

Data Assimilation in Variable Dimension Dispersion Models using Particle Filters

K. V. Umamaheswara Reddy
Dept. of MAE
University at Buffalo
Buffalo, NY, U.S.A.
venkatar@buffalo.edu

Yang Cheng
Dept. of MAE
University at Buffalo
Buffalo, NY, U.S.A.
cheng3@buffalo.edu

Tarunraj Singh
Dept. of MAE
University at Buffalo
Buffalo, NY, U.S.A.
tsingh@eng.buffalo.edu

Peter D. Scott
Dept. of CSE
University at Buffalo
Buffalo, NY, U.S.A.
peter@buffalo.edu

Abstract—Data assimilation in the context of puff based dispersion models is studied. A representative two dimensional Gaussian puff atmospheric dispersion model is used for the purpose of testing and comparing several data assimilation techniques. A continuous nonlinear observation model, and a quantized probabilistic nonlinear observation model, are used to simulate the measurements. The quantized model is used to simulate bar sensor readings of the concentration. Dispersion models usually lead to high dimensional space-gridded state space models. In the case of puff based dispersion models, this may be avoided by using puff parameters themselves as the states, but at the cost on nonlinearity and variable dimensionality. The potential of sampling based techniques is discussed in this context, with a particular focus on the Particle Filter approach, for which variable state dimensionality creates no difficulties. The performance of Particle Filter is compared with that of the Extended Kalman Filter, and its advantages and limitations are illustrated.

Keywords: Chem-Bio Dispersion, Gaussian Puff, Particle Filter, Bar Sensor, Variable State Dimensionality.

I. INTRODUCTION

Real-time detection, tracking, and prediction of chemical and biological releases is important for fast response to chemical and biological accidents and attacks. Knowledge about the releases is obtained by fusing data from an array of deployed and possibly mobile sensors and an atmospheric dispersion model. The use of atmospheric dispersion model along with the meteorological model is essential because of the very limited coverage of the sensors in time and space. Gaussian puff based models [1] are often used to make fast release concentration prediction, in which a series of Gaussian shaped puffs are released at the sources and propagated in the atmosphere. Sensors used in dispersion applications usually have imprecise bar readings. The bar readings provide limited information about the true concentration and increase the nonlinearity of the measurement process.

Data assimilation is the science and art of fusing the observations with the model predictions, to get a better estimate. Real-time and operational data assimilation for atmospheric release dispersion, need to deal with two main problems: high nonlinearity and high state dimensionality. Nonlinearities in the dispersion model and the observation model pose significant challenges for data assimilation techniques. Further, due to continuous release/emissions initially and potential

subsequent puff splitting, the number of puffs and therefore the length of the state vector is not constant, but increases with time. The state dimension may become so high that estimating all states from the sensor data is impossible.

The Extended Kalman Filter (EKF) [2] is one of the most widely used nonlinear filters. However, the EKF poses certain problems for the dispersion applications because of its low-order linearization approximation, computationally expensive Jacobian calculations and propagation and update of large state error covariance matrix, which is a memory intensive process. Unscented Kalman Filter [3] uses deterministic sampling techniques for better approximation in nonlinear filtering applications. Reduced-rank filters, which employ reduced-rank approximation to the full-rank covariance matrix, hold good potential for data assimilation in large scale models [4]. The Ensemble Kalman Filters (EnKF) [5], [6] approximate the error covariance matrix by the ensemble covariance around the ensemble mean. Each ensemble member is propagated through the nonlinear dynamics directly and the ensemble mean and covariance computed from the ensemble provides the necessary statistics used in the measurement update step. Because the ensemble size is much smaller than the state dimension, the computational complexity with the EnKF is significantly reduced than that of the EKF. It is also assumed that a relatively small number of ensemble members is adequate to capture the main characteristics of the release dispersion in the dominant subspace, which is of a lower dimension than the whole state space. However, the EnKF cannot be used with the puff based models due to the inherent variable state dimensionality in these models. At a measurement time, the ensemble members, the puffs of which may be emitted at different source locations (to account for source location uncertainty) and may have undergone different advection and splitting processes, may have different dimensions. For example, at a time instant one ensemble member may consist of ten puffs while another may consist of fifteen. Even if all the ensemble members have the same number of puffs, the correspondence among the puffs of different ensemble members is often unavailable. As a result, the ensemble mean and covariance as well as the cross correlation matrix between concentration and puffs cannot be computed and the measurement update cannot be carried out.

Particle Filters [7] provide a good alternative for such non-

standard dynamical models. Particle Filters provide estimates of the higher moments, and are mostly suitable for highly nonlinear and non-Gaussian models. The variable and high dimensionality of the state vector, which poses a problem to the standard nonlinear assimilation techniques, can be dealt with using Particle Filters. Daum, et al. [8] showed that a carefully designed Particle Filter should mitigate the curse of dimensionality for certain filtering problems. The efficiency of Particle Filter and Ensemble based filters, for data assimilation in a high dimensional linear diffusion model, is discussed in our earlier work [9]. Further, the Monte Carlo nature of Particle Filtering techniques provides the potential for predicting Multiple Hypotheses scenarios in case of large uncertainty, typical of such large models. Multiple Hypotheses are crucial in case of lethal releases and when significant uncertainties exist in the source and the atmospheric data.

In this paper, a representative nonlinear puff based dispersion model is designed for the purpose of data assimilation. This model incorporates the characteristic feature of rapid growth in state dimensionality with time. The performance of Particle Filter is studied, using two different nonlinear sensor models, and the various advantages of sampling based techniques are discussed. The representative dispersion model is introduced and its dynamics are described in Section II, and the two sensor models are described in Section III. The state space description is introduced in Section IV and the various data assimilation techniques used, are described in Section V. The results of data assimilation on the representative dispersion model, using the two sensor models are discussed in Section VI. The conclusions and further work are discussed in Section VII.

II. DISPERSION MODEL

The atmospheric dispersion model used is based on the RIMPUFF [1] (Riso Mesoscale PUFF) model which was designed to calculate the concentration and doses resulting from the dispersion of airborne particles. It is a Lagrangian mesoscale atmospheric dispersion puff model, which applies both to homogeneous and inhomogeneous terrain with moderate topography on a horizontal scale of up to 50 km, and responds to changing (non-stationary) meteorological conditions [1]. The model simulates the time changing release (emission) of airborne materials by sequentially releasing a series of Gaussian shaped puffs at a fixed rate on a specified grid. The amount of airborne materials allocated to individual puffs equals the release rate times the time elapsed between puff releases. At each time step, the model advects, diffuses and deposits the individual puffs according to local meteorological parameter values. This model is used as a basis for the present dispersion model, which can be used as a benchmark model for testing various data assimilation techniques. Also, to reduce the computational complexity of the simulations, the present model is restricted to the two dimensional horizontal spread of the plume, the vertical effects on dispersion being ignored.

A. Gaussian Puff Characteristics

The concentration distribution in each puff is Gaussian in the two dimensional space. The mean of the Gaussian represents the location of the puff center, and the standard deviations represent the size of the puff in various directions. For simplicity, the Gaussian is assumed circular in our case. The standard deviation σ_x in the downwind direction is used as a mathematical tool and is made equal to the standard deviation σ_y in the crosswind direction [1]. The standard deviation of this circular Gaussian is denoted σ_{xy} . This is illustrated in figure 1.

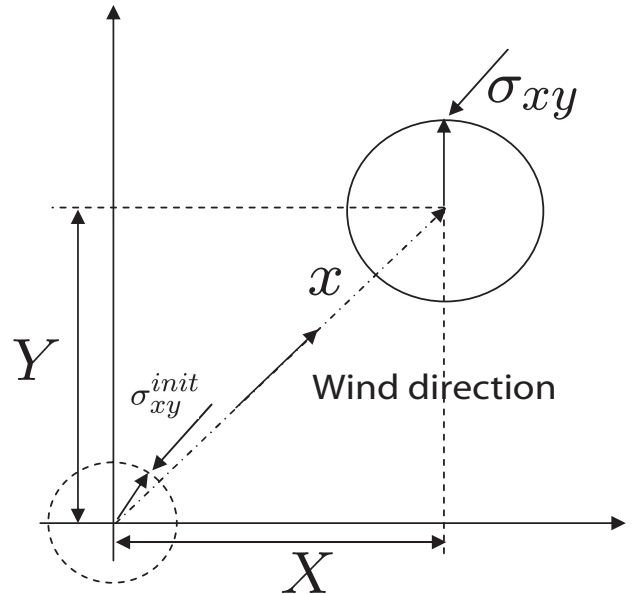


Fig. 1. Model Illustration

Each Gaussian puff has four parameters: $[X, Y, \sigma_{xy}, Q]$, where

- $[X, Y]$ = Centroid of the Gaussian puff
- σ_{xy} = Puff size (std. deviation)
- Q = Activity (mass) of the puff

B. Concentration Calculation

The concentration at a grid point $[x_g, y_g]$, at each time step, is calculated by summing the contributions of all the puffs at that instant.

$$C = \sum_{i=1}^N \frac{Q_i}{2\pi\sigma_{xy(i)}^2} \exp \left[-\frac{1}{2} \left(\frac{(X_i - x_g)^2 + (Y_i - y_g)^2}{\sigma_{xy(i)}^2} \right) \right] \quad (1)$$

where, N is the number of puffs.

Height (m)	Stability category	Diffusion coefficients			
		p_y	q_y	p_z	q_z
50	A	1.503	0.833	0.151	1.219
	B	0.876	0.823	0.127	1.108
	C	0.659	0.807	0.165	0.996
	D	0.640	0.784	0.215	0.885
	E	0.801	0.754	0.264	0.774
	F	1.294	0.718	0.241	0.662
100	A	0.179	1.296	0.051	1.317
	B	0.324	1.025	0.07	1.151
	C	0.466	0.866	0.137	0.985
	D	0.504	0.818	0.265	0.818
	E	0.411	0.882	0.487	0.652
	F	0.253	1.057	0.717	0.486
180	A	0.671	0.903	0.025	1.5
	B	0.415	0.903	0.033	1.32
	C	0.232	0.903	0.104	0.997
	D	0.208	0.903	0.307	0.734
	E	0.345	0.903	0.546	0.557

TABLE I
KARLSRUHE-JÜLICH DIFFUSION [1]

C. Puff Dynamics

The advection and diffusion of each puff takes place according to local meteorological parameter values. The advection of each puff is calculated according to the wind vector, \mathbf{u} at the puff center and the time step, ΔT is used to determine the next position of the puff center.

$$X_{k+1} = X_k + u_X \Delta T \quad (2)$$

$$Y_{k+1} = Y_k + u_Y \Delta T \quad (3)$$

$$\text{where, } \mathbf{u} = [u_X, u_Y]^T \quad (4)$$

Expansion with time of a single puff is fundamentally related to the relative diffusion process. It is computed from simultaneous measurements or specifications of the atmospheric turbulence intensity and/or stability in the dispersion area. For this model, standard plume dispersion information is used. Pasquill parameterisation, using a modified Karlsruhe-Jülich system, (table I) is used as a basis for the present model.

This parameterization is valid for limited cases of near ground level releases and dispersion over flat terrains. The dispersion categories are based on surface wind speed measurements, incoming insolation or cloud cover (see table II). The growth of each puff is described by

$$\sigma_{xy} = p_y x^{q_y} \quad (5)$$

where, x = downwind distance given by,

$$x_{k+1} = x_k + \sqrt{u_X^2 + u_Y^2} \Delta T$$

and p_y, q_y = stability dependent parameters

(Karlsruhe-Jülich diffusion coefficients)

Surface (10 m)	Daytime			Night-time	
wind speed	Incoming solar radiation			Cloudiness	
(m s ⁻¹)	Strong	Moderate	Slight	≥4/8	≤3/8
<2	A	A-B	B	-	-
2-3	A-B	B	C	E	F
3-5	B	B-C	C	D	E
5-6	C	C-D	D	D	D
>6	C	D	D	D	D
A : Extremely unstable conditions			B : Moderately unstable conditions		
C : Slightly unstable conditions			D : Neutral conditions		
E : Slightly stable conditions			F : Stable conditions		

TABLE II
PASQUILL STABILITY CATEGORIES [1]

Deposition is neglected for the present model and hence the mass of each puff remains constant during the dispersion process.

1) *Puff Splitting*: A puff splitting scheme is usually employed in the puff based dispersion models like RIMPUFF, to account for dispersion over complex terrains which might result in plume splitting. This scheme is included in our model to represent a similar structure for testing purpose. When an initially small puff is grown to the size comparable with the grid spacing of the flow model, the original puff is replaced with five new smaller puffs, under the following constraints [1]:

- 1) $\sum_{i=1}^5 Q_i = Q$
- 2) $\sum_{i=1}^5 Q_i (d_i^2 + \sigma_{xy(i)}^2) = Q \sigma_{xy}^2$
- 3) $\sum_{i=1}^5 C_{i(X,Y)} = C_{(X,Y)}$ for $i = 1$ to 5
- 4) $\sigma_{xy(i)} = \frac{1}{2} \sigma_{xy}$ (arbitrarily assigned)

The solution corresponding to these constraints is illustrated in figure 2. After splitting, the new puffs can set off in individual directions in response to the information of divergence now explicitly contained in the flow field.

The model described in this section, is used as a representative atmospheric dispersion model for demonstrating the data assimilation methodologies.

III. SENSOR MODELS

Two observation models are used for studying data assimilation performance. The first model is a continuous model, which measures concentration according to equation 1, and gives a noisy output with a certain variance R at each sensor location.

The second model is based on a model for a simple ion mobility sensor described in [10]. The outputs of the sensor are discrete numbers of bar readings ranging from 0 to 7. These bar readings indicate concentration magnitude [10]; the sensor displays $i = 0, \dots, 7$, when the “measured” continuous-valued concentration magnitude c_v is between thresholds T_i and T_{i+1} . As in [10], the thresholds are assumed to be exactly known. The measured concentration c_v is assumed

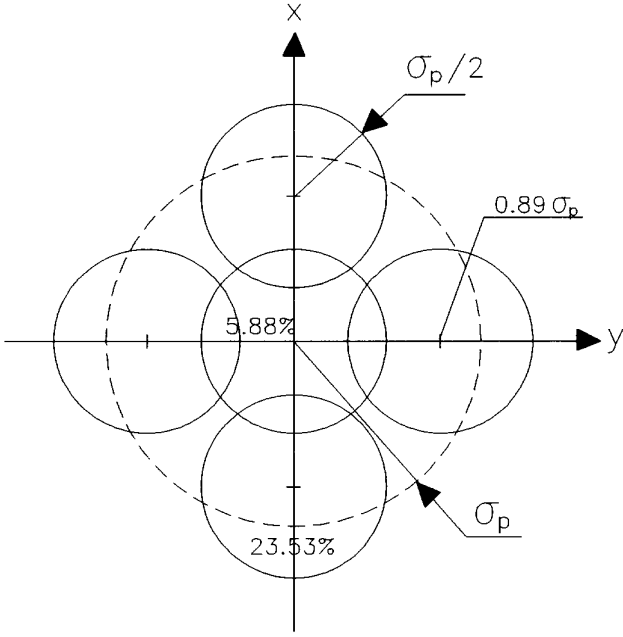


Fig. 2. Puff-Splitting Illustration [1]

normally distributed about the true concentration c obtained using equation 1. The variance R of the measurement error $c_v - c$ is given by

$$R = \alpha c + J \quad (6)$$

where α is the constant of proportionality and J accounts for the thermal motion of the electrons in the components.

The probability of a reading i given the true concentration c is then given by the following integral

$$P(i|c) = \frac{1}{\sqrt{2\pi R}} \int_{T_i}^{T_{i+1}} e^{-\frac{(c_v - c)^2}{2R}} dc_v \quad (7)$$

This probability is used in the update of the important weights of a Particle Filter.

IV. STATE SPACE MODEL DESCRIPTION

A general nonlinear stochastic state space description of the discrete time model and the measurements with additive zero mean noise, can be shown as below:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k) \quad (8)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \quad (9)$$

When such a stochastic description of model and measurements is available, it is possible to incorporate the measurements into the model to obtain an optimal estimate of the system theoretically. The standard assignment of states for the dispersion problem uses concentrations measured on a spatial grid. Using a puff model, the opportunity exists to dramatically reduce the dimensionality of the resulting state space. Here, the atmospheric dispersion model described in the previous section is interpreted as a state space model as follows:

Define a state vector for each Gaussian puff as $\mathbf{z}^i = [X^i \ Y^i \ \sigma_{xy}^i \ Q^i]^T$, which captures all the information of the i^{th} Gaussian. For the dynamical system, the state vector is given by \vec{Z} , which is a vector of all \mathbf{z}^i s, concatenated into a single vector. The state space model of the discrete time atmospheric dispersion model is then given by

$$\vec{Z}(k+1) = F(\vec{Z}(k), \vec{u}(k), \vec{w}(k)) \quad (10)$$

$\vec{u}(k)$, being the wind velocity vector at the instant k . The output of the dynamical system is given by the concentrations calculated, at various grid points of the dispersion region, according to equation 1.

V. ESTIMATION TECHNIQUES

Nonlinearities in the system dynamics model and the quantized (bar readings) nonlinear observation models lead to non-Gaussian state distributions. These factors pose significant challenges for data assimilation techniques. Further, due to continuous release/emissions initially and potential subsequent puff splitting, the length of the state vector is not constant, but varies with time. Also, with time, the state dimension becomes very high leading to difficulties in the full state estimation.

A. Extended Kalman Filter

The Extended Kalman Filter (EKF) is one of the most widely used nonlinear filters, and it involves model linearization around the most recent estimate. Varying state dimension necessitates block operations on covariance matrices at each time step.

At each step, if there is a new puff release, its covariance is initialized with \mathbf{P}_0 . The propagation equations for the puffs, when there is no splitting, are as follows:

$$\hat{\mathbf{x}}_{0|0} = \mathbf{x}_0 \quad (11)$$

$$\mathbf{P}_{0|0} = \mathbf{P}_0 \quad (12)$$

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{f}(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k) \quad (13)$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}_k \mathbf{P}_{k|k} \mathbf{F}_k^T + \mathbf{Q}_k \quad (14)$$

where, $F = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}$. When there is puff splitting, the state mean and covariance propagation occurs as above, and they change according to the following equations due to splitting:

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{A} \hat{\mathbf{x}}_{k+1|k,BS} \quad (15)$$

$$\mathbf{P}_{k+1|k} = \mathbf{A} \mathbf{P}_{k+1|k,BS} \mathbf{A}^T \quad (16)$$

where \mathbf{A} is the linear operator corresponding to the puff splitting process described in section II-C.1, $\hat{\mathbf{x}}_{k+1|k,BS}$ and $\mathbf{P}_{k+1|k,BS}$ are the propagated state mean and the error covariance before puff splitting.

The propagated state mean and covariances for all the puffs are then concatenated appropriately and are updated with the sensor measurements as follows:

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \mathbf{h}(\hat{\mathbf{x}}_{k+1|k})) \quad (17)$$

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T (\mathbf{H}_{k+1} \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T + R_{k+1})^{-1} \quad (18)$$

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{K}_{k+1} \mathbf{H}_{k+1} \mathbf{P}_{k+1|k} \quad (19)$$

where, $H = \frac{\partial \mathbf{h}}{\partial \mathbf{x}}$.

The implementation of EKF requires Jacobian calculation, which is an expensive computation. The state error covariance matrix rapidly increases in size due to increase in the state dimension with time. Further, a quantized nonlinear observation model, such as the bar sensor model, poses continuity problems during the measurement update in the EKF.

Particle Filters provide a good alternative for such non-standard dynamical and observation models. The Bootstrap Particle Filter is implemented on the representative atmospheric dispersion model using the two sensor models, discussed in the earlier section. The implementation of the Particle Filter, as applicable to our present models, is outlined in the following:

B. Particle Filter

In Particle Filters, the posterior distribution $p(\mathbf{x}_k | \mathbf{Y}_k)$ is approximated with N weighted particles $\{\mathbf{x}_k^{(i)}, w_k^{(i)}\}_{i=1}^N$, given by

$$P_N(d\mathbf{x}_k | \mathbf{Y}_k) \approx \sum_{i=1}^N w_k^{(i)} \delta_{\mathbf{x}_k^{(i)}}(d\mathbf{x}_k) \quad (20)$$

where $\mathbf{x}_k^{(i)}$ are the particles drawn from the importance function or proposal distribution, $w_k^{(i)}$ are the normalized importance weights, satisfying $\sum_{i=1}^N w_k^{(i)} = 1$, and $\delta_{\mathbf{x}_k^{(i)}}(d\mathbf{x}_k)$ denotes the Dirac-delta mass located in $\mathbf{x}_k^{(i)}$. We use \mathbf{X}_k and \mathbf{Y}_k to denote the state trajectory $\{\mathbf{x}_j\}_{j=0}^k$ and measurement history $\{\mathbf{y}_j\}_{j=1}^k$, respectively. The expectation of a known function $\mathbf{f}(\mathbf{x}_k)$ with respect to $p(\mathbf{x}_k | \mathbf{Y}_k)$ is then approximated by

$$\int \mathbf{f}(\mathbf{x}_k) p(\mathbf{x}_k) d\mathbf{x}_k \approx \sum_{i=1}^N w_k^{(i)} \mathbf{f}(\mathbf{x}_k^{(i)}) \quad (21)$$

For example, the approximation to the arithmetic mean of \mathbf{x}_k is $\sum_{i=1}^N w_k^{(i)} \mathbf{x}_k^{(i)}$.

Particle Filter updates the particle representation $\{\mathbf{x}_k^{(i)}, w_k^{(i)}\}_{i=1}^N$ in a recursive manner. A cycle of a generic Particle Filter includes [7]

- Sequential Importance Sampling
 - For $i = 1, \dots, N$, sample $\mathbf{x}_{k+1}^{(i)}$ from the importance function $q(\mathbf{x}_{k+1} | \mathbf{X}_k^{(i)}, \mathbf{Y}_{k+1})$
 - For $i = 1, \dots, N$, evaluate and normalize the importance weights

$$w_{k+1}^{(i)} \propto w_k^{(i)} \frac{p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}^{(i)}) p(\mathbf{x}_{k+1}^{(i)} | \mathbf{x}_k^{(i)})}{q(\mathbf{x}_{k+1}^{(i)} | \mathbf{X}_k^{(i)}, \mathbf{Y}_{k+1})} \quad (22)$$

- Resampling: Multiply/Discard particles $\{\mathbf{x}_{k+1}^{(i)}\}_{i=1}^N$ with respect to high/low importance weights $w_{k+1}^{(i)}$ to obtain N new particles $\{\mathbf{x}_{k+1}^{(i)}\}_{i=1}^N$ with equal weights.

It should be noted that the computation of the mean and covariance is not required for the process of the Particle Filter.

The importance function plays a significant role in the particle filter. One of the simplest importance function is given

by $q(\mathbf{x}_{k+1} | \mathbf{X}_k^{(i)}, \mathbf{Y}_{k+1}) = p(\mathbf{x}_{k+1} | \mathbf{x}_k^{(i)})$. This implementation is called the Bootstrap Particle Filter. The corresponding importance weights are $w_{k+1}^{(i)} \propto w_k^{(i)} p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}^{(i)})$. Sampling $\mathbf{x}_{k+1}^{(i)}$ from $p(\mathbf{x}_{k+1} | \mathbf{x}_k^{(i)})$ is equivalent to the dynamic propagation of $\mathbf{x}_k^{(i)}$ to time t_{k+1} .

The optimal importance function that minimizes the variance of the importance weight $w_{k+1}^{(i)}$ conditional upon $\mathbf{x}_k^{(i)}$ and \mathbf{y}_{k+1} is given by $q(\mathbf{x}_{k+1} | \mathbf{X}_k^{(i)}, \mathbf{Y}_k) = p(\mathbf{x}_{k+1} | \mathbf{x}_k^{(i)}, \mathbf{y}_{k+1})$. The corresponding importance weights are $w_{k+1}^{(i)} \propto w_k^{(i)} p(\mathbf{y}_{k+1} | \mathbf{x}_k^{(i)})$. The optimal Particle Filter, using the optimal importance function, gives the limiting performance of Particle Filters. Note that, when the state model and the observation model are linear Gaussian, all the above-mentioned probability density functions are Gaussian. In this case, the optimal Particle Filter reduces to the Kalman filter.

The measurement likelihood, in the Particle Filter, for the continuous observation model, is given by:

$$P(c_v | c) = \frac{1}{\sqrt{2\pi R}} e^{-\frac{(c_v - c)^2}{2R}} \quad (23)$$

For the case of bar sensor model, the likelihood is calculated using the following equation:

$$P(i | c) = \frac{1}{\sqrt{2\pi R}} \int_{T_i}^{T_{i+1}} e^{-\frac{(c_v - c)^2}{2R}} dc_v \quad (24)$$

where c is the predicted concentration at time $k+1$, and c_v is obtained by sampling from a Gaussian distribution with mean c , and the measurement noise variance R .

Particle Filter does not involve the calculation of Jacobians, the storage of large state error covariance matrices and associated block operations due to varying state dimensions. Each of the particles, in general, may have different state dimensions. Each particle corresponds to a particular distribution of concentration, whose mean concentration can be calculated for a single point estimate. Thus, Particle Filters provide a unified approach to the case of variable state dimensions.

VI. RESULTS AND DISCUSSION

The performance of the estimation techniques, is studied for the discussed models using simulated atmospheric dispersion and concentration measurements. The grid resolution over the dispersion region is $200m$ and sensors are located every $1km$ along both the axes as shown in figure 3.

The total simulation runtime is $3600s$ and sampling interval is $20s$. The source location and strength are uncertain. The release occurs for 20% of the simulation time, with a 10% standard deviation. This is implemented as a fraction of the total number of time steps and is modeled as a random variable with mean 0.2 and a standard deviation of 0.02. There is a puff release every three time steps. The nominal value of the wind speed is $5m/s$, with a standard deviation of 20%. The wind direction changes from 15 deg to 60 deg after $1800s$. The truth is simulated using a perturbed model, taking all these uncertainties into account. The observations are obtained from this truth, using the two sensor models described in section III.

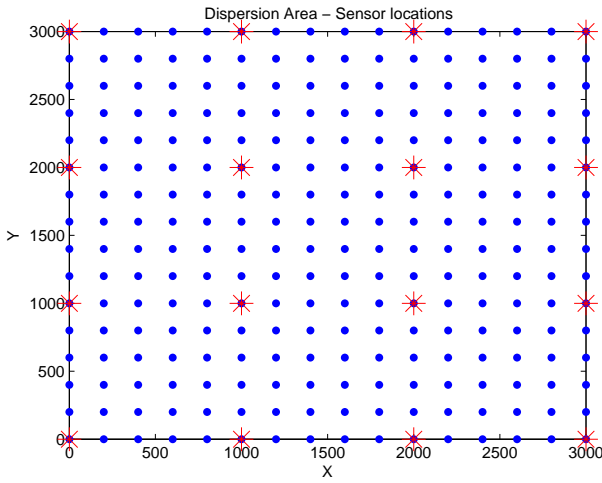


Fig. 3. Sensor locations on the grid

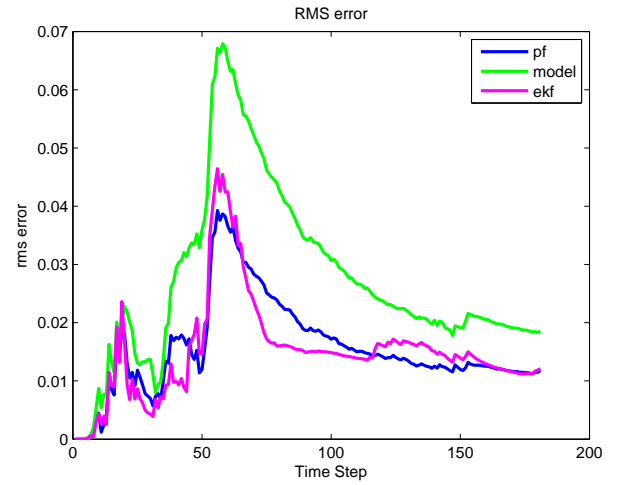


Fig. 5. RMS error in the concentrations for continuous sensor

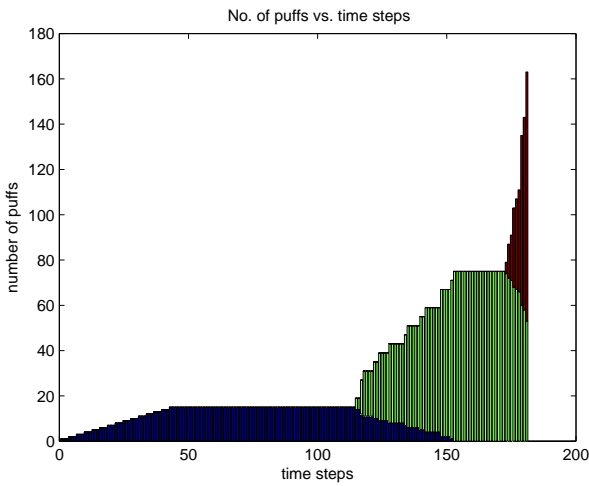


Fig. 4. Number of puffs

The variable state dimensionality is illustrated in figure 4, where the number of puffs is plotted against the time step, for the truth. The initial increase in the number of puffs due to source release can be seen till the time step of 42. There is again a sharp increase at two different later stages due to puff splitting.

For the continuous observation model, the variance at each sensor location is chosen to be 0.001. The nominal model is used for the purpose of testing the estimation algorithms using these observations. The Extended Kalman Filter and the Bootstrap Particle Filter are implemented on the representative 2D atmospheric dispersion model. The number of particles used in the Particle Filter, is 30. The results are averaged over 100 Monte Carlo simulations.

For error calculations, a grid domain of $15 \times 15 \text{ km}^2$ is used. In figure 5, the root mean squared error in concentrations across all grid points in the domain is calculated and is plotted

vs. time, for the Particle Filter as well as for the nominal model without data assimilation. The decrease in the rms error for the nominal model without data assimilation, can be explained by the dispersion of the plume over a larger area with time. The implementation of the EKF and the Particle Filter improves the estimates with time as more and more observations are assimilated. It can be seen that the Particle Filter performance is comparable to that of the EKF in the case of continuous observation model.

For the bar sensor model, the values of the parameters α and J are chosen to be 0.001 and 10^{-16} respectively. The probabilities of various bar readings for these parameters can be seen in figure 6. While the leftmost mode corresponds to a reading of 0 bar, the rightmost mode corresponds to a reading of 7. The bar sensor readings are simulated from the truth, using these parameters.

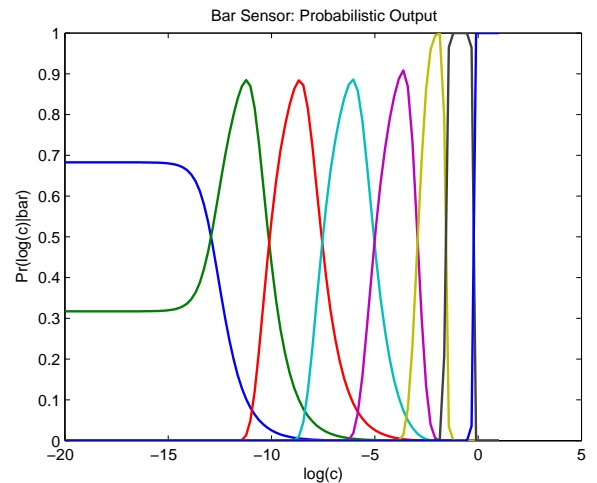


Fig. 6. Probabilistic Bar Sensor likelihood

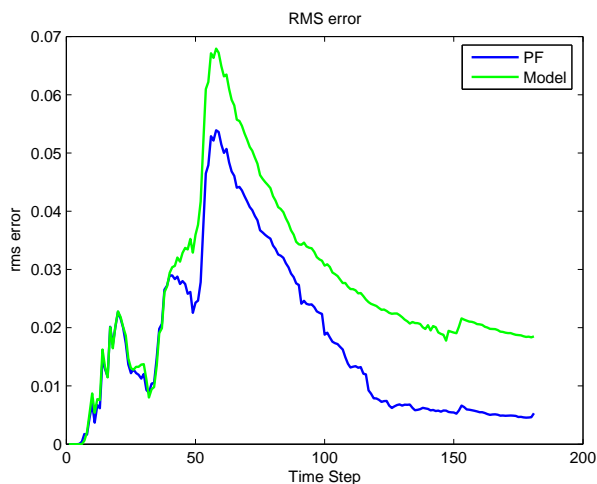


Fig. 7. RMS error in the concentrations for bar sensor

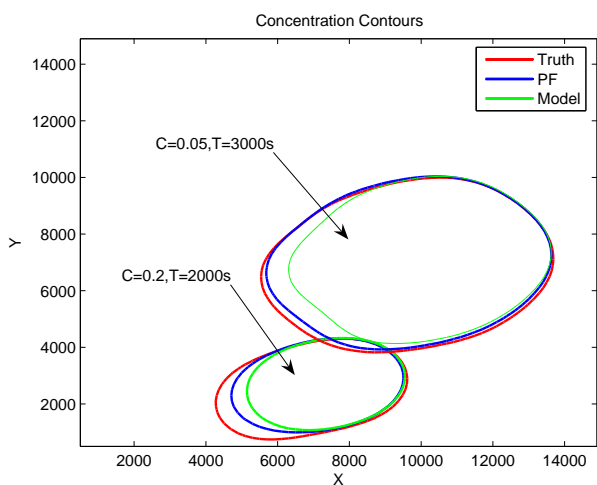


Fig. 8. Concentration Contours

Implementation of the EKF in the case of bar sensor model faces the problems of continuity. Further, calculation of the Jacobian matrix corresponding to the probabilistic bar sensor model is not meaningful, because there is no change in the bar reading for a small change in the local concentration. Linearization is not a valid approach for such observation models. The Particle Filter does not face this problem and its implementation, as can be seen from figure 7, improves the estimates with time as more observations are assimilated. The same can also be observed in figure 8, where the concentration contours are plotted at two different time instants, for different levels. It can be seen that the estimates after assimilation by the Particle Filter, compare favorably with the truth.

VII. CONCLUSION

A test bed is designed to assess the performance of various data assimilation techniques for puff based dispersion models. A two dimensional Gaussian puff based dispersion model is

used as a representative model for this purpose. The splitting of puffs which is a characteristic of more complex models is incorporated in the representative model. Two kinds of observation models are used to demonstrate the potential of sampling based filters. A continuous observation model is used to test the performance of Particle Filter against the widely used Extended Kalman filter. It was seen that while both the filters are comparable in terms of estimation performance, Particle Filter offers significant advantages in terms of ease of implementation and memory requirements. Further, Particle Filter can be effectively applied for a quantized observation model, as demonstrated using the probabilistic bar sensor model. The EKF, and other linearization based filters, fail with such sensor models. The application potential of Particle Filter, in the context of variable state dimensionality, high state dimension, non-linearities and quantized observations, is thus demonstrated.

The possibilities of incorporating the advantages of other sampling based filters will be explored in the future. The deterministic sampling techniques of the Unscented Filters and the reduced rank representation used in the Ensemble Kalman Filtering techniques hold good potential, towards improving the estimation performance while managing the computational requirements.

ACKNOWLEDGMENT

This work was supported by the Defense Threat Reduction Agency (DTRA) under Contract No. W911NF-06-C-0162. The authors gratefully acknowledge the support and constructive suggestions of Dr. John Hannan of DTRA.

REFERENCES

- [1] S. Thykier-Nielsen, S.Deme, and T. Mikkelsen, "Description of the atmospheric dispersion module rimpuff," Riso National Laboratory, P.O.Box 49, DK-4000 Roskilde, Denmark, Tech. Rep. RODOS(WG2)-TN(98)-02, 1999.
- [2] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. San Diego, CA: Academic Press, 1970, ch. 4-7.
- [3] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of the American Control Conference*, Seattle, WA, June 1995, pp. 1628-1632.
- [4] J. Lewis, S. Lakshminarayanan, and S. Dhall, *Dynamic Data Assimilation: A Least Squares Approach*. Cambridge, UK: Cambridge University Press, 2006.
- [5] G. Evensen, "The ensemble kalman filter: Theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, pp. 343-367, 2003.
- [6] —, "Sampling strategies and square root analysis schemes for the enf," *Ocean Dynamics*, vol. 54, pp. 539-560, 2004.
- [7] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Boston, MA: Artech House, 2004, ch. 3.
- [8] F. Daum and J. Huang, "Curse of dimensionality and particle filters," in *Proceedings of the IEEE Aerospace Conference*, vol. 4, Big Sky, MT, March 2003, pp. 1979-1993.
- [9] K. V. U. Reddy, Y. Cheng, T. Singh, and P. D. Scott, "Data assimilation for dispersion models," in *the Ninth International Conference on Information Fusion*, Florence, Italy, July 2006.
- [10] P. Robins, V. Rapley, and P. Thomas, "A probabilistic chemical sensor model for data fusion," in *the Eighth International Conference on Information Fusion*, Philadelphia, PA, August 2005.