Sensitivity analysis of the probability-based inverse modeling method for indoor contaminant tracking

Zhiqiang (John) Zhai* and Xiang Liu
Department of Civil, Environmental and Architectural Engineering, University of Colorado at Boulder, UCB 428, ECOT 441, Boulder, CO 80309-0428, USA

Abstract

Accurate and prompt identification of contaminant sources ensures that the contaminant sources can be quickly removed and contaminated spaces can be isolated and cleaned. The adjoint probability method shows great potential to identify indoor pollutant sources with limited pollutant concentration data from sensors. Application of the method to the reality with unideal conditions such as transient velocity and inaccurate measurement of contaminant concentration requires a sensitivity analysis of the method to these critical parameters. The study finds that with up to 90% of random errors in indoor air flow velocity, the inverse algorithm is still able to produce acceptable predictions, as long as the flow pattern remains the same. In a reasonable yet wide range of contaminant concentration accuracy ([0.01, 100] of the sensor accuracy), the measurement error will not influence the capability of the inverse algorithm to predict the correct source location. This paper further proposes an approach to prescribing the required but presumed contaminant mass range so that the algorithm is able to properly predict the source location.

Keywords: sensitivity analysis; adjoint probability method; indoor contaminant tracking

*Corresponding author: John.Zhai@Colorado.edu Received 3 May 2016; accepted 23 May 2016

1 INTRODUCTION

Indoor environment quality exacerbation and frequent building safety issues call for effective control and improvement measures for indoor environment. Accurate and prompt identification of contaminant sources ensures that the contaminant sources can be quickly removed and contaminated spaces can be isolated and cleaned. The source identification process is essentially an inverse modeling problem in that the results of an indoor air quality (IAQ) event—contaminant concentrations, are used to find the causes of the event—contaminant source(s). Inverse modeling methods have been developed for the identification of all kinds of source conditions [1–6], in which the adjoint probability method shows most promising that can identify potential indoor pollutant sources with limited pollutant concentration data from sensors [4–7].

Both numerical and experimental tests have verified the effectiveness and accuracy of the probability based algorithm [8, 9]. However, before the method finds successful applications in the real world, sensitivity issue has to be addressed. This issue emerges when parameters used in obtaining location probabilities in numerical and laboratory studies may be different from the actual ones, resulting in perturbed location probabilities, upon which predictions are made. This paper therefore investigates the potential changes of prediction due to the variations in input parameters between numerical model and the reality.

2 OVERVIEW OF PROBABILITY-BASED INVERSE MODELING METHOD

The literature review of existing inverse transport modeling methods [7] shows that most models merely deal with source release flux or history. For those dedicated to source location identification, potential source zones need be specified. The only exception is the adjoint probability method developed by...
Neupauer and Wilson [10, 11] for identifying groundwater pollution, which theoretically can find source locations, flux and activate time with no prior information. For most building indoor environment incidents, events often occur with completely unknown source conditions including locations, release intensity and activation time. Hence, the adjoint probability method appears promising to be adopted for indoor airborne pollutant tracking.

One of the key concepts of the adjoint probability method is the location probability. In a flow domain of interest, if a contaminant parcel is released from a point source at $\mathbf{x}_0$ and $t = 0$, the forward location probability of the parcel is defined as the probability that it reaches some location in the domain at a given time $t = T > 0$. Assume the pollution source releases instantaneous contaminants with a mass of $M_0$ into the domain at $t = 0$, which spread all over the domain at $t > 0$ due to both convection and diffusion flows. If the pollutant mass trapped in a small finite volume $\Delta V_1$ at $\mathbf{x} = \mathbf{x}_1$ at time $t = T$, the forward location probability at this finite volume $\Delta V_1$ at $t = T$ is

$$P(\Delta V_1|\mathbf{x} = \mathbf{x}_1; t = T, \mathbf{x}_0) = \frac{M_1}{M_0}$$

(1)

The location probability density function at $t = T$ is then defined as

$$f_s(\mathbf{x}_1; t = T, \mathbf{x}_0) = \frac{P(\Delta V_1|\mathbf{x} = \mathbf{x}_1; t = T, \mathbf{x}_0)}{\Delta V_1} = \frac{M_1}{M_0} = \frac{C_1}{M_0}$$

(2)

where $C_1$ is the resident concentration at $\mathbf{x} = \mathbf{x}_1$, defined as a measure of the mass of pollutant per unit volume of flow medium, or a volume-averaged concentration. Generalizing the definition to any location in the domain yields

$$f_s(\mathbf{x}; t = T, \mathbf{x}_0) = \frac{C(\mathbf{x}, T)}{M_0}$$

(3)

where $f_s(\mathbf{x}; t = T, \mathbf{x}_0)$ is the forward location probability density at $t = T, \mathbf{x}$ is the arbitrary location vector in the domain, $C(\mathbf{x}, T)$ is the distribution of resident contaminant concentration at $t = T$ due to the instantaneous release of the point source, $\mathbf{x}_0$ is the source location, $M_0$ is the total source release mass. For the cases with steady-state velocity field, there is a linear relationship between source release strength and resident concentration, which leads to

$$f_s(\mathbf{x}; t = T, \mathbf{x}_0) = \frac{dC(\mathbf{x}, T)}{dM_0} = \psi_0(\mathbf{x}; t = T, \mathbf{x}_0)$$

(4)

where $\psi_0(\mathbf{x}; t = T, \mathbf{x}_0)$ is the state sensitivity of resident concentration at $\mathbf{x}$ to the source mass $M_0$ at $\mathbf{x}_0$.

The forward location probability density function $f_s(\mathbf{x}; t = T, \mathbf{x}_0)$ describes the possibility of a contaminant parcel, originating from an instantaneous source at $\mathbf{x}_0$, to be at an arbitrary $\mathbf{x}$ after a fixed time $t = T$. From statistics, this forward location probability is equal to the possibility of the parcel found at $\mathbf{x}$ when $t = T$ to be at the source location $\mathbf{x}_0$ at $t = 0$ (or $\tau = T$ ago if a new backward time sign $\tau = T - t$ is defined), which is named as backward location probability $f_b(\mathbf{x} = \mathbf{x}_0; \tau = T, \mathbf{x})$. The backward location probability can be determined via

$$f_b(\mathbf{x} = \mathbf{x}_0; \tau = T, \mathbf{x}) = \psi^*(\mathbf{x}_0; \tau = T, \mathbf{x})$$

(5)

where $\psi^*(\mathbf{x}_0; \tau = T, \mathbf{x})$, termed as adjoint location probability, denotes the solution to an adjoint backward equation of the forward contaminant transport equation [7].

The forward fluid flow governing equations can be written in a general form:

$$\frac{\partial \Phi}{\partial t} + (\nabla \cdot \mathbf{V}) \Phi - \Gamma_\Phi \nabla^2 \Phi = S_\Phi$$

(6)

where $\Phi$ is $V_j$, which stands for the air velocity component in the $j$-direction, is 1 for mass continuity, is $T$ for temperature, is $C$ for species concentration. $\nabla$ is the velocity vector, $\Gamma_\Phi$ is the diffusion coefficient and $S_\Phi$ is the source term. The $\Phi$ can also stand for turbulence coefficient and $S_\Phi$ is the source term. The forward contaminant transport equation, together with the initial and boundary conditions can be further expressed as

$$\frac{\partial C}{\partial t} + \frac{\partial (\nu C)}{\partial x_j} = \frac{\partial}{\partial x_j} \left( \nu \frac{\partial C}{\partial x_j} \right) + (S + q_i C_i - q_i C)$$

$$C(\mathbf{x}, 0) = C_0(x)$$

$$C(\mathbf{x}, t) = g_1(t)$$

$$\left[ \nu C_j \frac{\partial C}{\partial x_j} \right] n_i = g_2(t)$$

$$V_j C - \nu C_j \frac{\partial C}{\partial x_j} n_i = g_3(t)$$

(7)

where $\rho$ is the air density, $C$ is the species concentration, $V_j$ is the air velocity at $X_j$-direction, $\nu C_j$ is the effective turbulent diffusion coefficient for $C$ at the $X_j$-direction, $q_0$ is the outflow (other than advection outflow through boundaries) rate per unit volume, $q_i$ is the inflow (other than advection inflow through boundaries) rate per unit volume, $C_i$ is the corresponding inflow concentration, $S_\nu$ is all other kinds of contaminant sources or sinks in cells, and $(S + q_i C_i - q_i C)$ is a combined term of all sources or sinks other than advection and dispersion. $C_0$ is the initial concentration, $g_1$, $g_2$ and $g_3$ are known functions, $I_1$, $I_2$ and $I_3$ are the domain boundaries, and $n_i$ is the outward unit normal vector in the $x_i$-direction.

To derive the adjoint equation for the forward contaminant transport equation (7), the sensitivity analysis approach of Sykes et al. [12] was employed. Liu and Zhai [7] described the detailed mathematical deduction procedure and provided the corresponding adjoint equation as
\[
\frac{\partial \psi}{\partial \tau} - \frac{\partial V \psi}{\partial x} = \frac{\partial}{\partial y} \left[ \nu_c \left( \frac{\partial \psi}{\partial y} \right) \right] + \left( -q \psi \right) + \frac{\partial h}{\partial C}
\]

\[
\psi_0(\vec{x}, 0) = 0
\]

\[
\psi_0(\vec{x}, \tau) = \Gamma
\]

\[
\left[ \frac{\partial}{\partial y} \frac{\partial \psi}{\partial y} + V \psi \right] n_i = 0 \quad \Gamma_1
\]

\[
\left[ \frac{\partial}{\partial y} \frac{\partial \psi}{\partial y} \right] n_i = 0 \quad \Gamma_2
\]

\[
\frac{\partial h}{\partial C} = \delta(\vec{x} - \vec{x}_w) \delta(\tau) \quad \text{for location probability} \quad (8)
\]

where \( \psi^* \) is the adjoint location probability, \( \tau \) is the backward time, \( \vec{x}_w \) is the location where the measurement is made, \( \partial h/\partial C \) is the load term and \( \delta(x) \) is the impulse function which equals 1 when \( x = 0 \) otherwise 0. Note that the adjoint of the first-type boundary condition is still first-type (boundary condition on \( \Gamma_1 \)); the adjoint of the second-type boundary condition becomes third-type (boundary condition on \( \Gamma_1 \)); and the adjoint of the third-type boundary condition becomes second-type (boundary condition on \( \Gamma_1 \)). The initial condition \( \psi^*(\vec{x}, \tau = 0) = 0 \) implies that the adjoint probability for observed pollutants to be from any potential source location is zero at the backward time \( \tau = 0 \). The boundary conditions constrain the adjoint probabilities at the boundaries. The load term represents a probability source term at the measurement point at \( \tau = 0 \).

Solving Equation (8) requires information of thermo-physical properties of fluid, a steady flow field, sensor location, boundary conditions and geometric characteristics of the flow domain. The solution provides the backward (and forward) location probability for one sensor observation. Lin [13] defined this type of backward probability based on one sensor location with no concentration reading as standard adjoint location probability (SALP).

The SALP equation (8) calculates the probability of source location based on one given sensor location and known source release time. A typical example is a smoke sensor in a firing house detecting fire smoke and triggering alarm when the threshold is reached. The SALP can help locate the fire origin with the sensor location, therefore eliminating or reducing the risk of on-site inspection of firefighters. Apparently, with minimum information (i.e. sensor location), SALP can only provide a rough estimate for possible source location that releases contaminants reaching the place where the sensor is located.

The SALP calculation can be refined with more alarm sensors. As suggested by Neupauer [14], the multiple-observation backward location probability can be calculated upon individual SALP simulation results for each sensor location, via

\[
f_x(\vec{x} \mid \vec{C}, \tau_0, \vec{x}_w) = \prod_{i=1}^{N} \frac{f_x(\vec{x} \mid \tau_0, \vec{x}_w)}{\int f_x(\vec{x} \mid \tau_0, \vec{x}_w, \tau_i) d\vec{x}}
\]

where \( N \) is the total number of sensor observations, \( \vec{x}_w \) and \( \tau_w \) \((k = 1 \text{ to } N) \) denote the sensor location and sampling time of sensor \( k \), respectively, \( \tau_0 \) is the known source release time in the backward time domain and \( f_x(\vec{x} \mid \tau_0, \vec{x}_w, \tau_w) \) is the SALP for the single sensor \( k \). In the simulation, the origin of the backward time domain can be set to the time when the last measurement was made. Since the time interval between each measurement can be obtained, \( \tau_w \) can be easily determined. The integral in the denominator of Equation (9) ensures that the total probability is unity. This sensing scenario gains more contaminant distribution information in that multiple sensors detect the existence of contaminants at different locations after a certain time period. As concentrations are not involved in the calculation, probabilities obtained with Equation (9) are called conditioned adjoint location probabilities without concentrations (symbolized as CALP-NC).

The probability prediction accuracy can be further refined if more contaminant dispersion information can be acquired, such as contaminant concentration. In practice, many contaminant sensors can detect, display and store contaminant concentration, either the current reading or the entire historical readings. Such quantitative information greatly improves the accuracy and efficiency of the source identification procedure. To incorporate the concentration information into the prediction, Lin [13] suggested the conditioned adjoint location probability with \( N \) measurements to be

\[
f_x(\vec{x} \mid \vec{C}, \tau_0, \vec{x}_1, \tau_1, \vec{x}_2, \tau_2, \ldots, \vec{x}_N, \tau_N) \approx \prod_{i=1}^{N} P(\hat{C} \mid M_0, \vec{x} \mid \tau_0, \vec{x}_i, \tau_i) f_x(\vec{x} \mid \tau_0, \vec{x}_i, \tau_i) dM_0
\]

where \( N \) is the number of measurements, \( \vec{x}_i, \tau_i \) and \( \hat{C}_i \) are, respectively, the measurement location, the backward measurement time \((\tau = \tau - \tau)\), and the measured concentration value of the \( i \)th measurement. \( \tau_i \) is the known source release time in the backward time domain and \( M_0 \) is the assumed instantaneous source release mass strength. \( f_x(\vec{x} \mid \tau_0, \vec{x}_i, \tau_i) \) is the SALP for the \( i \)th observation (without concentration measurement), \( P(\hat{C} \mid M_0, \vec{x} \mid \tau_0, \vec{x}_i, \tau_i) \) is the probability for the measured concentration conditioned on source mass \( M_0 \) and source location \( \vec{x} \). \( P(\hat{C} \mid M_0, \vec{x} \mid \tau_0, \vec{x}_i, \tau_i) \) follows a normal distribution as suggested by Neupauer and Wilson [15]

\[
P(\hat{C} \mid M_0, \vec{x} \mid \tau_0, \vec{x}_i, \tau_i) = N(M_0 f_x(\vec{x} \mid \tau_0, \vec{x}_i, \tau_i), \sigma_x^2)
\]

where \( \hat{C} \) is the possible concentration value of the \( i \)th measurement and \( \sigma_x^2 \) is the variance for the \( N \) measurements. As concentrations are involved in the calculation, probabilities obtained with Equation (10) are called conditioned adjoint location probabilities with concentrations (symbolized as CALP-C).

In Equation (10), the multiple measurements can be either one-time (static) concentration measures by multiple (or one) sensors or multi-time (dynamic) concentration measures by one (or multiple) sensors. Calculation of the conditioned adjoint location probability (CALP) with both Equations (9) and (10) gives the conditional probability distribution (CPD) of the source location. To incorporate the concentration information into the prediction, Lin [13] suggested the conditioned adjoint location probability with \( N \) measurements to be
and (10) requires first solving SALP for each observation location with Equation (8).

3 PARAMETERS FOR SENSITIVITY ANALYSIS OF THE INVERSE MODELING METHOD

Sensitivity analysis for the inverse modeling algorithm investigates how SALP, CALP-NC or CALP-C changes with input parameters. The proposed algorithm has been implemented in two different airflow simulation engines: computational fluid dynamics (CFD) and multi-zone airflow modeling. The required key input parameters for each tool are summarized in Table 1 in order to calculate various location probabilities.

The required input parameters can be obtained from different approaches. Air velocities can be calculated from computer simulation; turbulent diffusion coefficient may be determined by theoretical analysis; zone contaminant removal coefficient may result from on-site measurement; filter efficiency may come from product specifications. Consequently, these parameters can be quite different from the true values in reality, which in turn result in the predicted source information being deviated from the actual conditions. In order to quantify the variations in the predictions, the following sections discuss how the differences in the input parameters may influence or change the predicted adjacent location probabilities and the source location prediction.

CALP-NC is a function of SALP according to Equation (9), and thus depends on the accuracy of SALP prediction directly. For CALP-C, because the concentration readings are needed and provided either from real sensors or from numerical experiments, and are regarded as known and constant values in source prediction, CALP-C becomes a function of three variables: SALP, measurement error variance \( \sigma^2_e \), and source mass \( M_s \).

4 SENSITIVITY ANALYSIS FOR SALP

Although SALPs obtained from CFD and multi-zone models depend on different variables shown in Table 1, airflow velocities are the major determining and variant factor. Air velocities in reality can be quite different from those obtained from computer simulation because many assumptions and approximations used for air velocity simulation may not represent the reality, and not all factors impacting air velocities are taken into account in the simulation. Disparities between actual and simulated air velocities are hard to predict. This research therefore employs a random method for the SALP sensitivity study. The procedure of this method is listed below:

- SALPs are first determined with the inverse algorithm based on simulated air velocities.
- An array of random numbers between −1 and 1 is produced. Total number of the random numbers is equal to the number of velocity variables.
- An error weighting factor is determined. This parameter indicates the percent of variations of the actual velocities from simulated ones. Representative values for this parameter are 0.1, 0.5, 0.9, etc.
- A deviated actual velocity is calculated to represent the actual air velocity for each cell, zone or flow path using the equation: actual velocity = \((1 + \text{error weight} \times \text{random number}) \times \text{simulated velocity}\).
- With the actual velocities, new SALPs are retrieved again through the inverse modeling and compared with SALPs associated with simulated velocities.

To demonstrate the SALP sensitivity analysis for the CFD inverse modeling, a simplified 2D office space was simulated as illustrated in Figure 1. The office is 10 m long and 3 m high and houses a 70 W occupant, a 200 W computer, an adiabatic desk and a large window with incoming heat flux of 100 W. The conditioned air was continuously supplied at \( V = 0.1 \text{ m/s} \) and \( T = 20^\circ C \) from the upper left corner. A point contaminant source underneath the window released 100 units of contaminant at \( t = 0 \). A forward CFD simulation was first performed to provide the steady-state airflow field required by the inverse modeling as well as the dynamic contaminant concentration dispersions that can be used as sensor readings and fed to the inverse modeling program as inputs. The predicted source location from the inverse modeling can then be validated against the inputs to the forward simulation.

Before the new SALP values are calculated in the sensitivity analysis, the velocities were perturbed and the error weight factor was determined to be 10%, 50% and 90%, indicating up to 90% random errors have been brought into the velocity field. Figure 2 presents the resulted velocity field when the 90% random errors are applied. Compared to the original flow field in

![Figure 1. CFD office model and predicted airflow field.](https://example.com/CFD_office_model.jpg)
Figure 1, although magnitude of each velocity vector has changed, the basic flow pattern remains the same and three major vortices still exist.

For the three levels of random errors, Figures 3–5, respectively, show the calculated SALP distribution and magnitude. In this particular case, the predicted source location moves up, and the max SALP at the predicted source location drops as the random error increases. However, even with 90% of random error in the velocities, the predicted source location is still very close to the actual source. The conclusion therefore is, as long as the flow pattern stays the same, up to 90% of random errors do not change the prediction result significantly. The inverse algorithm still offers acceptable prediction.

5 SENSITIVITY ANALYSIS FOR CALP-NC

The same office case was used to demonstrate the sensitivity analysis on CALP-NC for the CFD inverse modeling. The same three levels of errors, 10%, 50% and 90%, were tested. Figures 6–8 show the predicted CALP-NC distribution and magnitude. Similar to the SALP study, the predicted source location moves up, and the max CALP-NC at the predicted source location drops as the random error increases. Still, even with 90% of random error in the velocities, the predicted source location is close to the actual one.
6 SENSITIVITY ANALYSIS FOR CALP-C

If contaminant concentration is used in source identification, Equations (10) and (11) indicate that SALP, the error variance $\sigma_e^2$ for each measurement, and the source mass range $M_s$ are three independent factors in obtaining CALP-C. Different measurements determine SALP, which is a function of specific sensor location and measurement time based on Equation (9). The second parameter, variance $\sigma_\varepsilon^2$ in measurement error, depicts the sensor accuracy. In reality, concentration reading from a sensor is usually not the true concentration value of the measurement point, due to the fact that sensors, no matter how accurate, always bear errors because of diverse factors, such as errors in the transmitter, the conversion of sensor signal, varying characteristics of sensor wires under actual environments, etc. The measurement error therefore signifies how significant the difference between the measured and true values can be—often depends on sensor type. This parameter stays fixed when readings come from the same type of sensor. Since the same category of sensors is usually adopted for the measurement of the same kind of contaminant in a space or a building, it is reasonable to assume that the same variance can be applied for various concentration readings in CALP-C calculation. The square root of the variance, $\sigma_\varepsilon$, is the sensor accuracy, which is usually expressed in terms of a percentage of actual reading. Because many practical sensors for contaminant concentration have an accuracy of 1%, $\sigma_\varepsilon$ is set to 1% of actual concentration reading in this study. The third parameter, source mass range $M_s$, dictates the estimated range of the released pollutant mass from the source. In the following section, sensitivity of CALP-C to the input parameters is studied by numerical experiment. Once at a time, only one parameter is varied by prescribed percentage with others unchanged, the variation in CALP-C is calculated to find how the source location prediction is influenced by the parameter changes.

6.1 Sensitivity analysis of CALP-C with respect to air velocity

As discussed in previous sections, SALPs vary with air flow velocity. The relationship between CALP-C and SALP is essentially the relation between CALP-C and air velocities. The office case was tested again with the same three levels of errors in velocities, 10%, 50% and 90%, respectively. Figures 9–11 show the CALP-C distribution and magnitude. Similar to previous studies related to air velocity variations, the predicted source location moves up, and the max CALP-C at the predicted source location drops as the random errors increases. 90% of random errors in the velocities do not change the predicted source location considerably.

6.2 Sensitivity analysis of CALP-C with respect to contaminant concentration

Different from the real world, in the numerical experiment, concentrations ‘recorded’ from each ‘sensor’ are reliable and accurate (depending on the accuracy of velocity filed though). These concentrations are consequently treated as an accurate inputs, representing the true contaminant concentrations. However, in reality, contaminant concentration measurement is always a big challenge due to the velocity instability and device resolution/accuracy, etc. To test the influence of this uncertainty, the contaminant sensor ‘readings’ from the numerical experiment are varied by certain percentage before they are used in the adjoint source identification algorithm. This percentage is related to the sensor accuracy expressed in a percentage of the adjusted sensor reading, i.e. the true concentration is within (sensed value $\pm$ sensor accuracy). The baseline sensor accuracy, $\sigma_\varepsilon$, is equal to 1% of the adjusted sensor reading.

The same office case was tested to illustrate the effect of the measurement error on CALP-C. One sensing scenario ($C_{\text{sensor-1}} = 7.1$ units/m$^3$ at $t = 15$ s) was studied. Measurement error was varied from (0.01*baseline) to (100*baseline). Predicted source locations or areas are presented in Figure 12.
When the measurement error is small, the predicted potential source area spreads out around the actual source location. The larger the error, the further away the centroid of predicted source area shifts from the actual source, and the lower the max CALP-C. However, in this case, even when the error is 100 times of the sensor accuracy, the predicted source location is still close enough to the actual source spot.

When the measurement error is infinitely large, which indicates the measured concentration is meaningless, CALP-C approaches SALP, which is consistent with Lin (2003)'s findings.

When the measurement error is small, the predicted potential source area becomes focused. But the predicted area or spot may shift away from the actual source. The smaller the error, the higher the max CALP-C. In this test case, even when the error is 1% of the sensor accuracy, the predicted location is still close enough to the actual source.

When the measurement error is infinitely small, which implies no difference between the measured and true values, the algorithm may fail to provide any prediction, indicating the need to include proper measurement error into the algorithm due to the inherent normal distribution assumption [15].

6.3 Sensitivity analysis of CALP-C with respect to contaminant mass range

In the calculation of CALP-C, the contaminant mass range is another parameter that needs to be presumed. Lin [13] found that if the range of source mass does not include the true source mass, the algorithm will produce an inexact characterization of possible source location. The physical interpretation of the finding is that the same concentration value can be the result of a weak source nearby (small SALP) or the result of a strong source far away (large SALP). When the mass range is wrong, the predicted source location may thus deviate from the actual one. Therefore, the goal of determining the proper mass range is to ensure the inclusion of the actual source strength.

To make sure the actual source strength is included in the presumed mass range, the concept of critical SALP proposed for the sensor network design [16] is adopted. For a given sensor location $x_o$ and contaminant release time $T$, any source with a total released mass of more than $M_{m0}$ at any location in the domain with SALP of more than $f_{x,cri} = C_{th}/M_{m0}$ can be identified by this sensor in less than $T$ time after the burst. $f_{x,cri}$ is called critical SALP. $C_{th}$ is the sensor sensing threshold that triggers either alarm or data reading. $M_{m0}$ is the theoretical maximum initially released contaminant mass. For any contaminant species and the space in study, there usually exists an exposure criteria $C_{cri}$ to prevent occupancies from being exposed to a high concentration of the contaminant for too long. For example, the National Institute for Occupational Safety and Health (NIOSH) has established a recommended exposure limit for CO of 35 ppm (40 mg/m$^3$) as an 8-h time-weighted-average, which means during the 8-h period, the averaged CO concentration in the studied space shall not be over 35 ppm (40 mg/m$^3$). In reality, after all contaminants are released by the instantaneous source at the beginning, due to the existence of HVAC systems and the air exchange induced by supply and exhaust air, total mass of released contaminants decreases with time because contaminants are constantly discharged out of the studied space with the exhaust air. Therefore, as long as the total mass of released pollutant from any possible source location is less than $C_{cri}V_{space}$, the time-averaged CO concentration of the space in 8 h would not raise above the exposure limit, and the space is deemed as safe. $M_{m0}$ therefore can be set to $C_{cri}V_{space}$.

The identification process requires that the SALP values calculated for all sensor locations in the response time are higher than the critical SALP, that is,

$$SALP = \frac{C_{measured}}{M_0} \geq SALP_{cri} \quad (12)$$

The source strength $M_0$ is found to be capped by the value dictated by the reformation of the above equation.
Therefore, the proper mass range can be determined to be \([0, M_{\text{max}}]\).

To test the effectiveness of the proposed algorithm to determine the mass range, a residential house case has been studied using the multi-zone airflow simulation technique \([17]\). Figure 13 illustrates its floor plan that contains one bedroom, one reading room, one bathroom, one kitchen and one living room. A centralized air handling unit (AHU) supplies conditioned air to the bedroom, the reading room, the living room, and also returns air from these rooms. The contamination condition tested was from an instantaneous CO source of 10 g released at 6 AM underneath the exterior window in the reading room. The other rooms did not have any CO contaminant sources.

The critical SALP for this case is the minimum critical SALP of all zones, 0.000177 (1/m³). One sensor is put in the return duct of the AHU system. One sensing scenario \((C = 14.2 \text{ mg/m}^3 \text{ at } t = 5 \text{ mins})\) is studied. The mass range for this case is determined to be \([0, 80226 \text{ mg} = (14.2 \text{ mg/m}^3 / 0.000177) \text{ 1/m}^3]\). The released mass in the tested source zone, the reading room, is 10,000 mg and is within the determined mass range. With this range, the source zone is identified to be the reading room, as shown in Figure 14. The same principle can be used in the determination of the mass range for the CFD inverse modeling.

7 CONCLUSIONS

This paper performs a sensitivity analysis on an indoor contaminant tracking algorithm with three different kinds of location probabilities: SALP, CALP-NC and CALP-C. Uncertainties on air flow velocity, concentration measurement error and contaminant mass range—three key inputs for the algorithm—have been studied to explore how the predicted location probabilities may change with these determining factors. It is found that even with up to 90% of random errors in the velocity, the inverse tracking algorithm is still able to produce acceptable predictions, as long as the main flow pattern remains the same. Very small measurement error may result in a failed prediction while extremely large error converts CALP-C back to SALP. The study reveals that within a fairly wide range \(([0.01, 100])\) of the sensor accuracy, the contaminant concentration measurement accuracy has little influence on the capability of the inverse algorithm to predict the right source location. Finally, this paper proposes an approach to determining the possible mass range to include the actual source strength such that the algorithm is able to correctly predict the source location.
REFERENCES

[1] Sohn MD, Sextro RG, Gadgil AJ, et al. Responding to sudden pollutant release in office buildings: 1. Framework and analysis tools. Indoor Air 2003;13:267–76.

[2] Vukovic V, Srebric J. Application of neural networks trained with multi-zone models for fast and accurate detection of contaminant sources in buildings. Paper presented at ASHRAE Annual Meeting, Long Beach, CA, USA, 2007.

[3] Zhang T, Chen Q. Identification of contaminant sources in enclosed environments by inverse CFD modeling. Indoor Air 2007;17:167–77.

[4] Zhai Z, Liu X. Principles and applications of probability-based inverse modeling method for finding indoor airborne contaminant sources. Build Simul 2008;1:64–71.

[5] Liu X, Zhai Z. Location identification for indoor instantaneous point contaminant source by probability-based inverse CFD modeling. Indoor Air 2008;18:2–11.

[6] Liu X, Zhai Z. Prompt tracking of indoor airborne contaminant source location with probability-based inverse multi-zone modeling. Build Environ 2009;44:1135–43.

[7] Liu X, Zhai Z. Inverse modeling methods for indoor airborne pollutant tracking: literature review and fundamentals. Indoor Air 2007;17:419–38.

[8] Zhai Z, Liu X, Wang H, et al. Experimental verification of tracking algorithm for dynamically-releasing single indoor contaminant. Build Simul 2012;5:5–14.

[9] Wang H, Zhai Z, Li Y, et al. Identifying index (source) patient source location of sars transmission in a hospital ward. HVAC & R Res 2012;18:616–25.

[10] Neupauer RM, Wilson JL. Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant. Water Resour Res 1999;35:3389–98.

[11] Neupauer RM, Wilson JL. Adjoint-derived location and travel time probabilities for a multi-dimensional groundwater system. Water Resour Res 2001;37:1657–68.

[12] Sykes JF, Wilson JL, Andrews RW. Sensitivity analysis for steady state groundwater flow using adjoint operators. Water Resour Res 1985;21:359–71.

[13] Lin R. Identification of groundwater contamination sources using probabilities conditioned on measured concentrations. M.S. Thesis. Dept. of Civil Eng., University of Virginia, Charlottesville, Virginia, USA, 2003.

[14] Neupauer RM. Receptor based modeling of groundwater contamination. Ph.D. Dissertation, New Mexico Institute of Mining and Technology, Socorro, New Mexico, 2000.

[15] Neupauer RM, Wilson JL. Backward probabilistic model of groundwater contamination in non-uniform and transient flow. Adv Water Resour 2002;25:736–46.

[16] Liu X, Zhai Z. Protecting a whole building from critical indoor contamination with optimal sensor network design and source identification methods. Build Environ 2009;44:2276–83.

[17] Stuart D, Walton GN. CONTAMW 2.0 user manual—multizone airflow and contaminant transport analysis software. Gaithersburg, MD: Building and Fire Research Laboratory, National Institute of Standards and Technology, 2002.