mechanism remains unclear. In this study, we first applied computational fluid dynamics (CFD) to simulate 160 hazardous chemical gas leakage scenarios under multi-directional wind conditions in two hypothetic scenes with a natural convection environment, creating an accident dataset. Subsequently, a mathematical model for sensor placement optimization was developed and applied to the dataset to generate a series of sensor layout solutions. Based on these layouts, 12,216 STE cases were calculated. By analyzing the error distribution of these cases, the relationship between sensor placement and STE performance was systematically investigated, and the most effective sensor layout optimization strategies were discussed. This study found that in scenarios with complex obstacles, increasing the average measured concentration of the sensor network can significantly reduce the errors in the STE algorithm.

Keywords: unexpected gas leak; sensor placement optimization; source term estimation; Bayesian inference; adjoint equation; simulated annealing

check for updates

Academic Editor: Zhibin Lin

Received: 22 January 2025
Revised: 23 February 2025
Accepted: 26 February 2025
Published: 12 March 2025

Citation: Tian, H.; Lang, Z.; Cao, C.; Wang, B. Optimizing Sensor Placement for Enhanced Source Term Estimation in Chemical Plants. *Processes* **2025**, *13*, 825. https://doi.org/10.3390/pr13030825

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In chemical plants, the unexpected release and dispersion of hazardous chemical materials pose a significant threat to both human health and the natural environment. For example, in India, the release of styrene chemical plants caused the hospitalization of more than 1000 individuals [1]. Gas sensors are commonly installed in chemical plants to detect gas leakages and prevent the escalation of accidents. These sensors are expected to (a) promptly detect the leak accidents and (b) provide preliminary estimations of the source location. However, uncertainty in the location of the leak source and the variation in wind conditions often lead to diverse leak consequences, presenting challenges to the effective detection of sensor networks. The determination of sensor locations is sophisticated and crucial.

The sensor placement optimization (SPO) problem aims at figuring out the optimal sensor placement with an objective under certain constraints, which are primarily the number of sensors and the candidate sensor locations. The objective is mainly influenced by practical requirements. Until now, various SPO models have been proposed with different objectives. In [2], the authors constructed a database comprising consequences resulting from various leak sources and proposed a mixed-integer linear programming (MILP) SPO

model with minimal expected detection time as the objective. Ref. [3] incorporated the unavailability of sensors and voting strategies into the MILP model to better align the method with real-world scenarios. Some other risk factors related to gas leaks such as minimal total gas concentration [4], minimal individual risk [4], minimal cumulative death probability [5], and minimal risk value [6] have also been considered as objectives of SPO in chemical plants. In addition to optimization algorithms, the development of gas sensor designation (e.g., chemoresistive gas detectors [7]) and the gas phase impurities measurement theorem [8] have also significantly enhanced the performance of hazardous gas detection in real scenarios.

Source term estimation (STE) aims to determine the parameters of the leak source, which typically include the location and the emission strength. The STE problem is recognized as an ill-posed inverse problem with high non-linearity [9]. Generally, there are two mainstream approaches in STE [10]: optimization-based methods and Bayesian inference. Optimization-based approaches usually define a cost function by evaluating the discrepancy between concentration values measured by sensors and those computed by a forward dispersion model. A single optimal solution that minimizes this cost function is determined as the result of STE. In contrast, the Bayesian inference STE evaluates the posterior probability of different source terms. By means of stochastic sampling methods, the posterior probability density function (PDF) of source terms is drawn. The outcome of Bayesian inference STE is the most probable parameters in terms of the posterior PDF. For either of these two methods, the source–receptor relationship is a vital component that provides the capability of quickly predicting the concentration measurements of sensors when a certain source appears. Some popular methods for modelling the source–receptor relationship include the Gaussian plume model, the advection–diffusion equation with its adjoint equation [11], and data-driven models [12]. Besides optimization-based methods and Bayesian inference, some studies have also developed STE algorithms based on deep learning. Ref. [13] developed an STE algorithm based on a convolutional neural network and concentration distribution images. Ref. [14] proposed a federated STE framework inspired by federated learning. The efficiency of these algorithms has been validated under chemical plant scenes.

In most cases, the concentration measurements of gas sensors form the foundation of STE. A well-designed sensor placement layout can effectively provide spatial–temporal information of the gas concentrations, thereby enhancing the accuracy of STE. Previous research has highlighted the influence of sensor placement on STE. Ref. [15] compared the STE errors among 10 different sensor placements in an urban neighborhood testing scene. Ref. [16] proposed an SPO model leveraging the concept of information entropy and demonstrated that the sensor group with higher capability of gathering information performs better on STE. These works extensively investigated the impact of sensor placement on STE. However, it is regrettable that the scales of the test scenes in these works are relatively small, and that validation has not been conducted in complex environments such as chemical plants.

Given the variability in wind condition and the presence of multiple potential leak sources in chemical plants, this research develops an SPO framework with the goal of detecting all potential leak scenarios and improving STE accuracy. The main objectives of this study focus on two aspects: Firstly, an analytical framework combining SPO and STE is established, which proposes a basis for assessing whether sensor placement can achieve accurate STE in a chemical plant. Secondly, an analysis regarding the correlation between sensor placement and STE accuracy is conducted, along with an exploration of the most suitable SPO strategy for STE.

The remaining sections of this paper are organized as follows: Section 2 introduces the analytical framework, the Bayesian inference STE, and the proposed SPO models. Section 3 explains the parameters of the numerical experiment. Section 4 exhibits the results of SPO and STE from the experiments and analyzes the impact of sensor placement on STE errors. Finally, conclusions and future work are presented in Section 5.

2. Methodology

This chapter begins by presenting the overall analytical framework of the methodology. Subsequently, the STE method employed in this study is introduced, and a detailed exposition of the proposed SPO model with several objectives is given.

2.1. Analytical Framework

Given the layout of a scene, this framework aims to find a rational SPO strategy for gas detection and enhanced STE utilizing computational fluid dynamics (CFD). The framework encompasses leak scenario identification, data preparation, SPO, and STE, as visually illustrated in Figure 1. The main processes listed as follows:

(1) Two test scenes are designed, in which variant potential leak sources and wind directions are defined to generate a range of gas leak scenarios. Subsequently, essential datasets of wind fields and gas concentration distributions for these scenarios are collected through CFD simulation. These datasets serve as the foundation for SPO and STE.

(2) Drawing on previous research, this study identifies two factors that might influence the STE accuracy and designs three SPO objectives accordingly. The mathematical model of SPO is solved independently for these objectives to obtain rational sensor placements.

(3) Based on the sensor placements optimized by SPO, STE is conducted for all scenarios. These STE results are summarized, and the impact of two factors identified in (2) on STE errors along with the optimal SPO objective are analyzed.

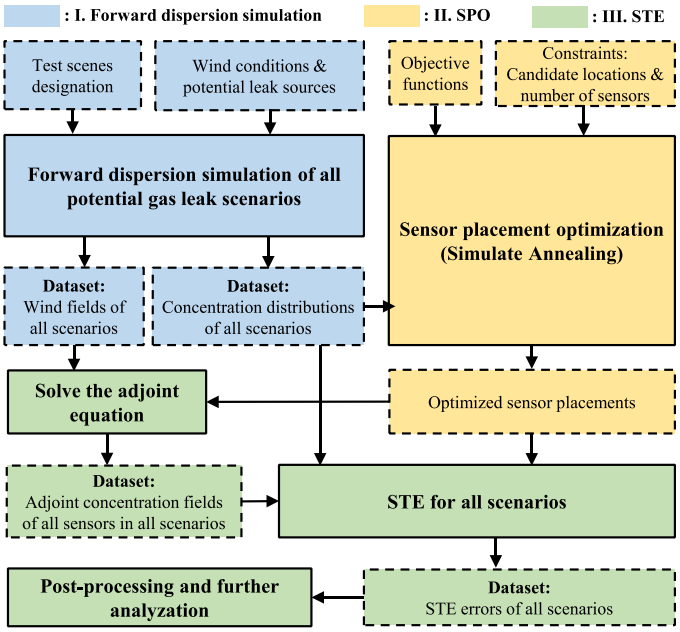


Figure 1. The complete workflow of this research, which can be divided into three parts: I. forward dispersion simulation, II. SPO process, and III. STE process.

2.2. Bayesian Inference STE

Bayesian inference STE with an adjoint equation, which is chosen as the STE method in this research, is briefly introduced here. This method comprises three components: Bayes' theorem, the source–receptor relationship (adjoint equation), and the stochastic sampling method.

A specified leak source can be expressed as

$$\mathbf{s} = (\mathbf{x}, q) \tag{1}$$

where **x** = (*x*,*y*,*z*) and q are the location and the emission strength of the source, respectively. The Bayesian inference STE attempts to draw the PDF of these parameters.

2.2.1. Bayes' Theorem

Bayes' theorem deduces the posterior probability of source parameters [11]. When the vector $\mathbf{D} = (D_i)$—which denotes the concentration measurements of all sensors, with i indicating the index for each sensor—the probability that **s** is the true source can be expressed as

$$p(\mathbf{s}|\mathbf{D}, I) = \frac{p(\mathbf{s}|I)p(\mathbf{D}|\mathbf{s}, I)}{p(\mathbf{D}|I)} \propto p(\mathbf{s}|I)p(\mathbf{D}|\mathbf{s}, I) \tag{2}$$

where the prior probability $p(\mathbf{s}|I)$ actually signifies the probability distribution of a leak source among the space domain. With $p(\mathbf{s}|I)$ chosen as a uniform distribution and the likelihood function $p(\mathbf{D}|\mathbf{s}, I)$ the Gaussian form, the posterior probability can be expressed as

$$p(\mathbf{s}|\mathbf{D}, I) \propto p(\mathbf{D}|\mathbf{s}, I) \propto \exp\left[-\frac{1}{2}\sum_i \frac{(D_i - R_i)^2}{\sigma_i^2}\right] \tag{3}$$

where R_i is the modelled concentration calculated through the source–receptor relationship model and σ_i^2 is the variance of the errors between D_i and R_i. In the STE algorithm, σ_i^2 presents an inverse correlation with the concentration of posterior PDF curve generated by stochastic sampling. According to previous research and performance in numerical experiments, this research sets the variance as $\sigma_i^2 = 0.1 \times D_i$.

2.2.2. Source–Receptor Relationship

The source–receptor relationship constructs a correlation between the leak source and the measured concentration of sensors. The adjoint equation, deduced from the advection–diffusion equation, is employed here for its adaptability to environments with multiple obstacles.

In a three-dimensional spatial domain Ω, consider a point source Q that continuously releases incompressible gas phase material at a steady rate q_s over time of $[0, T]$:

$$Q = q_s\delta(\mathbf{x} - \mathbf{x}_s)[H(t - T) - H(t - 0)] \tag{4}$$

the general form of advection–diffusion equation is

$$\begin{aligned} &\tfrac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla C - \nabla \cdot (K\nabla C) = Q \\ &s.t. \\ &\nabla_{\mathbf{n}} C = 0 \quad at\ \partial\Omega \\ &C(\mathbf{x}, t = 0) = \mathbf{0} \end{aligned} \tag{5}$$

Here, $\delta(\cdot)$ and $H(\cdot)$ are Dirac delta function and Heaviside unit step function, respectively. $C(\mathbf{x}, t)$ is the gas concentration distribution over the space–time domain $\Omega \times [0.T]$, **u**

is the wind velocity field, K is the eddy diffusivity, $\partial\Omega$ is the spatial domain boundary, and $\nabla_{\mathbf{n}}$ is a directional derivative normal to the boundary. The adjoint equation can be derived from Equation (5) as follows [17]:

$$-\frac{\partial C^*}{\partial t} - \mathbf{u} \cdot \nabla C^* - \nabla \cdot (K \nabla C^*) = h = \delta(\mathbf{x} - \mathbf{x_i}) \tag{6}$$

where $h = \delta(\mathbf{x} - \mathbf{x}_i)$ represents a sensor located at $\mathbf{x}_i$ and $C^*(\mathbf{x}, t)$ is the adjoint concentration (AC) field generated by this sensor. The steady-state advection–diffusion equation and the adjoint equation can be illustrated as follows:

$$\overline{\mathbf{u}} \cdot \nabla \overline{C} - \nabla \cdot \left(K \nabla \overline{C}\right) = Q \tag{7}$$

$$-\overline{\mathbf{u}} \cdot \nabla \overline{C^*} - \nabla \cdot \left(K \nabla \overline{C^*}\right) = h = \delta(\mathbf{x} - \mathbf{x_i}) \tag{8}$$

where $\overline{\mathbf{u}}$ is the mean velocity field obtained by taking the time average of $\mathbf{u}$, $\overline{C(\mathbf{x})}$ is the mean concentration field, and $\overline{C^*(\mathbf{x})}$ is the mean AC field obtained in the same manner. The adjoint equation simulates a hypothetical inverse dispersion process with the sensor as the leak source, which generates the AC field over the spatial domain. The modelled concentration of the sensor is proportional to the AC at the source location:

$$R_i = q_s \cdot \overline{C_i^*(\mathbf{x_s})} \approx D_i \tag{9}$$

Here, $\overline{C_i^*(\mathbf{x_s})}$ is the AC value of sensor i at the leak source. This adjoint equation serves as an alternative to the advection–diffusion equation but significantly reduces the computational cost by transforming the differential equation into a multiplication operation. After obtaining the AC fields for each sensor, it becomes possible to calculate the modelled concentrations of the sensors corresponding to different source parameters. This capability is very beneficial for the sampling process as it needs to repeatedly calculate the posterior probabilities for a large number of source parameter samples.

2.2.3. Stochastic Sampling Method

Given measured concentration $\mathbf{D} = (D_i)$, for each specified source parameter, the posterior probabilities can be calculated by Equation (3). It is computationally expensive to calculate posterior probabilities for all possible source parameter combinations. The stochastic sampling process, which samples a different $\mathbf{x}$ values from the spaces Ω and q, is leveraged to draw the posterior PDF with lower computational cost. Markov chain Monte Carlo (MCMC) [18] is one of the most popular sampling methods for its simplicity and availability. In this research, an improved MCMC with affine invariance ensemble samplers [19] was utilized.

2.3. Sensor Placement Optimization

2.3.1. Problem Statement and Objectives

When a gas leak occurs in chemical plants, the source location and the wind direction significantly influence the spatial concentration distribution of the species. As mentioned in Section 2.1 and Figure 1, a dataset of concentration distributions of all leak scenarios with variant source locations and wind directions is prepared in advance. Based on the dataset, the goal of SPO in this study is twofold: ensure the sensor placement is capable of (I) detecting all leak scenarios and (II) enhancing STE accuracy, especially that of the source location. It is obvious that (I) is the foundation of (II), as STE cannot be conducted if no sensor is activated by gas plumes. The impact of sensor placement on STE is still ambiguous. Ref. [16] illustrates the potential existence of the positive correlation between the number of sensors and STE accuracy. Here, two factors, the measured concentrations of

sensors (MC) and the count of activated sensors (AS), are treated as experimental objectives of SPO. After SPO and STE, an analysis is conducted to explore the impact of these two factors on STE. Three SPO objectives are designed as follows:

Objective 1: Highest Mean Measured Concentration (HMC)

HMC attempts to place sensors at locations with higher average concentration across all leak scenarios by calculating a score proportional to MC. If there is no sensor activated in one or more scenarios, penalties will be added to the score to ensure that the sensor placement is capable of detecting all leak scenarios. The mathematical formulation of HMC is shown in Equation (10a).

Objective 2: Most Activated Sensors (MAS)

MAS seeks to enhance the mean value of AS over all scenarios. Similar to HMC, a score, incorporating penalties, is computed to evaluate the sensor placement. The mathematical formulation of MAS is illustrated in Equation (10b).

Objective 3: Most Activated Sensors with Higher Measured Concentration (MAS-MC)

Both HMC and MAS focus on one single factor during optimization. MAS-MC evaluates the performance of sensors on both MC and AS in all scenarios. The result of MAS-MC may not excel in a particular factor but will strike a balance in the two. The mathematical formulation of MAS is expressed in Equation (10c).

2.3.2. Mathematical Model

The mathematical model of SPO is illustrated in Equation (10). The constraints of the model are the total number and candidate locations of sensors, and the objectives are HMC, MAS, and MAS-MC.

$$\mathrm{HMC}: \max \overline{\left(\mathrm{C}_{\mathrm{mean}(:)}\right)_{0.75m}} \tag{10a}$$

$$\mathrm{MAS}: \max \left[\left(\min_{1\le j\le m} \mathrm{A}_j + 0.1\right) \cdot \overline{\left(\mathrm{A}_{(:)}\right)_{0.75m}}\right] \tag{10b}$$

$$\mathrm{MAS-MC}: \max \left[10^{\left(\min\limits_{1\le j\le m} \mathrm{A_j}\right)} \cdot \overline{\left(\mathrm{C}_{(:)}\right)_{0.75m}}\right] \tag{10c}$$

$$s.t.$$

$$\varepsilon_l \in \{0,1\} \quad \forall l \in L \tag{10d}$$

$$\sum_{l\in L} \varepsilon_l = n \tag{10e}$$

$$\alpha_{l,j} = \begin{cases} \varepsilon_l, \ \text{if}\, c_{l,j} \ge thres \\ 0, \ \text{otherwise} \end{cases} \tag{10f}$$

where:

$$\mathrm{X}_{(:)} = \ \text{sorted_ascending}\,(\mathrm{X}) = \ \left(\mathrm{X}_{(1)}, \ldots, \mathrm{X}_{(m)}\right) \tag{10g}$$

$$\overline{\left(\mathrm{X}_{(:)}\right)_{0.75m}} = \frac{\sum\limits_{j=1}^{\lceil 0.75m\rceil} \mathrm{X}_{(j)}}{\lceil 0.75m\rceil} \tag{10h}$$

$$\mathrm{C}_j = \sum_{l\in L}\left(\alpha_{l,j}\cdot c_{l,j}\right) \tag{10i}$$

$$\mathrm{A}_j = \sum_{l\in L} \alpha_{l,j} \tag{10j}$$

$$\mathrm{C}_{\mathrm{mean}j} = \begin{cases} \frac{\mathrm{C}_j}{\mathrm{A}_j}, \text{ if } \mathrm{A}_j > 0 \\ -c_{pnt}, \text{ if } \mathrm{A}_j = 0 \end{cases} \tag{10k}$$

The notation used in SPO model Equation (10) is summarized in Table 1. This model selects n optimal locations from $\mathbf{L}$ to install sensors with guidance of the objective function. Equation (10a–c) are mathematical expressions for three objectives. It should be noted that when calculating the score, these objectives only consider the average performance over scenarios ranking in bottom 75% as illustrated by Equation (10g,h). This strategy aims to highlight the overall performance on the particular objective over m scenarios and mitigate polarization. Furthermore, Equation (10b,c) also consider the worst performance among m scenarios, which bring penalty if polarization exists. In Equation (10c), the minimum value of AS over all scenarios appears as an exponent with a base of 10 to accommodate the magnitude of MC, as the order of magnitude of MC is typically around $10^{-3} \sim 10^{-5}$. When the minimum AS increases, the score of this objective increases by an order of magnitude of 10. Equation (10d,e) define the constraint on the total number of sensors. Equation (10f) determines whether a sensor located at l is activated in scenario j by comparing the concentration measurement with an activation threshold. This threshold is set to 1×10^{-6} kg/m^3. Equation (10i) calculates the sum of concentration measurements in scenario j, Equation (10j) counts the number of activated sensors in scenario j, and Equation (10k) assesses the average concentration measurement among activated sensors in scenario j.

Table 1. Problem notation.

Symbols	Meaning
n	Number of candidate sensors
m	Number of all potential leak scenarios
$\mathbf{L} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N}\}$	Set of candidate sensor locations
ε_l	Binary variable indicating if a sensor is installed at location l
$thres$	Detection threshold of sensors
$\alpha_{l,j}$	Binary variable indicating if the sensor at l is activated by gas leak in scenario j
A_j	Total number of activated sensors in scenario j
$c_{l,j}$	Concentration measurement at location l under scenario j
c_{pnt}	A penalty to concentration if no sensor is activated in a scenario
C_j	Sum of concentration measurement value of all activated sensors in scenario j
$\mathrm{C}_{\mathrm{mean}j}$	Average value of concentration measurement over all activated sensors in scenario j

In this model, the candidate locations $\mathbf{L}$ are determined by the scene layout, and the number of sensors n is specified artificially. In practical engineering applications, the necessary number of sensors can be determined according to [2]. In this research, to obtain adequate data for STE result analysis, the number of sensors is incrementally increased from the minimum number required to detect all scenarios. SPO is conducted individually for each objective and each specified number of sensors.

2.3.3. Simulate Annealing

This SPO model is solved by the simulated annealing (SA) algorithm [20]. In the algorithm, $maxf(\mathbf{p})$ is set as the objective function where $\mathbf{p} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x_n}\}$ is the vector of coordinates for n sensors, N is the maximum iteration, and W is the maximum refusion count for each iteration. Figure 2 shows the detailed process of the algorithm.

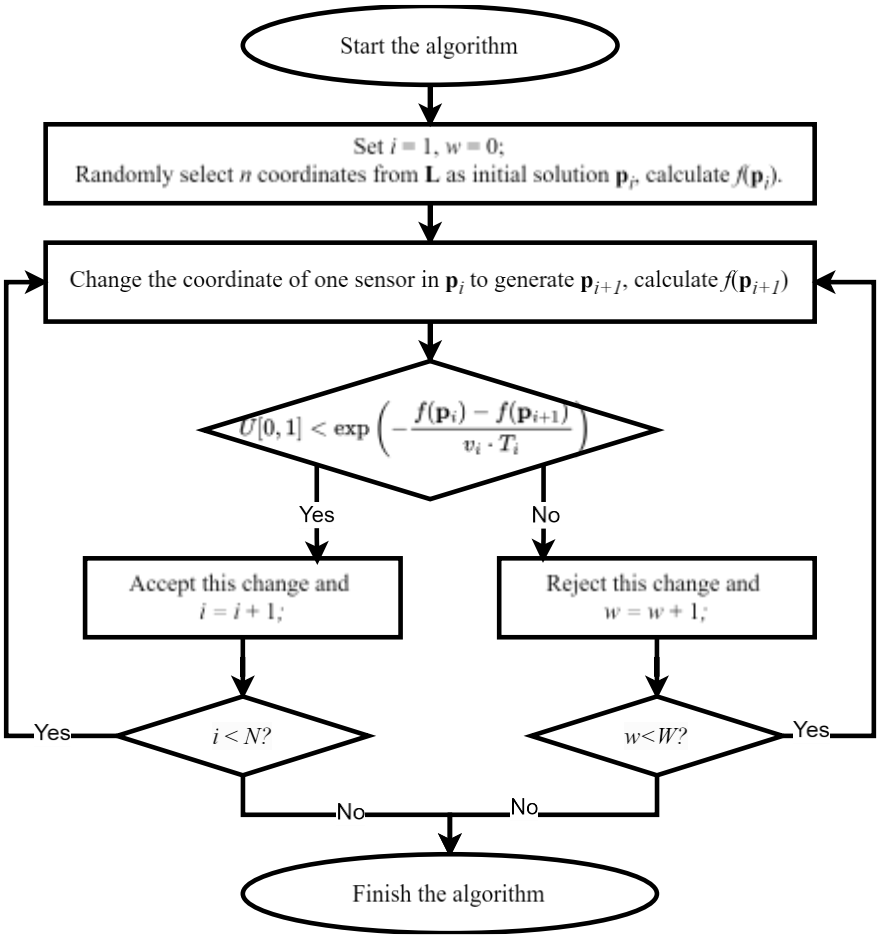


Figure 2. Block diagram of SA algorithm.

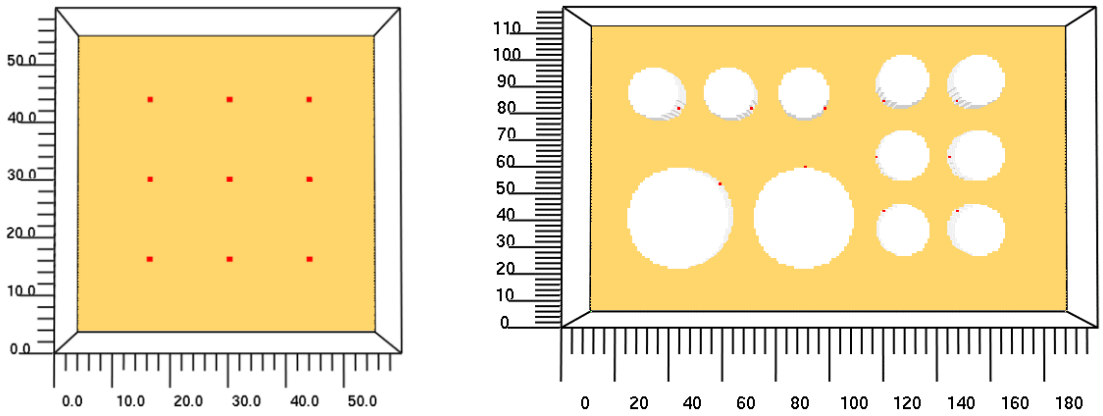
During one iteration in SA, if the number of unaccepted changes hits W times, it can be considered that no better solution than the current one could easily be found. $U[0, 1]$ is a random variable conforming to a uniform distribution between 0 and 1. T_i is the virtual temperature of ith iteration, which decreases gradually according to $T_i = T_{i-1} \cdot k$. This setting allows for the algorithm to accept a worse solution with a certain probability, which is beneficial to the global search. As i increases, T_i decreases and this probability tends to 0. In this research, the initial temperature T_0 is set to 5000, the cooling coefficient k is set to 0.9, N is set to 20,000, and W is set to 20,000.

3. Numerical Experiment Settings

3.1. Test Scenes and Leak Scenarios Design

Two scenes were designed for algorithm validation. The first scene, with flat terrain, denoted as *Scene1*, has a calculation domain measuring 60 m × 60 m × 20 m in length, width, and height, respectively. The primary aim of *Scene1* is to assess the algorithm's effectiveness; hence, this scene is smaller and contains no obstacles. Nine hypothetical leak sources are uniformly located inside *Scene1* at the height of $z = 10$ m. The layout of *Scene1* is shown in Figure 3a. The second scene, mimicking the layout of a chemical plant, is denoted as *Scene2*. The calculation domain of *Scene2* spans 180 m × 120 m × 30 m, which is significantly larger than that of *Scene1*. Eleven cylindrical storage tanks with heights of 20 m are located, and eleven hypothetical leak sources are defined on each storage tank at a height of $z = 10$ m in *Scene2*. Inhomogeneous wind fields would be formed under the impact of these obstacles. Figure 3b exhibits the layout of *Scene2*. The locations of leak sources in these scenes are listed in Table 2. The emission strength of all leak sources is

set to a constant 0.5 kg/(m^3·s). The leaked species is propane, which is a commonly used chemical with a molar mass slightly greater than that of air.



(**a**) (**b**)

Figure 3. Layouts of test scenes. The red points in each scene represent hypothetical leak sources. (**a**) Layout of *Scene1*. (**b**) Layout of *Scene2*.

Table 2. Locations of leak sources.

Scene1				*Scene2*			
Index	**x (m)**	**y (m)**	**z (m)**	**Index**	**x (m)**	**y (m)**	**z (m)**
1	15	15	10	1	38	84	10
2	30	15	10	2	68	84	10
3	45	15	10	3	98	84	10
4	15	30	10	4	55	53	10
5	30	30	10	5	90	60	10
6	45	30	10	6	122	87	10
7	15	45	10	7	152	87	10
8	30	45	10	8	119	64	10
9	45	45	10	9	149	64	10
				10	122	42	10
				11	152	42	10

Compared to wind speed, the direction of wind more significantly influences the concentration distribution when a leak source is activated. Constrained by computational resources, this study only generates scenarios under varying wind directions while the wind speed is constant. Eight wind directions are chosen uniformly within the range of 0° to 359°, namely, $\theta_k = (k-1) \cdot 45^\circ$, where θ_k is the kth direction of the wind. Here, the angle specifies where the wind blows from. A direction of 0° represents the wind blowing from north to south, and 90^o represents it blowing from east to west. The wind speed is set at a constant 2 m/s in Scene1 and 4 m/s in Scene2.

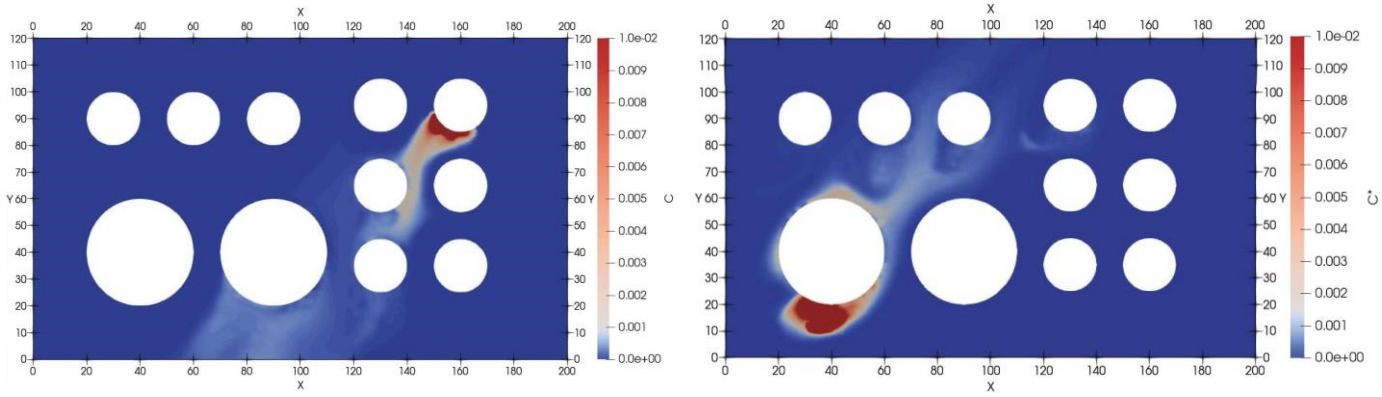
A leak scenario is defined by a specified leak source and a determined wind direction. For each scene, the forward dispersion simulation of all scenarios formed a dataset for SPO, and then STEs were conducted for each scenario. If a scene contains m' potential leak sources, then $m = 8 \times m\prime$ leak scenarios will be generated from this scene. Hence, 72 scenarios and 88 scenarios were generated for *Scene1* and *Scene2*, respectively.

3.2. CFD Software and Configuration

Given that the source–receptor relationship in STE is derived from the steady-state advection–diffusion equation, we utilized this equation to simulate the gas dispersion

in order to mitigate discrepancies between the forward dispersion model and adjoint equations. The wind field is critical for both forward dispersion and an adjoint equation. Large eddy simulation (LES), which is capable of providing a representation of turbulent structures and dynamics, was utilized to generate wind fields for all leak scenarios.

The open-source software program fire dynamic simulator (FDS) v6.7.9 [21] was chosen to generate a wind field with effect of gas dispersion for different scenarios. In this experiment, FDS simulates the wind field by solving the Navier–Stokes equation with LES modelling the small-scale eddies. Subgrid-scale eddy coefficients are modeled by Deardorff's model [22]. After averaging the obtained wind fields over the time domain, the open-source software program OpenFOAM v1912 [23] was selected to simulate the forward dispersion by solving the steady-state advection–diffusion equation (Equation (8)), without the consideration of viscosity and turbulence. This operation is carried out to reduce the error between mathematical models of forward dispersion and the adjoint equation. The AC fields of sensors under different wind conditions were also calculated by OpenFOAM by solving adjoint equation expressed as Equation (9). The FDS and OpenFOAM calculations were conducted on a workstation running a Linux OS with an Intel® Gold 6242R 3.1 GHz CPU and 192 GB of memory. The volume of grids for CFD was approximately 1 m^3. It generally costs about 1.2 h and 3 h for FDS to finish a single scenario in *Scene1* and *Scene2*, respectively. The calculations of OpenFOAM were quite fast because the steady-state equations are easy to solve. Figure 4 illustrates some calculation results from *Scene2*.



(**a**) (**b**)

Figure 4. Forward dispersion and adjoint equation results for *Scene2*. (**a**) The concentration distribution of the leak from Source 7 under a northeast wind at a height of z = 10 m. (**b**) AC field of a sensor located at $\mathbf{x} = (38, 18, 10)$ under the same wind condition.

3.3. Candidate Sensor Locations, SA, and STE

Considering the balancing of SPO performance and computational limitations, 360 and 600 available candidate locations were determined uniformly in the plane at z = 10 m for *Scene1* and *Scene2*, respectively. Details of the candidate locations are illustrated in Appendix A. To enhance the robustness of the optimization, SA was independently conducted five times for each of the three objectives. The solution yielding the highest score among these five runs was then selected as the optimized outcome for the specific objective.

After SPO, STE was conducted for all scenarios based on the optimized sensor placement. To accelerate the STE algorithm, during the MCMC process, the range of q is fixed as $q \in [0,\ 1]$ (the actual value of q is 0.5 kg/s). The STE result is identified as the point at which the PDF is maximized for the dimensions x, y, z, and q. The location error $Error_{x,j}$ was defined as the Euclidean distance between estimated and true location, while the error

of leak strength $Error_{q,j}$ was calculated as the absolute difference. The formulations of these indices are expressed as Equation (11):

$$\begin{aligned} x_j &= \underset{x}{argmax} p_j(x|D,I) \\ y_j &= \underset{y}{argmax} p_j(y|D,I) \\ z_j &= \underset{z}{argmax} p_j(z|D,I) \\ q_j &= \underset{q}{argmax} p_j(q|D,I) \\ Error_{x,j} &= \sqrt{(x_j - x_s)^2 + (y_j - y_s)^2 + (z_j - z_s)^2} \\ Error_{q,j} &= |q_j - q_s| \end{aligned} \tag{11}$$

The SA and STE algorithms were executed on the same workstation introduced in Section 3.2. The time consumed by SA increases proportionally with the number of candidate sensors, which varies from 10 min to approximately 1 h. The time required for STE depends on the scale of the scene and the number of sensors. For Scene1, concurrent STE computation for 72 scenarios takes around 1 to 3 h to complete. For Scene2, this time extends to approximately 2 to 6 h.

4. Experiment Results

For both scenes, this section presents the SPO and STE results, along with an analysis of the impact of AS and MC on the error of STE and the exploration on the most suited SPO objective. For ease of reference, the abbreviations used in this section are listed in Table 3.

Table 3. Abbreviation reference.

Abbreviation	Meaning
SPO	Sensor placement optimization
STE	Source term estimation
AS	Factor: number of activated sensors
MC	Factor: measured concentrations of sensors
HMC	SPO objective: highest mean measured concentration
MAS	SPO objective: most activated sensors
MAS-MC	SPO objective: most activated sensors with higher measured concentrations

4.1. Results in Scene1

4.1.1. SPO Results in *Scene1*

In Scene1, the number of candidate sensors for SPO is represented by an increasing sequence. This sequence begins at 8, increases by 1 until it reaches 25, and then increases by 5 until reaching 65. In total, 26 elements are contained in this sequence. For each optimized sensor placement, we evaluated the average values of AS and MC over all scenarios, as well as the 75% average of these factors calculated by Equation (10h). The variations of these two factors with changes in the number of sensors are illustrated in Figure 5. We checked the results and ensured that all optimized sensor placements were capable of detecting all leak scenarios. Additionally, the results show the following:

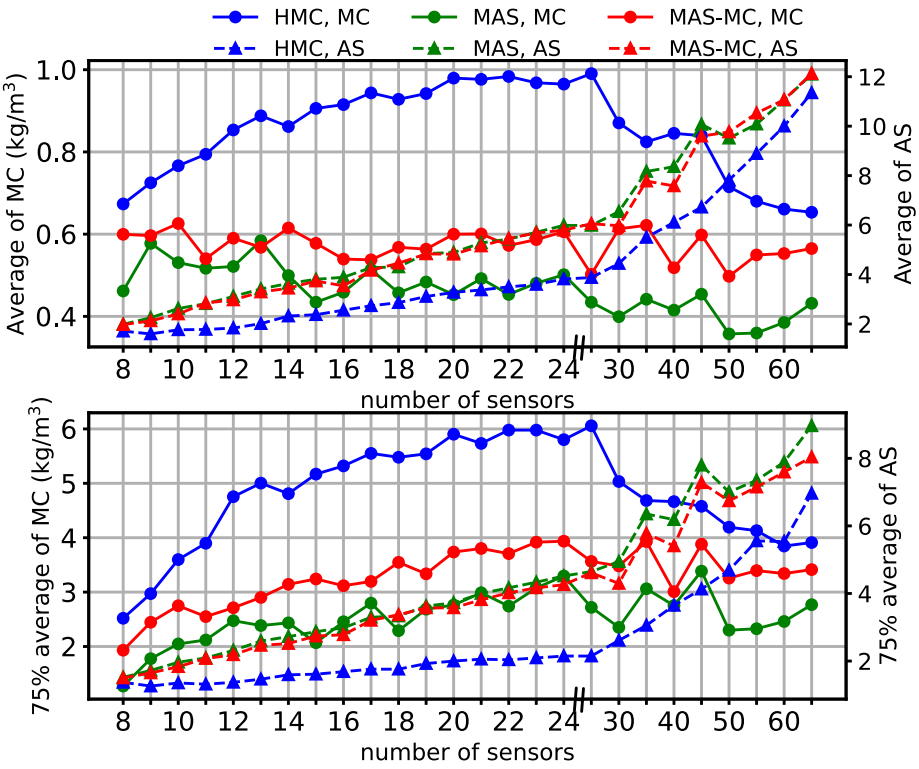


Figure 5. Results of SPO for different numbers of candidate sensors in *Scene1*. (**Upper**): performance on average MC (left) and AS (right). (**Lower**): performance on 75% average MC (left) and AS (right).

(1) As the number of candidate sensors increases, average AS increases proportionally for three objectives. Among the three objectives, HMC exhibits the lowest average AS with a noticeable gap. The average AS of MAS-MC is comparable to that of MAS but slightly lower.

(2) In these three objectives, the average MC of HMC surpasses that of MAS-MC, which, in turn, exceeds that of MAS. Additionally, as the average AS reaches a certain threshold, the average MC seems to exhibit a negative correlation with it. This is explainable, as when there are numerous candidate sensors, from an overall perspective across all scenarios, the average AS is bound to increase, while the average MC of each sensor will inevitably experience a certain degree of decline. Generally speaking, these results align with the expectation of these optimization objectives.

4.1.2. STE Results in *Scene1*

The 78 sensor placements generated by different candidate sensor quantities and three optimization objectives, along with the 72 leak scenarios, constitute a total of 5616 distinct STE situations. The box figures and line plots in Figure 6 illustrate location and q errors of these STE situations with different quantities of candidate sensors. Overall, as the quantitation of candidate sensors increases, the STE errors decrease, and the error distributions are more concentrated around mean values. When comparing among three objectives, we noticed that HMC leads to the worst overall STE accuracy, both on location and q, and that the accuracy of MAS-MC is better than that of MAS. Combined with the SPO result shown in Figure 5, it can be inferred that the STE accuracy has a proportional correlation with AS, while that correlation with MC is not obvious.

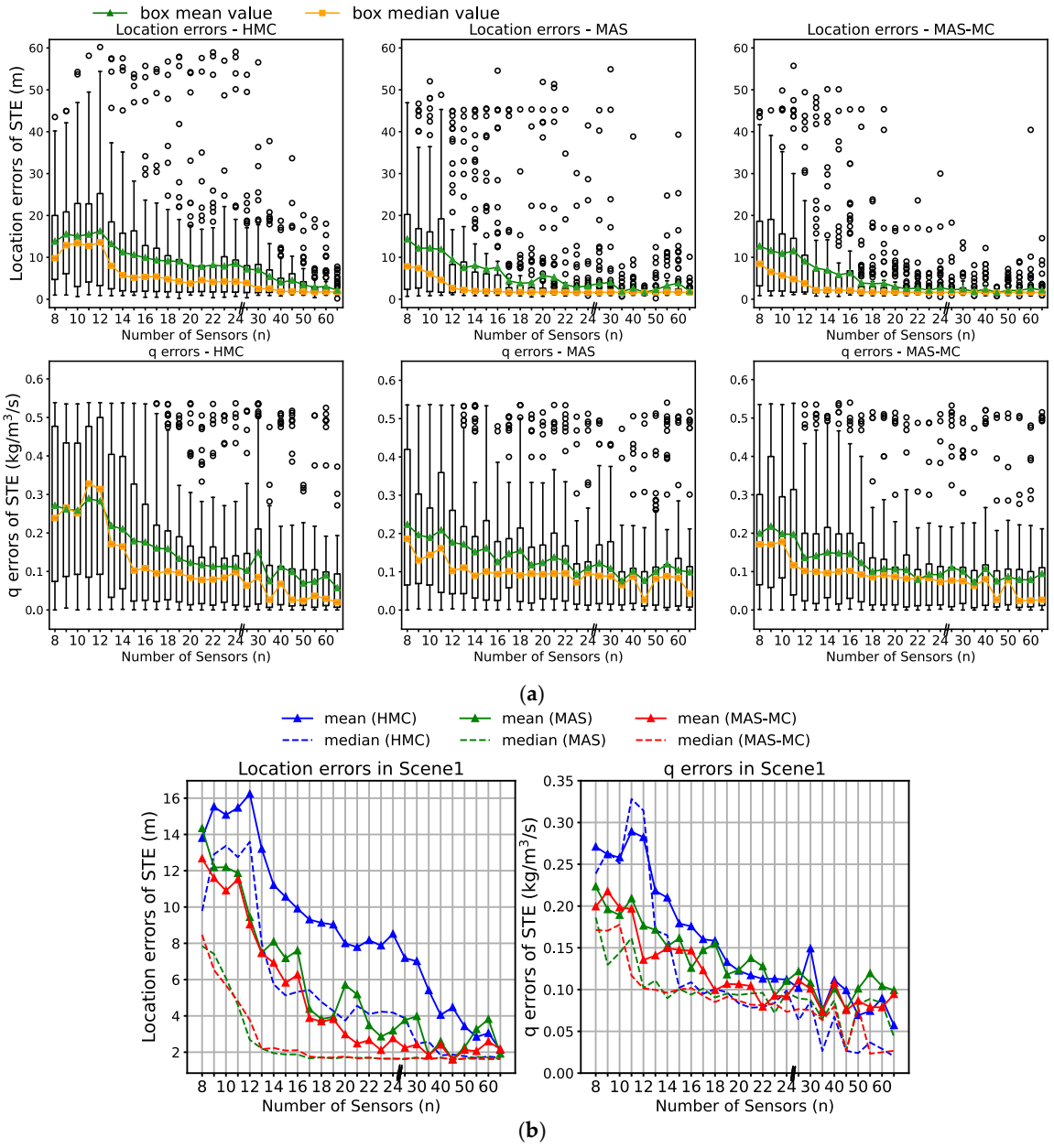


Figure 6. STE errors in *Scene1*. (**a**) STE errors over different numbers of sensors and three optimization objectives. (**Upper row**): location errors. (**Lower row**): q errors. (**b**) Mean and median values of STE errors over different numbers of sensors and three optimization objectives. (**Left**): location errors. (**Right**): q errors.

Figure 7 summarizes all 5616 STE situations and divides them into several intervals according to the values of AS and MC, then examines the average and median STE errors of these intervals. The values on the axes in Figure 7 are the lower limits of each interval. For example, the number 6 on the AS-axis of Figure 7b (II) is the average STE location error of scenarios with AS values between 6 and 10. From (I) and (III) in Figure 7a, it is evident that as AS increases, the decrease in the error of STE is noticeable, both in terms of position error and q error. However, (II) and (IV) in the same figure fail to exhibit a monotonic relationship between STE error and MC, and, in particular, the relationship between the q error and MC is ambiguous. When horizontally comparing different columns along the AS-axis in (II) and (III) of Figure 7b, it illustrates that the location error decreases significantly as AS increases. When the AS value is very low (e.g., the values of AS as 1 and

2), a negative correlation between the location error and MC occurs; however, under other conditions, this correlation is not noticeable. (IV) and (V) in the same figure show similar results; the only difference is that MC has almost no effect on the STE error of q.

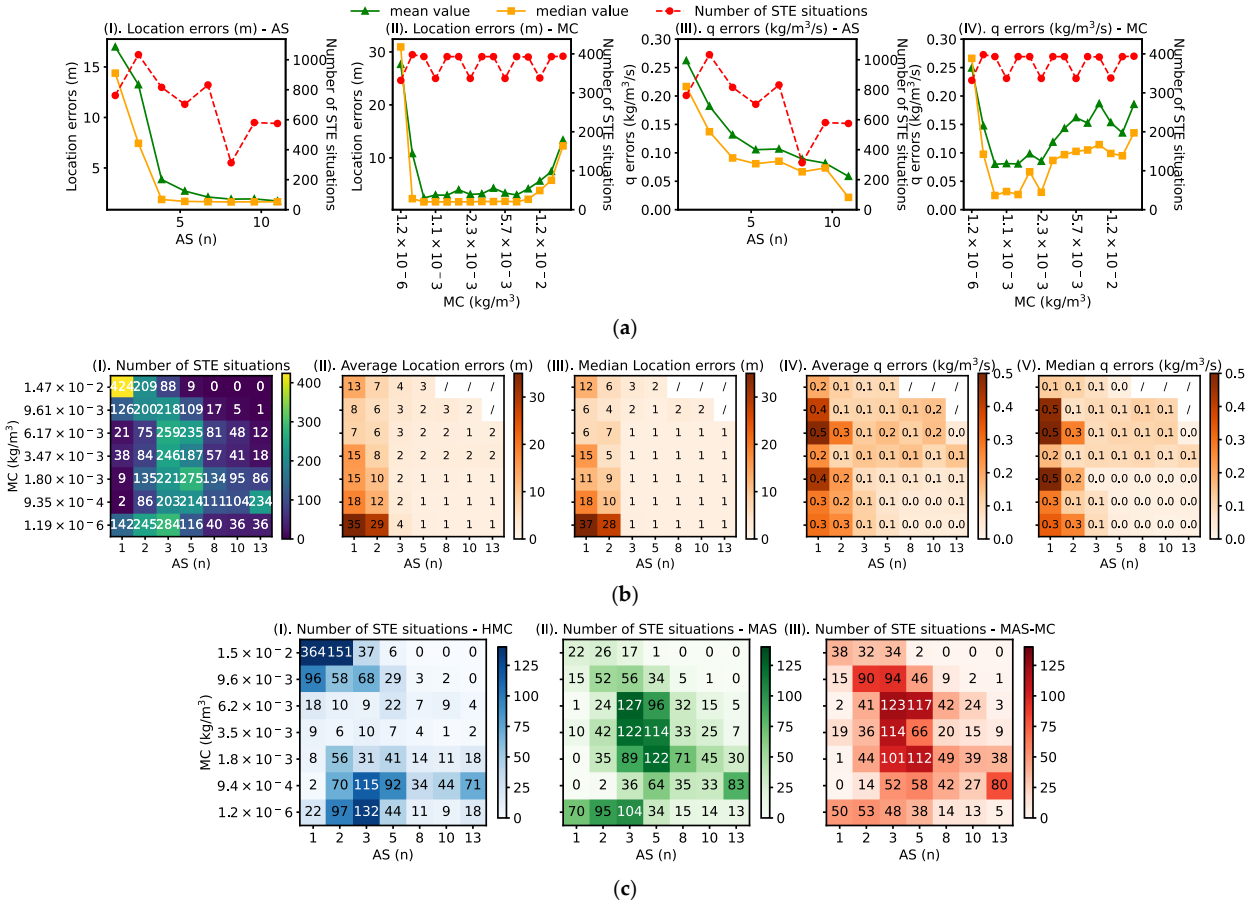


Figure 7. Summary on relationship between STE errors and sensor placements in *Scene1*. (**a**) Mean and median values of STE errors over different intervals of AS and MC. Green and yellow lines represent mean and median values, respectively. Red dashed lines represent quantities of STE conditions. (**b**) Heatmaps illustrating the STE errors in different AS and MC intervals. The numbers in (**II**,**III**) are rounded to the nearest integer, and the numbers in (**IV**,**V**) are presented with one decimal place. (**c**) Quantities of STE situations over different intervals of AS and MC under different SPO objectives.

Figure 7c categorizes all STE situations according to the SPO objectives associated with the sensor placements and draws the distribution in different intervals. From these heatmaps, it can be read that in this simple scene, MAS and MAS-MC achieve a balance between AS and MC, while HMC leads to extreme intervals and thus increases STE errors.

4.2. Results in Scene2

4.2.1. SPO Results in *Scene2*

Scene1 is a very simple scene that has no obstacles, while *Scene2* is a much more sophisticated one at larger scale and with several cylindrical obstacles inside. These huge obstacles significantly impact the homogeneity of the wind field and greatly extend the coverage area of gas concentration. The numbers of candidate sensors in this scene comprise an increasing sequence that begins at 8, increases by 2 until 20 and then increases by 5 until 110. Similar to *Scene1*, the average and 75% average AS and MC are evaluated for each optimized sensor placements, and the results are shown in Figure 8. The SPO results in this scene are roughly similar to those of *Scene1*, while some new features are also presented:

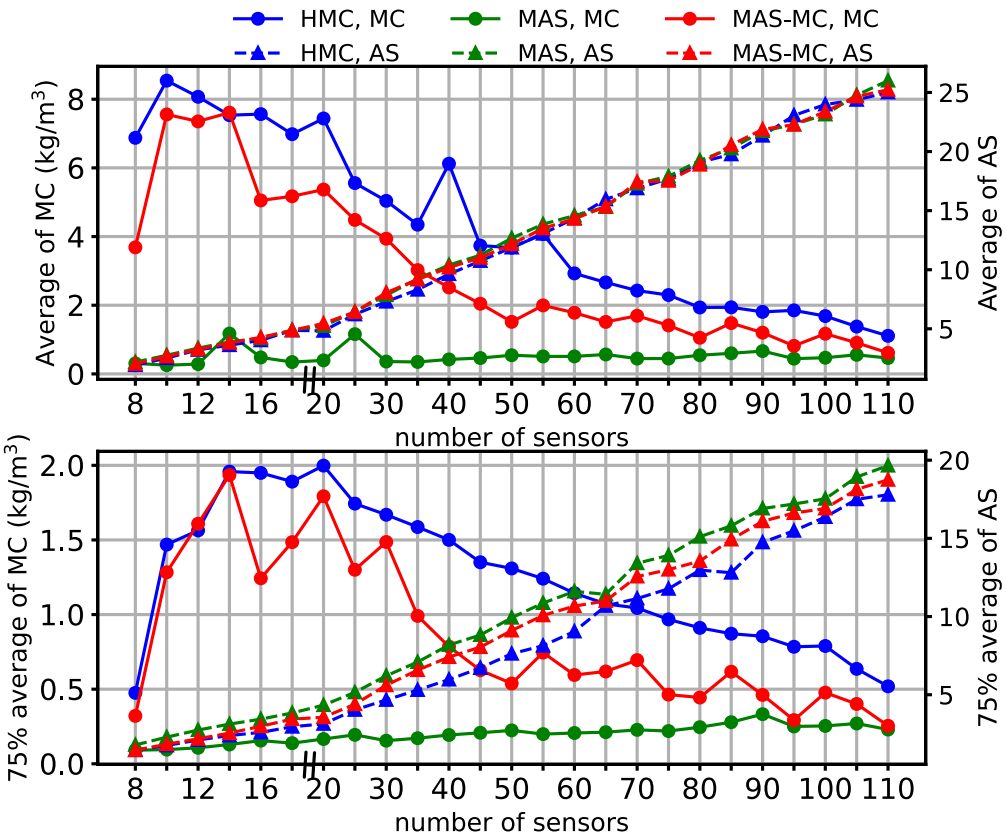


Figure 8. Results of SPO for different quantities of candidate sensors in *Scene2*. (**Upper**): performance on average MC (left) and AS (right). (**Lower**): performance on 75% average MC (left) and AS (right).

(1) The average AS of three objectives shows little disparity. Although the 75% average AS of HMC still lags behind the other two objectives, the differences between HMC and other objectives are no longer that substantial. This is due to the escalation in the gas coverage area, which makes the deployed sensors much more likely to detect the gas.

(2) The average MC of MAS-MC is lower than that of HMC, while that of MAS is significantly lower than that of the other two objectives. The average MC shows a negative correlation with the average AS. Overall, the SPO results for MAS and MAS-MC are close to expectations. However, the optimization results for HMC exceed expectations as the AS values are not significantly low, even though AS is not included in the objective function of HMC.

4.2.2. STE Results in *Scene2*

The 75 sensor placements generated by different candidate sensor quantities and three optimization objectives, along with the 88 leak scenarios, constitute a total of 6600 distinct STE situations. Figure 9 shows the location and q errors of all STE situations. Due to the increased scene scale, the overall location error is notably higher than that in *Scene1*. Also, the correlation between q error and quantities of candidate sensors is unrecognizable. Therefore, the location errors of STE are emphasized here. Among these three objectives, MAS location errors become the highest. When the numbers of candidate sensors are relatively low (lower than 14), the number of errors in MAS-MC is lower than in HMC. In other cases, the location errors in HMC are the lowest.

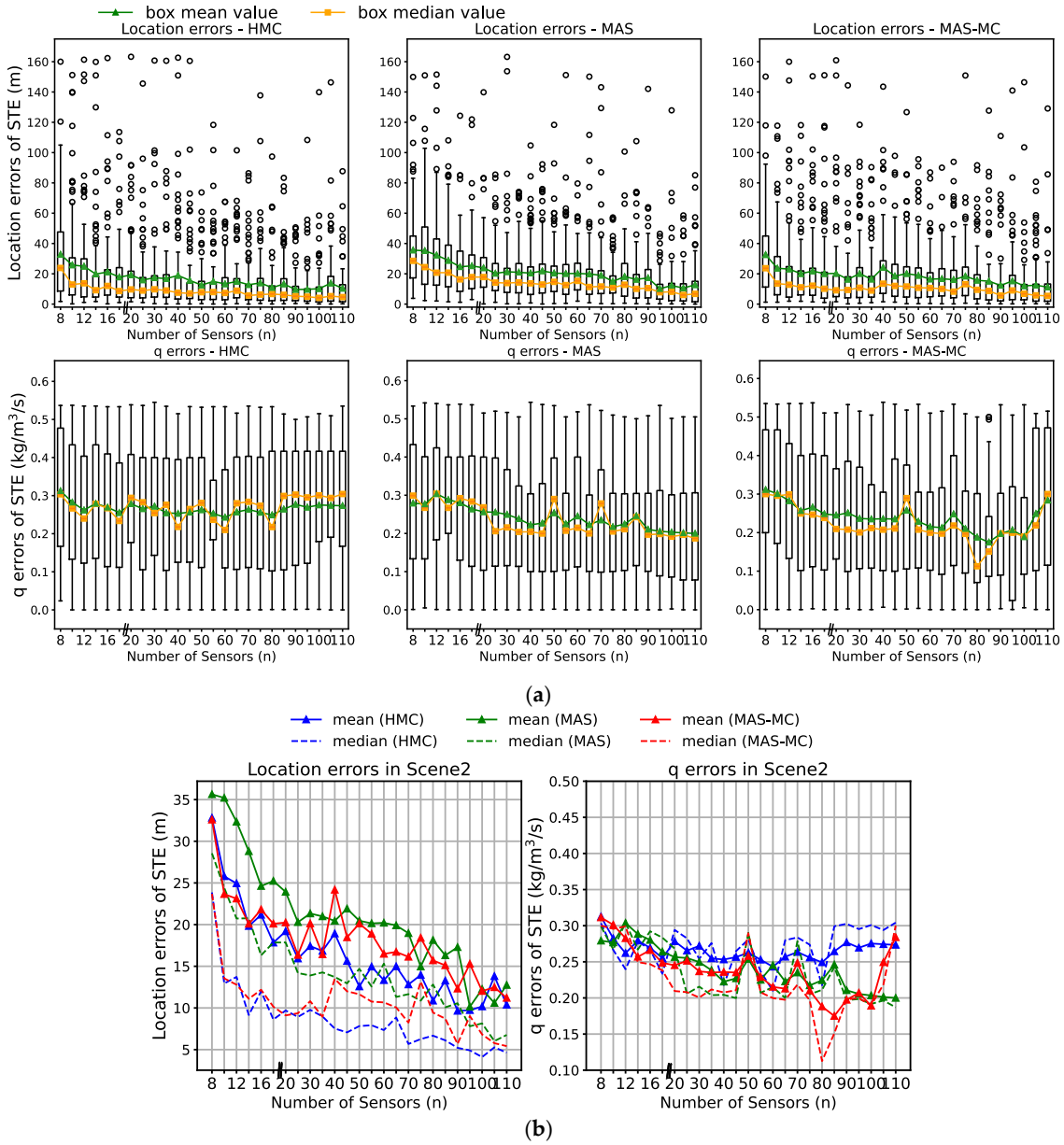


Figure 9. Errors under STE conditions in Scene 2. (**a**) STE errors over different numbers of sensors and three optimization objectives. (**Upper row**): location errors. (**Lower row**): q errors. (**b**) Mean and median values of STE errors over different numbers of sensors and three optimization objectives. (**Left**): location errors. (**Right**): q errors.

The 6600 STE results summarized in Figure 10 to some extent align with the findings in *Scene1* but also reveal some novel phenomena. From (I) and (II) in Figure 10a, the average and median location errors gradually decrease as AS increases, and a sharp decline occurs when AS reaches around 30. When it comes to MC, we observe a significant decline when MC is relatively low (lower than approximately 1.73×10^{-4} kg/m^3). Then, as MC increases, the location errors decrease gradually. (III) and (IV) of the same figure fail to reveal any relationship between the q error and AS or MC. In (II) and (III) of Figure 10b, the colors are darkest in the bottom-left corner, indicating that it is nearly impossible to conduct STE if the values of AS and MC are both low. Additionally, it is noted that the colors in the upper halves of both images are slightly lighter than those in the lower halves, which means lower STE location errors are detected when MC values are higher. The location errors are quite low when AS is extremely high (higher than around 31); however, this is usually not

cost-effective for enterprises. The leftmost rows in (IV) and (V) illustrate very high q errors when only one sensor is activated. Apart from this, little representative information can be extracted from these two subplots.

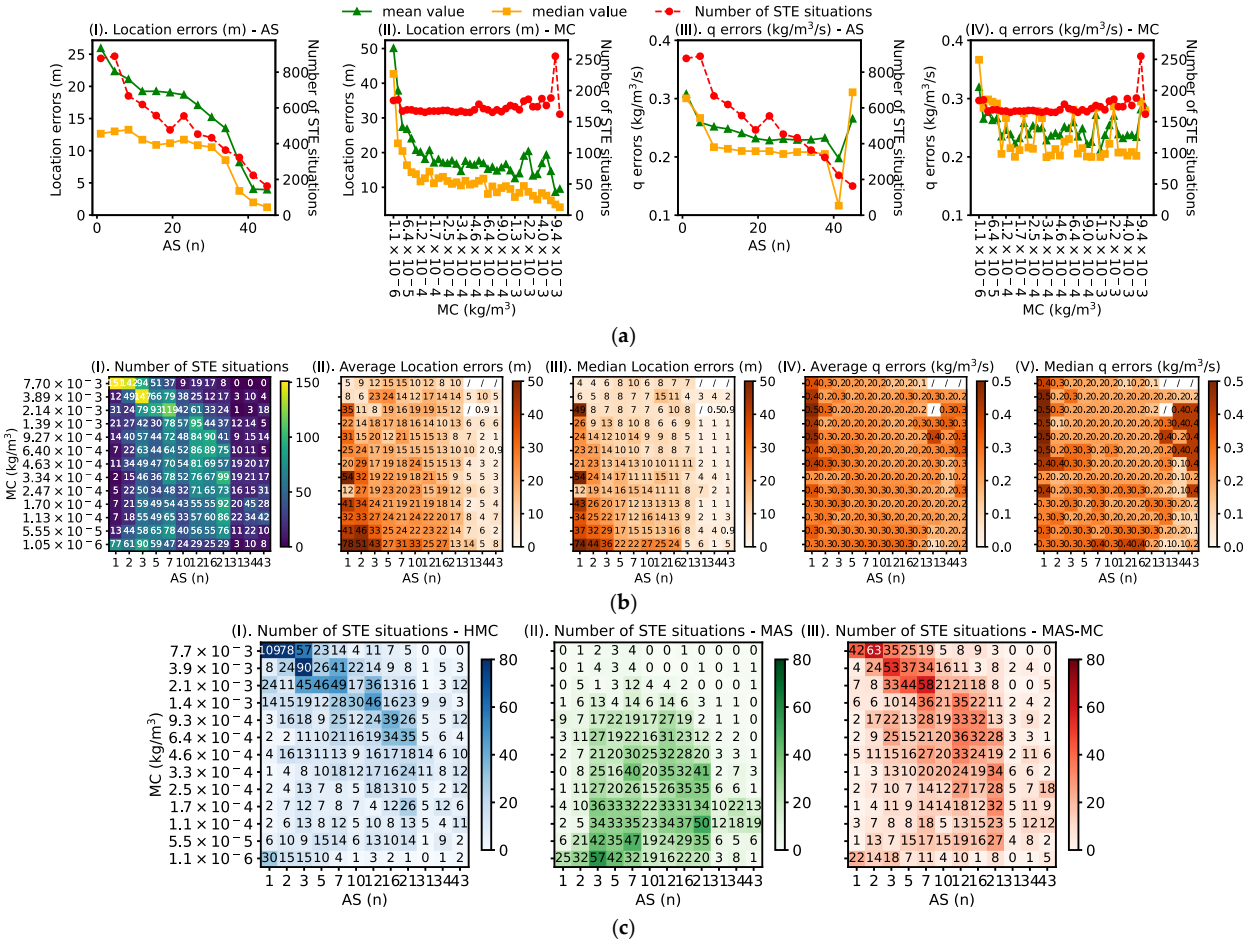


Figure 10. Summary of relationship between STE errors and sensor placements in Scene 2. (**a**) Mean and median values of STE errors over different intervals of AS and MC. Green and yellow lines represent mean and median values, respectively. Red dashed lines represent quantities of STE conditions. (**b**) Heatmaps illustrating STE errors in different AS and MC intervals. The numbers in (**II**,**III**) are rounded to the nearest integer, and the numbers in (**IV**,**V**) are presented with one decimal place. (**c**) Distribution of quantities of STE situations over different intervals of AS and MC under different SPO objectives.

From Figure 10c, it can be observed that although there is some clustering in the upper-left corner interval of the graph, the STE situations resulted by HMC do not exhibit the excessively low values of AS observed in *Scene1*. The distribution of MAS is relatively concentrated in the lower intervals of the MC-axis, indicating lower MC, while the distribution of MAS-MC is similar to that of HMC. The results reveal that in such complex scenarios, MC should be a key consideration for SPO, while AS can be given less emphasis.

4.3. Discussion

The SPO and STE conducted on *Scene1* and *Scene2* yielded rich outcomes, including diverse SPO results and a total of 12,216 STE results. These results also demonstrate that the complexity of the scene influences the SPO and STE results. Compared to *Scene1*, the expansion of scale and the complexity of obstacles in *Scene2* result in numerous new features in the experimental outcomes. Due to the closer resemblance of *Scene2* to a real

chemical plant, we consider the results obtained in this scene to be more representative. From the results, some regularities can be observed:

(1) A negative correlation between the STE location error and the value of AS is observed in both two scenes. However, the STE location errors in *Scene2* only show a noticeable reduction when AS reaches approximately 20. In practice, the quantity of candidate sensors is a primary constraint of SPO, and thus the AS values are usually not that large. In such conditions, placing sensors in locations with higher MC is more cost-effective for SPO.

(2) Although a negative correlation between q error and AS is observed in *Scene1*, the results in *Scene2* fail to provide supporting evidence. Actually, since q merely serves as a multiplier in STE as Equation (10), it is relatively challenging to make precise estimates of it through stochastic sampling.

(3) Among the three SPO objectives, HMC leads to the best overall STE performance in *Scene2* even though it results in the worst performance in *Scene1*. The scale of *Scene2* and the complexity of obstacles make it relatively easy for sensors to be activated in a gas leak scenario. This prevents AS from being excessively low, despite HMC not incorporating AS as part of the optimization objective. Therefore, when applying the SPO method described in this paper in complex scenes such as chemical plants, it suffices to consider only the MC of sensors to achieve a reduction in STE location error, without emphasizing the values of AS.

(4) The discrepancy between results of *Scene1* and *Scene2* cautions us to mind the influence of scene complexity on the analysis results when analyzing sensor placements utilizing this analytical framework. Conclusions drawn from simple scenes may not necessarily be generalizable to complex scenes such as chemical plants.

(5) In *Scene1* (Figure 6), the downward trend of STE location errors slows down when the number of sensors exceeds 17 (with SPO objective as MAS-HC). In Scene2 (Figure 9), such performance requires that the number of sensors exceeds around 45 (with SPO objective as HMC). In practical scenarios, the total number of sensors also acts as an objective, considering the economic cost.

5. Conclusions and Future Work

This paper proposed an analytical framework to determine sensor locations for enhanced STE and analyzed the SPO strategy in complex scenes such as chemical plants through investigating the impact of sensor placement on STE errors. To obtain rational sensor placements, a mathematical model of SPO is proposed in Section 3 with HMC, MAS, and MAS-MC as objectives. All these objectives ensure that the sensor placements possess the capability of detecting all potential gas leak scenarios. Subsequently, STE calculations and post-processing are conducted based on sensor placements gained by SPO. In experiments, two test scenes were created, in which CFD, SPO, and STE were sequentially conducted. A total of 12,216 STE results indicated that both AS and MC have a negative correlation with the STE location errors. Specifically, a significant reduction in this error was observed when AS is large. In *Scene1*, AS has a greater impact on STE errors compared to MC, and MAS-MC as the objective of SPO yields optimal STE accuracy. In *Scene2*, the significant reduction in STE location errors requires AS to be larger than around 20, which is hard to achieve under the constraints of the number of candidate sensors. Moreover, results in *Scene2* revealed that in chemical plant scenes, the objective of SPO should focus on MC rather than AS for enhanced STE. In this experiment, HMC as the objective resulted in lowest overall STE errors. Additionally, the different conclusions drawn from the two scenes discerned the lack of generalizability across scenes with distinct scales and complexities.

The current work still has some limitations. Firstly, although this paper extensively analyzed the impact of AS and MC on STE errors, more research is needed to explore the underlying principles of this impact from a mathematical perspective. Secondly, this study temporally focused on the influence of AS and MC in sensor placement; however, some potential factors like the distance between sensors, and that between sensor and the leak source, might also be critical to the accuracy of STE. It would be meaningful to explore further and propose an improved SPO model. Thirdly, more realistic factors should be included in further experiments, for instance, the settling of species dispersion, the expansion of experimental scene, and the expansion or limitations of candidate sensor locations for SPO. Lastly, the factors affecting the STE errors of q need more exploration. These issues remain to be addressed in future research.

Author Contributions: Conceptualization, H.T., Z.L. and B.W.; Methodology, H.T., Z.L., C.C. and B.W.; Software, H.T.; Validation, H.T.; Formal analysis, H.T.; Investigation, B.W.; Resources, C.C. and B.W.; Writing—original draft, H.T.; Writing—review and editing, Z.L., C.C. and B.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (62394343, 62394345, 62103151), the Programme of Introducing Talents of Discipline to Universities (the 111 Project) under Grant B17017 and Fundamental Research Funds for the Central Universities (222202517006).

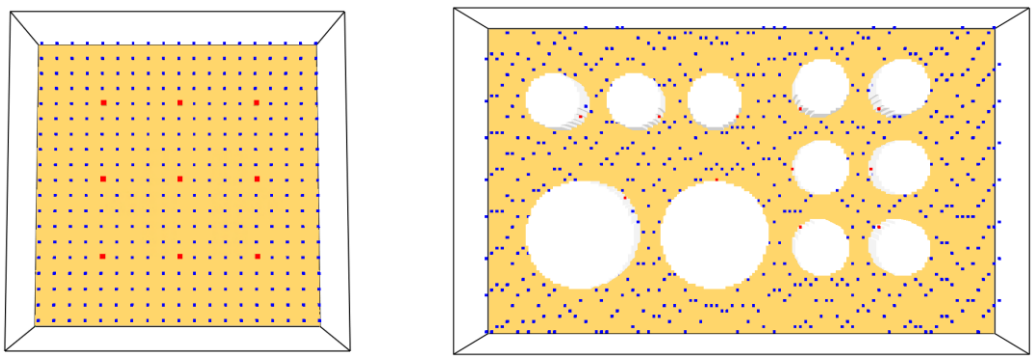
Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Details of candidate sensor locations and some of the SPO results are listed in this appendix.

The candidate sensor locations are determined mainly by the height of leak sources and the vertex coordinates of CFD meshes. As mentioned in Sections 3.1 and 3.3, for consistency with adjoint equation, the forward gas dispersion is calculated by adjoint–diffusion equation, which weakens the settling of leaking species. Therefore, the candidate locations are selected at the same height of leak sources $z = 10$ m. At this height, 352 and 600 candidates are uniformly selected by program for the SA algorithm; Figure A1 shows these candidate locations (blue dots) in *Scene1* and *Scene2*, respectively. The optimized sensor locations for some cases in *Scene1* and *Scene2* are shown in Figures A2 and A3, respectively.



(**a**) (**b**)

Figure A1. Leak sources (red dots) and candidate locations for sensor installation (blue dots). (**a**) *Scene1*; (**b**) *Scene2*.

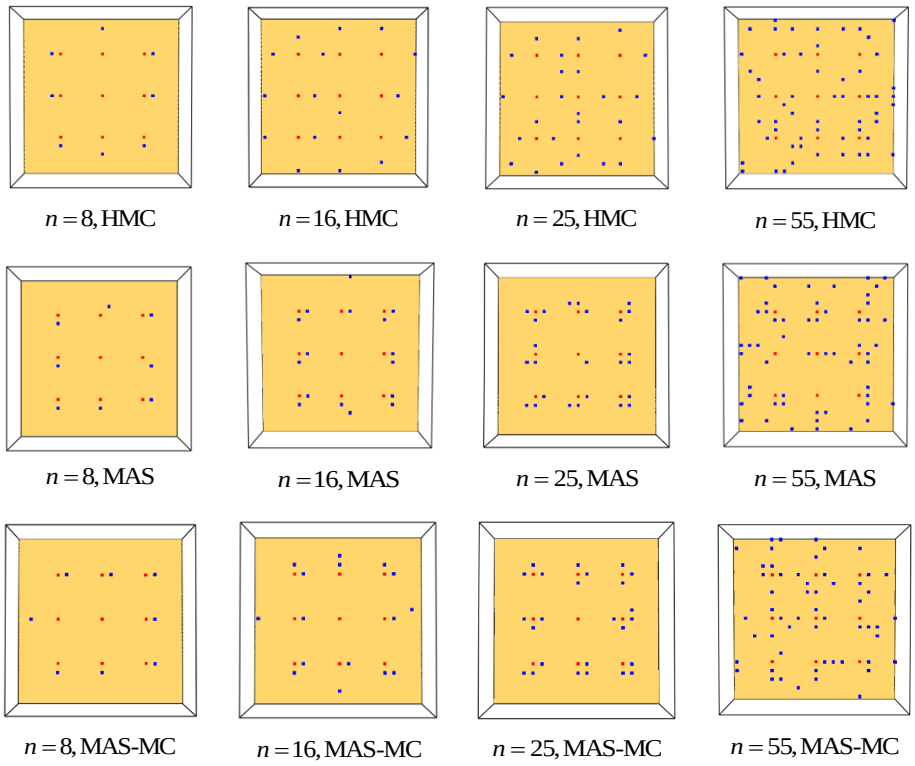


Figure A2. Leak sources (red dots) and optimized sensor placements (blue dots) for some cases in *Scene1*.

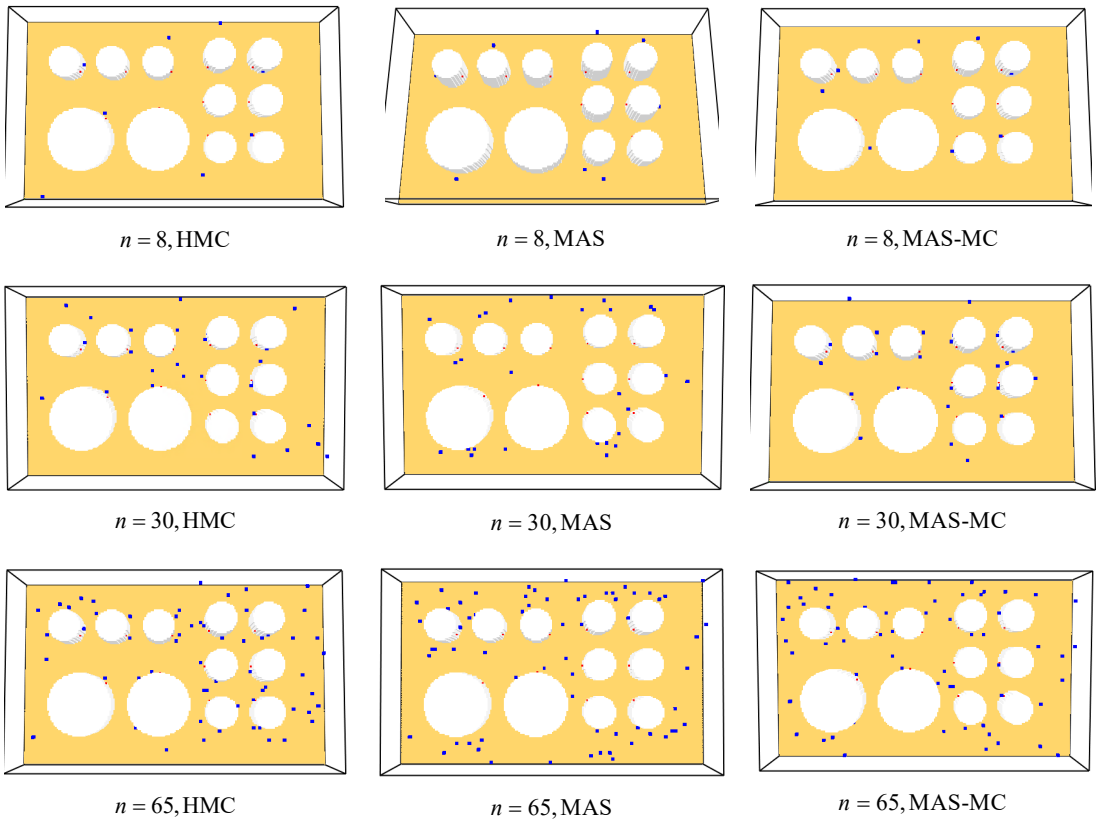


Figure A3. Leak sources (red dots) and optimized sensor placements (blue dots) for some cases in *Scene2*.

References

1. Sivaraman, S.; Tauseef, S.M.; Siddiqui, N.A. Investigative and probabilistic perspective of the accidental release of styrene: A case study in Vizag, India. *Process Saf. Environ. Prot.* **2022**, *158*, 55–69. [CrossRef]
2. Legg, S.W.; Benavides-Serrano, A.J.; Siirola, J.D.; Watson, J.P.; Davis, S.G.; Bratteteig, A.; Laird, C.D. A stochastic programming approach for gas detector placement using cfd-based dispersion simulations. *Comput. Chem. Eng.* **2012**, *47*, 194–201. [CrossRef]
3. Benavides-Serrano, A.J.; Legg, S.W.; Vázquez-Román, R.; Mannan, M.S.; Laird, C.D. A stochastic programming approach for the optimal placement of gas detectors: Unavailability and voting strategies. *Ind. Eng. Chem. Res.* **2013**, *53*, 5355–5365. [CrossRef]
4. Liu, Y.; Zhang, B.; Mu, C. A comparative study of optimization models for the gas detector placement in process facilities. *Comput. Chem. Eng.* **2020**, *143*, 107095. [CrossRef]
5. Zhou, C.; Zhang, B.; Mu, C.; Chu, Z.; Sun, L. Multi-objective optimization considering cost-benefit ratio fo r the placement of gas detectors in oil refinery installations. *J. Loss Prev. Process Ind.* **2019**, *62*, 103956. [CrossRef]
6. Rad, A.; Rashtchian, D.; Badri, N. A risk-based methodology for optimum placement of flammable gas detectors within open process plants. *Process Saf. Environ. Prot.* **2017**, *105*, 175–183. [CrossRef]
7. Astolfi, M.; Zonta, G.; Gherardi, S.; Malagù, C.; Vincenzi, D.; Rispoli, G. A Portable Device for I–V and Arrhenius Plots to Characterize Chemoresistive Gas Sensors: Test on SnO_2-Based Sensors. *Nanomaterials* **2023**, *13*, 2549. [CrossRef]
8. Chen, Z.; Yao, J.; Saifutdinov, A.I.; Kudryavtsev, A.A.; Yuan, C.; Ma, G.; Dou, Z.; Cao, J.; Ma, M.; Zhou, Z. Determination of organic impurities by plasma electron spectroscopy in nonlocal plasma at intermediate and high pressures. *Plasma Sources Sci. Technol.* **2022**, *31*, 107001.
9. Singh, S.; Sharan, M.; Issartel, J.P. Inverse modelling methods for identifying unknown releases in emergency scenarios: An overview. *Int. J. Environ. Pollut.* **2015**, *57*, 68–91. [CrossRef]
10. Hutchinson, M.; Oh, H.; Chen, W.H. A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. *Inf. Fusion* **2017**, *36*, 130–148. [CrossRef]
11. Keats, A.; Yee, E.; Lien, F.S. Bayesian inference for source determination with applications to a complex urban environment. *Atmos. Environ.* **2007**, *41*, 465–479. [CrossRef]
12. Lang, Z.; Wang, B.; Wang, Y.; Cao, C.; Peng, X.; Du, W.; Qian, F. A novel multi-sensor data-driven approach to source term estimation of hazardous gas leakages in the chemical industry. *Processes* **2022**, *10*, 1633. [CrossRef]
13. Chen, S.; Du, W.; Peng, X.; Cao, C.; Wang, X.; Wang, B. Peripheric sensors-based leaking source tracking in a chemical industrial park with complex obstacles. *J. Loss Prev. Process Ind.* **2022**, *78*, 104828. [CrossRef]
14. Xu, J.; Du, W.; Xu, Q.; Dong, J.; Wang, B. Federated learning based atmospheric source term estimation in urban environments. *Comput. Chem. Eng.* **2021**, *155*, 107505. [CrossRef]
15. Liu, Z.; Li, X. Sensor layout strategy for source term estimation of external pollution sources in urban neighbourhoods. *Build. Environ.* **2022**, *220*, 109276. [CrossRef]
16. Jia, H.; Kikumoto, H. Sensor configuration optimization based on the entropy of adjoint concentration distribution for stochastic source term estimation in urban environment. *Sustain. Cities Soc.* **2022**, *79*, 103726. [CrossRef]
17. Pudykiewicz, J.A. Application of adjoint tracer transport equations for evaluating source parameters. *Atmos. Environ.* **1998**, *32*, 3039–3050. [CrossRef]
18. Hastings, W.K. Monte carlo sampling methods using markov chains and their applications. *Biometrika* **1970**, *57*, 97–109. [CrossRef]
19. Goodman, J.; Weare, J. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.* **2010**, *5*, 65–80. [CrossRef]
20. Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671–680. [CrossRef]
21. McGrattan, K.; Forney, G.; Floyd, J.; Hostikka, S.; Prasad, K. *Fire Dynamics Simulator-User's Guide Version 3*; National Institute of Standards and Technology (NIST): Gaithersburg, MD, USA, 2001; p. ir6784.
22. Deardorff, J.W. Stratocumulus-capped mixed layers derived from a three-dimensional model. *Bound.-Layer Meteorol.* **1980**, *18*, 495–527. [CrossRef]
23. Weller, H.G.; Tabor, G.; Jasak, H.; Fureby, C. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Comput. Phys.* **1998**, *12*, 620–631. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.