

# Detection and Localization of Material Releases With Sparse Sensor Configurations

Emily B. Fox, *Student Member, IEEE*, John W. Fisher, III, *Member, IEEE*, and Alan S. Willsky, *Fellow, IEEE*

**Abstract**—We consider the problem of detecting and localizing a material release utilizing sparse sensor measurements and formulate the problem as one of abrupt change detection. Methods which rely on single-sensor detection require dense deployment to achieve adequate coverage; costly sensors preclude such approaches. Furthermore, localization requires the fusion of multiple sensor measurements. Fusion in sparse sensor configurations is dependent on the knowledge of the dynamics of particle dispersion, which is, itself, problematic due to the inherent randomness on the wind field. We consider the efficacy of using an approximate dynamic model with coarse parameter estimates for the detection and localization of material releases. Specifically, we consider propagation models consisting of diffusion plus transport according to a Gaussian dispersion model. Assuming a known wind field, unconstrained intersensor communication, and a centralized processor, we derive optimal inference algorithms and provide a hybrid detection-localization hypothesis-testing framework with linear growth in the hypothesis space. We then analyze the probability of detection, time-to-detection, and localization performance as a function of the number of sensors. Furthermore, we examine the impact on performance when the underlying dynamical model deviates from the assumed model. This detailed analysis provides the basis for the design of more sophisticated algorithms for 1) performing robust detection followed by refined nonlinear parameter estimation which provides enhanced localization, and 2) distributed architectures aimed at conserving communication resources in which detections within local clusters are used to trigger more intensive intercluster communication to improve detection and localization.

**Index Terms**—Distributed detection, localization, sensor networks.

## I. INTRODUCTION

WE consider the problem of detecting and localizing a material release utilizing sparse sensor measurements and cast the problem as one of abrupt change detection. The underlying objective is one which appears in a variety of applications. In any distributed sensor network application, questions arise as to how the network should be configured to monitor the given environment as well as how the sensor measurements should be combined in an informative manner. In addition, as the number of sensors increase with the size of the region being monitored, one needs to address how to distribute computation and reduce costly communication.

It is frequently the case that only a sparse set of sensors can be deployed to monitor an area. Typically, these sensors are sensitive to local regions and as such yield only a myopic view of

the environment when used in isolation. These issues complicate the monitoring problem and impact the ability to rapidly detect releases. However, by exploiting a dynamic model it is possible to combine measurements from several sensors so as to improve detection, time-to-detection, and localization beyond that obtained using a single sensor.

As our goal is primarily detection and localization, we do not attempt to model the full complexity of the system dynamics. Instead, we strive to model only as much of the physics as is necessary to achieve reasonable performance. We use coarse approximations to the true system parameters to derive optimal inference algorithms assuming this parametrization is accurate and then utilize the sensor measurements to provide enhanced estimation of these parameters. In our simulation experiments, we characterize the performance for a set of sensor configurations and analyze how much deviation from the assumed coarse model can be tolerated without detrimental effects to our performance criteria. This analysis imparts a method for assessing how accurately we need to know the parameter set for initial detection. We use the initial detection as a trigger for refined nonlinear parameter estimation with the goal of enhanced localization performance.

Additionally, our analysis provides the foundation for a distributed architecture in which clusters of sensors perform optimal local fusion to monitor subregions. Local processing of subregions is then aggregated across clusters in order to monitor a wider geographic region. However, this costly communication between sensor clusters only occurs after a local detection rather than at every time step as in the centralized version of the inference algorithm. Because missed detections are very costly in this application, we set the cluster thresholds to produce a higher probability of detection and, thus, higher false alarm rates, and rely on the subsequent more accurate global processing to filter some of these detections by having a more conservative threshold. Our base analysis provides simple, approximate methods for determining local and global thresholds.

Previous methods considered Bayesian approaches to biological or chemical release detection. Nofsinger and Smith describe a method of using ensemble Kalman filtering for solving the inverse problem of localizing in 2-D a chemical plume occurring at a known release time in the absence of wind [1]. Nofsinger and Cybenko explore detecting and tracking a chemical plume in a known wind field using multiple hypothesis tracking (MHT) combined with a customized plume predictor in [2]. Zhao and Nehorai discuss plume modeling and detecting and localizing a moving source using a nonrecursive dynamic model and generalized likelihood ratio test (GLRT) [3]. Beyond the methods of previous work, we develop a recursive inference algorithm with

Manuscript received April 26, 2006; revised June 29, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jonathan H. Manton.

The authors are with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Digital Object Identifier 10.1109/TSP.2006.890914

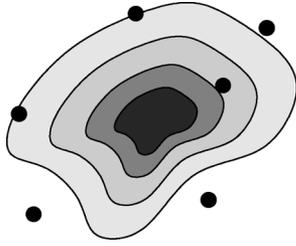


Fig. 1. Sparse sensors, indicated by  $\bullet$  symbols. Contours indicate particle density following a localized release at a single time instant.

robust detection and enhanced localization performance through a communication sensitive hierarchical architecture triggered by early detection. Our approach is described for the case of single releases, but the system is capable of handling multiple releases (see Section VI).

## II. PROBLEM FORMULATION

The problem of detecting a localized release is illustrated in Fig. 1. Particle sensors, indicated by black circles, are distributed sparsely over a region of interest where the measure of sparsity is relative to the rate of diffusion. Particle concentrations at time  $t > t_o$  are indicated by shaded contours where  $t_o$  is the time of the release. In a sparse sensor network, it is generally the case that there is a delay in observability of the event.

The goal, then, is to intelligently combine sensor measurements distributed both temporally and spatially to first detect the release in a timely manner and secondarily to localize the release in space and time. The two goals are distinct in that reasonable detection may be obtained from early measurements while localization generally relies on measurements aggregated over a longer time period. Bayesian filtering approaches provide a framework for this type of problem, presuming that the model parameters are available and intersensor communication is possible. We adopt a standard abrupt change methodology [4]–[6] where the standard state space equations are modified to include a term representing a localized event whose effects persist over time. This results in a hybrid detection-localization hypothesis framework.

Given a sparse sensor configuration and uncertainty regarding the underlying dynamics (e.g., rate of diffusion, knowledge of the wind field), it is generally the case that optimal inference under an assumed parametrization degrades in both detection and localization performance as compared to the matched case. Our results indicate that inference using a mismatched model yields reasonable detection performance at the cost of precise localization. This detection robustness allows for a more sophisticated inference procedure in which initial detections are used to trigger nonlinear least squares estimation of parameters so as to improve localization performance.

When we increase the size of the sensor network under examination and place costs on communication, it is advantageous to consider a distributed approach comprised of clusters performing local fusion along with a global decision maker. Our analysis provides the basis for a) understanding how to distribute computation in order to reduce costly communication and b) exploring the resulting performance tradeoffs.

### A. Bayesian Approaches for Localization

Given a sparse sensor network, it is appropriate to consider a model of a comparable level of spatial granularity. In particular, we use a spatially and temporally discrete model in which we partition space into a set of cells and define the state of the system at time  $t \in \mathbb{Z}_+$  to be the vector of mean particle concentrations in the set of cells. A certain subset of these cells have sensors which provide noisy observations of the particle concentration in these locations.

Accurately describing the spatial and temporal distribution of particles released into an environment is a challenging problem due to the inherent randomness of the wind velocities in turbulent flow. The concentration of the particles,  $c$ , satisfies the continuity partial differential equation (PDE) [7].

We simplify the dynamic model by idealizing the abrupt change to be an instantaneous point source which leads to a common approximation to solving the continuity PDE, namely the Gaussian puff model [7]

$$\langle c(x, y, \Delta t) \rangle = \frac{c_o}{4\pi\Delta t(K_{xx}K_{yy})^{\frac{1}{2}}} e^{-\frac{(x-x_0)-\bar{u}\Delta}{4K_{xx}\Delta}} e^{-\frac{(y-y_0)-\bar{v}\Delta}{4K_{yy}\Delta}}. \quad (1)$$

Here,  $\langle c(x, y, \Delta t) \rangle$  is the mean concentration at location  $(x, y)$  at time  $\Delta t = t - t_o \in \mathbb{R}_+$ ,  $(\bar{u}, \bar{v})$  are the  $(x, y)$  components of the mean wind velocity at time  $\Delta t$ ,  $c_o$  is the initial release concentration,  $(x_0, y_0)$  are the coordinates of the release location, and  $(K_{xx}, K_{yy})$  are the diffusion coefficients which account for the randomness of the wind field (see [7] for a complete derivation of this approximation).

We derive our elementary algorithm assuming that  $(\bar{u}, \bar{v})$  are known, then examine the algorithm's robustness relative to errors in our knowledge, and finally devise a nonlinear estimation algorithm to refine our estimates of these quantities.

Our state,  $x(t)$ , is taken to be the vector of cell particle concentrations. We consider  $t \in \mathbb{Z}_+$ , a discrete time index, and form a mapping,  $A(t; \mathcal{V}_w)$ , of mean particle concentration from one time step to the next parameterized by the mean wind field  $\mathcal{V}_w = (\bar{u}, \bar{v})$ . The concentration mapping and wind field are determined by the cell-localized Gaussian diffusion kernels  $\langle c(x, y, t) \rangle$  and wind vector  $(\bar{u}, \bar{v})$ , respectively. We write the system equations as

$$x(t+1) = A(t; \mathcal{V}_w)x(t) + w(t) + \beta f_i(t, \phi) \quad (2)$$

$$y(t) = C(t)x(t) + v(t). \quad (3)$$

As described in [5] and [6], for example, the term  $f_i(t, \phi)$  is inserted into the standard state space equations to account for an abrupt change to a system. In this application, we take  $f_i(t, \phi)$  to be an impulse of particles in cell  $i$  at discrete time  $\phi$  of magnitude  $\beta$ . Other parameters of these equations include  $w(t)$  the process noise,  $v(t)$  the measurement noise, and  $C(t)$  a binary matrix with ones in sensor cells to relate the state to the noisy particle concentration measurements  $y(t)$ . Both the process and measurement noise are assumed to be Gaussian and independent of each other over all time.

By adding the abrupt change term  $f_i(t, \phi)$ , we incorporate knowledge of the localized nature of the release as opposed to accounting for the change by merely increasing the variance on

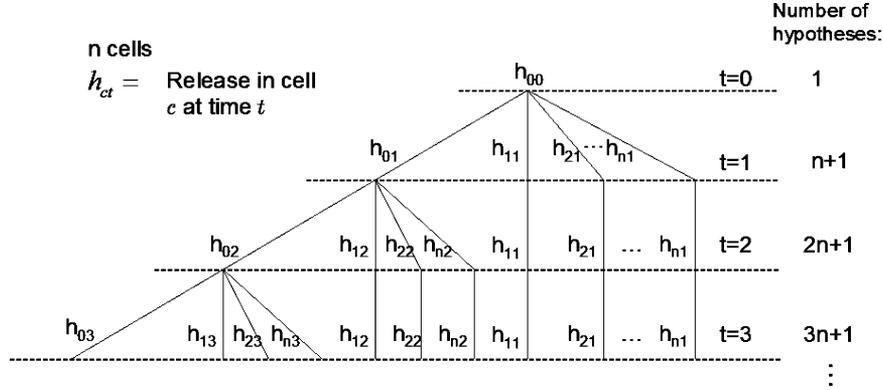


Fig. 2. Hypothesis tree for a region with  $n$  cells at time step 3. Hypothesis  $h_{ct}$  represents a release in cell  $c$  at time  $t$ .

the dynamic uncertainty. The advantage to this approach is that it allows one to simultaneously model a release as a localized event in time and space while capturing its persistent effects. This leads to an efficient hypothesis testing approach. The equations presented in this section follow the development of [5].

The linearity of the system allows us to decompose the state into a term  $x_0(t)$  due the background process and a term  $\gamma_i(t, \phi)$  which captures the persistent effect of the release  $f_i(t, \phi)$

$$x(t) = x_0(t) + \beta\gamma_i(t, \phi). \quad (4)$$

The utility of the decomposition is that the problem reduces to one of detection in correlated noise. The detection formulations are simplified by whitening the data, yielding

$$\nu(t) = \nu_0(t) + \beta\rho_i(t, \phi) \quad (5)$$

where  $\nu_0(t)$  is the white sequence resulting from  $x_0(t)$  and  $\rho_i(t, \phi)$  is the signature of  $\gamma_i(t, \phi)$  once it has been passed through the whitening filter (specifically a Kalman filter [8]). Thereafter, we process the resulting innovations sequence  $\nu(t)$ .

### B. Hypothesis Testing Framework

While the preceding framework has simplified the dynamic model, we are still left with the task of testing various release hypotheses enumerated over space (cell index) and time. Additionally, the model is parameterized by the unknown release amount  $\beta$ . We address this issue using a generalized likelihood ratio (GLR) [8], [9] in which one substitutes the ML estimate of  $\beta$  for each hypothesis to approximate the Bayes optimal joint estimation of all parameters. This allows us to formulate and analyze a set of enumerated hypotheses which both detect and localize a release.

Referring to (5), we can construct an indexed set of hypotheses,  $H_{i,\phi}$ , based on release cell location  $i$  and time  $\phi$ , to which we compare the null hypothesis of no release,  $H_0$

$$H_0 : \nu(t) = \nu_0(t) \quad (6)$$

$$H_{i,\phi} : \nu(t) = \beta\rho_i(t, \phi) + \nu_0(t). \quad (7)$$

It is well known that the sufficient statistics for this particular hypothesis test are the correlation between the signal  $\rho_i(t, \phi)$

and the output  $\nu(t)$  normalized by the inverse covariance  $V^{-1}$  of the residual  $\nu(t)$  [9]

$$d_i(t, \phi) = \sum_{\tau=\phi}^t \rho_i^T(\tau, \phi)V^{-1}(\tau)\nu(\tau) \quad (8)$$

and the normalized energy of the signal

$$a_i(t, \phi) = \sum_{\tau=\phi}^t \rho_i^T(\tau, \phi)V^{-1}(\tau)\rho_i(\tau, \phi) \quad (9)$$

resulting in a log-likelihood of

$$l_i(t, \phi) = \beta d_i(t, \phi) - \frac{1}{2}\beta^2 a_i(t, \phi). \quad (10)$$

See Appendix A of [10] for a derivation.

We simplify the analysis in the following way. We model hypotheses restricted to there being one and only one release, which leads to linear growth in the hypothesis tree rather than exponential. Even though we have simplified the hypothesis space under consideration, the formulation still allows for detection in the multiple release situation because the inference procedure will try to explain the measurements by *some* hypothesis. In Section VI, we discuss the scenarios in which our current formulation (with minor extensions) can handle multiple detections and how this procedure is simplified by considering a distributed formulation.

The resulting hypothesis tree after three time steps is shown in Fig. 2. Because of the growing size of the tree, in practice, a sliding window is used to implement the inference algorithm. This is well justified for two reasons. The first is that it takes some time for releases to propagate to the nearest sensor. The second is that, after a period of time, additional measurements do not contribute significantly to the probability of a detection in a given cell at a specific time. Therefore, there is little need to look further back in time. Specifically, the window enforces that only hypotheses of a release occurring a maximum of  $M$  timesteps in the past are considered.

Conditioned on hypothesis  $H_{i,\phi}$ ,  $d_i(t, \phi)$  is Gaussian with mean  $\beta a_i(t, \phi)$  and variance  $a_i(t, \phi)$  [10]. Consequently, the ML estimate of  $\beta$  is

$$\hat{\beta}_i(t, \phi) = \frac{d_i(t, \phi)}{a_i(t, \phi)} \quad (11)$$

and the resulting log-likelihood ratio is

$$l_i(t, \phi) = \frac{1}{2} \frac{d_i^2(t, \phi)}{a_i(t, \phi)}. \quad (12)$$

If one takes  $\hat{\phi}$  and  $\hat{i}$  to be the values that maximize  $l_i(t, \phi)$  and define

$$l_{\hat{i}}(t, \hat{\phi}) \triangleq \max_{i, \phi} l_i(t, \phi) \quad (13)$$

and denotes  $H_1$  as the hypothesis that a release has occurred regardless of time or location, then the likelihood ratio test for detection is simply

$$l_{\hat{i}}(t, \hat{\phi}) \underset{H_0}{\overset{H_1}{\geq}} \eta \quad (14)$$

where the value of  $\eta$  is set using Neyman–Pearson criterion to specify a fixed probability of false alarm,  $P_F$ .

### III. EMPIRICAL RESULTS

We present a series of experiments in which we examine the utility of using multiple sensors. The first set of experiments uses a known mean wind field to characterize probability of detection, time to detection, and localization in time and space as a function of the number of sensors. We consider the case of pure diffusion (zero mean wind field) in these experiments in order to illustrate some of the localization results. Note that the performance of the inference procedure would be similar in the presence of a *known* wind bias. The other two sets of experiments examine the efficacy of the inference procedure as the model deviates from that assumed by the inference procedure. In the second experiment, a sequence of increasing wind biases are incorporated into the dynamic model, but not into the inference procedure. In the third experiment, random variations are generated on the mean wind field provided to the simulation while the inference procedure only uses knowledge of the mean wind field. For both of these model mismatch scenarios, we analyze the impact on detection probabilities for four and 16 sensor configurations as well as localization performance.

The conclusions of this analysis are twofold. First, we see that detection is robust to inaccuracies in our knowledge of the wind field although localization can be poor. This conclusion leads to the modification of our inference procedure to include a non-linear least squares parameter estimation stage after detection to refine our knowledge of the system parameters so as to improve localization (see Section IV). Second, our analysis shows that in the sparse network configuration, groupings of four sensors performing local fusion are sufficient for detection and localization. This suggests a clustered approach to distributed processing over wider geographic regions. We discuss such an approach in Section V.

#### A. Simulation Parameters

We consider a release in the center of a room under various sensor configurations, consisting of one, two, four, and 16 sensors, as given in Fig. 3. This release location is the point whose distance to the nearest sensor is the same in all sensor configurations. The symmetry of the configurations serves to simplify the analysis of the results.

Fig. 3 shows the sensors overlaid on a  $13 \times 13$  cell area of interest. However, to reduce edge effects, the data is simulated

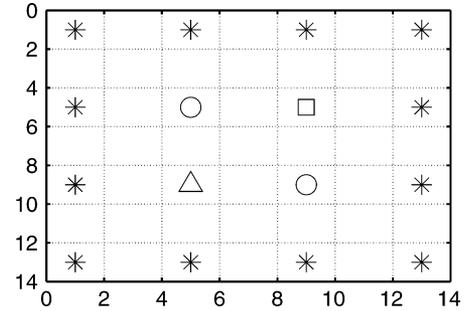


Fig. 3. Sensor configurations for 1  $\{\Delta\}$ , 2  $\{\Delta, \square\}$ , 4  $\{\Delta, \square, \circ\}$ , and 16  $\{\Delta, \circ, \square, *\}$  sensors.

over an area of  $25 \times 25$  cells. A sensor is assumed to observe the average particle concentration throughout the entire cell in which it is located. The standard deviation of the noise on the measurements is ten particles per cell area.

The particle concentrations are propagated using a Gaussian kernel with a diffusion coefficient of  $K_{xx} = K_{yy} = 0.5$ . This choice of diffusion kernel makes the sensor network sparse because the diffusion is slow relative to the sensor density (see simulation snapshots in Fig. 4). The process noise standard deviation is taken to be 100 particles per cell area. Initially, we restrict ourselves to a pure diffusion model ( $\mathcal{V}_\omega = \bar{0}$ ) in order to avoid the effects of wind which could blow the particles towards or away from the nearest sensor. In the case where there is wind, the utility of having multiple sensors to better cover the region will be apparent. For every experimental scenario, 100 Monte Carlo simulations were performed.

For all calculations involving probability of detection reliant on the 100 samples (see Figs. 5, 10, and 17), we can use an approximation based on the central limit theorem (CLT) [11] to bound the probability of a given discrepancy between our sample estimate and the true value. Take  $X_i$  to be a Bernoulli random variable indicating a detection for Monte Carlo sample  $i$ . Let  $p$  be the true probability of detection and  $P_D$  the sample estimate based on  $n$  samples. We can upper bound  $P(|P_D - p| \geq \epsilon)$  by assuming  $P_D - p$  has the largest possible variance, namely  $1/(4n)$  when  $p = 1/2$ . This assumption results in  $P(|P_D - p| \geq \epsilon) \approx 2(1 - \Phi(2\epsilon\sqrt{n})) = 0.05$  for  $n = 100$  and  $\epsilon = 0.098$ , where  $\Phi(z) = P(Z \leq z)$  and  $Z$  has a normal distribution. That is to say, with 95% confidence our sample estimates  $P_D$  are within 0.098 of  $p$ .

The size of the sliding window,  $M$ , was determined to be 14 time steps by analyzing the minimum time from the onset of the release by which the accrual of additional information was insignificant in the case of pure diffusion. In a sensor network which is distributed as a regular grid, as in the 16 sensor configuration used in this paper, the pure diffusion scenario will be a worst case analysis because any wind would blow particles more rapidly towards a sensor.

The log-likelihood ratio threshold was determined experimentally to achieve a false alarm rate of 0.01 on 1500 Monte Carlo simulations under benign conditions (i.e., no release) for the four and 16 sensor configurations. The threshold values are 9.55 and 10.85, respectively. These values were used in all experiments conducted.

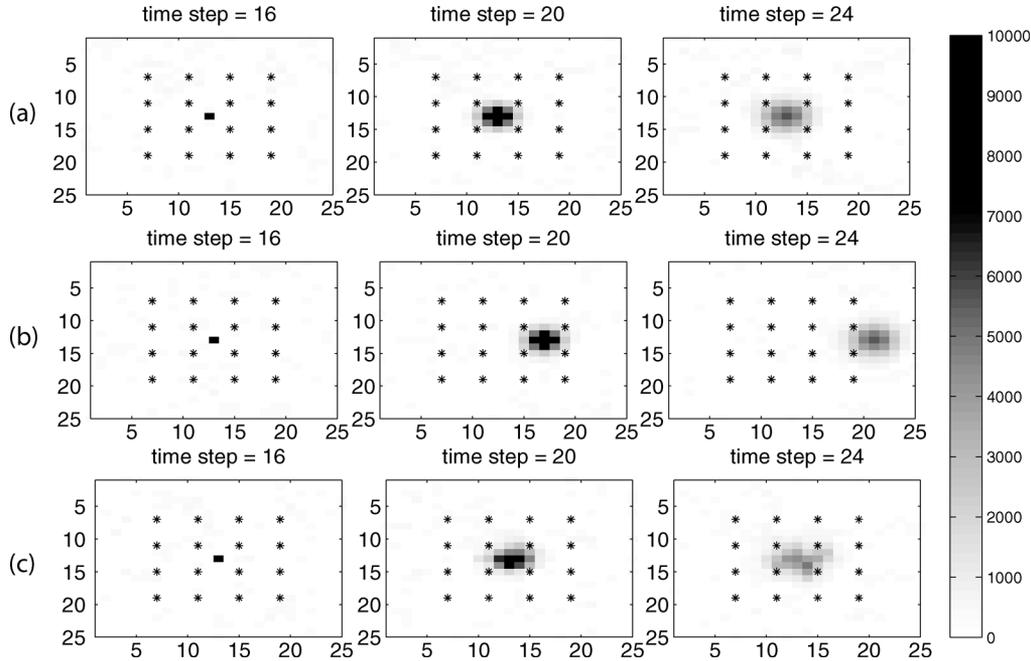


Fig. 4. Screenshots of a simulation of an aerosol release at time steps (left) 16, (middle) 20, and (right) 24 under (a) pure diffusion  $(\bar{u}, \bar{v}) = (0, 0)$ , (b) wind bias  $(\bar{u}, \bar{v}) = (1, 0)$ , and (c) random variation on the wind field  $(\bar{u}, \bar{v}) = (1, 0) + \mathcal{N}(0, 0.3162 * I_2)$ . The \* symbols indicate sensor locations in the 16 sensor configuration. Parameters: release amount =  $1e5$  particles; release time = 16; release location = (13,13); st. dev of dynamic uncertainty = 100 particles per cell area.

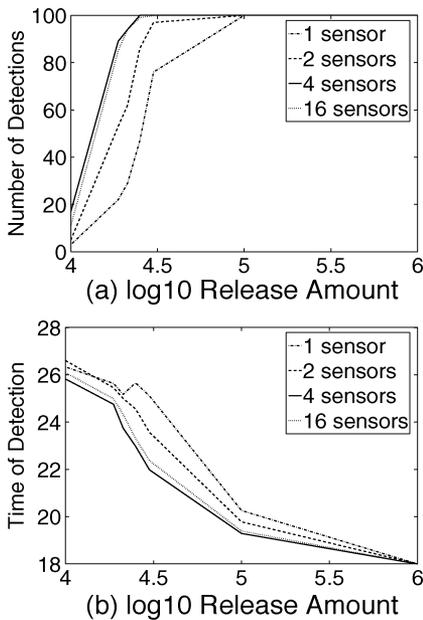


Fig. 5. (a) Number of detections per 100 simulation data sets and (b) time of first detection versus release amount. Actual release was at  $t_0 = 16$  s.

**B. Detectability**

Fig. 5 shows the number of detections achieved out of 100 simulations as a function of release magnitude for various sensor configurations. The figure demonstrates that significant gains are made in ability to detect small release amounts going from one to four sensors while there is only a marginal difference in performance between four and 16 sensors. For a sufficient

release amount, all sensor configurations are able to reliably detect.

The corresponding time to first detection for the same experiment is plotted in Fig. 5. The figure shows the time of detection of a release occurring at time index 16. For smaller release amounts, there is some performance gain obtained by using a larger number of sensors. Note that the discrepancy in this trend for small release amounts in the one sensor case is statistically insignificant due to the small number of detections. The convergence for large release amounts is indicative of the time required for the release to propagate to the nearest sensor and is purely a function of the relative distance between the release and the sensor and, thus, sensor density relative to the rate of diffusion. We conclude that the inner four sensors can provide sufficient evidence of a release in both the four and 16 sensor configurations earlier and more often than one or two sensors which have not aggregated enough information.

It is important to note in Fig. 5 that, for the release scenario described, four sensors actually marginally outperform 16 for a fixed probability of false alarm. The reason for this is that under pure diffusion with a release in the center, at least four sensors will see the release in both configurations. Almost regardless of the release amount, on average if the inner four sensors do not detect a release the outer ring of 12 sensors will not accrue substantial evidence to trigger a detection because of the lower signal to noise ratio of the measurements when the release has diffused to the outer ring. In addition, because the inference procedure hypothesizes releases at every location, not just the center, and the GLR tries to explain measurements by *some* hypothesis, adding sensors increases the probability of false alarm and, thus, requires a higher detection threshold. By only considering releases at the center, we do not observe the benefits of the

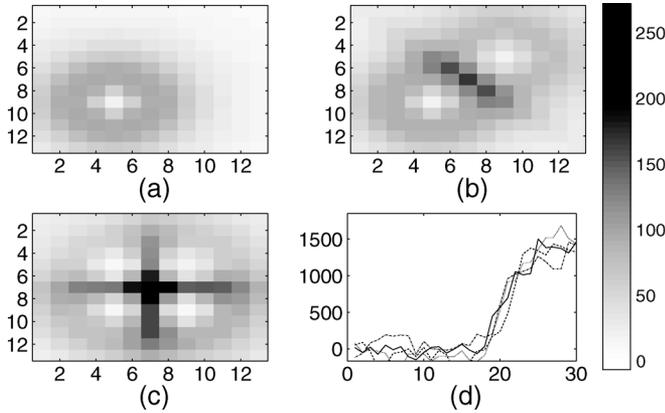


Fig. 6. (a)–(c) Maximum likelihood ratio at every cell in region of interest for one, two, and four sensor configurations. (d) The measurement sequence for the four sensor configuration.

16 sensor scenario, namely more accrued information or better spatial coverage, over that of the four sensor configuration while still having a significantly higher threshold. With a release in a random location, the 16 sensor configuration will have better performance over that of fewer sensors because the release is likely to be further from the nearest sensor in sparser configurations. Also, in many wind conditions the benefit of using more sensors becomes apparent.

C. Localization Uncertainty

In Fig. 6, we examine the hypotheses considered by the process at time  $t = 30$ , when the release occurred at  $t_o = 16$  s. At any given time, there are multiple hypotheses of releases in a given cell, each differentiated by various times of releases. The maximum likelihood ratio over all the hypotheses for each cell is plotted for various sensor configurations. The actual sensor measurements provided to the inference algorithm for the four sensor case are plotted in the lower right-hand figure. The measurements provided for each of the other sensor configurations were a subset of those provided in the four sensor case.

This image gives a graphical representation of the likelihood that a release occurred in any of the cell regions and, thus, illustrates the degree of localizability achievable with different sensor configurations. The results follow our intuition considering the simulations are under pure diffusion with a symmetric sensor configuration, and as such these plots serve to validate our model. In the one sensor case, the manifold is approximately circular indicating that the location of the release cannot be distinguished along a ring around the sensor. The two sensor and four sensor cases provide progressively better localization. As expected, two sensors have difficulty distinguishing along a line while four sensors are able to localize to a point. The degree of shading in the plots indicates the likelihood ratio value and, hence, the detection probability.

Note that the symmetry seen in Fig. 6 is due to the combination of the release being in the center of the sensor configurations and the model being pure diffusion. With wind or off-center release locations, the localization manifolds are asymmetric.

Each hypothesis has an associated time of release and maximum likelihood release amount, which is conveyed in Fig. 7.

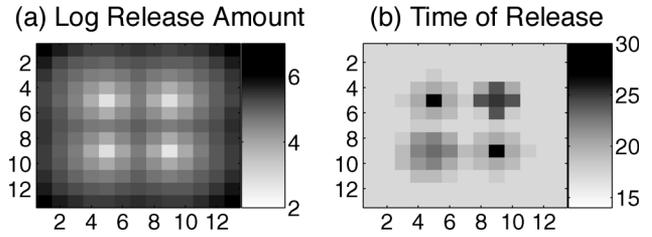


Fig. 7. (a) Map of maximum likelihood estimate of the log release amount associated with maximum LR hypothesis, four sensors. (b) Map of time of release associated with maximum LR hypothesis, four sensors.

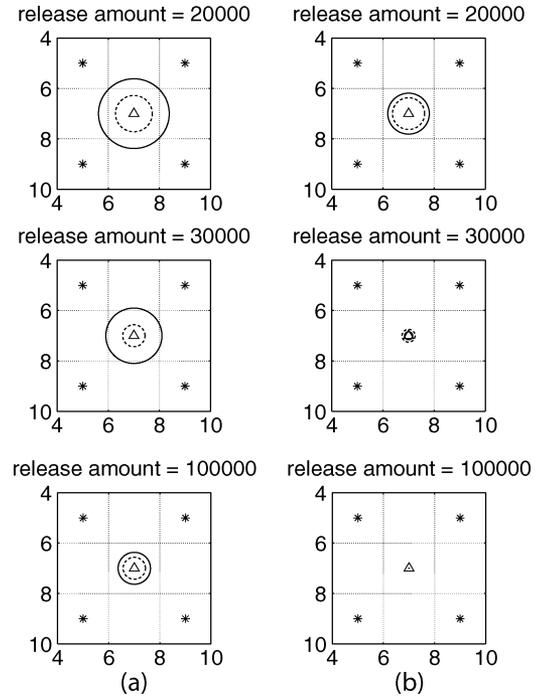


Fig. 8. Mean and covariance of localization for various release amounts conditioned on first detection based on information (a) at time of first detection and (b) accrued until the hypothesis is almost out of the window. Dashed lines = 16 sensor configuration, solid lines = 4 sensor configuration.

Consistent with the diffusion model, hypothesized releases further from the true release location had to occur earlier in time and be of a larger magnitude in order to appear similar to the actual sensor measurements.

Fig. 8 shows the localization accuracy achieved using four and 16 sensors for various release amounts. The circles illustrate the mean and covariance of the hypothesized location of the detected release. We show the plots under two detection schemes. The plots on the left are estimates produced at the time of first detection. The plots on the right are produced under a system where our goal is to alarm at the time of first detection and then wait an allotted period of time until producing localization results. This scheme involves accruing information until the detected hypothesis is almost outside the hypothesis window. We define the term almost to be a fixed lag parameter,  $n$ , which enforces that we stop accruing information when the window in which hypotheses of a release  $n$  time steps before the time corresponding to the initial best hypothesis are still considered.

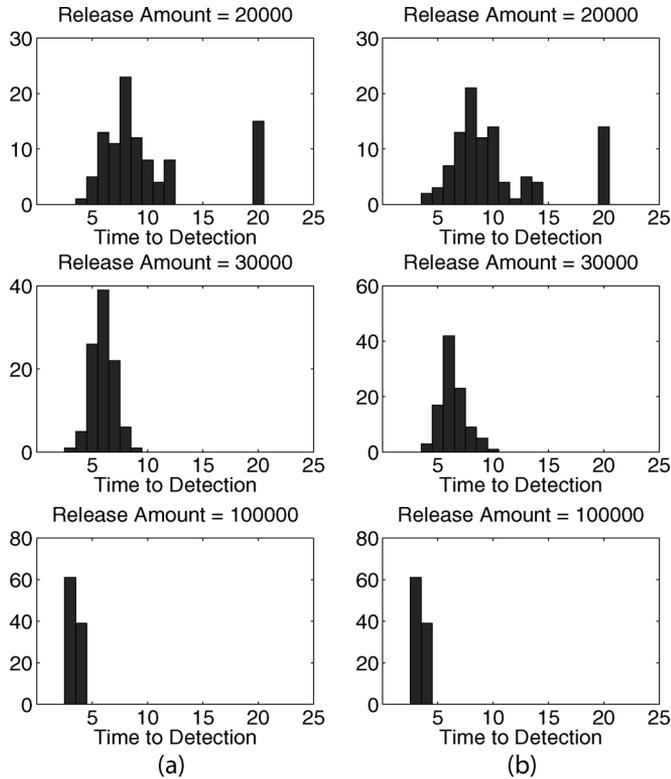


Fig. 9. Time to detection histograms for (a) four sensors and (b) 16 sensors for various release amounts. Missed detections were assigned a value of 20 s as a point of comparison between the release amounts.

This is done because after the added time of gathering information, a hypothesis at an earlier time step (and, thus, ahead in the hypothesis window) may be better than the initially detected hypothesis. The acceptable amount of time to wait is application dependent. Thus, the plots in Fig. 8 are indicative of the bounds on localization performance.

Not surprisingly, the trend indicates that larger release amounts can be localized better and regardless of release amount, 16 sensors outperform four sensors. The one and two sensor cases were excluded due to the poor localization results. For smaller release amounts, 16 sensors do not gain much by waiting to accrue additional information. However, both sensor configurations achieve significantly better localization in the case of larger release amounts. These results can be explained by the fact that the signal to noise ratio at later time steps or further from the release location is lower. Thus, the added information accrued by the outlying sensors in the 16 sensor case for small releases will be negligible.

In Fig. 9, histograms of the time from release until detection are plotted for the various release amount scenarios for four and 16 sensors. Not surprisingly, for larger release amounts, the mean and variance of the time to detection is smaller than that of small release amounts. Because the threshold in the 16 sensor case is higher than in the four sensor configuration, the time of detection distribution is skewed towards slightly longer detection times in the 16 sensor cases. There are scenarios in which it is possible for a release to be detected by the four sensor configuration and not by the 16 sensor configuration. This is purely a result of the fact that in order to maintain a  $P_F = 0.01$ , the

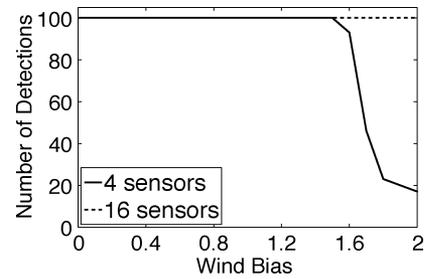


Fig. 10. Number of detections versus strength of  $x$  direction wind bias for (solid) four sensors and (dashed) 16 sensors.

likelihood ratio threshold is more conservative in the 16 sensor case and in some outlying cases this will result in lower detection though it is not the overall trend. As a point of comparison between the release amount plots, missed detections were assigned a value of 20 s.

#### D. Effect of Model Mismatch

To analyze the effects of errors in the assumed known wind field, the inference algorithm assumed a pure diffusion model (i.e., zero mean wind field) while the data was produced under various wind biases and variances (i.e., when our knowledge of the mean wind field is biased or there are statistical variations around the assumed mean.) The wind bias was always taken to be straight to the right such that it tends to transport the release between the two right-hand sensors in the four sensor configuration. The wind variance was an additive spatially and temporally white Gaussian random variation on the underlying zero mean wind field. We examine the trends for a material release of size  $1e5$ .

*a) Probability of detection:* From our experimental analysis we conclude that probability of detection is fairly robust against inaccuracies in the assumed mean wind field. Fig. 10 illustrates this for the case of an unknown wind bias. We see that the probability of detection drops off with increased wind bias, but this trend is solely a function of sensor density relative to rate of diffusion and size of wind field rather than an algorithmic deficiency. That is to say, a strong wind bias will push the cloud of material between the sensors before it has dispersed enough to be registered. Note that following the trends of Fig. 5, it is likely that lower release amounts would yield lower detection probabilities. Although no plot is given, for all simulated wind variance cases, 100% of the releases were detected.

*b) Localization performance:* Fig. 11 summarizes the localization performance for the four and 16 sensor configurations in three cases: nominal, unmodeled wind bias, and unmodeled wind variance. The plots indicate the mean and standard deviation of the  $x$  direction localization error. In the nominal and wind variance scenarios, this is nearly identical to the statistics in the  $y$  direction because the experiments are run under a zero mean wind field. In the wind bias scenario, this direction represents the principal axis of the covariance (same direction as that of the wind). The first two bars in each grouping indicate the estimates produced at the time of first detection for four and 16 sensors, respectively. The second two bars are estimates for four and 16 sensors after the system has accrued information.

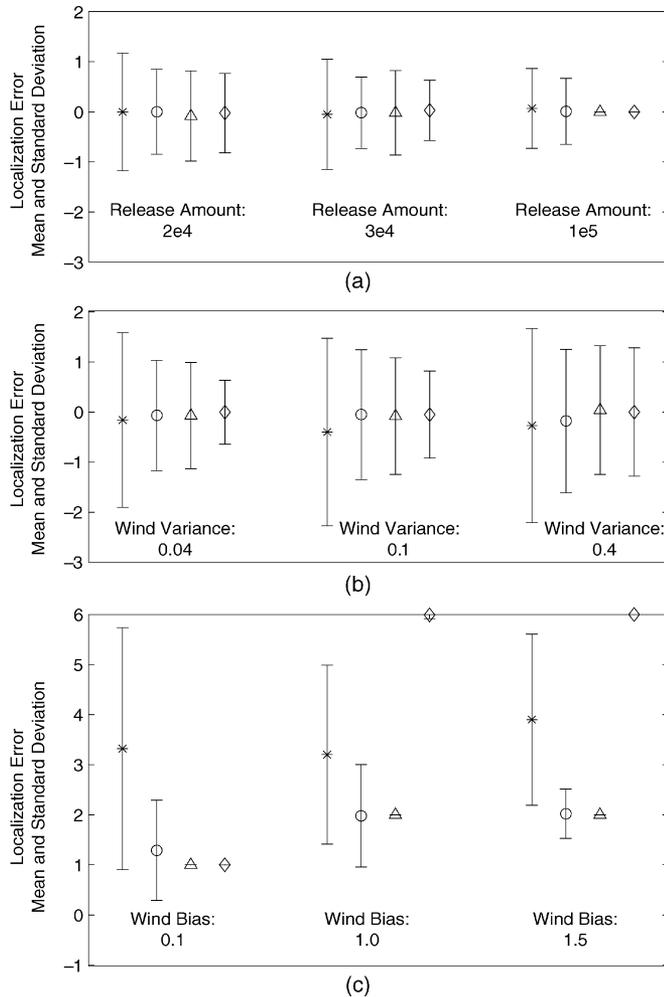


Fig. 11. Mean and standard deviation of localization error for 4 and 16 sensors at time of first detection ( $*$  and  $\circ$ ) and after accruing information ( $\Delta$  and  $\diamond$ ) for three values of (a) the release amount, (b) the unmodeled wind variance, and (c) the unmodeled wind bias.

The plots of the nominal case convey that larger release amounts can be localized better and that 16 sensors outperform four sensors. In addition, in most scenarios better localization is achieved after accruing more information. These plots can be used as a point of comparison for the model mismatch results presented below.

Localization estimates for the unmodeled wind bias case indicate that for smaller wind biases, accruing information helps localization significantly. However, for larger wind biases the 16 sensor configuration localizes the release to the edge of the room. This is because the inner four sensors only measure a small amount of the release as it passes through while the outer ring measure a much larger amount of material. The inference procedure assumes pure diffusion, so the most likely case is that the release was far from the inner four sensors. As a whole, these results highlight the point that localization performance suffers more quickly than detection performance when the wind field is mismatched to the inference procedure.

By comparing the wind variance case with the corresponding  $1e5$  release amount of the nominal case, one can see that localization performance degrades with increased unmodeled wind un-

certainty. In all wind variance cases, accruing more information provides better localization results. As with the matched model, the 16 sensor configuration outperforms the four sensor configuration. The disparity becomes less pronounced in the high variance with accrual situation. Comparing these results with those from an unmodeled wind bias, we see that the trends for randomness on the wind field better mirror those of the nominal matched case. This is because the disturbances on the propagation of particle concentrations caused by noise on the wind field are more similar to the type of errors modeled in the dynamic equation than the effects caused by a bias on the wind field.

#### IV. NONLINEAR LEAST SQUARES PARAMETER SMOOTHING

Throughout this paper, we have assumed knowledge of the mean wind field. In practice, the wind field itself presents a challenging estimation problem. We have explored empirically how performance degrades with increasing errors in the assumed mean wind field and seen that detection is robust against wind field modeling errors while localization suffers more rapidly, especially in the case of an unknown wind bias. Good detection performance can be explained by the fact that with sufficient signal to noise ratio at the nearest sensor, one sensor is adequate for detection (see Section III-B). These results motivate the work described in this section on nonlinear least squares refining of parameter estimates.

We consider the case of an unknown (and, thus, unmodeled) wind bias on the presumed mean wind field  $\mathcal{V}_w$ , where we have seen that detection is likely while localization performance is poor. In order to improve our estimate of release location, time, and amount, as well as the underlying wind field, we wait to accrue more sensor measurements until we have a window of length  $T$  containing data before and after the time of detection and then use a Levenberg–Marquardt nonlinear least squares fitting [12] of the sensor measurements to the dynamics described by (1). We use the estimate provided by the GLR approach as an initialization for the nonlinear optimization problem. See Fig. 12 for a block diagram of the overall process.

The dynamic equation is nonlinear in the release location, release time, and wind field parameters and is linear in the release concentration parameter. We use a coordinate descent approach to minimize the difference between our parameterized Gaussian puff model (1) and the data. We first estimate the nonlinear parameters  $(x_0, y_0, \bar{u}, \bar{v})$  and then use linear least squares to determine the release concentration which best fits the data given the current parametrization.

Through our modeling choice of a windowed hypothesis space and our analysis of window size, we have calculated that the effects of a release more than  $M$  time steps in the past will have negligible impact on the probability of detection of that release. Therefore, at the time of detection,  $t_d$ , we only consider hypotheses of releases occurring a maximum of  $M$  time steps in the past. For this reason, and the fact that the exact time of the release is not of utmost importance, we run  $M$  parallel nonlinear least squares estimators each parameterized by a release time  $\phi \in \{t_d - M, \dots, t_d - 1, t_d\}$ . This reduces the search space to four nonlinear parameters: release location  $(x_0, y_0)$  and wind field  $(\bar{u}, \bar{v})$ , and one linear parameter: release concentration  $\beta$ . We choose the parameter estimates associated

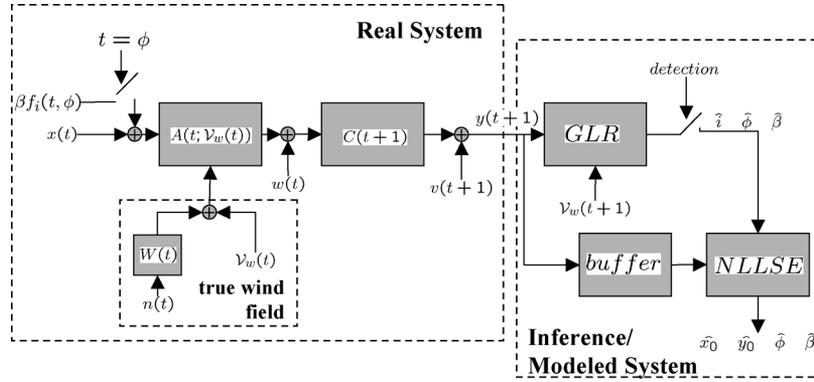


Fig. 12. Block diagram of the overall system. A release of  $\beta$  particles is inserted into the system at time  $\phi$ . The physical system is driven by the known wind field  $\mathcal{V}_w$  plus noise (not necessarily white nor Gaussian) while the inference procedure only has knowledge of  $\mathcal{V}_w$ . The GLR estimates of release time and location are used as an initialization for the nonlinear least squares estimator (NLLSE) block which takes in a buffer of sensor data in order to perform parameter smoothing.

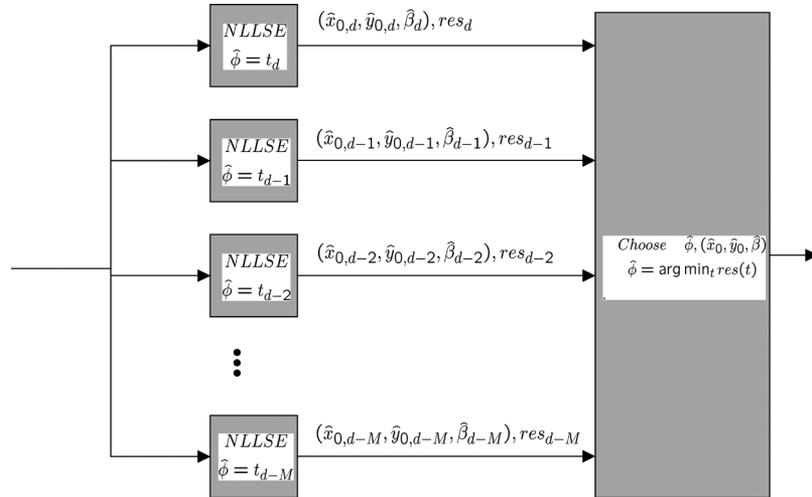


Fig. 13. Nonlinear least squares estimator (NLLSE) process is comprised of  $M$  parallel NLLSE blocks each parameterized by a release time running from  $\{t_d - M, \dots, t_d\}$ , where  $M$  is the size of the hypothesis window and  $t_d$  is the time of detection. We choose the parameters associated with the NLLSE block with minimum  $\mathcal{L}_2$  error residuals.

with the block with minimum  $\mathcal{L}_2$  error residuals. See Fig. 13 for a diagram of the *NLLSE* block in Fig. 12.

#### A. *NLLSE* Experimental Results

We now examine the performance improvement by using the nonlinear least squares stage. We first note that in the experiments we present we are fitting data simulated using a discretized Gaussian kernel and dynamic uncertainty (i.e., process noise) to the continuous dynamic model of (1). This will naturally create errors between the simulated sensor measurements and the projected noise free sensor measurements generated by providing (1) with the actual simulation parameters. The normalized magnitude of these errors is given as True:  $\mathcal{L}^2(\text{residual})/\mathcal{L}^2(\text{data})$  in Table I and  $\mathcal{L}^2(\text{data})$  is the  $\mathcal{L}^2$  energy of the data sequence and  $\mathcal{L}^2(\text{residuals})$  is that of the errors between the simulated and the projected noise free sensor measurements.

Fig. 14 shows a plot of the sensor measurements in the 16 sensor configuration for a wind bias of 0.2 in the  $x$  direction. The inference procedure assumes pure diffusion. We have seen

that in this scenario probability of detection is high. In the simulation presented in this figure, detection occurs despite the mismatched model. The projected noise free sensor measurements generated by providing the continuous dynamic model with parameter estimates of release location, time, and amount from the GLR stage are shown in Fig. 14. The figure also shows the corresponding plot for the refined NLLSE parameter estimates. Together, the plots in Fig. 14 provide a visualization of the improvement in parameter estimation obtained by using the combined standard GLR inference procedure along with the nonlinear least squares estimation stage.

The values of the parameter estimates from the GLR and NLLSE stages are shown in Table I. This table also contains the real parameters provided to the simulator which generated the synthetic sensor measurements. From the table we see that the re-estimation of the wind field reduces the error in our  $x$  direction wind bias estimate,  $\bar{u}$ , by 98% and results in a total wind velocity  $\mathcal{L}_2$  error of 0.007. Our release location estimates are also improved by using the NLLSE stage while there is negligible change in the release amount error.

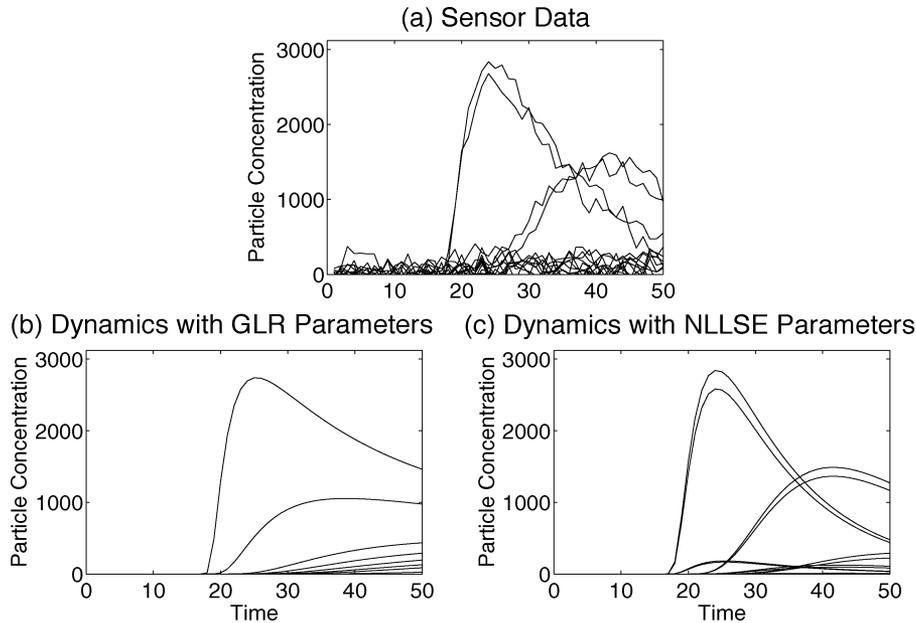


Fig. 14. (a) Sensor measurements for the 16 sensor configuration. Beneath, plots of projected noise free sensor measurements generated by providing the continuous dynamic model with (b) GLR parameter estimates and (c) NLLSE smoothed parameter estimates of release amount, time, and location as well as wind field. Data was generated with an unknown wind bias of 0.2 in the  $x$  direction while the inference procedure assumed pure diffusion.

TABLE I  
FOR AN UNMODELED WIND BIAS OF 0.2 IN THE X-DIRECTION THIS TABLE  
LISTS: 1) VALUES OF TRUE PARAMETERS, 2) PARAMETER ESTIMATES FROM  
THE GLR STAGE, AND 3) PARAMETER ESTIMATES FROM NLLSE STAGE

	True	GLR	NLLSE
$\hat{\phi}$	16	17	15
$\hat{x}_0$	13	14	12.694
$\hat{y}_0$	13	13	12.984
$\hat{\beta}$	1e5	1.17e5	0.986e5
$\hat{u}$	0.2	0	0.2004
$\hat{v}$	0	0	-0.0057
$\frac{\mathcal{L}^2(\text{residuals})}{\mathcal{L}^2(\text{data})}$	0.2169	1.0515	0.2073

A similar set of plots and data table are provided for an unknown wind bias of 0.5 in the  $x$  direction. See Fig. 15 and Table II.

## V. DISTRIBUTED PROCESSING

The detection and localization formulation described in the preceding sections suggests a distributed implementation. Under the worst case time to detection scenario of pure diffusion and a release in the center of a regular grid of sensors, we have, in Section III, contrasted the performance of the one, two, four, and 16 sensor configurations. It is clear that in order to have reasonable localization, the one and two sensor configurations are not sufficient. We have also demonstrated that there is only a marginal increase in performance by having 16 versus four sensors with the given rate of diffusion. The true benefits of 16 sensors is in better coverage of the region being monitored as well as improved localization results in the presence of an unknown wind when used in conjunction with the nonlinear least squares stage. We use these conclusions in designing a distributed processing network.

Assume that there is a large regular grid of sensors. We will also assume that in such a configuration it is infeasible to do centralized processing of the measurements because of both constrained intersensor communication and computation. We divide the region into overlapping four sensor clusters where each cluster has processing capability at a chosen sensor (see Fig. 16). In this framework, each cluster is performing the same procedure as the four sensor configuration described in the previous sections. That is to say, at every time step, each local cluster performs joint detection-localization by formulating local hypotheses and then comparing

$$l_i(t, \hat{\phi})_{\text{loc}} \triangleq \max_{i, \phi} l_i(t, \phi)_{\text{loc}} \geq \eta_{\text{loc}} \quad (15)$$

where  $l_i(t, \hat{\phi})_{\text{loc}}$  is the local hypothesis with greatest likelihood and is indicative of a release at time  $\phi$  in cell  $i$ . If  $l_i(t, \hat{\phi})_{\text{loc}}$  is greater than the local detection threshold  $\eta_{\text{loc}}$ , this cluster declares a local detection. Note that each cluster maintains hypotheses of releases within its local region and at every time step only three measurements must be transmitted to the fourth sensor with the processing capability.

This overlapping network of distributed clusters fully covers the region of interest. If a release occurs anywhere within this large region, it will occur within at least one local cluster. There are two cases to consider: when the release occurs in a region of overlapping clusters or in a region solely within a single cluster. In the case of a release in a region of overlap, there may be multiple detections (from hypotheses of releases in the overlapping clusters). In the other case, the release is local to a single cluster, but the effects the release will propagate to neighboring clusters. However, it can be shown that with high probability the detection will occur in the local cluster first and that the local likelihood statistic will be greater than that of the neighboring clusters. This is because the neighboring clusters will be attempting

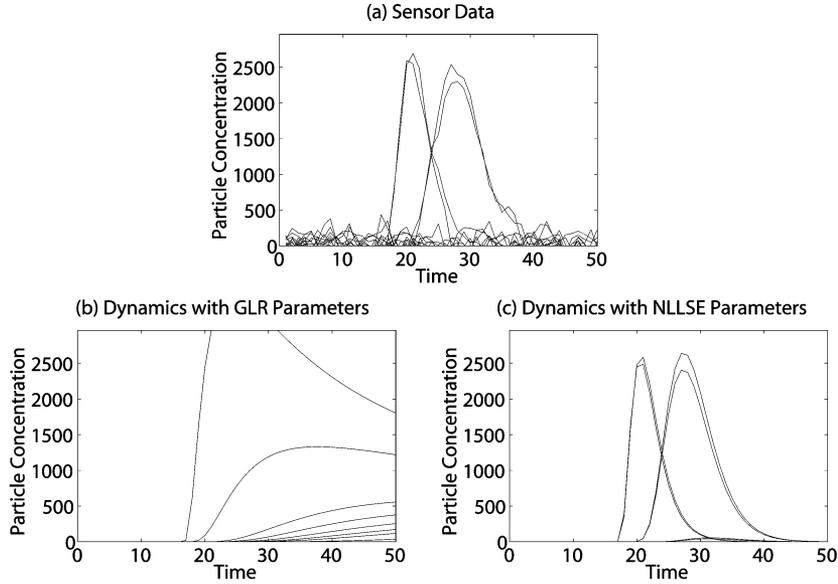


Fig. 15. Corresponding plot to Fig. 14 for an unknown wind bias of 0.5 in the x direction. (a) Sensor data; (b) dynamics with GLR parameters; (c) dynamics with NLLSE parameters.

TABLE II  
SAME AS TABLE I, BUT FOR A WIND BIAS OF 0.5

	True	GLR	NLLSE
$\hat{\phi}$	16	16	15
$\hat{x}_0$	13	14	12.114
$\hat{y}_0$	13	13	12.947
$\hat{\beta}$	1e5	1.477e5	0.954e5
$\hat{u}$	0.5	0	0.5383
$\hat{v}$	0	0	0.0119
$\frac{\mathcal{L}^2(\text{residuals})}{\mathcal{L}^2(\text{data})}$	0.3226	2.0649	0.3136

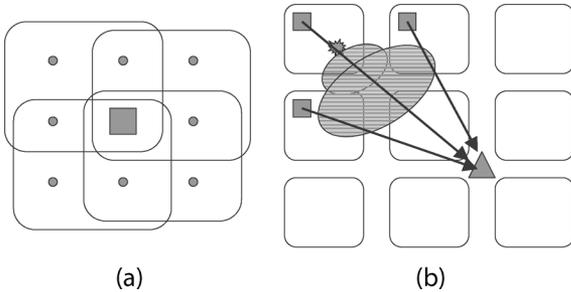


Fig. 16. (a)  $3 \times 3$  grid of sensors (circles) grouped into four clusters with one processing sensor (square). (b) A release in one cluster resulting in three local detections being transmitted to a central decision maker (triangle). Each rounded square is a cluster, separated for clarification.

to explain the release by local hypotheses which do not include the true location.

To provide a consolidated picture when detections occur in multiple clusters, it is necessary to combine information from each cluster at a centralized decision node. There are many possible fusion architectures for how information from the distributed processing clusters should be transmitted to the centralized decision maker. A simple scheme is local tripwire detection in which clusters are performing local fusion and if a local

detection occurs, the distributed processors transmit all recent sensor data to the centralized processor and the system reverts to performing inference over the entire network in order to improve localization accuracy and further filter false alarms. This results in event-driven rather than continual incurring of costly communication.

An even less communication and computationally complex approach is for only the likelihood statistic of all locally detected releases to be sent to the global decision maker, as illustrated in Fig. 16 (right). The resources consumed in transmitting this information to the centralized processor are insignificant when the probability of false alarm and the prior probability of a release are both low. The centralized decision maker could then choose the largest of the set of detected releases

$$(\hat{i}, \hat{\phi}) = \arg \max_{i, \hat{\phi}} l_i(t, \hat{\phi})_{loc}. \quad (16)$$

However, under this scheme, the centralized decision maker is making a detection and localization decision off of solely four sensors. For the rest of this section, we assume that the first architecture presented is feasible and analyze its performance relative to purely centralized processing.

We note that in order to maintain a fixed global probability of false alarm, the truly optimal local thresholds  $\eta_{loc}$  depend on the details of overlap between local clusters and, thus, may not be the same for all regions. However, we ignore these second order effects in order to infer the performance of the global system from that of the local clusters. Assume that the probability of local detection,  $P_{D,loc}$ , for a single cluster is independent from that of all other clusters. Also, assume that the probability of centralized detection,  $P_{D,c}$ , is independent of the local detections even though the decision is based off of the same data. We then have that the overall probability of false alarm is given by

$$P_D = P_{D,c} * (1 - (1 - P_{D,loc})^N) \quad (17)$$

$$P_{FA} = P_{FA,c} * (1 - (1 - P_{FA,loc})^N) \quad (18)$$

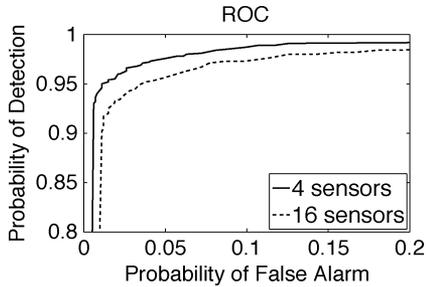


Fig. 17. Receiver operator characteristic (ROC) curve for 4 sensors (solid) and 16 sensors (dashed) with a release amount of  $2e4$ . Each data point corresponds to the probability of detection and false alarm from 1000 background simulations and 1000 release simulations for a known release amount and time and thresholds between 4.5 to 7.5 in the four-sensor case and from 6 to 9 in the 16-sensor case.

where  $N$  is the total number of local clusters. We can use the receiver operator characteristics (ROC) of the system for the local and centralized sensor cluster sizes (see Fig. 17) and (17) and (18) to help us approximate how we should set the local and centralized thresholds to maintain a specified global detection/false alarm rate.

The four sensor local and 16 sensor centralized cluster ROC curves are shown in Fig. 17 for a *known* release amount of  $2e4$  and *known* release time. We note that the apparently better performance of the four sensor configuration over that of 16 sensors can be explained by the fact that the 16 sensor configuration hypothesizes over a region of 169 cells whereas the four sensor configuration hypothesizes over a region of 25 cells. For a release of moderate size (such as  $2e4$ ) in the center of the sensor configurations under a zero mean wind field, the outer ring of 12 sensors does not increase the probability of detection significantly over that of the inner four sensors because by the time the release has diffused to the outer ring of sensors there is low signal to noise ratio. However, due to the larger number of hypotheses, the probability of false alarm increases relative to that of the four sensor scenario.

Because missed detections are very costly in this application, we can set the local cluster thresholds to produce a higher probability of detection and, thus, higher false alarm rates, and rely on the subsequent more accurate global processing to filter some of these detections by having a more conservative threshold. However, a higher probability of detection increases the rate of costly communication. The appropriate balancing of probability of detection versus rate of costly communication is application dependent and is a function of the bandwidth and computation capacity of the system, as well as the severity of missed detections.

In the centralized version, every measurement is transmitted to the central processor at every time step. The majority of the time these measurements will be of the background/benign situation which does not contain useful information. In the distributed case, we assume that local intracluster sensor transmissions are not very costly, but intercluster transmissions are costly but rare. In this case, only local measurements are transmitted to local processors at every time step and in the rare scenario of a local detection, either a single likelihood statistic or a buffer of sensor data is transmitted along the costly link to the centralized decision maker.

One can see the significant communication savings of the decentralized schemes versus the centralized version by the following back of the envelope calculation. Let the term *costly communication* refer to the communication incurred in transmitting sensor data of benign conditions (i.e., no release) to the centralized processor. In the distributed processing architecture, this only occurs when there is a local false alarm. As an example, we consider the scenario where there are 16 sensors and the region is divided into four nonoverlapping local clusters. The probability of local false alarm,  $P_{FA,loc}$ , is set to 0.1 (the centralized processor further filters these false alarms to produce a lower global rate). From the ROC plot, we see that this yields a local probability of detection,  $P_{D,loc}$  of 0.99. We aim for a global probability of false alarm,  $P_{FA}$ , of 0.01, so this requires that the centralized false alarm rate,  $P_{FA,c}$ , be set to 0.03 when there are four four-sensor clusters. The corresponding probability of detection,  $P_{D,c}$ , is 0.94

$$\begin{aligned} \text{centralized} & \begin{cases} P_D = 0.95 \\ P_{FA} = 0.01 \\ P(\text{costly communication}) = 1 \end{cases} \\ \text{distributed} & \begin{cases} P_D \cong 0.94 \\ P_{FA} \cong 0.01 \\ P(\text{costly communication}) \\ = 1 - (1 - P_{FA,loc})^N = 0.34. \end{cases} \end{aligned}$$

In the case of a larger sensor network with a multiscale hierarchical distributed processing structure, we would see even greater computational saving.

The actual choice of sensor cluster size and decision maker schema is application dependent and based on aspects such as total number of sensors, rate of diffusion, cost of local intracluster communication, cost of global intercluster communication, and assumptions about the likelihood and number of potential releases.

## VI. DISCUSSION

We have modeled the system to hypothesize one and only one release, which leads to a linear growth in the hypothesis tree rather than exponential. Even though we have simplified the hypothesis space under consideration, the formulation still allows for detection in the multiple release situation because the inference procedure will try to explain the measurements by *some* hypothesis.

However, one may be interested in a more precise detection and localization of multiple releases. If the releases are isolated in space (i.e., one release has negligible effects on the other release), then our current formulation is sufficient for localizing each release separately. Even though our hypothesis tree does not directly model a release in cell  $i$  followed by a release in cell  $j$ , if the releases are isolated then this scenario is approximately the same as the superposition of a release in cell  $i$  plus benign conditions (no release) followed by a release in cell  $j$ , which is a modeled element of our hypothesis space. All that is then required is for the decision maker architecture to declare detections of multiple hypotheses which exceed the threshold.

If the effects of the multiple releases interact, we can consider this scenario to be the same as a single spatially distributed release. Simulations we have run indicate that spatially and tem-

porally close releases are detected and localized to the center of the time-space mass of the releases. In order to improve localization estimates of these releases, a hierarchical approach could be implemented.

Although the distributed algorithm has to approximate the optimal centralized inference procedure, there are some gains to considering a distributed formulation. The distributed approach directly allows for maintaining hypotheses of multiple releases, each occurring in separate sensor clusters. How the global fusion center handles multiple local detections is simply an issue of architecture design. One could imagine a scheme where the centralized decision maker compared hypotheses of neighboring clusters but allowed multiple detections if the regions were disjoint or separated by enough distance that the effects of a release in one cluster would have dispersed enough to be similar to the background level in the other cluster.

## VII. CONCLUSION

We have presented a Bayesian state estimation approach to the problem of detection and localization of an aerosol release. Our approach allows for integration of measurements from multiple sensors over time. We have demonstrated the utility of this formulation and characterized the performance of a set of sensor configurations with regard to detection, time to detection and localization performance. We have also investigated some aspects of model mismatch due to incorrect wind field assumptions. From the experiments conducted, we see that model mismatch impacts localization performance more than detectability. This result motivates our work on estimating the true wind field and refining release parameter estimates after a detection has occurred using nonlinear least squares techniques.

The formulation presented provides a framework for answering questions such as the interaction between release amount, release location, sensor density, and sensor placement. The appropriateness of this formulation for a given application depends on the validity of the modeling assumptions we have made. These assumptions include that the mean wind field is coarsely known (nonlinear least squares techniques can be used to further estimate the true wind field) and that the dynamic model can be described as linear when conditioned on the wind field. The linearity of a diffusion plus transport dynamic model makes this framework suitable in many scenarios.

Our base algorithm integrates sensor measurements at a centralized processor. In a communication constrained sensor network, this approach may become infeasible as the number of sensors increases. By arranging sensor nodes into groups, our approach provides the basic building blocks for a distributed processing configuration. Our experimental analysis has motivated the fact that overlapping local clusters of four sensors are sufficient for detection. Only in the event of a detection, rather than at every time step as in the centralized approach, are sensor measurements transmitted along the costly intercluster communication channels. We have shown an estimation of the significant communication cost savings of such an approach.

## REFERENCES

- [1] G. Nofsinger and K. Smith, "Plume source detection using a process query system," presented at the SPIE Defense and Security Symp., Apr. 2004.

- [2] G. Nofsinger and G. Cybenko, "Distributed chemical plume process detection," presented at the Military Communications Conf., Oct. 2005.
- [3] T. Zhao and A. Nehorai, "Detection and localization of a moving biochemical source in a semi-infinite medium," presented at the IEEE SAM Workshop, 2004.
- [4] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [5] A. Willsky and H. Jones, "A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems," *IEEE Trans. Autom. Contr.*, vol. 21, no. 1, pp. 108–112, Feb. 1976.
- [6] P. Sanyal and C. Shen, "Bayes' decision rule for rapid detection and adaptive estimation scheme with space applications," *IEEE Trans. Autom. Contr.*, vol. 19, no. 3, pp. 228–231, Jun. 1974.
- [7] J. Seinfeld and S. Pandis, *Atmospheric Chemistry and Physics*. New York: Wiley, 1997.
- [8] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.
- [9] H. Van Trees, *Detection, Estimation and Modulation Theory, Part I*. New York: Wiley, 1968.
- [10] E. Fox, "Detection and localization of aerosol releases from sparse sensor measurements," M.Eng. thesis, Massachusetts Inst. Technol., Cambridge, MA, 2005.
- [11] D. Bertsekas and J. Tsitsiklis, *Introduction to Probability*. Nashua, NH: Athena Scientific, 2002.
- [12] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM J. Appl. Math.*, vol. 11, pp. 431–441, 1963.



**Emily B. Fox** (S'06) received the S.B. degree in 2004 and the M.Eng. degree in 2005 from the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge. She is currently pursuing the Ph.D. degree in the Stochastic Systems Group, MIT.

She is an NDSEG and NSF Fellowship recipient. Her research interests are in statistical signal processing, sensor networks, and nonparametric Bayesian methods.



**John W. Fisher, III** (M'90) received the Ph.D. degree in electrical and computer engineering from the University of Florida, Gainesville, in 1997.

He is currently a Principal Research Scientist in the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology (MIT), Cambridge, and the Laboratory for Information and Decision Systems, MIT. Prior to joining MIT, he was with the University of Florida, as both a faculty member and graduate student since 1987, during which time he conducted research in the areas of

ultrawideband radar for ground penetration and foliage penetration applications, radar signal processing, and automatic target recognition algorithms. His current area of research focus includes information theoretic approaches to signal processing, multimodal data fusion, machine learning, and computer vision.



**Alan S. Willsky** (S'70–M'73–SM'82–F'86) joined the Massachusetts Institute of Technology (MIT), Cambridge, in 1973.

He is the Edwin Sibley Webster Professor of Electrical Engineering at MIT. He was a Founder of Alphatech, Inc., and Chief Scientific Consultant, a role in which he continues at BAE Systems Advanced Information Technologies. From 1998 to 2002, he served on the U.S. Air Force Scientific Advisory Board. He has delivered numerous keynote addresses and is the coauthor of the text *Signals*

*and Systems* (Prentice-Hall, 1983 and 1997). His research interests are in the development and application of advanced methods of estimation and statistical signal and image processing.

Dr. Willsky has received several awards, including the 1975 American Automatic Control Council Donald P. Eckman Award, the 1979 ASCE Alfred Noble Prize, the 1980 IEEE Browder J. Thompson Memorial Award, the IEEE Control Systems Society Distinguished Member Award in 1988, the 2004 IEEE Donald G. Fink Prize Paper Award, and the Doctorat Honoris Causa from the Université de Rennes in 2005.