

A TWO-STEP MODEL TO INVERSELY IDENTIFY A TEMPORARILY RELEASED POLLUTANT SOURCE WITH TWO SENSORS

Hongbiao Zhou, Tengfei (Tim) Zhang, and Shugang Wang
School of Civil Engineering, Dalian University of Technology (DUT)
2 Linggong Rd, Dalian 116024, China

ABSTRACT

This investigation proposed a two-step model to inversely identify release location and temporal rate profile of an indoor pollutant source in an arbitrary release form. The first step is to run inverse solution of the release rate profiles based on Tikhonov regularization for all possible source locations with concentration information provided by one sensor. The second step is to interpret occurrence probability of each solution obtained in the first step according to the Bayesian probability model, by matching the concentration at the other sensor. To test the proposed model, a tracer-gas source released along with the respiration of a passenger in an aircraft cabin mockup is identified. The results show that the two-step model can correctly identify the pollutant source location and the temporal release rates. The performance of the proposed inverse model is highly subject to the sensor placement locations. As a general guideline, both sensors shall be placed in the direct down stream as close as possible to the actual release sources.

INTRODUCTION

In case there is an accidental release of an airborne pollutant, it is critical to promptly infer the pollutant source location and temporal release rate profile. Sensors must be deployed to monitor whether an accidental release of airborne pollutants has occurred or not. Current sensors may provide temporal pollutant concentration but cannot tell the source location and its temporal release rates. Mathematical modeling is required to determine where and how a pollutant has been released. Identification of pollutant source information forms the framework of inverse modeling, i.e., to infer the source information based on certain detected consequence.

There have been couple of researches addressing inverse modeling to identify release locations of indoor pollutant sources. The strategies can be generally divided into two categories (Zhang et al., 2013): (1) solving forward transport equations to search and match appropriate release sources, and (2) directly reversing the transport equations.

The first categorized strategy must enumerate all possible source locations and temporal release rate profiles and then search for a solution that best

matches the monitored concentration. Forward governing equations are solved to obtain the cause-effect relation between a pollutant source scenario and its exhibited concentration response. Sohn et al. (2002) developed a Bayesian probability model to match a pollutant source scenario in a pre-established database based on multi-zone flow models. A similar interpretative strategy was applied to establish a systematic sensor network to infer pollutant source information (Sreedharan et al. 2006; 2007). Vukovic et al. (2010) and Bastani et al. (2012) proposed to locate an indoor pollutant source based on an artificial neural network. In these studies, the release forms of pollutant sources are fixed and assumed to be known, so the pollutant sources can be timely interpreted. However, the release form of a pollutant source is unknown before identification and can be quite complicated. It is extremely difficult to enumerate all of the possible pre-event scenarios to establish a comprehensive database.

The second categorized strategy directly reverses the pollutant transport equation. The pollutant concentration distribution or discrete concentration information must be provided by sensors as known inputs. Zhang and Chen (2007a; 2007b) developed both the quasi-reversibility and pseudo-reversibility models to successfully locate an instantaneous gaseous source and later on a particle source (Zhang et al., 2012). Liu and Zhai (2008) proposed to locate an instantaneous point source by solving the adjoint probability equation to infer the likelihood of a pollutant source location. This method has further shown being able to locate a continuously released source, but the release profile must be provided in advance (Zhai et al., 2012).

As compared to identification of a pollutant source location, researches on inverse quantification of temporal release rate profiles are few. For a gaseous pollutant source released in an instantaneous form, linear scaling method (Zhang and Chen, 2007b) can be applied to estimate the total release rate once a temporal concentration profile is provided by a sensor. As for a continuously released source with more complicated profiles, Zhang et al. (2013) proposed an inverse matrix method based on the unsteady concentration provided by a single sensor. The cause-effect relation between the source release rates and the exhibited concentration responses at a

sensor is expressed into a matrix. Through matrix inversion and improvement to the matrix condition, the release rate profile can be solved in a steady flow field. However, the pollutant source location must be known in advance.

Till date, no inverse modeling is reported being able to identify both pollutant source location and the complicated temporal release rate profiles simultaneously. It seems that the search and match strategy and the direct reversion strategy can be combined together for pollutant source identification. For example, as the first step, all possible source locations can be assumed to be known and each location is assigned a prior probability as a pollutant source, and then the direct inversion can be implemented to solve for the corresponding rate profiles based on the temporal concentration provided by a sensor. The second step is to evaluate the feasibility of all inverse solutions in the first step by matching the exhibited concentration responses at the other sensor. In this way, a two-step model that combines a direct inversion strategy and a forward match strategy can be established. Illustration of the above solution roadmap forms the research objectives of this investigation.

METHODOLOGY

The proposed model contains two submodels: the submodel for inverse quantification of temporal release rate profiles, and the submodel for identification of pollutant source location. The temporal concentration profiles monitored by sensors at two different locations must be provided as known inputs. Each submodel would require one set of temporal concentrations provided by a sensor.

Quantification of temporal release rate profile based on one sensor

Suppose the steady flow field is available, and the concentration response of an arbitrary gaseous source satisfies the following relation,

$$c = Aq \quad (1)$$

A is a matrix that describes the cause-effect relation between the release rate and the exhibited pollutant concentration, in terms of the passive scalar transport equation. If the pollutant source location is fixed and the diffusion coefficient is proportional to the effective viscosity, the concentration varies linearly with the release strength in a steady flow field, i.e., A is a linear matrix.

If linear least squares optimization is used to solve for the source release rate profile, the task is to find a q that can minimize the concentration residual function,

$$\text{Min } z(q) = \|Aq - c\|_2^2 \quad (2)$$

Because the matrix A cannot be directly inverted, this paper adopts the well-known Tikhonov regularization

(Tikhonov and Arsenin, 1977) to convert the ill-conditioned matrix into a well-conditioned one as,

$$\text{Min } z(q) = \|Aq - c\|_2^2 + \lambda^2 \|Lq\|_2^2 \quad (3)$$

By differentiating the right hand side of Equation (3) with respect to q and setting it equal to zero, the corresponding q to minimize the concentration residual becomes,

$$q = (A^T A + \lambda^2 L^T L)^{-1} (A^T c) \quad (4)$$

The matrix $(A^T A + \lambda^2 L^T L)$ is a well-conditioned matrix, so it can be inverted. The most popular form of L is of the second-order time derivative like format as (Hansen, 1997),

$$L \in R^{(n-2) \times n} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & -2 & 1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & -2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & -2 & 1 \end{bmatrix} \quad (5)$$

One effective method to find an optimal λ is by applying the L curve method (Hansen, 1992; Hansen and O'Leary, 1993). The L curve presents the second norm of the regularized term (Lq) versus the second norm of the concentration residual ($Aq - c$) in the log-log coordinates, for a sequence of regularized parameters λ .

Thus far, an explicit expression for the matrix A is needed. In a linear system, the concentration response of an arbitrary source can be expressed as the convolution integral between the temporal release rate and the response factor (the concentration response of a unit impulse release). The expression to calculate concentration response becomes,

$$c(t) = Aq(t) = \int_{-\infty}^{+\infty} q(\tau) [A\delta(t - \tau)] d\tau \quad (6)$$

The discrete format of the concentration response at a certain point can be expressed as,

$$c_{t_n} = \sum_{k=0}^{\infty} q_{t_n-k} F_{t_k} = q_{t_n} F_{t_0} + \cdots + q_{t_n-k} F_{t_k} + \cdots + q_{t_n-\infty} F_{t_n} \quad (7)$$

For convenience, the unit impulse release is defined as a unit isosceles triangle release, which spans a time interval of $2\Delta T$ (Ishida and Kato, 2008; Hiyama et al., 2008). The total rate of the unit isosceles triangle release becomes 1 unit $\cdot \Delta T$. The matrix A can be expressed in terms of concentration response factors.

Once the flow mode and the pollutant source location are set, the matrix A can be determined by running the CFD program. One just needs to define a unit isosceles triangle release, and then obtain the response factors. As the response factors are only

subject to the flow field and source location, the response factors can be solved in advance and stored in a database before inverse identification.

Identification of pollutant source location based on the other sensor

In case there are N locations where pollutant sources can be located and the coordinates for these N locations are known, the first-step submodel is required to run N times to solve for the corresponding release rate profiles. Based on the results provided in the first step, both release locations and rate profiles are known. However, only one location is correct in these N possible locations, if only one single source released pollutant over the domain. The second-step submodel is to identify the occurrence probability of these N solutions obtained in the first step. The temporal pollutant concentration information at the other sensor is required as the known input.

According to the Bayesian probability model, the posterior probability for source location Y_k based on the measurement M is,

$$p(Y_k|M) = \frac{p(Y_k)L(M|Y_k)}{\sum_{i=1}^N L(M|Y_i)p(Y_i)} \quad (8)$$

where $p(Y_k)$ is the prior probability for source location Y_k , which does not rely on the measurement. For example, if all possible locations are assigned identical probability at each location, then $p(Y_k)=1/N$. $L(M|Y_k)$ is the likelihood to acquire measurement M for a source at location Y_k .

For unbiased measurements with errors in normal distribution, the likelihood to acquire measurement M for a source at Y_k can be expressed into,

$$L(M|Y_k) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left[\frac{c_M - c_{Y_k}}{\sigma}\right]^2\right) \quad (9)$$

where c_M is the measured concentration by a sensor, c_{Y_k} is the predicted concentration at the sensor placement location for a source located at Y_k . σ^2 is the error variance of the measurements, which can account for errors in both sensor instruments and model comparison (Sohn et al., 2002). A better match between the measured and predicted concentrations implies a higher probability of the source at this location.

Solution procedure

Figure 1 illustrates the whole solution flow chart to identify both release location and its rate profiles. The steady flow field must be solved in advance and the temporal concentration at two locations should be provided as the known input. It may take some time to establish the matrix A , because the unsteady solution to the passive scalar equation must be executed for each possible source location. It should also be noted that multiple λ s may be provided by the

L curve method and hence an optimal λ has to be selected by the users.

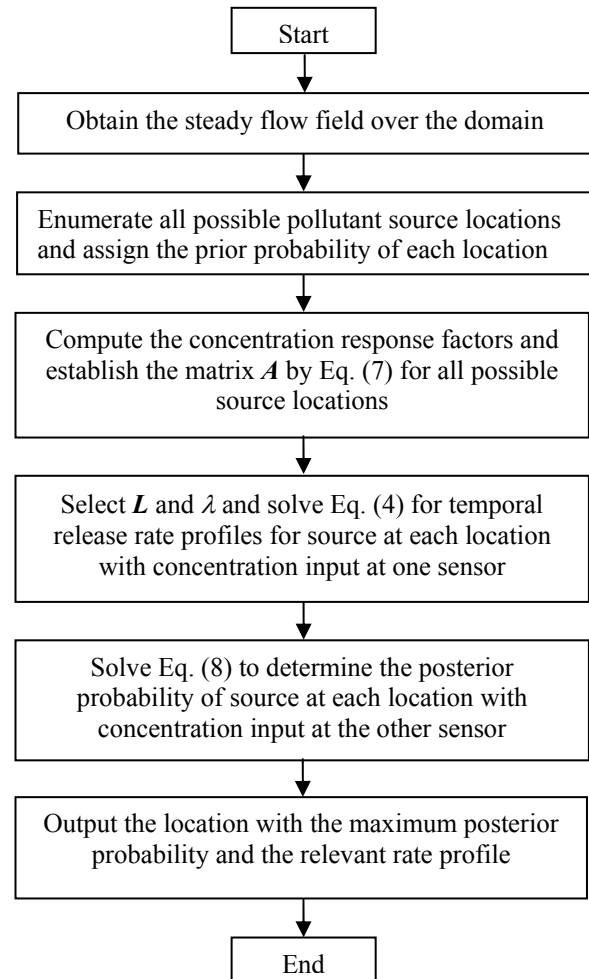


Figure 1 Solution flow chart to inversely identify pollutant source location and the release rate profiles

A DEMONSTRATION CASE

To evaluate the proposed inverse model, the above solution procedure was implemented to identify a pollutant source in a three dimensional aircraft cabin as shown in Figure 2. Totally 21 passengers are seated in a twin aisle aircraft cabin section in three rows. Conditioned air is supplied into the cabin by both symmetric slot inlets on the ceiling and contaminated air is extracted by both outlets near the floor level. Assume that the passengers seated in the second row (numbered by P1 to P7 in Figure 2(b)) are those who possibly release a tracer gas pollutant along with their respiration. There are two sensors to monitor the temporal pollutant concentration: one is near the ceiling highlighted by S1 in Figure 2(a), the other is S2 near the left outlet. Suppose these two sensors can respond simultaneously once the pollutant release starts.

The total air supply rate to the aircraft cabin is 0.248 m³/s, in which the average share by each passenger is 11.8 l/s. The air supply temperature is 19.7 °C, maintaining an averaged temperature of 25 °C in the occupied zone. The passenger surface temperature is 33 °C and around 24.3 °C on the whole cabin walls.

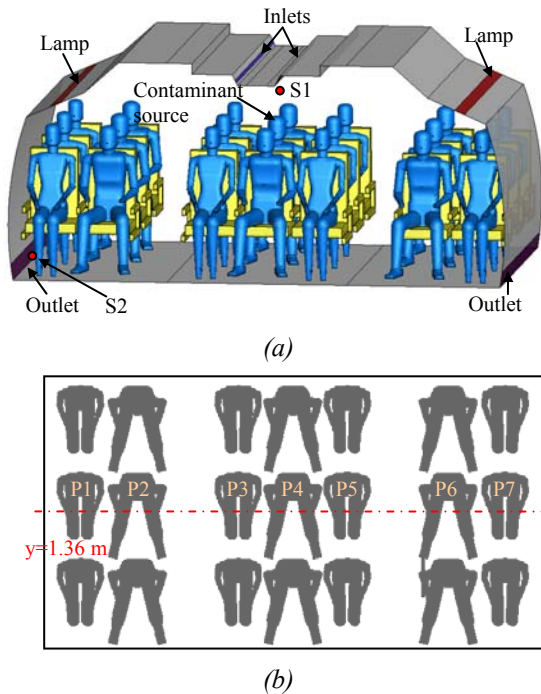


Figure 2 A demonstration case to identify pollutant source: (a) overview of the geometric model, (b) plan view and possible source locations in the second row.

The CFD was applied to solve the steady flow distribution by employing the RNG $k-\epsilon$ model for turbulence approximation. Figure 3 illustrates the airflow pattern in the mid cross section at $y=1.36$ m, which is right across the thighs of the passengers seated in the second row. The strong jet-induced flow is directed to both sides of the cabin and then moves upward in the central recirculated region. Nearly symmetric flow patterns are created inside the cabin section.

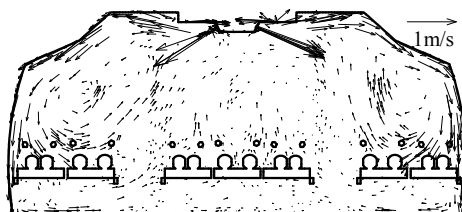


Figure 3 Steady airflow pattern in the mid cross section, i.e., at $y=1.36$ m

This investigation mimics the sensor response to an accidental pollutant release again by the CFD simulation. Accompanying with human respiration, a

tracer gas is assumed to release out by the central passenger P4. Figure 4(a) presents the temporal release rate profile, which contains two cycles of pollutant release. The release rate profile is close to an intermittent sinusoidal wave. Figure 4(b) shows the simulated pollutant concentration response by S1 sensor. Because S1 is just above the pollutant source, the sensor can respond to the pollutant release quickly. Two concentration peaks can be detected due to the two cycles of pollutant release. However, sensor S2 can sense only one peak in the concentration profile and the concentration values are much smaller due to the dilution effect.

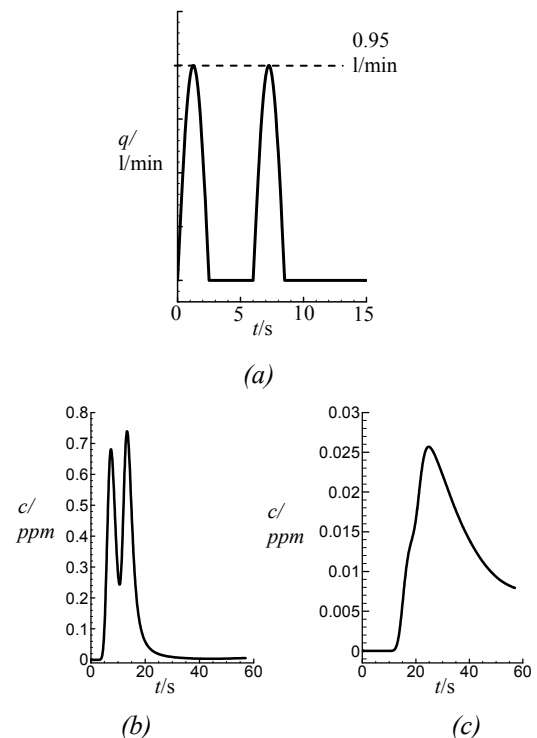


Figure 4 A pollutant release and its resulted concentration response at two sensor locations: (a) release rate profile at P4, (b) temporal concentration at S1, (c) temporal concentration at S2

The steady airflow as shown in Figure 3 and the two sensor concentration responses in Figure 4(b) and (c) are the known input for inverse modeling. In addition, the seven passengers seated in the second row are known as those who would possibly release pollutant.

RESULTS AND DISCUSSION

By following the solution flow chart as shown in Figure 1, the steady flow distribution is obtained in advance, with an example flow pattern shown in Figure 3. The seven passengers seated in the second row are set into candidate pollutant sources. Each passenger is attributed with equal prior probability of 14.3% as the pollutant source. The forward simulation was run seven times by adopting a unit isosceles triangle release at these seven passengers,

and then the corresponding response factors were extracted to establish the linear matrix A . Equation (5) was chosen as the regularization matrix to stabilize the inverse solution for the temporal release rate profile. By plotting the L curve, an appropriate regularized parameter was selected for each candidate source to inversely compute the temporal release rate profiles with Equation (4).

When solving Equation (4), temporal concentration at one sensor is required; and when solving Equation (8), temporal concentration at the other sensor is required. In principle, the roles of these two sensors can be switched. Hence, at least two sets of solution can be provided when switching the sensor information provided to Equations (4) and (8).

Figure 5 presents the inversely solved temporal release rate profiles with the temporal concentration provided by sensor S1 (Figure 4(b)). Table 1 summarizes the adopted regularized parameters when solving Equation (4). More details on the L curve method can refer to Zhang et al. (2013). The inversely solved release rate profile at P4 as shown in Figure 5(a) is quite close to that of the actual release. At the concentration peaks the maximum relative errors are less than 2.4%. However, wiggles in small amplitudes are also found after each concentration peak. As shown in Figure 5(b), the release rate profiles at the other positions present quite different shapes. The release rates can even approach four orders of magnitude.

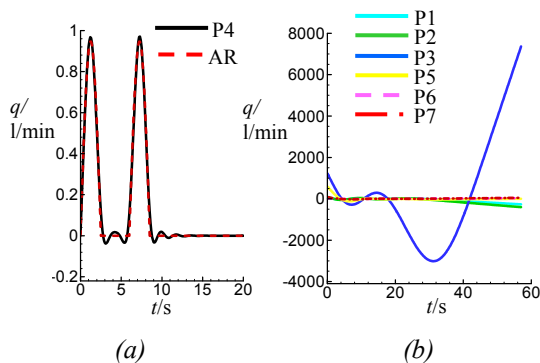


Figure 5 Inversely solved release rate profiles based on the input sensor information at S1: (a) for source at P4 and the comparison with the actual release (highlighted by AR), (b) for sources at P1 to P7 other than P4

Table 1

Selected regularized parameters based on the input sensor information at S1

Source	P1	P2	P3	P4	P5	P6	P7
$\lambda \times 10^4$	1.2	0.7	4.5	1.6	8.9	4.7	4.6

Let us assume all inversely solved rate profiles presented in Figure 5 are possible. By solving

Equation (8) based on Figure 4(c), the likelihood of each solution can be quickly obtained. The square root of the error variance σ is set into 0.003. The probability evolving with time for each source is presented in Figure 6. The initial probability for these seven sources is identical to 14.3% when $t < 2$ s, during which no concentration is detected by sensor S2. When $t > 2$ s, source P1 is eliminated. This is because according to the provided concentration profile in Figure 4(c), no concentration is sensed until $t = 11.8$ s. Any source that would result pollutant concentration at S2 when $t < 11.8$ s is excluded. Sources P1 to P3 are hereby eliminated quickly. After $t = 14.8$ s, sources P5 to P7 cannot result in similar concentration response at sensor S2, and thus they are not the sources.

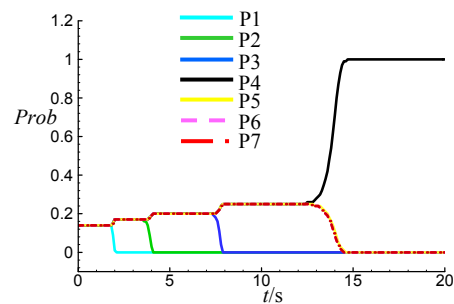


Figure 6 The probability of each pollutant source interpreted by sensor S2

The above solution identifies that P4 is the only pollutant source with 100% probability. The inversely solved release rate profile is in excellent agreement with that of the actual release.

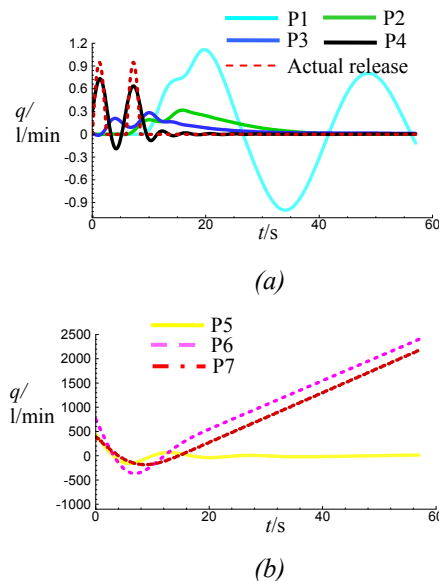


Figure 7 Inversely solved release rate profiles based on the input sensor information at S2: (a) for sources at P1 to P4 and the comparison with the actual release, (b) for sources at P5 to P7

The inverse procedure can also be implemented by quantifying the release rates based on sensor S2 and identification of location probability by sensor S1. Figure 7 shows the inversely solved release rate profiles based on the sensor information at S2. Table 2 outlines the adopted regularized parameters by the L curve method. For candidate sources P1 to P4, the shapes of the inversely solved release rate profiles are quite different. The rates range from -1 to 1.2 l/min, which do not deviate too much from the actual release as shown in Figure 7(a). This is because pollutants released from P1 to P4 can be easily transported to sensor S2 at the left exhaust. However, the inversely solved rates at P6 to P7 exceed 1500 l/min. Due to strong dilution effect, the pollutants released at P5 to P7 must be large enough to result in the sensor response at S2 (Figure 4(c)).

Table 2
Selected regularized parameters based on the input sensor information at S2

Source	P1	P2	P3	P4	P5	P6	P7
$\lambda \times 10^4$	0.39	6.6	1.6	6.3	6.7	3.7	1.2

Again by assuming all solutions provided in Figure 7 are possible, the sensor information at S1 (Figure 4(b)) is used to further identify the likelihood of each solution. Figure 8 presents the probability of each source interpreted with time. The square root of the error variance σ is set into 0.003. The initial probability for each source is 14.3%. Once sensor S1 collects non zero concentration, pollutant source P4 can be quickly identified as the only source. This is because other sources can not provide the sensor response as that shown in Figure 4(b) except by P4.

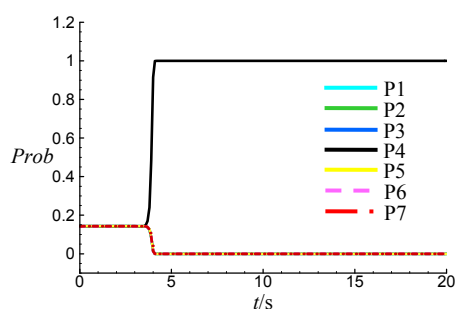


Figure 8 The probability of each pollutant source interpreted by sensor S1

The above solution identifies that P4 is the only pollutant source and the release rate profile shown as in Figure 7(a). By comparing with the actual release profile, it can be seen although the inversely solved rate profile resembles that of the actual release, they are not fairly identical. A closer look at Figure 7(a) illustrates there are some wiggles in the inversely solved profiles at P4. The relative errors in the profile

peaks can reach 32.9%. The responsible reasons are the ill-conditioned matrix A and the regularized operation. Further research can be implemented to improve the condition of the matrix and the inverse solution accuracy.

The results also show that after switching the roles of these two sensors it is still able to correctly identify the pollutant source location. However, the inversely solved release rate profiles are not accurate enough. This implies that the performance of the proposed inverse model is highly subject to the sensor placement locations. As a general guideline, the sensor providing information to inversely quantify the temporal release rates shall be as close as possible to the actual release sources.

CONCLUSION

This investigation proposed a two-step model to inversely identify release location and temporal rate profile of an indoor pollutant source in an arbitrary release form. The known conditions must be provided to the inverse modeling include a steady flow field and temporal concentration information at two sensor locations. By adopting the inverse model to identify a pollutant source in a three dimensional aircraft cabin, the results show that the model is able to correctly determine the pollutant source location and reasonably quantify the temporal release rates.

The sensor information can be switched between the one provided to the submodel of release rate quantification and the other provided to the location probability identification. The demonstration application shows that although the location probability evolving with time is different after switching the sensor information input, the model is still able to correctly identify the pollutant source location. However, the switch may lead different accuracy in quantifying the temporal release rate profiles. The performance of the proposed inverse model is highly subject to the sensor placement locations. As a general guideline, both sensors shall be placed in the direct down stream as close as possible to the actual release sources, to respond quickly and prevent too much dilution to the pollutant concentration.

NOMENCLATURE

A	=	linear matrix
$c(t)$	=	monitored concentration series by a sensor
c_n	=	monitored discrete concentration at time of $n\Delta T$ by a sensor
c_M	=	measured concentration by a sensor
c_{Y_k}	=	predicted concentration at the sensor placement location for a source located at Y_k .
F_n	=	local concentration at time of $n\Delta T$ for a unit impulse release
$L(M Y_k)$	=	likelihood to obtain measurement M for a source at location Y_k

L	=	regularized operator matrix
M	=	measurement
N	=	total number of possible locations where pollutant sources can be located
$p(Y_k M)$	=	posterior probability for source location Y_k based on the measurement M
$p(Y_k)$	=	prior probability for source location Y_k
q	=	temporal release rate series of the contaminant source
q_n	=	discrete release rate at time of $n\Delta T$
Y_k	=	the k^{th} possible source location
$z(q)$	=	concentration residual function
$\delta(t)$	=	Dirac delta function
λ	=	regularized parameter
σ^2	=	error variance of the measurements
ΔT	=	time interval for the unit triangle release

ACKNOWLEDGEMENT

The research presented in this paper was financially supported by the National Key Basic Research and Development Program of China (the 973 Program) through grant No. 2012CB720100.

REFERENCES

- Bastani, A., Haghghi, F., Kozinski, J.A. 2012. Contaminant source identification within a building: Toward design of immune buildings. *Building and Environment* 51, 320-329.
- Hansen, P.C., 1992. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review* 34, 561-580.
- Hansen, P.C., O'Leary, D.P., 1993. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM Journal of Scientific Computation* 14, 1487-1503.
- Hansen, P.C., 1997. Rank-deficient and discrete ill-posed problem: Numerical aspects of linear inversion. SIAM, Philadelphia.
- Hiyama, K., Ishida, Y., Kato, S., 2008. Coupling 3D transient pollutant transport in a room into a flow network model with concentration response factor method. *ASHRAE Transactions* 114(2), 119-129.
- Ishida, Y., Kato, S., 2008. Method for coupling three-dimensional transient pollutant transport into one-dimensional transport simulation based on concentration response factor. *ASHRAE Transactions* 114 (1), 259-272.
- Liu, X., Zhai, Z. 2008. Location identification for indoor instantaneous point contaminant source by probability-based inverse computational fluid dynamics modeling. *Indoor Air* 18 (1), 2-11.
- Sohn, M.D., Reynolds, P., Singh, N., Gadgil, A.J. 2002. Rapidly locating and characterizing pollutant releases in buildings. *Journal of Air Waste Management Association* 52, 1422-1432.
- Sreedharan, P., Sohn, M.D., Gadgil, A.J., Nazaroff, W.W. 2006. Systems approach to evaluating sensor characteristics for real-time monitoring of high-risk indoor contaminant releases. *Atmospheric Environment* 40, 3490-3502.
- Sreedharan, P., Sohn, M.D., Nazaroff, W.W., Gadgil, A.J. 2007. Influence of indoor transport and mixing time scales on the performance of sensor systems for characterizing contaminant releases. *Atmospheric Environment* 41, 9530-9542.
- Tikhonov, A.N., Arsenin, V.Y., 1977. Solutions of ill-posed problems. Halsted Press, Washington, DC.
- Vukovic, V., Tabares-Velasco, P.S., Srebric, J. 2010. Real-time identification of indoor pollutant source positions based on neural network locator of contaminant sources and optimized sensor networks. *Journal of Air Waste Management Association* 60, 1034-1048.
- Zhang, T., Chen, Q. 2007a. Identification of contaminant sources in enclosed environments by inverse CFD modeling. *Indoor Air* 17, 167-177.
- Zhang, T., Chen, Q. 2007b. Identification of contaminant sources in enclosed spaces by a single sensor. *Indoor Air* 17, 439-449.
- Zhang, T., Li, H., Wang, S. 2012. Inversely tracking indoor airborne particles to locate their release sources. *Atmospheric Environment* 55, 328-338.
- Zhang, T., Yin, S., Wang, S. 2013. An inverse model based on CFD to quantify the temporal release rate of a continuously released pollutant source. *Atmospheric Environment*. (under 2nd revision)
- Zhai, Z., Liu, X., Wang, H., Li, Y., Liu, J. 2012. Experimental verification of tracking algorithm for dynamically-releasing single indoor contaminant. *Building Simulation* 5, 5-14.