Estimation of pollutant sources in multi-zone buildings through different deconvolution algorithms

Mo Li¹, Fei Li¹ (✉), Yuanqi Jing¹, Kai Zhang¹, Hao Cai¹, Lufang Chen¹, Xian Zhang², Lihang Feng³ (✉)

1. College of Urban Construction, Nanjing Tech University, Nanjing, 210009, China

2. Institute of Defense Engineering, Academy of Military Science, PLA, Beijing, 100850, China

3. College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing, 210009, China

Abstract

Effective identification of pollution sources is particularly important for indoor air quality. Accurate estimation of source strength is the basis for source effective identification. This paper proposes an optimization method for the deconvolution process in the source strength inverse calculation. In the scheme, the concept of time resolution was defined, and combined with different filtering positions and filtering algorithms. The measures to reduce effects of measurement noise were quantitatively analyzed. Additionally, the performances of nine deconvolution inverse algorithms under experimental and simulated conditions were evaluated and scored. The hybrid algorithms were proposed and compared with single algorithms including Tikhonov regularization and iterative methods. Results showed that for the filtering position and algorithm, Butterworth filtering performed better, and different filtering positions had little effect on the inverse calculation. For the calculation time step, the optimal T_r (time resolution) was 0.667% and 1.33% in the simulation and experiment, respectively. The hybrid algorithms were found to not perform better than the single algorithms, and the SART (simultaneous algebraic reconstruction technique) algorithm from CAT (computer assisted tomography) yielded better performances in the accuracy and stability of source strength identification. The relative errors of the inverse calculation for source strength were typically below 25% using the optimization scheme.

Keywords

pollutant source;
measurement noise;
inverse algorithm;
indoor air;
algorithm ranking

Article History

Received: 28 March 2021

Revised: 10 July 2021

Accepted: 26 July 2021

© Tsinghua University Press and
Springer-Verlag GmbH Germany,
part of Springer Nature 2021

1 Introduction

In recent years, people have paid more attention to indoor air quality. Studies by several agencies, including the EPA (2017), suggest that indoor air pollution levels can be two to five times higher than outdoor levels. Poor indoor air quality can lead to a range of respiratory problems, such as allergies, asthma and sick building syndrome (Tsantaki et al. 2020). Certain infectious diseases, such as SARS (Li et al. 2005) and COVID-19 (Chan et al. 2020), also pose a great threat to human health. The pathogens that cause these diseases can spread in the form of bioaerosols (Yan et al. 2018). Additionally, pollutants such as dust from industrial production are highly carcinogenic and have a serious impact on worker health and productivity (Bhatti et al. 2011; Hancock

et al. 2015). Therefore, accurate estimation of the emission rates of such sources is essential for the subsequent designation of ventilation measures and evacuation of personnel.

For source emission rate measurement, the tracer gas method has been extensively studied in 80s and 90s (Sherman et al. 1980; Etheridge and Sandberg 1996). Different tracer gas methods for measuring airflow rates in buildings or rooms have been investigated by Sandberg and Blomqvist (1985). Based on the measured airflow rates, the emission rates of indoor gaseous (Chen et al. 2018) and particulate (Liang and Yang 2013) sources could be calculated by solving the continuity equation directly. In these studies, the source position was always known, and the source was usually placed in a well-mixed space equipped with sensors. However, in some real world applications, the pollutant source location

is unknown, or the source must be located away from the sensors for technical reasons. Experimental estimation of this unknown source through measurement of its response (the concentration) at sensor locations is an “inverse problem”.

The inverse calculation methods can be divided into the following three categories: forward, backward, and probability methods. The forward method requires pre-simulations of the potential source and then describes the matching degree between the measured and simulation results using a residual function. When the residual reaches a minimum, the optimal source characteristics are determined. The backward method is used to directly solve forward transport equations with negative time steps. However, this method is ill-conditioned and exhibits poor stability. The adjoint probability method can obtain the source location or time probability distribution by solving adjoint equations (Wang et al. 2017). In addition, the robot olfaction method is also often used to search the pollutant source in indoor environments (Feng et al. 2019; Yang et al. 2019). Accurate estimation of the source strength is a prerequisite for the above methods. While the source temporary release strength estimation is an ill-posed deconvolution inverse problem which can amplify high frequencies from the measurement noise. Therefore, it is important to enhance the stability of the inverse calculation and reduce the error in the deconvolution process. The regularization and iteration methods are commonly used to improve the solution stability.

Tikhonov regularization adds a regularization term to the objective function so that the value of the solution is kept within a certain range, thus reducing and avoiding the oscillation or divergence of the numerical solution (Yin 2011). Zhang et al. (2013) combined computational fluid dynamics (CFD) with Tikhonov regularization and least-squares optimization to quantify the time-release rate of a known fixed gas pollution source. Results showed that this method could effectively and accurately identify the source release rate, and the choice of regularization parameter had a strong influence on the inverse results. Zhang et al. (2013) used the L-curve algorithm to identify the pollutant sources in a 2D cavity and found that choosing appropriate regularization parameters with this method could accurately determine the source release rate. Liu et al. (2019) proposed a comprehensive inverse model called the Markov chain with regularization and Bayesian inference (MCRB) to identify the source location and dynamic release rate under time-varying airflow. In their study, the effects of sensor positions and regularization parameter selection algorithms were also investigated, and they found that the generalized cross-validation (GCV) algorithm performed better than the L-curve and quasiopt (quasi-optimality) algorithms. However, when the response matrix is large and sparse, the regularization

methods do not tend to work well. For example, if response matrix is a million by million matrix, it will require approximately 80 gigabytes (GB) of storage. The iterative method is often used to solve a large system of linear equations, particularly in computed tomography (Aster et al. 2005). Yu (2006) compared the maximum likelihood expectation maximization (MLEM) algorithm with the least squares with nonnegative constraint (LSNC) algorithm and found that the MLEM algorithm yielded marked advantages in concentration tomography. Li et al. (2020a) proposed an improved tomographic imaging algorithm (LSTR) based on Tikhonov regularization and the least-squares method, and compared its performance with that of LSNC and least squares with QR decomposition (LSQR) for gas concentration distribution reconstruction. These studies are mostly based on simulated data and single inverse algorithms, and the performances of hybrid algorithms in real conditions are not clear. Besides, the inverse algorithms for source strength calculation have not been quantitatively compared and evaluated.

Additionally, for source strength identification, flow field information and sensor reading errors have strong influences on the inverse results. Pang et al. (2014) and Wang and You (2015) considered the influence of sensor measurement noise while studying a single sensor to identify a constant source and found that the sampling time interval and sensor measurement error affected the source strength inverse results. Li et al. (2020b) considered the influence of measurement noise and explored methods to reduce measurement noise impact during source strength estimation. They found that properly increasing the time step could reduce the effects of noise, and the performances of filters were different under different time steps. In previous studies, they only qualitatively analyzed the effects of measurement noise and denoising approaches, and did not provide quantitative measures, such as the optimal calculation time step and filtering position.

This study presents an optimization scheme for the deconvolution process in the source strength inverse calculation in terms of the denoising method and inverse algorithm. We investigate and optimize the calculation time resolution, filtering algorithm and filtering position to minimize the impacts of measurement noise. Then, we introduce nine inverse algorithms, including hybrid and two tomographic algorithms, and compare their performances on source strength inverse reconstruction under experimental and simulated conditions. Their performances are scored from the aspects of accuracy and stability, and the best algorithms under different conditions are presented. The above methods could improve the accuracy of the source strength estimation which is the basis for source identification in real world application.

2 Methodology

The mass transfer process for specific pollutant in air can be expressed as a partial differential equation (PDF) as follows:

$$\frac{\partial(\rho C)}{\partial t} + \text{div}(\rho \mathbf{u} C) = \text{div}[\Gamma \text{grad}(C)] + S \quad (1)$$

where ρ is the air density, C is the instantaneous pollutant concentration at the source, $\mathbf{u}$ is the air velocity vector, t is the time, Γ is the effective mass diffusion coefficient of the pollutant and S is the source term.

When the flow field is fixed, the release strength of source forms a linear relationship with its concentration distribution. Therefore, the concentration response of an arbitrary source can be expressed as the convolution integral between the temporal release rate and the response factor (concentration response per unit pulse release), show as:

$$C(t) = \int_{-\infty}^{+\infty} q(\tau) F[\delta(t - \tau)] d\tau \quad (2)$$

where $q(\tau)$ is the time-dependent release rate profile, and $F[\delta(t - \tau)]$ is the response factor, i.e., the concentration response per unit pulse release at $t = \tau$. The discretization of Eq. (2) can be transformed into the following form (Hiyama et al. 2010):

$$C_{t_n} = \sum_{i=1}^n q_{t_{n-i}} F_{t_n} = q_{t_n} F_{t_0} + \dots + q_{t_{n-i}} F_{t_i} + \dots + q_{t_0} F_{t_n} \quad (3)$$

The above equation can be converted to $\mathbf{C} = \mathbf{M}(\mathbf{A})\mathbf{q}$, where $\mathbf{C}$ is the concentration vector monitored at different moments; $\mathbf{A} = [F_{t_0}, \dots, F_{t_i}, \dots, F_{t_n}]^T$ is the response vector; $\mathbf{M}(\mathbf{A})$ is a Toeplitz matrix defined for $\mathbf{A}$, i.e., response matrix; $\mathbf{q}$ is the release rate vector of the pollutant at different moments. This equation can be expanded into a matrix relationship as follows:

$$\begin{bmatrix} C_{t_0} \\ \vdots \\ C_{t_i} \\ \vdots \\ C_{t_n} \end{bmatrix} = \mathbf{M}(\mathbf{A})\mathbf{q} = \begin{bmatrix} F_{t_0} & 0 & 0 & \dots & 0 \\ F_{t_1} & F_{t_0} & 0 & \dots & 0 \\ F_{t_2} & F_{t_1} & F_{t_0} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{t_n} & F_{t_{n-1}} & F_{t_{n-2}} & \dots & F_{t_0} \end{bmatrix} \begin{bmatrix} q_0 \\ \vdots \\ q_i \\ \vdots \\ q_n \end{bmatrix} \quad (4)$$

The solution for the source release rate $\mathbf{q}$ can be converted to a least squares problem as follows.

$$\min_{\mathbf{q}} \|\mathbf{M}(\mathbf{A})\mathbf{q} - \mathbf{C}\|_2^2 \quad (5)$$

If the source release rate $\mathbf{q}$ is solved directly to the above equation, the result may be incorrect. Because the property of the coefficient matrix $\mathbf{M}(\mathbf{A})$ is unknown, it may be pathological or non-pathological. Therefore, further analysis

of the nature of $\mathbf{M}(\mathbf{A})$ is required.

The degree of sickness of the matrix $\mathbf{M}(\mathbf{A})$ will affect the solution result, and the degree of sickness of the matrix can be expressed as the condition number. The condition number can effectively represent the sensitivity of the matrix calculation to input noise and error. If the condition number of the response matrix $\mathbf{M}(\mathbf{A})$ is large, a smaller change in $\mathbf{C}$ in Eq. (5) will cause a larger change in the solution $\mathbf{q}$ and will be less stable numerically (Young et al. 2008). Hansen (2010) mentioned that Eq. (2) is the form of the first type of Fredholm integral equation, where the higher frequencies are filtered out when $\mathbf{C}$ is calculated with known $\mathbf{q}$ and $\mathbf{M}(\mathbf{A})$ forward, so that the curve of $\mathbf{C}$ is smoother. However, when $\mathbf{q}$ is calculated inversely with known $\mathbf{C}$ and $\mathbf{M}(\mathbf{A})$, it is a deconvolution process. The deconvolution process could amplify the high frequencies in sampling data and the higher the frequency, the greater the amplification, so the curve of $\mathbf{q}$ is more perturbed. In other words, source identification by a sequence $\mathbf{C}$ of pollutant concentrations from sensors in an actual building will result in distortion of the final inverse source strength due to measurement noise ill-conditioned response matrix. Therefore, an optimization scheme was proposed to calculate the source strength inversely and reduce the measurement noise from the sensor.

Figure 1 shows the procedure and configuration of the proposed optimization scheme. Firstly, we measure and obtain the concentration response matrices and pollution concentration. Then, based on the matrix dimension and

Fig. 1 Flow chart of the proposed optimization scheme (green parts: denoising methods; orange parts: inverse algorithms; *: including Butterworth and sliding filters)

concentration data resolution, we select different time resolutions. Subsequently, different inverse algorithms are used to calculate the source strength combined with different filtering positions and filtering algorithms. Finally, we compare the performances of different inverse algorithms and get the optimal algorithm. The green parts of Figure 1 represent the denoising methods, in which the time resolution is proposed to analyze the effect of different calculation time steps. The orange parts represent the deconvolution inverse algorithms, in which the hybrid algorithms are derived and investigated.

2.1 Denoising methods

Adding a digital filter to the sensor data can reduce the effects of measurement noise (Kaiser and Reed 1977). The sliding window filter and low pass filter have been shown to markedly increase sensor sampling quality and compensate for the limitations of low-cost sensors (Li et al. 2018). Therefore, this study uses these two filtering algorithms to investigate the effects of different filtering positions and calculation time steps.

Sliding filtering reduces data fluctuations by calculating the average value within the window width. The filtering effect depends on the window width, and the wider the window is within a certain range, the more marked the filtering effect will be. The principle formula of sliding filter is as follows:

$$y_{i+\frac{m-1}{2}} = \frac{\sum_{j=i}^{i+m-1} x_j}{m}, \quad i = 1, 2, 3, \dots \quad (6)$$

where m is the window width, x_j is the data that needs filtering, and $y_{i+\frac{m-1}{2}}$ is the filtered data. The window width affects the filtering effect, and the filtering effect becomes evident as the window width increases. In this study, m was set equal to five suggested by Shi (2012).

Butterworth filtering is one type of low-pass filtering, and it can also be used to eliminate high-frequency noise. Butterworth filters are high-pass, low-pass, band-pass, band-stop and other filters, and under the same order, its filtering results inside and outside the pass frequency are typically stable compared to other filters. For a linear filter, the input $X(f)$ and output $Y(f)$ are dominated by frequency and related by a transfer function H_j , which can be expressed as follows:

$$Y(f) = H(f)X(f) \quad (7)$$

An analog low-pass Butterworth filter of order N has a transfer function:

$$H_N(f) = \frac{1}{a_0 + a_1(i2\pi f) + \dots + a_N(i2\pi f)^N} \quad (8)$$

where a_0 to a_N are coefficients; and detailed principles are explained by Manal and Rose (2007).

Different filtering positions may have effects on measurement noise reduction. Three filtering positions were analyzed: pre-filtering (i.e., filtering the monitored pollutant concentration data); post-filtering (i.e., filtering the inverse source strength); and double filtering (i.e., using both pre-filtering and post-filtering).

The time step is the data selection interval of the sensor data and determines the dimension of the response matrix. The monitoring concentration was selected based on the time step to form a concentration sampling sequence, and this sequence was used to inversely calculate the source strength. Liu et al. (2019) noted that the influence of sensor measurement noise can be reduced by adjusting the calculated time step. However, they did not confirm the specific time step that has the best noise reduction performance quantitatively. In this study, to compare the accuracy of source strength identification under different calculation time steps, we defined the time resolution

$Tr = \frac{\Delta t}{T}$, where Δt is the selected calculation time step, and T is the total calculation time.

2.2 Inverse algorithms

The inverse solution is ill-posed and unstable, and we introduce three algorithms to solve the inverse problem: Tikhonov regularization is the most common and well-known method for stabilizing the inverse solution by adding a regularization term; the iterative algorithm generates a sequence of trial solutions that converges to a final solution; and the hybrid algorithm is a combination of Tikhonov regularization and the iterative algorithms.

2.2.1 Tikhonov regularization

Tikhonov regularization avoids instability of the inverse solution of Eq. (5) by minimizing the weighted combination of the residual and marginal constraint terms, which can be expressed as the following least squares problem:

$$\min Z(q) = \|M(A)q - C\|_2^2 + \lambda^2 \|L(q - q_0)\|_2^2 \quad (9)$$

where L is the regularization matrix and λ is the regularization parameter. In Eq. (9), the first term on the right of the equal sign is the residual norm. The second term is the regularization term, where λ controls the weight for the minimization of the regularization term. Given the regularization matrix, regularization parameter, response

matrix and concentration, the source strength can be inversely calculated using filter factors and singular value decomposition (SVD) (Hansen 1994).

It is important to select the appropriate regularization parameter to determine the source strength. Based on a previous study (Kathirgamanathan et al. 2004), the GCV method has a better accuracy and more stable performance than other methods. The primary principle of the GCV method is to draw the regularization parameter function graph based on cross-validation and select the minimum GCV value. This principle is shown in the following formula:

$$G = \frac{\|M(A)q - C\|_2^2}{(\text{trace}(I_m - M(A)M(A)^T))} \quad (10)$$

where $M(A)^T$ is a matrix that produces the regularized solution q_{reg} when multiplied by C (i.e., $q_{\text{reg}} = M(A)^T C$). The denominator in this equation can be computed in $O(n)$ operations if the bidiagonalization algorithm is used.

2.2.2 Iterative methods

When the response matrix $M(A)$ is large and sparse, the iterative method is typically used. The LSQR algorithm is a common iterative method that is similar to the principle of the conjugate gradient (CG) method for solving the least squares problem, and it is based on the bidiagonalization procedure. The solution is derived from the updated QR decomposition of the bidiagonal matrix. Compared to the CG algorithm, the LSQR algorithm has faster iterative convergence and more stable solution results (Paige and Saunders 1982). The nonnegative least squares (NNLS) algorithm is a constrained version of the least squares method. Constrained least squares problems are common in physics, statistics, mathematics, and economics. NNLS requires that all coefficients cannot be negative, and the Kuhn-Tucker conditions characterize the solution vector. Solution details are described by Lawson and Hanson (1995).

In addition to these iteration methods, the computed tomography (CT) algorithm is often used to reconstruct the concentration or construction distributions and also used to inversely calculate the source strength in this study. Two most common CT algorithms are used: the SART and MLEM algorithms. The SART algorithm is one of the classical algebraic iterative reconstruction methods (Andersen and Kak 1984) that improves upon the ART and SIRT algorithms, and reduces the problems of salt-and-pepper noise caused by the ray sequence correction and slow convergence of the SIRT algorithm (Cheng 2015). The SART algorithm introduces back-projection error correction and avoids the inaccurate reconstruction results caused by a certain ray measurement error. The algorithm principle is shown as:

$$x_j^{(k+1)} = x_j^{(k)} + \frac{\sum_{i=1}^N \frac{w_{ij}}{\sum_{j=1}^N w_{ij}} \left(p_i - \sum_{j=1}^N w_{ij} x_j^{(k)} \right)}{\sum_{i=1}^N w_{ij}} \quad (11)$$

where k represents the number of iterations, and $x_j^{(0)}$ is the initial iteration value. The second item on the right side of Eq. (11) is the correction value of the back-projection error, where p_i represents the real measured projection value, and w_{ij} is the weighting factor. The x_j from each iteration is the input value of the next iteration until the iteration converges. For the source strength calculation, x_j in Eq. (11) represents the source strength sequence q , w_{ij} represents the response matrix $M(A)$, and p_i is the pollutant concentration sequence C monitored by the sensor.

The MLEM algorithm is another popular statistical iterative method that establishes the maximum likelihood estimation function and then obtains the maximum value of the likelihood function (Shepp and Vardi 1982; Cheng 2015). The solution corresponding to the maximum value is the required value, and the specific formula is as follows:

$$x_j^{(k+1)} = \frac{x_j^{(k)} \sum_i w_{ij} \frac{p_i}{\sum_j w_{ij} x_j^{(k)}}}{\sum_i w_{ij}} \quad (12)$$

where $\sum_j w_{ij} x_j^{(k)}$ is the estimated projection value. The calculated estimated value is compared to the real measured projected value p_i , and the ratio is back-projected on the reconstruction area by the weighting factor w_{ij} . The estimated value is revised repeatedly until the optimal solution is obtained. For the source strength calculation, the meanings of x_j , w_{ij} and p_i in Eq. (12) are the same as those in the SART algorithm.

2.2.3 Hybrid algorithms

The hybrid algorithm synthesizes the advantages of two or more algorithms to solve the inverse problem. The combination of the total variation regularization method and Tikhonov regularization method can markedly improve the reconstruction effect of resistivity imaging (Han et al. 2012). The combination of the Lanczos method and GCV method can overcome the semi-convergence of the Lanczos method, thus making the solution less sensitive to the number of iterations (Chung et al. 2008). As mentioned above, the Tikhonov regularization algorithm can effectively reconstruct the smooth data solution; however, it becomes difficult to solve the problem with large sparse matrix. Whereas, the iterative method is suitable for the solution of a large sparse matrix, and the solution is stable after several iterations. Therefore, this paper combines the Tikhonov regularization algorithm with the iteration method to study the performances of four

hybrid algorithms: GCV+LSQR, GCV+NNLS, LSQR+GCV and NNLS+GCV.

The principle of the GCV+LSQR algorithm is to calculate the source strength q_0 from the LSQR algorithm, and then q_0 is used as the initial value in Tikhonov regularization (Eq. (9)) based on the GCV method. The principle of the GCV+NNLS algorithm is similar to that of the GCV+LSQR algorithm except q_0 comes from the NNLS algorithm.

For LSQR+GCV, an alternative way to use LSQR is to solve the Tikhonov regularization (Eq. (9)) by applying LSQR to an equivalent form of Eq. (9):

$$\min \left\| \begin{bmatrix} M(A) \\ \lambda L \end{bmatrix} q - \begin{bmatrix} C \\ \lambda L q_0 \end{bmatrix} \right\|^2 \quad (13)$$

where the regularization parameter is also calculated by the GCV method. The principle of the NNLS+GCV algorithm is similar to that of LSQR+GCV except for its nonnegative constraint.

2.3 Evaluation index on the performances of inverse algorithms

A comprehensive evaluation index combining the threshold ratio (THR) with the coefficient of variation (CV) was used to evaluate the accuracy and stability of each algorithm in different cases.

2.3.1 THR for accuracy evaluation

We introduce the relative error to quantify the difference between the inverse solution and the real source release rate:

$$\text{error} = \frac{|N_{\text{IR}} - N_{\text{AR}}|}{N_{\text{AR,Peak}}} \quad (14)$$

where N_{IR} is the source inverse release rate, N_{AR} is the real release rate, $N_{\text{AR, peak}}$ is the peak value of the real release rate, and $N_{\text{AR, peak}}$ is 5 L/min in the simulation and 10 L/min in the experiment. Referring to the study of Emmerich et al. (2003) and Sreedharan et al. (2006), a threshold of 25% relative error was chosen. For comparing the performances of different filtering algorithms and inverse algorithms under experimental and simulated conditions, a unified threshold value (25%) was used. The ratio of the number of time nodes greater than this threshold to the total number of time nodes is defined as THR. The choice of THR is only related to the error between the inverse calculated source strength and the actual source strength, and not to the form of the source release. A smaller THR indicates that the inverse source strength is closer to the real source release rate.

2.3.2 CV for stability evaluation

CV is used to evaluate the robustness of the algorithms in the source-strength inversion (Cui et al. 2019) and is defined as:

$$\text{CV} = \frac{\mu_{ij}}{N_{\text{AR,peak}}} \quad (15)$$

where i represents the studied cases that will be introduced below, j represents the nine inverse algorithms, and μ_{ij} represents the standard deviation between the inverse source strength and the real release rate. Similarly, a smaller CV indicates better stability.

2.3.3 Comprehensive evaluation index

To analyze the accuracy and stability of source inverse algorithms in simulation and experimental cases, we comprehensively scored the inverse algorithms based on THR and CV. Because the ranges of THR and CV are not consistent, they must be scaled to the same dimension. The maximum difference normalization method (Cui et al. 2019) is used to normalize the ranges, and the ranges is converted to 0–1. The conversion formula is as follows:

$$x'_{(ij),n} = (x_{(ij),n} - \min(x_{(ij),n})) / (\max(x_{(ij),n}) - \min(x_{(ij),n})) \quad (16)$$

where n is THR or CV; $x_{(ij),n}$ is the THR or CV value of a certain case, and $x'_{(ij),n}$ is its dimensionless value; $\min(x_{(ij),n})$ and $\max(x_{(ij),n})$ are the minimum and maximum values of THR or CV. The accuracy and stability of THR and CV are considered in the comprehensive scoring as follows:

$$S_j = \omega_1 \times x'_{(ij),\text{THR}} + \omega_2 \times x'_{(ij),\text{CV}} \quad (17)$$

where ω represents the weight ratio of each part. Because the accuracy and stability of the algorithms were equally important when using different algorithms for source strength deconvolution calculations, $\omega_1 = \omega_2 = 0.5$ in this study (Mao et al. 2020). Because the comprehensive evaluation index is based on THR and CV, a smaller S_j indicates better accuracy and stability of the inverse algorithms.

2.4 Case study

Three simulated cases and two experimental cases are used to investigate the performances of the filtering and inverse algorithms. The cases were based on an apartment unit (Cases 1, 2, 3) and an entire floor of the office building (Cases 4, 5). Figure 2(a) shows the layout of the apartment unit in the simulation. The positions of the sensors, pollution source and fan are shown in the figure. The windows in all rooms were closed except for Balcony 2 and Bedroom 2.

CO₂ was released as a pollutant source at a rate of 5 L/min in Bedroom 2, and the window in Bedroom 2 was used as a fan inlet ventilation to mix the indoor pollutants. The results of the pollutant release concentration prediction by CONTAM simulation agree with the experimental data (Liu et al. 2019), and the sensor concentration response factors for Bedroom 2 can be obtained from the CONTAM software simulation as shown in Figure 2(b). Three simulation cases (Cases 1, 2 and 3) are set up for source identification analysis.

Details of the simulation cases are shown in Table 1. In this study, calculation time steps of 1 s, 2 s, 5 s, 10 s, 15 s and 20 s were analyzed. Additionally, the concentration data from sensors were added with uniform noise to represent measurement error.

Figure 2(c) shows the layout of the scaled floor of the office building in the experiment. The multi-zone model is made of an acrylic sheet that is 1.35 m long, 1.2 m wide and 0.2 m thick. Four sensors (Telaire 6615, GE company, accuracy:

Fig. 2 Basic information for the study cases: (a) multi-zone layouts for the simulation case (Liu et al. 2019; reprinted with permission ©2019 Elsevier), (b) sensor response factors for the source in simulation, (c) multi-zone layouts for the experiment case (Zhuang et al. 2021; reprinted with permission ©2021 Elsevier), (d) sensor response factors for the source in experiment

Table 1 Details of the cases

Case	Source release form	Source release rate (L/min)	Source release time (s)
Case1	Constant	5	1500
Case2	Simulated cases	Periodic (600 s/period)	1500
Case3*		Periodic (600 s/period)	1500
Case4	Experimental cases	Constant	300
Case5		Periodic (100 s/period)	300

* Indicates that the flow field of this case is unsteady, and the rest cases are steady flow.

± 75 ppm or 10% of reading) are arranged inside the model, as shown in Figure 2(c), and CO₂ (carbon dioxide) is used as a tracer gas for this experiment; thus, an additional sensor outside the model is used to detect the background concentration of CO₂ in the air. The background concentration was subtracted from the pollutant concentrations monitored by the sensors inside the multi-zone model to avoid the impact of background concentration fluctuation. The five sensors are powered by a 24-V power supply, and concentration data are transmitted to a computer by an Ethernet device (WISE-4051-AE, Advantech Co., Ltd.). Doors are provided between the zones within the model, and windows are provided in the surrounding zones. The door opening conditions between the zones are shown in Figure 2(c). Two fans (F-105, PCCOOLER Co. Ltd.) were installed at the windows of Zones 8 and 12 to simulate mechanical ventilation in the zone, and for Zones 8 and 12, the windows from the fan were 28.77 m³/h and 9.96 m³/h, respectively.

First, the impulse response method was used to obtain the response matrix. Pollutants were released in Zone 21 at a known release rate q_{i0} for an impulse period of 2 s. The response factor (F_{tk} , $k = 1, \dots, n$) is obtained from the monitored concentrations (C_{tk} , $k = 1, \dots, n$) from four sensors. The impulse response factor can be expressed as: $F_k = C_{tk}/q_{i0}$. Figure 2(d) shows the corresponding time response vector **A** for Zone 21 obtained from the different sensors. Since only Sensor A and Sensor B have significant concentration responses, the response matrix **M(A)** was constructed using these two sensor concentration responses for source identification. The data sampled from the pulse experiments were also analyzed using the statistical analysis software G.Power 3.1, yielding a minimum sample size of seven per group of experiments required to ensure statistical power of more than 0.9. The sensor error and response measurement uncertainty were both considered in the analysis, and the specific method can be found in Zhuang et al. (2021). For the sensor time delay, because the response vector measurement had considered the effect of sensor time delay, in turn, when using the response vector to estimate the source strength, the effect of sensor time delay would be offset. Then, we set up two cases (Cases 4 and 5) for source identification. Details of the experimental cases are shown in Table 1. Because the computational time step could not be less than the impulse period (2 s), the calculation time steps selected in experiment cases were 2 s, 3 s, 4 s, 10 s and 15 s.

3 Results

3.1 Effects of filtering positions

Digital filters are used to filter the measurement noise from sensors, and filtering positions may have effects on the

inverse results. Three filtering positions (pre-filtering, post-filtering, double filtering) are analyzed combining with sliding and Butterworth filters. The GCV algorithm was chosen for the filter position analysis, because the inverse values through GCV algorithm were most sensitive to the perturbation. The averaged THR for different filtering positions are shown in Table 2.

From Table 2, with the sliding filter, the post-filtering yields lower THR and performs well on both simulated and experimental data. With the Butterworth filter, the performances of different filtering positions are inconsistent between the simulations and experiments. In the simulation, the post-Butterworth filtering has the lowest THR, while in the experiment, the double-Butterworth filtering yields the best performance. This may be because the Butterworth filter is more sensitive to noise in the measured data, and the noise in the experimental data is more complicated than that of the simulated data. The THR does not differ much for the different filter positions, especially for the experimental cases. This implies that different filtering positions have little effect on the results of source strength calculation. In addition, for other time steps, we also found that the filter position had little effect. Therefore, the post-filtering, which performed relatively better, was chosen for the subsequent study in this paper.

3.2 Effects of time resolutions

Based on the results regarding post-filtering position, this section further analyzes the accuracy of inverse source strength under different time steps. In a previous study, it was found that increasing the calculation time step properly could reduce the measurement noise effect. However, decreasing the calculation time step would ensure a higher resolution for the inverse source strength (Wang and You 2015; Chata et al. 2017). Therefore, the optimal calculation time step should balance the inverse accuracy and resolution. In this study, Tr representing the ratio of the calculation time step to the total calculation time of the source strength was employed; and we used THR as the error analysis index. The point at which the THR reached its minimum or a stable value is the optimal Tr value. The optimal Tr

Table 2 THR for different filtering positions

Filtering position	Simulation (%)	Experiment (%)
Pre-sliding	48.95	45.87
Post-sliding	34.77	42.28
Double-sliding	45.27	45.10
Pre-Butterworth	40.58	43.55
Post-Butterworth	32.43	46.14
Double-Butterworth	37.34	42.94

analysis for the simulation and experiment are shown in Figures 3 and 4.

In Figure 3(a), the THR curves of Cases 2 and 3 reach their minimum when $Tr = 1\%$ when no filtering is used. In Figure 3(b), the THR curves are marginally different when using the sliding filter. And when using the Butterworth filter (Figure 3(c)), the THR curves change markedly: the

minimum or stable points move forward to $Tr = 0.667\%$, and the THR decreases compared with the values without filtering. This is because the deconvolution process for the source strength calculation can amplify the high frequency noise in the measurement data, and the Butterworth low-pass filter can filter out the high frequency noise by setting the passband and stopband frequencies. Therefore, the Butterworth filter could improve the performance of inverse calculation, i.e., the accuracy and resolution of inverse source strength. In addition, from Figure 3, the THRs for Case 3 are much larger than those of Case 1 and Case 2. This might be because the flow fields of Case 1 and Case 2 were steady, while the flow field of Case 3 was unsteady and periodic. The complicated flow field increased the ill-conditioned degree of reverse calculation.

Figure 4 shows that both filters can markedly reduce errors in the inverse source strength, and the Butterworth filtering performs better under smaller Tr ($Tr < 1.33$). In Case 4 (constant source) (Figure 4(a)), with Butterworth filtering, the minimum THR value is $Tr = 1.33\%$. Although the THRs with the sliding filter are lower when $Tr = 5\%$, the larger calculation step can yield lower solution resolution

Fig. 3 THR for source strength identification at different calculation Tr in simulation: (a) no filter, (b) adding sliding filter, (c) adding Butterworth filter

Fig. 4 THR for source strength identification with different calculation Tr in experiment: (a) Case 4, (b) Case 5

and may not be suitable to identify the strength of a highly dynamic source. In Case 5 (periodic source) (Figure 4(b)), with Butterworth filtering, the minimum THR is also at $Tr = 1.33\%$. From Figure 4, it is found that both larger and smaller Tr will lead to larger THRs. This may be because when the Tr is too large, the sampling size of the input monitoring concentration will become small and difficult to reflect the time-varying characteristic of the pollutant source strength. While as Tr becomes too small, the condition number of the response matrix will become large, and the deconvolution calculation will be more sensitive to the perturbation of measurement noise. The optimal Tr should balance these two aspects. Based on the analysis above, $Tr = 0.667\%$ and 1.333% with the Butterworth filter are the optimal Tr for the simulation and experiment, respectively.

3.3 Comparison of single and hybrid algorithms

Based on the post-Butterworth filtering and optimal Tr value, this section compares the accuracy and stability of nine inverse algorithms, including single and hybrid algorithms. Additionally, algorithms were scored based on their performances, and the highest scores were selected for both simulated and experimental cases.

3.3.1 Accuracy and stability of source strength identification

THR and CV are used to evaluate the accuracy and stability of source recognition, respectively.

In Figure 5(a), the inverse source strengths in the simulated cases (Cases 1, 2, and 3) have slightly lower THR than those in the experimental cases. The hybrid algorithms show no obvious improvement in the accuracy of the source strength inversion. In the constant source cases (Cases 1 and 4), the GCV, SART and MLEM algorithms yield the best performances in source identification. Among those three algorithms, the THRs of the SART algorithm for the experiment (Case 1) and simulation (Case 4) are 2.98% and 17.88%, respectively. In Case 4, the THRs of the hybrid algorithms are all higher than 60%, which indicates that these hybrid algorithms are not suitable for constant source identification. For the periodic source cases (Cases 2, 3, and 5), the NNLS and NNLS + GCV algorithms perform well along with the SART and MLEM algorithms. In addition, as shown in Figures 4(b) and 5(a), the THRs of Case 5 with different algorithms are almost larger than 40%. As mentioned in Section 2, the inverse calculation is a deconvolution process, and it will amplify the larger concentration fluctuation due to the periodic source. The signal amplification characteristics of the inverse convolution process was the essential reason for higher THRs. In summary, the SART and MLEM algorithms perform well in both experiments and simulations,

Fig. 5 Accuracy and stability of source strength recognition with different algorithms: (a) THR, (b) CV (the white part: constant source; the yellow part: periodic source)

particularly in the experimental cases, where the uncertainty of external perturbation is large.

In Figure 5(b), the CVs of the NNLS, GCV+NNLS and GCV+LSQR algorithms are markedly higher than those of the other algorithms, which indicates that the inversion based on these algorithms is not stable, particularly in the experimental cases (Cases 4 and 5). The inverse results from other algorithms, such as GCV, SART and MLEM, are relatively stable in all cases. Combining the THR and CV data from Figures 5(a) and (b), the algorithms are scored based on their accuracy and stability performances in the next section.

3.3.2 Algorithm ranking

After the dimensionless processing of THR and CV based on Eq. (16), the scores of the nine algorithms are given by Eq. (17). Lower scores mean better performance, and Table A1 (in the Appendix) shows the ranking of the algorithms from best to worst. The three best performing algorithms are also shown in Table 3.

The scoring results in Table 3 show that the SART algorithm ranks first in most cases, indicating that this

Table 3 Algorithm scoring results

Source release form	Constant source		Periodic source		
Case	Case1	Case4	Case2	Case3	Case5
Three best algorithms	SART	SART	GCV	SART	SART
	MLEM	MLEM	GCV+LSQR	NNLS	MLEM
	NNLS+GCV	GCV	NNLS	MLEM	NNLS+GCV

algorithm has the best adaptability. This may be because SART algorithm can suppress the amplification of the noise signal in the deconvolution calculation by point-by-point correction which corrects the accumulation of the projection errors from all monitored values. It is more suitable for the calculation of large and sparse matrices like $M(A)$. As shown in Table A1, the hybrid algorithms do not necessarily perform better than single algorithms. For example, the GCV+LSQR and GCV+NNLS algorithms perform worst in most cases, and these two algorithms are not suitable for source identification. This is probably because the convergence of LSQR algorithm is delayed due to the influence of the finite precision arithmetic, and this algorithm would compute some “ghost” eigenvalues and singular values. For the NNLS algorithm, although it maintains the non-negativity of the results, it cannot reconstruct the source strengths vary accurately, because its results are rough and contain many zero values. Therefore, when the prior source strengths calculated by the LSQR or the NNLS algorithm are brought into the GCV algorithm as the initial values, the influence of the small singular values is not filtered out and the bias is even amplified, resulting in deviation of the calculated source strengths. The NNLS+GCV and LSQR+GCV algorithms also show more variability in performance across different cases. In summary, the regularization approach does not perform as well as the iterative algorithm, and the hybrid algorithms combining the regularization and iterative algorithms also do not perform better. Therefore, the SART algorithm yields the best performance among the inverse algorithms examined in this study.

4 Discussion

In order to compare the inverse source strengths through different algorithms more visually, the results of Case 2 and Case 5 through algorithms with good, medium and bad scores (SART, LSQR+GCV and GCV+NNLS) are shown in Figure 6.

From Figures 6(a) and (b), the SART algorithm has better results in source strength estimation compared to LSQR+GCV and GCV+NNLS algorithms, and GCV+NNLS algorithm performs poorly in both experiment (Case 5) and simulation (Case 2). These findings are consistent with

Fig. 6 Estimated source strengths of (a) Case2 (from Sensor 2), (b) Case5 (from Sensor A)

those in Table 3 and confirm the advantages of SART algorithm. The SART algorithm can be used for both source strength identification and reconstruction of pollutant concentration distributions, as well as for medical imaging, demonstrating its general applicability. The algorithm can reduce the high frequency error amplification problem caused by deconvolution process through point-by-point correction. However, the SART algorithm also requires a sequence of response factors F to be obtained in advance, and the

quality of its reconstruction is related to the quality of the response matrix and input data.

In this study, the introduction of iterative methods commonly used in CAT into the source strength inverse calculation provides a new perspective on the pollutant source identification field. And the application of sensor noise reduction methods is not only limited to the field of source identification, but also can be used wherever digital signal processing is involved. We conduct the experiment in a scaled multi-zone building model instead of a real office building, because, for the model comparison and analysis, well-controlled and simple cases are usually selected. For example, a lot of research has assessed various turbulence models for airflow prediction in built environment. These studies almost conducted airflow measurement in controllable test cabins, but rarely in real and complex buildings. This is because the thermal and flow boundary conditions of real buildings are complex, and it is difficult to obtain reliable data to evaluate the performance of different models. Therefore, the experimental data from the controllable scaled building model are used in this study. In addition, for the scaled multi-zone building, there are 31 zones and 41 airflow paths, which can form more than 31×2^{41} cases; it is thus impractical to calculate every case. Because we only analyzed the inverse calculation of the source strength, we selected representative experimental cases to investigate the filtering and inverse algorithms. For the more sophisticated cases, the measurement method for the response matrices and the inverse algorithms are almost the same as the methods used in this study. Therefore, the results of this study have important reference value for more sophisticated cases. We will further validate the findings of this paper in more complicated source release and ventilation scenarios in the future.

The sensors used in the simulation are ideal sensors that are not subject to interference from external factors; however, the sensors in the experiment have errors in monitoring concentration data. Therefore, the optimal Tr is not consistent between the experiments and simulations, and the error of source strength identification in the experiment is higher than that in the simulation. Additionally, this study does not consider the effect of sensor location, and the data from the sensor with the most significant response are used to analyze filtering and inverse algorithms.

5 Conclusions

This study explored the optimization scheme for the deconvolution process in source strength inverse calculation. Different measures to reduce the effects of measurement noise were investigated. On this basis, the performances of single and hybrid inverse algorithms under experimental

and simulated conditions were studied, and all algorithms in different scenarios were scored and compared. The main conclusions are as follows:

1) Different filtering positions had little effect on the inverse calculation, and the Butterworth filter was more effective than the other types. Based on the post-Butterworth filtering, the minimum error was found at $Tr = 0.667\%$ for the simulated cases and at $Tr = 1.333\%$ for the experimental cases.

2) The error analysis and algorithm scoring of the nine inverse algorithms showed that the source identification of hybrid algorithms was not necessarily better than that of single algorithms. For example, the GCV+LSQR and GCV+NNLS algorithms could even amplify the error and distort the results completely. The SART algorithm, which is widely used for image reconstruction, yielded good results with source identification, and its superiority was even more evident during experimental conditions. Therefore, iterative algorithms such as the SART algorithm are recommended for future research.

3) Using optimization scheme proposed in this paper, the relative errors between the inverse source strength and the actual source strength were mostly less than 25% in both experimental and simulated cases. However, the errors of source strength identification in the experiment were higher than those in the simulation.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (No. 51708286), the Natural Science Foundation of Jiangsu Province (No. BK20180701) and Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. SJCX20_0325).

References

- Andersen AH, Kak AC (1984). Simultaneous algebraic reconstruction technique (SART): A superior implementation of the art algorithm. *Ultrasonic Imaging*, 6: 81–94.
- Aster RC, Borchers B, Thurber CH (2005). *Parameter Estimation and Inverse Problems*. Oxford, UK: Academic Press.
- Bhatti P, Newcomer L, Onstad L, et al. (2011). Wood dust exposure and risk of lung cancer. *Occupational and Environmental Medicine*, 68: 599–604.
- Chan JF-W, Yuan S, Kok K-H, et al. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. *The Lancet*, 395: 514–523.
- Chata F, Belut E, Maillet D, et al. (2017). Estimation of an aerosol source in forced ventilation through prior identification of a convolutive model. *International Journal of Heat and Mass Transfer*, 108: 1623–1633.

- Chen C, Zhao Y, Zhao B (2018). Emission rates of multiple air pollutants generated from Chinese residential cooking. *Environmental Science & Technology*, 52: 1081–1087.
- Cheng L (2015). Iterative tomographic algorithms of gas diffusion distribution reconstruction based on incomplete projection data. Master Thesis, Hefei University of Technology, Chinese. (in Chinese).
- Chung J, Nagy JG, O’Leary DP (2008). A weighted-GCV method for Lanczos-hybrid regularization. *Electronic Transactions on Numerical Analysis*, 28: 149–167.
- Cui J, Lang J, Chen T, et al. (2019). Investigating the impacts of atmospheric diffusion conditions on source parameter identification based on an optimized inverse modelling method. *Atmospheric Environment*, 205: 19–29.
- Emmerich SJ, Nabinger SJ, Gupta A (2003). Comparison of measured and predicted tracer gas concentrations in a townhouse. NIST Interagency/Internal Report (NISTIR), 7035. Gaithersburg, MD, USA: National Institute of Standards and Technology.
- EPA (2017). Report on the environment about indoor air quality: What are the trends in indoor air quality and their effects on human health? U.S. Environmental Protection Agency. Available at <https://www.epa.gov/report-environment/indoor-air-quality>.
- Etheridge DW, Sandberg M (1996). Building Ventilation: Theory and Measurement: Chapters 12&13. Chichester, UK: John Wiley & Sons.
- Feng Q, Cai H, Chen Z, et al. (2019). Experimental study on a comprehensive particle swarm optimization method for locating contaminant sources in dynamic indoor environments with mechanical ventilation. *Energy & Buildings*, 196: 145–156.
- Han B, Dou Y, Ding L (2012). Electrical resistivity tomography by using a hybrid regularization. *Chinese Journal of Geophysics*, 55(3): 970–980. (in Chinese)
- Hancock DG, Langley ME, Chia KL, et al. (2015). Wood dust exposure and lung cancer risk: A meta-analysis. *Occupational and Environmental Medicine*, 72: 889–898.
- Hansen PC (1994). REGULARIZATION TOOLS: A Matlab package for analysis and solution of discrete ill-posed problems. *Numerical Algorithms*, 6: 1–35.
- Hansen C (2010). Discrete Inverse Problems: Insight and Algorithms. Lyngby, Denmark: The Society for Industrial and Applied Mathematics (SIAM).
- Hiyama K, Kato S, Ishida Y (2010). Thermal simulation: Response factor analysis using three-dimensional CFD in the simulation of air conditioning control. *Building Simulation*, 3: 195–203.
- Kaiser JF, Reed WA (1977). Data smoothing using low-pass digital filters. *Review of Scientific Instruments*, 48: 1447–1457.
- Kathirgamanathan P, McKibbin R, McLachlan RI (2004). Source release-rate estimation of atmospheric pollution from a non-steady point source at a known location. *Environmental Modeling & Assessment*, 9: 33–42.
- Lawson CL, Hanson RJ (1995). Solving Least Squares Problems. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).
- Li Y, Duan S, Yu ITS, et al. (2005). Multi-zone modeling of probable SARS virus transmission by airflow between Flats in Block E, Amoy Gardens. *Indoor Air*, 15: 96–111.
- Li J, Li H, Ma Y, et al. (2018). Spatiotemporal distribution of indoor particulate matter concentration with a low-cost sensor network. *Building and Environment*, 127: 138–147.
- Li F, Cai H, Xu J, et al. (2020a). Gas distribution mapping for indoor environments based on laser absorption spectroscopy: Development of an improved tomographic algorithm. *Building and Environment*, 172: 106724.
- Li F, Liu X, Liu J, et al. (2020b). Solutions to mitigate the impact of measurement noise on the air pollution source strength estimation in a multi-zone building. *Building Simulation*, 13: 1329–1337.
- Liang W, Yang X (2013). Indoor formaldehyde in real buildings: Emission source identification, overall emission rate estimation, concentration increase and decay patterns. *Building and Environment*, 69: 114–120.
- Liu X, Li F, Cai H, et al. (2019). Dynamical source term estimation in a multi-compartment building under time-varying airflow. *Building and Environment*, 160: 106162.
- Manal K, Rose W (2007). A general solution for the time delay introduced by a low-pass Butterworth digital filter: An application to musculoskeletal modeling. *Journal of Biomechanics*, 40: 678–681.
- Mao S, Lang J, Chen T, et al. (2020). Impacts of typical atmospheric dispersion schemes on source inversion. *Atmospheric Environment*, 232: 117572.
- Paige CC, Saunders MA (1982). LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software*, 8: 43–71.
- Pang L, Wang W, Qu H, et al. (2014). Approach to identifying a sudden continuous emission pollutant source based on single sensor with noise. *Indoor and Built Environment*, 23: 955–970.
- Sandberg M, Blomqvist C (1985). A quantitative estimate of the accuracy of tracer gas methods for the determination of the ventilation flow rate in buildings. *Building and Environment*, 20: 139–150.
- Shepp LA, Vardi Y (1982). Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1: 113–122.
- Sherman MH, Grimsrud DT, Condon PE, et al. (1980). Air infiltration measurement techniques. Berkeley. In: Proceedings First Air Infiltration Centre Conference, Windsor, UK.
- Shi Y (2012). The application of the Butterworth low-pass digital filter on experimental data processing. In: Proceedings of the International Conference in Electrics, Communication and Automatic Control Proceedings.
- Sreedharan P, Sohn MD, Gadgil AJ, et al. (2006). Systems approach to evaluating sensor characteristics for real-time monitoring of high-risk indoor contaminant releases. *Atmospheric Environment*, 40: 3490–3502.
- Tsantaki E, Smyrnakis E, Constantinidis TC, et al. (2020). Indoor air quality and sick building syndrome in a university setting: A case study in Greece. *International Journal of Environmental Health Research*, <https://doi.org/10.1080/09603123.2020.1789567>
- Wang L, You X (2015). Identification of indoor contaminant source location by a single concentration sensor. *Air Quality, Atmosphere & Health*, 8: 115–122.

- Wang H, Lu S, Cheng J, et al. (2017). Inverse modeling of indoor instantaneous airborne contaminant source location with adjoint probability-based method under dynamic airflow field. *Building and Environment*, 117: 178–190.
- Yan J, Grantham M, Pantelic J, et al. (2018). Infectious virus in exhaled breath of symptomatic seasonal influenza cases from a college community. *Proceedings of the National Academy of Sciences of the United States of America*, 115: 1081–1086.
- Yang Y, Feng Q, Cai H, et al. (2019). Experimental study on three single-robot active olfaction algorithms for locating contaminant sources in indoor environments with no strong airflow. *Building and Environment*, 155: 320–333.
- Yin S (2011). Quantitatively identify unsteady gas pollutant releases in indoor environment by inverse CFD modeling. Master Thesis, Dalian University of Technology, China. (in Chinese)
- Young DL, Tsai CC, Chen CW, et al. (2008). The method of fundamental solutions and condition number analysis for inverse problems of Laplace equation. *Computers & Mathematics with Applications*, 55: 1189–1200.
- Yu B (2006). Open Path Fourier Transform Infrared Spectroscopy (OP-FTIR) for Atmospheric Environment Monitoring. Master Thesis, Nanjing University of Science and Technology, Nanjing. (in Chinese).
- Zhang TT, Yin S, Wang S (2013). An inverse method based on CFD to quantify the temporal release rate of a continuously released pollutant source. *Atmospheric Environment*, 77: 62–77.
- Zhuang J, Li F, Liu X, et al. (2021). An experiment-based impulse response method to characterize airborne pollutant sources in a scaled multi-zone building. *Atmospheric Environment*, 251: 118272.

Appendix

Table A1 Algorithm score summary

	Case									
	Case1		Case2		Case3		Case4		Case5	
Algorithmic scoring ranking	0.015	SART	0.077	GCV	0.122	SART	0.090	SART	0.164	SART
	0.018	MLEM	0.080	GCV+LSQR	0.177	NNLS	0.094	MLEM	0.169	MLEM
	0.023	NNLS+GCV	0.084	NNLS	0.186	MLEM	0.114	GCV	0.194	NNLS+GCV
	0.025	GCV	0.090	LSQR	0.187	NNLS+GCV	0.156	LSQR	0.250	LSQR+GCV
	0.033	LSQR	0.099	NNLS+GCV	0.203	GCV	0.183	NNLS	0.253	LSQR
	0.035	LSQR+GCV	0.114	SART	0.209	GCV+LSQR	0.313	NNLS+GCV	0.253	NNLS
	0.050	NNLS	0.114	LSQR+GCV	0.234	LSQR	0.385	LSQR+GCV	0.332	GCV
	0.070	GCV+LSQR	0.117	MLEM	0.253	GCV+NNLS	0.450	GCV+LSQR	0.434	GCV+NNLS
	0.198	GCV+NNLS	0.121	GCV+NNLS	0.256	LSQR+GCV	0.781	GCV+NNLS	0.436	GCV+LSQR