Lawrence Berkeley National Laboratory

Title
Towards improved characterization of high-risk releases using heterogeneous indoor sensor systems

Permalink
https://escholarship.org/uc/item/30g5q7c0

Author
Sreedharan, Priya

Publication Date
2011

Towards improved characterization of high-risk releases using heterogeneous indoor sensor systems

Priya Sreedharan [a,b,†], Michael D. Sohn [b,*], William W Nazaroff [b,c], Ashok J. Gadgil [b,c]

[a] Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740 USA

[b] Indoor Environment Department, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA

[c] Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720-1710 USA

[*] Corresponding author: Michael D. Sohn, Indoor Environment Department, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mail Stop 90R3058, Berkeley, CA 94720, USA. Tel.: +1 510 486 7610; Fax: +1 510 486 6658; e-mail: MDSohn@lbl.gov

[†] Present address: Climate Protection Partnerships Division, US Environmental Protection Agency, Washington, DC, USA

June 2010

Towards improved characterization of high-risk releases using heterogeneous indoor sensor systems

Priya Sreedharan [a,b,†], Michael D. Sohn [b,*], William W Nazaroff [b,c], Ashok J. Gadgil [b,c]

[a] Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740 USA

[b] Indoor Environment Department, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA

[c] Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720-1710 USA

[*] Corresponding author: Michael D. Sohn, Indoor Environment Department, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mail Stop 90R3058, Berkeley, CA 94720, USA. Tel.: +1 510 486 7610; Fax: +1 510 486 6658; e-mail: MDSohn@lbl.gov

[†] Present address: Climate Protection Partnerships Division, US Environmental Protection Agency, Washington, DC, USA

Abstract

The sudden release of toxic contaminants that reach indoor spaces can be hazardous to building occupants. For an acutely toxic contaminant, the speed of the emergency response strongly influences the consequences to occupants. The design of a real-time sensor system is made challenging both by the urgency and complex nature of the event, and by the imperfect sensors and models available to describe it. In this research, we use Bayesian modeling to combine information from multiple types of sensors to improve the characterization of a release. We discuss conceptual and algorithmic considerations for selecting and fusing information from disparate sensors. To explore system performance, we use both real tracer gas data from experiments in a three-story building, along with synthetic data, including information from door-position sensors. The added information from door-position sensors is found to be useful for many scenarios, but not always. We discuss the physical conditions and design factors that affect these results, such as the influence of the door positions on contaminant transport. We

highlight potential benefits of multisensor data fusion, challenges in realizing those benefits, and opportunities for further improvement.

1 Introduction

The ultimate goal of the research described in this paper is to develop approaches for the design and implementation of sensor systems that can detect and diagnose — in real time — chemical or biological contaminant releases that pose acute hazards for building occupants. Such sensor systems should perform robustly — for example, identifying the location and mass of the release and suggesting safe evacuation routes — under diverse release circumstances, for different building operating conditions, and with limited and imperfect sensor information.

As building scale increases, the physical processes that govern indoor contaminant transport and fate become more complex and simple engineering rules of thumb may be insufficient for designing effective sensor networks. Instead, an integrated-design approach that considers contaminant transport processes, sensor characteristics, and sensor interpretation algorithms can help to develop a system that performs well in diverse conditions. The present paper builds on our previous efforts to develop system design methods and model-based algorithms capable of reconstructing and estimating the unobserved parameters of a release event [1-3].

Our previous research explored data interpretation based solely on contaminant measurements. The present paper considers fusing information from multiple sensor types. First, contaminant sensors are relatively expensive, so that adding different types of sensors may decrease the system cost without compromising performance. Second, some inexpensive-to-detect aspects associated with the state of the building, such as the operating mode of the air handling system or window and door positions, can play a significant role influencing contaminant transport. In general, the designer would wish to consider incorporating any

4

information that would improve the predictive power or robustness of the detection and diagnostic system.

In this paper, we present elements of a conceptual framework and an algorithmic approach to integrate information from heterogeneous sensors into a diagnostic sensor system. Under ideal conditions — perfect information and a perfect interpretive model — many types of information would help to improve the event reconstruction, albeit to varying degrees. Under nonideal conditions, however, can a type of information improve event reconstruction in some cases, but confound it in others, and why? In this paper, we introduce an approach for multisensor system design and illustrate its application in a case study that combines door-position sensors with contaminant concentration sensors.

This research contributes to an emerging literature on indoor contaminant sensor systems. In recent related efforts, Liu and Zhai employed the adjoint probability method in combination with multizone modeling and computational fluid dynamics (CFD) to locate indoor contaminant releases [4,5]. Zhang and Chen developed CFD and probabilistic techniques for locating contaminant sources in enclosed spaces [6]. Probabilistic inverse modeling methods have been used, also, to reconstruct atmospheric releases [7]. The above efforts explored sensor networks composed solely of contaminant sensors.

More broadly, this paper addresses issues relevant to information fusion and environmental system modeling. While heterogeneous (or multisensor) fusion has been applied to problems ranging from target tracking to medical diagnostics [8], the development of heterogeneous sensor systems for environmental applications is relatively new [9]. The concepts we introduce relate to general issues of model accuracy, parsimony, and complexity [10,11]. This paper highlights concepts and findings that are reported in greater detail elsewhere [12].

5

2 Background

2.1 Algorithm overview

We interpret sensor data through application of a two-stage Bayes Monte Carlo (BMC) algorithm to estimate release parameters and future concentrations, given that a contaminant sensor system has detected an event [1]. We briefly review here the key concepts. In the first stage, prior to an event, a library of model realizations spanning all possible release conditions is generated, with input parameters based on expert judgment. In our application of the BMC algorithm, this library contains predicted concentrations, where Y_k denotes the predicted concentrations from release scenario k.

The second stage occurs once a release is detected. The prior probability distributions of release conditions and concentration profiles ($P(Y_k)$) are revised by applying a sequential Bayesian updating technique as sensor measurements (O) become available. In the case of a network comprising contaminant concentration sensors, Bayes' rule calculates $P(Y_k|O)$, i.e., the probability that the array of observed concentrations O corresponds to scenario k and its corresponding modeled concentration values, Y_k, as follows:

$$P(Y_k \mid O) = \frac{p(O \mid Y_k)P(Y_k)}{\sum_{i=1}^{N} p(O \mid Y_i)P(Y_i)} \tag{1}$$

where $p(O|Y_k)$ is the likelihood function, $P(Y_k)$ is the prior probability of Y_k, and N is the number of realizations in the library. The likelihood of observing a set of concentrations (O), for an assumed release condition and its associated concentrations (Y_k), is influenced by both sensor and modeling error. The updated probabilities of the library realizations contain useful information, for example, that there is a less than a specified probability that the release originated in a particular room.

Bayesian updating is not restricted to interpreting contaminant information. Brand and

6

Small [13] describe how the BMC algorithm can reduce the uncertainties of outputs from interconnected risk models, providing a natural framework for fusing disparate sensor information. Sohn et al. [14] used hydraulic head and chemical sensors to reduce the uncertainties of coupled groundwater flow and contaminant transport models. In the present work, the fate and transport model predicts airflows, and so, for example, heating, ventilating and air conditioning (HVAC) duct airflow measurements could be used to help determine the relative probability of each scenario in the library. Other models could be linked with a multizone model, such as an energy model that better captures the influence of a building's thermal characteristics on airflow conditions. These would provide additional sensor fusion opportunities.

2.2 Case study

Fan pressurization and tracer gas experiments were conducted in a three-story building (Figure 1) for which a multizone pollutant dispersion model was applied [15]. In this building, a single mechanical air-handling unit (AHU) recirculates the air to the first and second floors. Air exchange with the outdoors occurs by infiltration, and intrazonal flows are induced by a combination of mechanical and natural convective forces. Tracer experiments consisted of an interior 20 g pulse release of an inert gas from a single point, after which concentrations in each room and staircase were sampled at regular intervals.

The case study AHU differs from a typical HVAC system for a commercial building. A typical HVAC system uses several sensors for operational control, many of which are unavailable in the test unit. However, hypothetical sensor outputs can be simulated for the purposes of evaluating multisensor fusion. For example, known door positions and AHU operating status during the tracer experiment can be used to simulate door position sensors and

7

control-system input signals, respectively.

3 Conceptual framing of sensor fusion

3.1 Identifying useful information

A sensor fusion algorithm potentially can use any information that can be monitored and is mathematically represented in a contaminant transport and fate model. Environmental and building operating conditions such as temperatures, pressure differences, door positions, damper positions, duct static pressure, and HVAC operating conditions influence the transport and fate of contaminants and are therefore candidates for inclusion in the sensor system. Building conditions should be considered for incorporation into the design of a sensor network depending on their information richness, i.e., their relative importance influencing contaminant transport and fate in the building. For example, infiltration and natural convective-driven airflow are likely to be small in a well-sealed building maintained at uniform temperatures under mild weather conditions. In this case, outdoor temperature and wind information may not be helpful, since the airflows are mainly driven by HVAC system operation. In contrast, window position in a building with operable windows may have significant influence on airflow, and sensors providing this information would be good candidates for inclusion.

3.2 Mapping measurements to the model

Identifying how a measurement maps to a contaminant transport and fate model is helpful in implementing an appropriate fusion algorithm. Although our present focus is on the BMC algorithm, these concepts are more broadly relevant.

Sensor information can correspond to a model input, a model output, or a model parameter. For example, concentration measurements map to model outputs; door positions and temperatures map to model inputs. When contaminant transport is simulated with a series of

8

models, a measurement may map to an output of one model and to an input of another. In this research, airflow and pressure differences are outputs of airflow models and serve as inputs to a contaminant transport model.

How a measurement maps to a model has computational implications. Model inputs are used in the BMC algorithm to generate the contaminant transport library prior to a release event. Typically a range of possible values of each model input is used to generate the library. If measurements of model inputs are available when an event occurs, they can be used to filter realizations in the library, narrowing the space that must be searched during the release event. Consolidating the solution space reduces the number of computations that need to be completed in real time, and serves to reduce the uncertainty of the release parameters. Alternatively, model input information can be applied towards the generation of a library that more closely represents actual conditions.

Some measurements, such as airflows, map to one model's input and another's output. Airflow measurements can be compared to and synthesized with predicted airflows from the airflow model, using the BMC algorithm, or they can substitute entirely for the predicted airflow. The number of calculations executed by the BMC algorithm in real time increases with the number of model output types interpreted. So, mapping the measurement to a model input may reduce the real-time computations if the information is interpreted prior to a release; however, the appropriateness of interpreting the information prior to a release depends on the dynamic state of that measurement.

3.3 Dynamic state of the measurement

The dynamic state of a measurement, in combination with its model mapping, can help inform how best to process the information. One can classify building conditions that pertain to

its dynamic state as "static," "pseudo-static" or "dynamic."

Static quantities rarely or never change and commonly map directly to model parameters. Examples are the leakage characteristics of the building envelope and the speed of a constant-drive fan. Static measurements are treated as constants in the contaminant fate and transport model and in the BMC algorithm. Repeated or real-time measurements of a static quantity are unnecessary.

Pseudo-static quantities change over time, but are expected to remain approximately constant over the time scale of an acute release event. Pseudo-static quantities can be treated as constants in the model and inversion algorithm, but because they change with time, updated measurements are needed for monitoring beyond the time scale of single release events. Although the sensor interpretation scheme can treat pseudo-static quantities as constants, it must also consider multiple possible values. Window position, air temperature and variable supply fan speed are examples of pseudo-static quantities.

In the BMC algorithm, measurements of pseudo-static model inputs can be used to generate a library in real time, as the pseudo-static values change. This approach may offer computational advantages over generating a comprehensive library at an earlier time and processing the larger library during a release. Alternatively, pseudo-static inputs can be interpreted at the start of a release event by a prior-filtering algorithm. If a pseudo-static measurement maps to a model output as well (e.g., airflow), then factors such as sensor accuracy, simulation time and model fidelity can help determine if the measurement should be processed as a model output — such as how concentration information is interpreted — or as an input — which may involve substituting the measurement for the model predictions.

Dynamic quantities change on a time scale that is comparable to or faster than the duration

of a release event. Elevator operation and poorly tuned HVAC systems can generate potentially important dynamic quantities, such as unstable duct damper positions or air handling unit control valves. The interpretation of sensor information for dynamic model inputs is challenging since the algorithm must process the model inputs at a rate comparable to the rate at which the variable changes. Although the BMC algorithm easily incorporates dynamic model inputs, it does not easily incorporate measurements of dynamic quantities that are model inputs, since future values of these measurements are unknown when the library is generated. However, if the building condition changes slowly in comparison to the time scale of a release event or if the dynamic element has little influence on airflow and the relative concentrations measured by the network, accounting for it in the algorithm may not be necessary.

4 Example: Combining door position and contaminant sensors

This section further explores heterogeneous sensor fusion concepts by evaluating sensor networks that utilize both door-position and contaminant-concentration sensors. The investigations use experimental and simulated data sets, for which the release conditions span varying door positions, air handling unit (AHU) operation, source release location, and release duration. We anticipate that door-position sensors are likely to be more informative when the AHU is off, because under that condition door positions have a larger influence on airflow. We are particularly interested in understanding how and why the performance improves or degrades with factors such as the release conditions, network configuration, and use of simulated vs. field data. We are also interested in the consistency of these results. For the purposes of this paper, the system performance is based on the ability of the network to identify the release location to a high probability and to estimate the release mass with a narrow confidence interval.

The analyses and results are reported in two sections: mechanically driven airflow

11

conditions (AHU on) and naturally driven airflow conditions (AHU off). First, though, we describe the approach for utilizing door-position sensor data along with contaminant concentration data.

4.1 Door position fusion algorithm: derivation and implementation

For the purposes of this paper, we classify door positions as pseudo-static, binary model inputs (i.e., the door is either open or closed). (We do not consider here the effect that different extents of openness may have on interzonal airflow.) Information from real-time measurements of door-position sensors can be integrated into the BMC algorithm by adjusting the prior probabilities of the library realizations.

A prior-filtering algorithm was developed to determine the prior probability adjustment in which library realizations with simulated door positions that more closely match the actual door positions are preferentially weighted. The algorithm assumes uniform initial probabilities of each realization. The degree of fit, d_i, measures how well the actual door positions match those assumed for realization i. Equation 2 calculates the degree of fit of each realization.

$$d_i = \frac{n_{match}}{n_{sensors}} \tag{2}$$

where n_{match} is the number of sensors in the network that match the positions of realization i, given a total number of sensors, $n_{sensors}$. A realization is determined to have a favorable match if d_i is greater than a user defined threshold, δ. A second user defined parameter, γ, reflects the confidence that is assigned to the measurement-to-model fit and influences the adjustment to a realization's prior probability. The user-defined parameters are constrained such that $0 \leq \delta, \gamma \leq 1$.

Equations 3 and 4 combine d_i, γ and δ to determine the adjustments to be applied to the prior probability of each realization with a favorable match, where θ_{max} is the upper limit of the adjustment and ensures that $\sum P_i$ for $\delta > d_i$ over all realizations in the library does not exceed 1.

12

$$\theta_{\max} = \frac{1 - \dfrac{N_{d_i \geq \delta}}{N}}{\displaystyle\sum_{i=1, d_i \geq \delta}^{N} d_i} \qquad (3) \qquad\qquad \theta = \gamma \theta_{\max} \qquad (4)$$

In these equations, $N_{d_i \geq \delta}$ is the number of realizations with $d_i \geq \delta$.

The prior probabilities of each realization with a favorable match are adjusted using equation 5a. All other realization priors are adjusted using equation 5b, which was derived by setting $\sum P_i = 1$ over all realizations in the library. For clarification, equations 5a and 5b adjust the prior probabilities of the realizations in the library, based on the door-position measurements. They do not reflect the sensitivity of the model to the door positions.

$$d_i \geq \delta : P_i = \theta \times d_i + \frac{1}{N} \qquad (5a) \qquad d_i < \delta : P_i = \frac{1 - \displaystyle\sum_{i=1, d_i \geq \delta}^{N} P_i}{N - N_{d_i \geq \delta}} \qquad (5b)$$

When $d_i = 1$, all sensors present in the system match the values of the ith realization. As $\delta \rightarrow 1$ and $\gamma \rightarrow 1$, the probabilities of all realizations with exact matches are maximally weighted and the prior probabilities of all other realizations approach 0. By adjusting γ, the degree to which the prior probabilities (P_i) are shifted from a uniform distribution is varied. The parameters δ and γ provide flexibility in treating the sensor measurements, collectively accounting for door position sensor error and model uncertainty (owing to the limited ability of the model to accurately capture the physical effect of door position on contaminant transport).

In this paper, all the door positions during a release are treated as hypothetical door sensor information available to the sensor system, simulating a maximum possible number of door sensors. Both adjustable parameters are set to unity, that is $\gamma = \delta = 1$. By assuming no model specification and sensor error from the simulations, there is greater chance of performance enhancement, but also, an increased risk of performance deterioration.

4.2 Mechanically driven airflow conditions

4.2.1 Library and likelihood function

The library and likelihood function from Sreedharan et al. [3], which were developed for the same field experiments, are used in this section. Table 1 reviews the model input sample distributions that were used to generate the library. Although there are seven doors (Figure 1), each pair of interior doors separating the rooms on levels one and two are assumed to be in one position, reducing the number of independent door settings to five, and thereby yielding $2^5 = 32$ possible door configurations.

Equation 6, which represents the empirical likelihood function for instantaneous or short-duration releases [3], was applied to the "Experiment 1" data in which 20 g of tracer gas was instantaneously released at the AHU return. The standard deviation of the error (units: g/m^3) is obtained by multiplying the normalized parameters from equation 6 by the idealized well-mixed concentration (assuming instantaneous, complete mixing throughout the building).

$$\sigma(t) = \left(A e^{-Ct} + B \right) \frac{t}{\tau_{release}} \qquad\qquad t \le \tau_{release} \qquad\qquad (6a)$$

$$\sigma(t) = A e^{-Ct} + B \qquad\qquad t > \tau_{release} \qquad\qquad (6b)$$

where $A = 1.74$, $B = 0.46$, $C = 0.37$ min^{-1}, and $\tau_{release}$ is the release duration in minutes.

Throughout this paper, the release time is treated as an additional unknown parameter that is estimated using the BMC algorithm. The library is augmented to represent realizations with release times that begin 1 s, 1 min, 2 min and 3 min prior to when the first datum is received.

4.2.2 Results using experimental data

In Experiment 1, the AHU was on and all interior doors were open. We compare the performance of the data fusion algorithm on hypothetical sensor networks both with and without

14

door position sensors. To make our conclusions more general, we explore a wide design space

by comparing every possible configuration of contaminant sensor, with 1 to 11 sensors in the

building, and with each room having either one sensor or none.

In Figure 2, the left-hand frames compare the ability of networks with and without door

position sensors to correctly identify the release location at time $t = 10$ min after the release.

Each frame reports the mean and standard deviation of the performance difference, Δ, which is

defined as the probability of the network finding the release location *with* door sensors minus the

probability *without* door sensors. The right-hand frames of Figure 2 show network performance

for identifying either the return intake or the room containing the intake (Room 1.2a) as the

release location — a less strict, but still useful metric.

In this example, where door positions have relatively low influence on airflow,

door-position sensors improve the identification of the release location on average by ~ 0–10%.

They are helpful for many, but not all, 2-sensor networks (left-hand frames of Figure 2). Using

the less strict outcome metric, fewer networks are negatively impacted (right-hand frames of

Figure 2). Networks that contain 6 contaminant sensors show similar results, although the

average performance advantage attributable to door sensors is less than for the 2-sensor

networks. For some 6-sensor and 2-sensor networks, door sensors degrade system performance.

There is little-to-no impact from using door sensors for networks with 10 contaminant sensors.

Overall, while we observe an average performance improvement from augmenting

contaminant sensor networks, there is considerable variability in the performance among

sensor-network configurations, and even degradation for some networks. The performance

improvement occurs when the AHU is on, a condition in which door positions should exhibit

relatively small influence on airflow and contaminant transport. For these networks, sensor

fusion may improve performance simply by reducing the library scale to a more optimal subset. Model error is likely to account for the networks that exhibited degraded performance. The addition of physically irrelevant parameters (and values) into the library may act as a buffer and can camouflage modeling error; this buffer is decreased when the library space is consolidated through sensor fusion. On a practical level, information on parameter values may not be helpful when one is utilizing a model that is inherently inaccurate or in which error structures are poorly characterized.

4.2.3 Generation of simulated data, likelihood function and updating

To investigate system performance over a broader range of conditions, we augmented the small number of actual tracer-gas experiments by generating concentration data for simulated releases. An additional advantage of using simulated data is that the error structure is more easily characterized, eliminating the confounding factor of model-specification error. The COMIS model (which was used to generate the library) was exercised to model spatially and temporally varying concentrations from a hypothetical 20 g pulse release originating in Stairwell 2. All doors were assumed to be open and the AHU fan was assumed to be on. The modeled concentrations were perturbed to simulate sensor error that was assumed to follow a lognormal distribution (equation 7).

$$f(X) = \frac{1}{\sqrt{2\pi}\,\ln(GSD)X}\exp\left\{-\frac{1}{2}\left(\frac{\ln X - \ln(GM)}{\ln(GSD)}\right)^2\right\} \tag{7}$$

where the geometric mean, GM, and geometric standard deviation, GSD, are set to the modeled concentration, and to $e^{0.2} = 1.22$, respectively. For each model prediction, a random variate was generated by sampling from $f(X)$. In cases where the modeled concentration was 0, the sensor measurement was also set to 0. The perturbed concentrations were adjusted to reflect the sensor

16

resolution and detection limits of the tracer data.

Because the likelihood function represents the assumed error structure of the model-data fit, equation 7 also serves as the appropriate basis for the likelihood function. The likelihood of "measurement" O_i given modeled concentration $Y_{k,i}$, $p(O_i/Y_{k,i})$, is calculated using equation 8. Modeled and sensor concentrations of zero were set to the sensor detection limit.

$$p(O_i \,|\, Y_{k,i}) = \frac{1}{\sqrt{2\pi}\,\ln(GSD) \times O_i} \exp\left(-\frac{1}{2}\left(\frac{\ln O_i - \ln Y_{k,i}}{\ln(GSD)}\right)^2\right) \tag{8}$$

4.2.4 Results using simulated data

Because the AHU does not directly serve the stairwell, the stairwell door positions were expected to exhibit strong influence on the rate at which the Stairwell 2 release disperses. Modeling indicated that the concentrations in the stairwell are impacted as the door positions are altered, although concentrations in many zones are less affected. A full set of hypothetical sensor network data streams, ranging from 1 to 11 contaminant sensors, was generated from the simulated data. Each network was investigated with and without a full set of door sensors.

Figure 3 compares the ability of the sensor systems to identify the release location at 5 min and 10 min after the release. Door-position sensors help the system to identify the release location to a higher probability on average ~ 0–3% at 5 min and ~0–0.5% at 10 min, but with significant variability and not consistently, as indicated by the placement of the points both above and below the diagonal lines. The networks for which door sensors impose a penalty appear to be those with moderately poor sensor placements, as indicated by their relatively low probabilities in the "without door sensors" cases. In some cases, door sensors may exacerbate the effects of poor contaminant sensor placement.

17

4.3   Naturally driven airflow conditions

Door-position sensors were also investigated for naturally driven airflow conditions. They are expected to be more useful here, and consistently so, because door positions exhibit greater influence on contaminant transport in the absence of forced flow through ducts.

4.3.1 Library, data generation and likelihood function

A library was generated as described in Section 4.2.1 for release conditions with the AHU fan off. Hypothetical release data were generated using the methods described in Section 4.2.3. In each case, 20 g of contaminant was released and various door positions, release locations and durations were simulated.

The first analysis evaluates performance of a single "typical" contaminant sensor network against each release scenario. The second analysis evaluates the performance of a combinatorial set of networks against releases originating in Stairwell 2. Table 2 lists the release conditions and networks evaluated. The likelihood function follows the description in Section 4.2.1.

4.3.2 "Typical" network performance

In this case, we deployed five contaminant sensors in the network. No sensor was placed in the release zone. Sensors were deployed on each floor, but no two sensors were placed in adjacent zones. If a floor had two sensors, one sensor was located in the stairwell and one in the main interior space (Table 2).

Figures 4 and 5 illustrate the time-dependent capability of the networks to locate the release and to estimate the released mass. As shown, door-position sensors improve the source identification by ~ 2–20% and reduce the uncertainty of the release mass estimated by ~ 40–85%. Because there are many release conditions that might explain the few sensor readings, the uncertainty bounds are not immediately narrowed.

For the Room 1.1 release, the system has difficulty locating the release, either with or without door-position sensors, suggesting that the contaminant sensor placements are suboptimal for this release location. However, both networks characterize the released mass with reasonable confidence intervals within 10 min, likely because releases of similar magnitude originating from the different locations generate similar concentrations at the sensed locations. This result is also consistent with the notion that information pertaining to the release mass persists even after information about the release location dissipates [3].

For each of the other simulated experiments, door-position sensors help to correctly identify the release location to higher probabilities and to estimate the release mass to narrower confidence intervals for a given monitoring time. The performance disparity among individual networks underscores the need to evaluate multiple networks against multiple release conditions.

Incorporating multiple sensor types may improve estimation of past and future concentrations. Figure 6 shows the concentration profiles for the Room 2.2 release for networks with and without door-position sensors, prior to receiving data and after receiving 5 min of data. Concentrations in one zone from each floor are plotted, none of which had sensors.

The left frames illustrate the prior confidence intervals of the concentrations in each of the selected zones. The "measured" concentrations for the release zone (Room 2.2) are not within the confidence interval. This outcome is reasonable given that there are eleven other equally likely release zones possible. After 5 min of updating, both networks (see central and right frames) are able to determine with narrow confidence intervals that the past and future concentrations in Stair 3 are small. The wider confidence intervals in Rooms 1.2a and 2.2 suggest that either zone may have been proximate to the release source. The networks with door sensors further reduce the uncertainty of the concentrations in these rooms. The distribution

without door sensors, however, is preferable for Room 2.2, as it fully spans the data; still, the reconstruction with door position sensors is an improvement over the prior uncertainty distribution.

4.3.3   Combinatorial network analyses

Combinatorial network analyses were performed to assess the consistency and variability of the performance benefit (and penalty) from fusing information from door-position sensors with data from contaminant-concentration sensors. We evaluated all networks containing either 3 or 8 contaminant sensors, with and without a full set of door-position sensors, against the two Stairwell 2 releases.

Excepting a few anomalous networks, door-position sensors consistently improve the identification of the release location, by an average of ~ 6–7% for three-sensor networks and ~ 2–3% for eight-sensor networks as illustrated in Figure 7. The average performance improvement in these scenarios is higher than what was observed with the AHU on, and there are fewer scenarios for which fusing door-position information imposes a penalty. However, the results also suggest that chemical sensor placement is the more significant design factor, and that door sensors do not compensate for poor contaminant-sensor placement. Similarly, door sensors are of limited value for cases in which chemical sensors are optimally located; there is little advantage to knowing the door positions when the information content of the chemical sensors is rich.

Quickly identifying the release location is important because information pertaining to the release location erodes as the interzonal transport time scale is exceeded. If a sensor-system goal is to locate the release to a minimum probability of 0.9, then several multisensor networks have a clear advantage on the order of minutes (Figure 8). Many three-sensor networks without door

20

sensors are unable to identify the source to at least a 0.9 probability within 10 min.

Door-position sensors are more consistently valuable under (simulated) naturally driven airflow conditions, where door positions have a relatively greater effect on airflow. A concept that we term the "degree of influence" may help to identify which sensors (in addition to contaminant sensors) may be useful in a hybrid sensor network. To be useful for sensor-system design, this concept should ideally be embodied in a metric that would quantify the extent to which a measurable model input affects contaminant transport over a range of operating and release conditions. Modeling analysis can help identify high-value measurements; however, these explorations would be limited by the fidelity of the model to the underlying physical processes. Field experiments can also help identify influential physical factors, although conducting experiments over a broad range of conditions is expensive.

A key challenge is that the actual or realized influence of any factor for a release event — as exhibited by the empirical results in this study — depends on specific release conditions, such as the magnitude and location. For example, the door positions may significantly affect the relative concentrations among sensed zones for some release scenarios, but less so for others. Furthermore, leveraging the degree of influence for model-based sensor fusion relies on the ability to model the effects of the parameter and an ability to characterize the inherent error structures. This task is much simpler when using simulated data. However, adequate fidelity of the simulation model for describing the physical system is needed if simulated data are to be used to design reliable sensor networks.

5 Conclusions

In this paper, we have explored how combining information from different types of sensors might improve the diagnosis of an acute contaminant release event that impacts indoor air. We

discuss a conceptual framework for sensor fusion and present an algorithmic approach. The conceptual framework is based on how the measurement is mapped to the model, its relative role in influencing contaminant fate and transport, and the dynamic state of the building or environmental condition it represents.

Heterogeneous sensors are more valuable when they characterize a model input that exerts larger influence on contaminant transport, as suggested by the illustrative example using door-position sensors. Several contaminant-monitoring networks were evaluated in this study, with and without door-position sensors, for conditions with the AHU on and off, using experimental and synthetic tracer-gas data. Door-position sensors improved the reconstruction of the release event for some scenarios with the AHU on, but did so more consistently for conditions with the AHU off for which there were far fewer cases of degraded performance. Penalties are more likely to occur when the parameter exhibits low influence on the contaminant concentrations at the sensed locations. In such cases, fusing information on what is a meaningless parameter may produce a reduced library that is insufficient to accurately reconstruct the event. One possible way of preventing library insufficiency problems is to dynamically generate the library using available model input information.

Specifically, this study indicates that while door-position sensors may help the sensor system to characterize the release more efficiently, and with less computational effort, they cannot compensate for poor contaminant sensor placement nor for an insufficient number of sensors. Further research should develop a metric for the degree of influence concept, to help identify and rank useful information in advance. A key challenge to developing this concept is that the actual influence depends on unknown release characteristics, such as the release location and mass, and will vary by location. In implementation, the designer may choose to optimize the

network for diagnosing certain release conditions, such as for detecting a release that originates in a particular room. Doing so narrows the scope of the technical challenge.

Although conceptually simple, the door-position sensor example illuminates some important aspects of the conceptual framing and suggests that disparate sensors, when used appropriately, could complement contaminant sensors in systems that are designed to characterize high-risk pollutant releases. Other sensor types, such as pressure and control-system signals, could be evaluated in future work using the concepts introduced and methods employed here. A more sophisticated implementation of the BMC algorithm, or other algorithms, may be required to ensure robust sensor system performance in dynamic conditions.

Acknowledgements

References

[1] Sohn MD, Reynolds P, Singh N, Gadgil AJ. Rapidly locating and characterizing pollutant releases in buildings, Journal of the Air & Waste Management Association 2002; 52: 1422-1432.

[2] Sreedharan P, Sohn MD, Gadgil AJ, Nazaroff WW. Systems approach to evaluating sensor

characteristics for real-time monitoring of high-risk indoor contaminant releases, Atmospheric Environment 2006; 40: 3490-3502.

[3] Sreedharan P, Sohn MD, Nazaroff WW, Gadgil AJ. Influence of indoor transport and mixing time scales on the performance of sensor systems for characterizing contaminant releases, Atmospheric Environment 2007; 41: 9530–9542.

[4] Liu X, Zhai Z. Prompt tracking of indoor airborne contaminant source location with probability-based inverse multi-zone modeling, Building and Environment 2009; 44: 1135-1143.

[5] Liu X, Zhai Z. Protecting a whole building from critical indoor contamination with optimal sensor network design and source identification methods, Building and Environment 2009; 44: 2276-2283.

[6] Zhang T, Chen Q. Identification of contaminant sources in enclosed spaces by a single sensor, Indoor Air 2007; 17: 439-449.

[7] Senocak I, Hengartner NW, Short MB, Daniel WB. Stochastic event reconstruction of atmospheric contaminant dispersion using Bayesian inference, Atmospheric Environment 2008; 42: 7718-7727.

[8] Hall DL, Llinas J. An introduction to multisensor data fusion. Proceedings of the IEEE 1997; 85: 6-23.

[9] Sanderson AC, Hombal V, Fries DP, Broadbent HA, Wilson JA, Bhanushali PI, Ivanov SZ, Luther M, Meyers S. Distributed environmental sensor network: Design and experiments. IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, 2006, pp. 79-84.

[10] Beck MB, Ravetz JR, Mulkey LA, Barnwell TO. On the problem of model validation for predictive exposure assessments. Stochastic Hydrology and Hydraulics 1997; 11: 229-254.

[11] Oreskes, N. Evaluation (not validation) of quantitative models, Environmental Health Perspectives 1998; 106: 1453-1460.

[12] Sreedharan P. Bayesian based design of real-time sensor systems for high-risk indoor contaminants. PhD Dissertation, University of California, Berkeley, 2007.

[13] Brand KP, Small MJ. Updating uncertainty in an integrated risk assessment: Conceptual framework and methods. Risk Analysis 1995; 15: 719-731.

[14] Sohn MD, Small MJ, Pantazidou M. Reducing uncertainty in site characterization using Bayes Monte Carlo methods, Journal of Environmental Engineering ASCE 2000; 126: 893-902.

[15] Sextro RG, Daisey JM, Feustel HE, Dickerhoff DJ, Jump C. Comparison of modeled and measured tracer gas concentrations in a multizone building. Proceedings of the 8th International Conference on Indoor Air Quality and Climate – Indoor Air 99, Edinburgh, Scotland, 1999, vol. 1, pp. 696-701.

Table 1 Parameter values used to simulate the library of 5000 contaminant releases.

Parameter	Values
Source location	Twelve locations: any room, or stairwell and return duct. Each location is equally probable.
Source duration	1 s to 5 min; log-uniform distribution.
Source amount	10 to 100 g; log-uniform distribution.
Door position	32 possible door configurations, based on interior doors on each level and stairwell doors acting independently.

Table 2 Description of simulated release conditions with AHU off.

Release location	Internal door position	Release duration	Sensor locations
Typical network analyses			
Room 1.1	Stairwell doors closed, other interior doors open	1 s	Rooms 1.2b, 2.3, 3.1, Stairwells 1 and 2
Room 2.2	All doors open	1 s	Rooms 1.3, 2.3, 3.1, Stairwells 1 and 2
Stairwell 3	All doors open	1 s	Rooms 1.3, 2.3, 3.1, Stairwells 1 and 2
Room 2.2	All doors open	120 s	Rooms 1.3, 2.3, 3.1, Stairwells 1 and 2
Combinatorial network analyses			
Stairwell 2 [a]	All doors open	1 s	All possible [b]
Stairwell 2	Stairwell doors closed, other interior doors open	1 s	All possible [b]

[a] This release scenario is identical to that investigated in Section 4.2.4 (where a release originating in Stairwell 2, but with the AHU was on, was explored).
[b] The combinatorial analyses considers all possible network configurations ranging from 1 to 11 sensors.

Figure Captions

Fig 1 Layout indicating location of interior doors for the case study.

Fig 2 Experiment 1: probability of correctly identifying the release at 10 min (a) return intake and (b) return intake or Room 1.2a, for networks containing 2, 6 and 10 contaminant sensors, with and without door sensors. ($\Delta \equiv$ probability with door – probability without door.) The experiment involved release and sampling of tracer gas data with the AHU on.

Fig 3 Effect of door sensors for correctly identifying Stairwell 2 as the site of the contaminant release, using simulated event data. The air handling unit was on in the simulation. Probabilities are plotted for 5 and 10 min of updating for networks containing 3, 5, 7 and 9 contaminant sensors.

Fig 4 Time-dependent probability of locating a release with and without door position sensors for the "typical" network. The analysis is based on simulated data with the air handling unit off.

Fig 5 Time-dependent estimates of mass released for "typical" network, with and without door-position sensors. The true released mass is 20 g. Plotted are the median (dashed line), and the 10th and 90th percentiles (solid lines). The analysis is based on simulated data with the air handling unit off.

Fig 6 Confidence intervals of time-dependent concentrations before receiving data and after 5 min of updating for networks with and without door sensors. Median (dashed line), 10th and 90th percentiles (solid lines), and (simulated) concentration data for pulse release in Room 2.2 (points) are displayed. The analysis is based on simulated data with the air handling unit off.
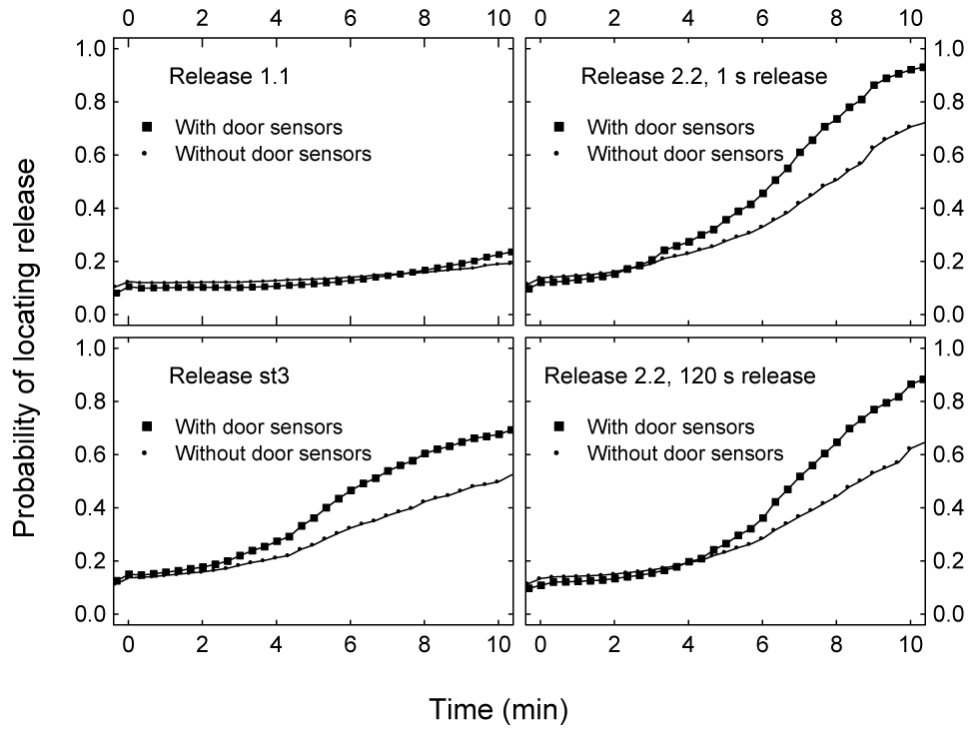
Fig 7 Effect of door-position sensors for identifying the location of releases that occurred in Stairwell 2. Probabilities are shown of being correct after 10 min of updating for networks containing 3 and 8 contaminant sensors. The analysis is based on simulated data with the air handling unit off.

Fig 8 Time to locate release to 0.9 probability for releases occurring in Stairwell 2 for networks with and without door-position sensors. A negative time indicates that a probability of 0.9 was not reached by 11 min. The analysis is based on simulated data with the air handling unit off.
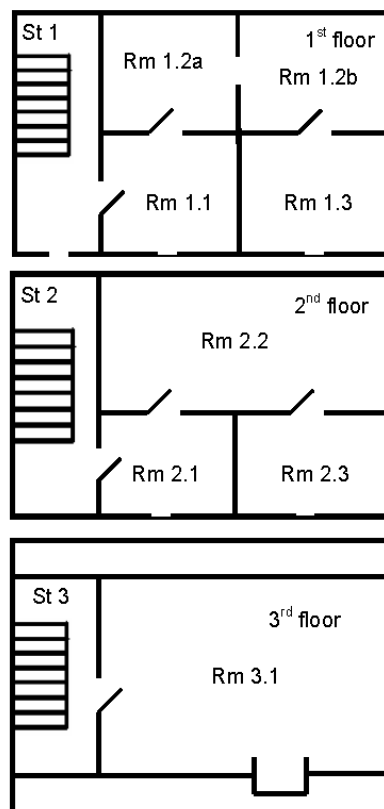
Fig 1
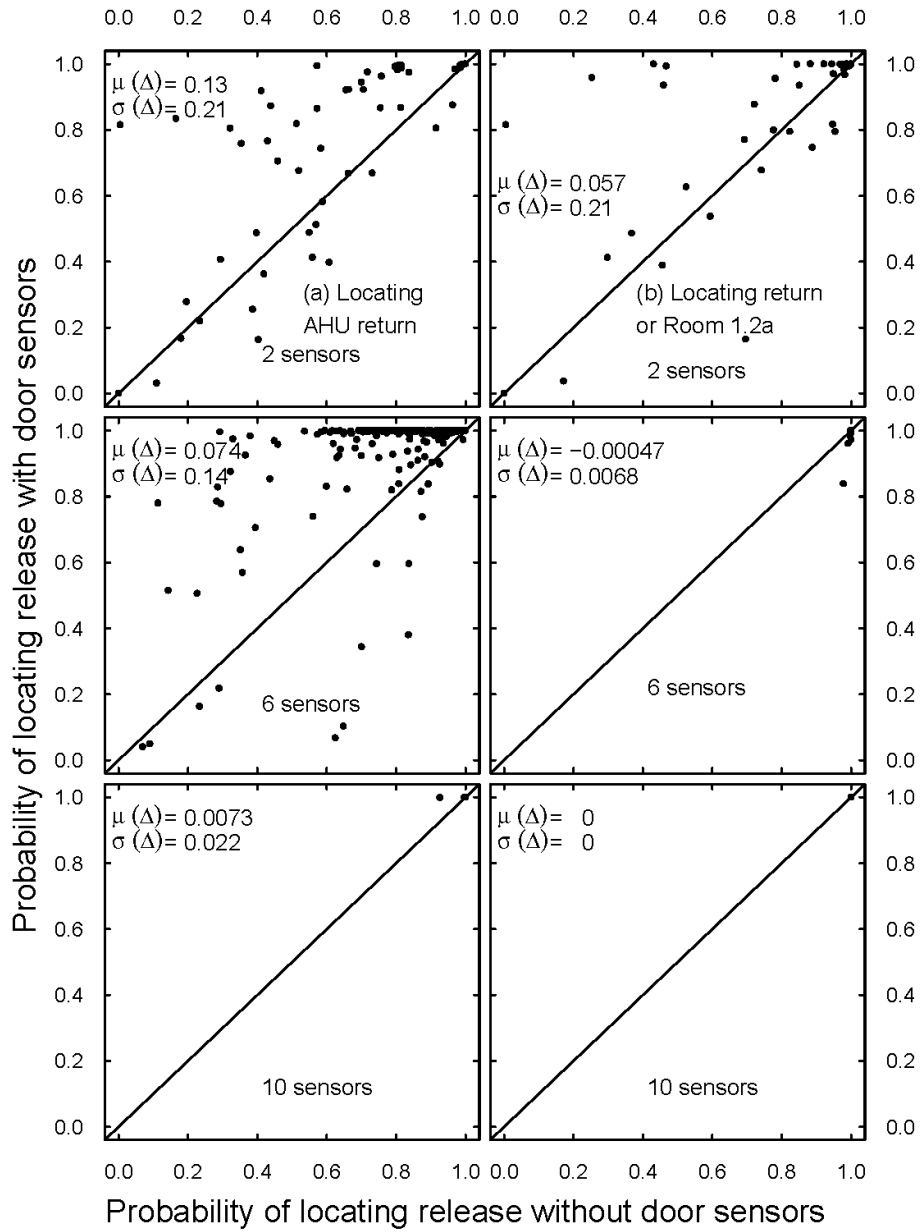
Fig 2
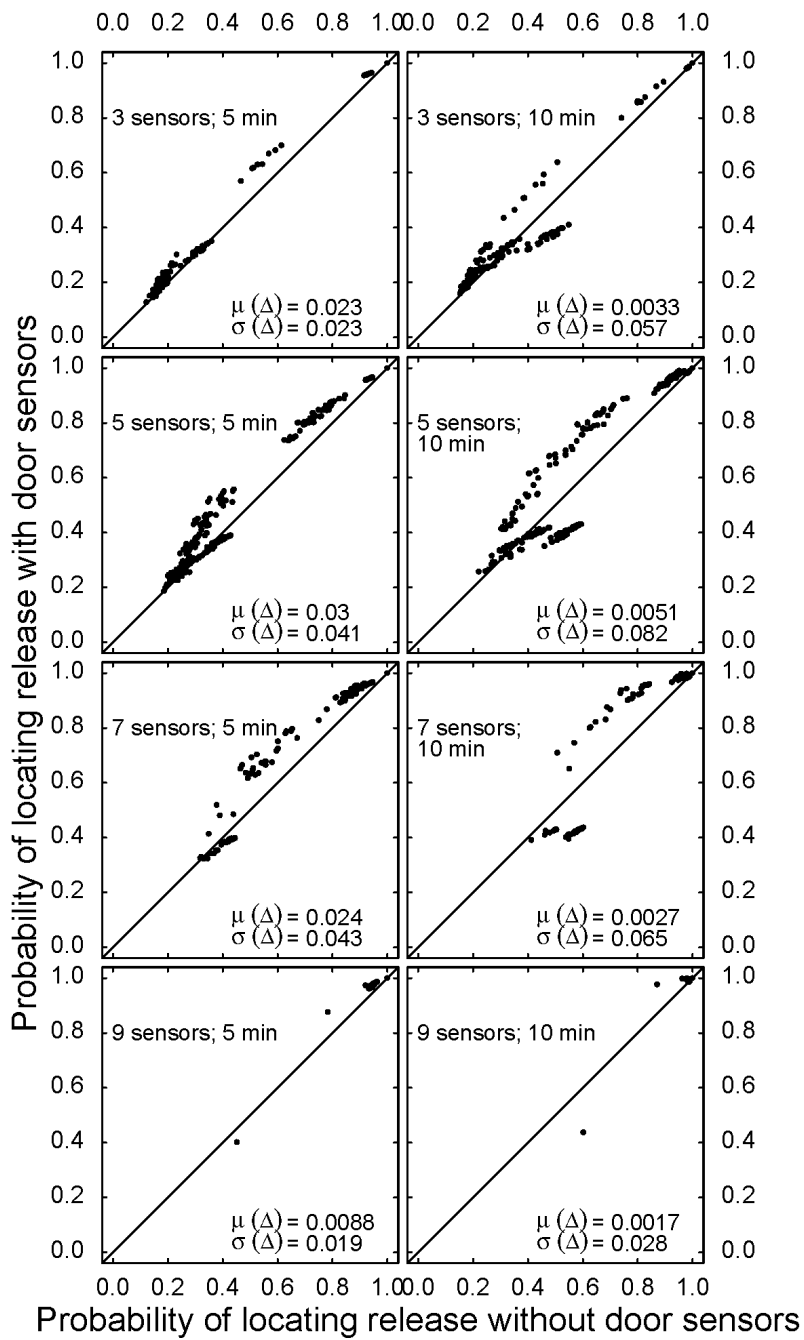
Fig 3

Fig 4

Fig 5

Fig 6

Fig 7

Fig 8