H14-211

AUTOMATED SOURCE PARAMETER AND LOW LEVEL WIND ESTIMATION FOR ATMOSPHERIC TRANSPORT AND DISPERSION APPLICATIONS

Francois Vandenberghe[1], Paul E. Bieringer[1], Ian Sykes[2], John Hannan[3], Jonathan Hurst[1], George Bieberbach[1], Steve Parker[3], and Luna Rodriguez[1]

[1]National Center for Atmospheric Research, Boulder, CO, USA
[2]Sage Management, Princeton, NJ, USA
[3]Defense Threat Reduction Agency, VA, US

Abstract: Accurate simulations of the atmospheric transport and dispersion rely heavily on the source term parameters necessary to characterize the initial release and meteorological conditions that drive the downwind dispersion. The source parameters are in many cases not known and consequently based on rudimentary assumptions. This is particularly true of accidental releases, and the intentional releases associated with terrorist incidents. Often the available meteorological observations are not representative of the conditions at the location of the release, which can result in significant errors in the hazard assessments downwind of the sensors even when the source parameters are accurately characterized. In this presentation we describe a computationally efficient algorithm that utilizes variational data assimilation techniques to produce a refined downwind hazard assessment by using all available observations to characterizing the release source parameters and the low-level winds. The underlying algorithm consists of a combination of modeling systems, including the Second order Closure Integrated PUFF model (SCIPUFF), its corresponding Source Term Estimation (STE) model, a hybrid Lagrangian-Eulerian Plume Model (LEPM), its formal adjoint, and the software infrastructure necessary to link them. SCIPUFF and its STE model are used to calculate a "first guess" source estimate based on the available chemical plume and meteorological observations. The LEPM and corresponding adjoint are then used to iteratively refine the SCIPUFF based STE estimate using variational data assimilation techniques. The entire process from beginning to end is completely automated and requires no human intervention. This algorithm has undergone testing using virtual "single realization" plume release data sets from the Virtual THreat Response Emulation and Analysis Testbed (VTHREAT) and data from the FUSION Field Trials 2007 (FFT07). An end-to-end prototype of this system has been developed to illustrate how it could potentially be deployed within the United States (US) Joint Effects Model (JEM). The STE prototype will be demonstrated in this presentation using VTHREAT generated chemical observations from point chemical detection systems. Preliminary results suggest that this concept provides an efficient means to better utilize CB and meteorological observations to provide a more accurate hazard assessment.

Key words: *Atmospheric, Chemical, Biological, Transport, Dispersion, Source, Parameter, Estimation, Variational Assimilation, Adjoint*

INTRODUCTION

More accurate source parameters in conjunction with atmospheric observations are key elements required to more accurately model the transport and dispersion (T&D) of chemicals in the atmosphere. However, a significant challenge for the atmospheric T&D community is the ability to make better use of chemical observations through data assimilation and fusion techniques in a way that is physically consistent with the environmental observations. These source parameters are in many cases not known and consequently based on rudimentary assumptions. This in turn leads to errors in the subsequent T&D simulations. This work utilizes variational data assimilation techniques in conjunction with a Gaussian puff dispersion model and an inverse plume modelling method to better characterize these source parameters and improve the accuracy of the subsequent plume dispersion solution. The algorithm is designed to be run on a laptop computer and provide a set of source parameters from seconds to several minutes after the observations are provided to the algorithm. The technique is suitable for any atmospheric T&D application where concentration observation and meteorological data are available and one or more of the release source parameters are not known.This methodology is particularly applicable for emergency response applications involving the dispersion of hazardous materials where a T&D solution is required as soon as possible following the collection of the observations. The overall source term estimation (STE) algorithm design and components are described in the first 5 sections. In the following section we discuss the implementation of this algorithm in an operational hazard assessment tool and some preliminary validation results based on outdoor dispersion field trial programs, and synthetically generated virtual observations from the Virtual THreat Response Emulation and Analysis Testbed (VTHREAT) (Bieringer *et al* 2008). In the final section we end with a summary discussion and concluding remarks on the direction of future work in this area.

SOURCE TERM ESTIMATION (STE) ALGORITHM SYSTEM DESIGN

Sensor data fusion (SDF) and data assimilation are techniques by which observations can be incorporated into a description of the environment using a model that constrains the problem using physical properties and time evolution (Courtier et al., 1994). Here, time varying observations of airborne chemical or biological (CB) materials are fused with available meteorological observations to characterize the release source parameters for these materials and then used to produce a refined hazard assessment or dispersion pattern that best matches the observations. Backward trajectory or trajectory inversion techniques have been extensively used for similar applications for many years (Hannan *et. al.* 2000). Although these techniques are capable of providing reasonable estimates of the source parameters, they can be computationally expensive and the computational burden increases significantly as one increases the number observation locations and times. For operational situations where CB observational data are available in a continuous fashion this technique alone may not be adequate to fully utilize the available data. Furthermore, backward trajectory methods tend to identify the intersection of relevant variables along the backward trajectory, and typically do not consider how closely a forward model solution matches the observations. One exception to this statement are the inverse plume modeling methods. In these cases the inverse of the plume model can be considered an adjoint and the solution does seek to match the observations through the dispersion model. Bayesian approaches are another means to explore the release source parameter space and have been demonstrated to be effective for this application. (Delle Monache *et. al.,* 2008) This approach, however, is also computationally expensive due to the need to perform numerous forward simulations, and, given the current computational capabilities of a laptop computer, can not meet the sub five minute computational requirement to characterize the source in an operational situation.

Variational data assimilation techniques have been used for over two decades to incorporate atmospheric observations into numerical weather prediction models and have been used extensively for other applications involving the minimization of differences between observations and model solutions. Variational data assimilation techniques are also computationally efficient and capable of meeting the operational requirements described above. This method, however, is not without weaknesses. Variational techniques require "at a minimum" a suitable first guess estimate and will find only the local minimum (between the model solution and observations) based on this first guess.

This work, sponsored by the Defense Threat Reduction Agency – Joint Science and Technology Office (DTRA-JSTO) is designed to address the need for timely source parameter estimates in their emergency response and hazard assessment modeling systems. This algorithm leverages the relative strengths of the both the backward trajectory and variational approaches to address the atmospheric CB release source estimation problem. It is a computationally efficient system that utilizes the Second order Closure Integrated PUFF model (SCIPUFF) and variational data assimilation techniques to characterize CB source parameters and provide a refined hazard assessment. The algorithm consists of a pre-processing step, a technique for making a first guess for the source type, (SCIPUFF), its corresponding source term estimation (STE) model, Gaussian Puff/plume models, the numerical adjoints, and the software infrastructure necessary to link them. SCIPUFF and its STE model are used to calculate a "first guess" source estimate based on the available CB and meteorological observations and source type estimation. The gaussian models and corresponding adjoints are then used to iteratively refine the SCIPUFF based STE estimate using variational data assimilation techniques. The entire process from beginning to end is completely automated and requires no human intervention. The major components of this algorithm, depicted in Figure 1, will be described in greater detail below.

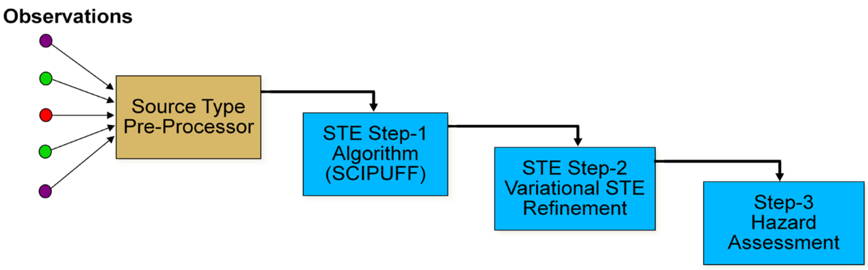

Figure 1: Steps and components of the source term estimation algorithm.

STE SOURCE TYPE PRE-PROCESSOR COMPONENT

The first component utilizes time-series observations (if available) in a source-type pre-processor algorithm to estimate the release type (Instantaneous vs. Continuous). Figure 2 illustrates a sample time-series from the FUsing Sensor Information from Observing Networks (FUSION) Field Trial 2007 (FFT07) chemical dispersion experiment, and the corresponding output from the source-type pre-processor. The algorithm uses the length of time that a polynomial fit curve of the concentration time-series is above the sensor noise floor, normalized by the wind speed, to attempt to characterize if the release type can be classified as a short/instantaneous release vs. a longer continuous release.

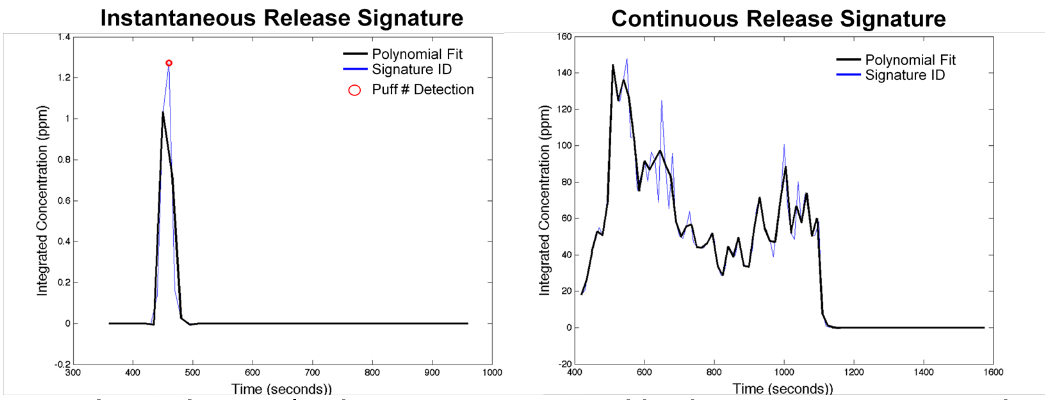

Figure 2. Sample intermediate output from the source-type pre-processor module in the NCAR/Sage Management STE Algorithm.

SCIPUFF-BASED STE ALGORITHM COMPONENT

The second component of this algorithm uses a backward trajectory-based methodology to produce a first guess estimate of the release parameters. This component utilizes the Second-order Closure Integrated PUFF (SCIPUFF) transport and dispersion model, developed by Sage Management (Sykes et al., 2008). SCIPUFF is a Lagrangian puff model, which uses a collection of three-dimensional Gaussian puffs to represent an arbitrary time varying concentration field. Each individual puff is transported and diffused by the ambient mean wind and turbulent fields, in a Lagrangian framework, similar to standard Lagrangian particle models. Turbulent diffusion is parameterized using second-order turbulence closure techniques devised by Donaldson (1973) and Lewellen (1977), which use available velocity statistics to predict the associated dispersion rates. SCIPUFF supports the modeling of various material types, including gases, particles, and liquid droplets and their associated size distributions. The first guess source parameters are computed through an inverse plume modeling process, where the winds are reversed and a series of chemical releases are made from the chemical sensor observation locations. Inverse plume solutions from the sensor locations with valid detections are handled as separate tracers and then combined in a post-processing step that triangulates the release location, time, and corresponding release mass. Figures 3A-D graphically illustrates the inverse SCIPUFF STE methodology. Figure 3-A depicts a chemical release scenario in which a plume crosses a network of sensors where some of the sensors detect the plume and are triggered while other sensors are outside the plume and provide null detections. Figure 3-B illustrates the first step involving the SCIPUFF STE process where the winds are reversed and tracer gas releases are made from the triggered sensor locations. The release rate/mass of the tracer gas are dictated by the measured sensor concentrations. Figure 3-C illustrates step 2 of the SCIPUFF STE process where again the winds are reversed and tracer gas releases are made from the null sensor locations. Figure 3-D illustrates the final step in the process where the inverse tracer gas release information from the triggered and null sensors are combined to triangulate the source parameters.

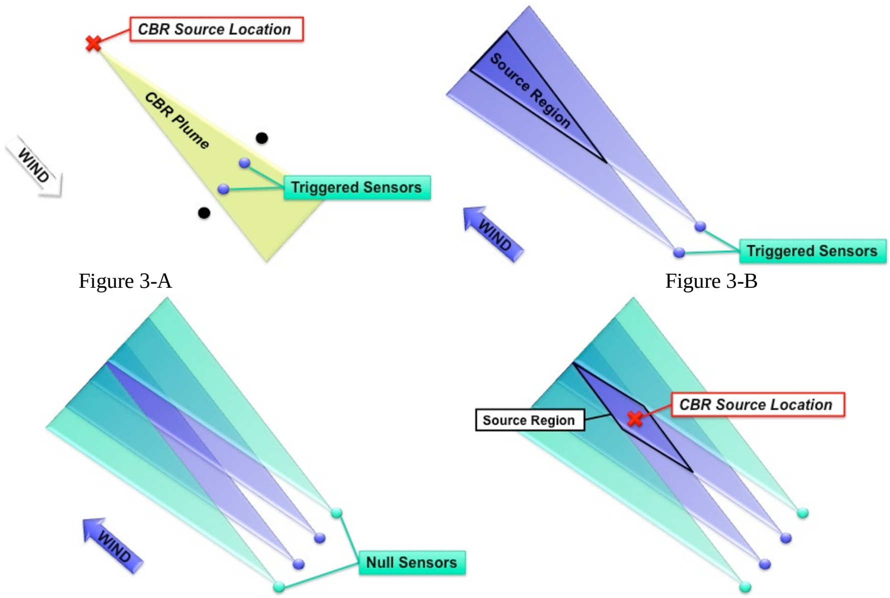

Figure 3-A Figure 3-B

Figure 3-C Figure 3-D

Figure 3 A-D. The inverse SCIPUFF STE process.

VARIATIONAL ITTERATIVE REFINEMENT STE ALGORITHM COMPONENT

The third component of the STE algorithm uses variational data assimilation techniques to iteratively refine the first guess solution provided by the inverse SCIPUFF STE component. This component makes use of two simple Gaussian models, one for instantaneous (puff) and one for continuous (plume) release. For simplicity, the puff model is presented here.

***Gaussian puff model*:** The Gaussian puff model is given by the following equation:

$$c(x,y,z,t) = \frac{q_s}{\sqrt{2\pi}^3 \sigma_x \sigma_y \sigma_z} \exp\left[-\frac{(x - x_s - u_e \cdot (t - t_s))^2}{2\sigma_s^2} - \frac{(y - y_s - v_e \cdot (t - t_s))^2}{2\sigma_s^2} \right] \times \left(\exp\left[-\frac{(z - z_s)^2}{2\sigma_z^2} \right] + \exp\left[-\frac{(z + z_s)^2}{2\sigma_z^2} \right] \right) \quad (1)$$

Where q_s is the source rate emission in kg/s, t_s is the emission time, (x_s, y_s, z_s) are the coordinates of the source, (u_e, v_e) are the horizontal wind components and $(\sigma_x, \sigma_y, \sigma_z)$, are the dispersion coefficients. In practice, the dispersion coefficients are provided by SCIPUFF, which provides the calculations of the puff/plume dispersion characteristics.

Variational assimilation: The implementation of the source estimation procedure follows the one described in Krysta *et al.* 2007. Our approach differs, however, on several points. First, the sensors and source are assumed to be at the same level and the dispersion coefficients constant. Second, the cost function that is minimized includes an additional background term:

$$J = \frac{1}{2}\left[S - S^b\right] B^{-1} \left[S - S^b\right] + \frac{1}{2} \sum_{t=1}^{t=N} \left[C^m(t) - C^s\right] R^{-1} \left[C^m(t) - C^s(t)\right] \quad (2)$$

Where $S = [q_s, x_s, y_s, t_s, u_e, v_e]$ is the control vector, , i.e. the vector of parameters that are adjusted during estimation process, z_s is provided by reverse SCIPUFF is assumed to be known. S^b is the background. It contains the source parameters S with the estimated values provided by application of reverse SCIPUFF during step two of the algorithm. S^b is also the first guess used to initialize the variational minimization. $C^s(t)$ is the observation vector made of all the sensor readings at time t, $t = 1,\ldots, N$. $C^m(t)$ is a vector made of model predicted concentrations at the sensors locations at time t by application of the model (equation (1)) to the source terms parameters S. R is the sensor observational error covariance matrix and B is the source background error covariance matrix. The determination of the matrices R and B are described in the following sub-section. Third, the estimation procedure incrementally minimizes the value of the cost function J by adjusting the source vector S., (Courtier *et al.* 1994). The solution of the estimation is the source vector *S* that minimizes the cost function *J* and at its minimum, the derivative of the cost function must vanish:

$$\frac{\partial J}{\partial S} = B^{-1}\left[S - S^b\right] + \sum_{t=1}^{t=N}\left[\frac{\partial C^m}{\partial S}\right]^T R^{-1}\left[C^m(t) - C^s(t)\right] = 0 \tag{3}$$

Where $[\partial C^m/\partial S]^T$ is the transpose matrix of the linear operator tangent to the model at point *S*, commonly denoted as the *adjoint* model. Here an incremental approach is used because of the non-linearity of the Gaussian puff model (Equation (1)). Since C^m is a function of *S* through the Gaussian puff model, equation (3) can be numerically solved using an iterative procedure. For our application, the minimization of *J* is conducted as a gradient descent using the LBGF algorithm. For better conditioning, the minimization is performed on the decimal logarithm of concentrations.

Error covariance matrices: There's little information available on sensor measurement errors. Krysta *et al.* 2007 assumed a 10% error. Such a choice of an error proportional to the measurement value is consistent with a log distribution of the measurements. It introduces, however, heterogeneities as the measurement error varies from sensor to sensor at a given time. For sake of simplicity and due to the lack of better information, we adopted a constant error value for all sensors. This value is equal to half the difference between the highest and lowest readings. Sensor error correlations in time and space are neglected. The observational error covariance matrix *R* is therefore proportional to the identity matrix. The background error covariance matrix *B* is also taken as diagonal with the following squared values on the diagonal: $\delta q = 1kgs^{-1}$, $\delta x_s = 500m$ $y_s = 500m$, $t_s = 45s$, $u_e = 1.4ms^{-1}$ $v_e=1.4ms^{-1}$. The system has also been built so that either wind components or wind speed and direction can be estimated. Errors values of $2ms^{-1}$ and $20°$ are used instead, when the option to work with wind speed and direction as control variables is selected.

Results: Application of the algorithm is demonstrated on Trial 37 of the FFT07 field campaign. The release was instantaneous and the puff travelled for about 4 minutes across the 1km x 1km sensors arrays. For this demonstration data were available from 16 fast response propylene sensors (see the network's geometry on figures 4) and one wind tower at the center of the network. Sensor data were sampled at 50 Hz but data every 20s and only readings above the 10^{-17} threshold were assimilated. This represents 6 time levels (92s, 112s, 132s, 152s, 172s and 192s) in the summation of the cost function of equation (3). Wind measurements indicate strong winds: $6ms^{-1}$, coming from Southeast (109°). Applications of the reverse SCIPUFF during the first stage of the algorithm resulted in an overestimation of the emission rate of $1kgs^{-1}$ and an estimated release point 470m from actual true source (left panel of Figure 4). These poor performances of reverse SCIPUFF raised some doubts about the accuracy of the wind measurements, which reverse SCIPUFF does not have the capability to correct. Reverse SCIPUFF, however, did correctly estimate the release time, but in the following experiments we artificially added a 36s shift in the first guess to illustrate the wind/time estimation procedure. In a first test, we apply the simple Gaussian model and its adjoint without wind adjustment. Table 1 provides the results of this no wind adjustment experiment with the truth and includes the truth values, estimated SCIPUFF first guess and the estimated values. Figure 4 graphically illustrates corresponding results for the release location. Here the release parameters are closer to the actual source than the first guess, but the location is still 243m upwind. The estimated release mass has been considerably reduced: 1.1kg instead of 4.0 kg, and the estimated release time is still 19s late. In a second test, the wind speed and direction are free to vary during the assimilation experiment. Table 2 and Figure 4 provide results of this test. When the wind is allowed to change, the estimated release point moves much closer to the actual source, they are now separated by 66m instead of 243m point. The estimated mass is not as low as previously 2.9kg instead of 1.1kg, when the true value is 4.0kg and the release time is also much closer to the truth (+8s). Comparisons between the results of the two experiments clearly show that the reported winds were not consistent with the dispersion that was observed by the sensors. It is, therefore, key for source term estimation procedures to allows wind to vary. This can be a challenge, however, as it's clear from Equation (2) that different combinations of locations, winds and release times can produce similar concentrations, or mathematically stated, the problem of recovering those three parameters distinctively from concentration measurements is ill posed. The background term in the cost function has a regularization effect and effectively breaks down the under-determination. Users can influence the process through the specification of the error covariance matrix *B*. For example the prescription of large wind errors will reduce the contribution of wind differences in the cost function and therefore the system will have more latitude to adjust the wind during the minimization.

Table 1: Estimation results of the no wind adjustment experiment

	q_s kgs^{-1}	x_s m	y_s m	speed ms^{-1}	dir degree	release time s
First Guess	5	3013	1436	6	109	36
Estimated	1.1	2828	1729	6	109	19
Truth	4.0	2826	1972	?	?	0

Table 2: Estimation results of the wind adjustment experiment

	q_s kgs^{-1}	x_s m	y_s m	speed ms^{-}	dir degree	release time s
First Guess	5	3013	1436	6	109	36
Estimated	2.9	2800	1911	4.2	116	8.2
Truth	4.0	2826	1972	?	?	0

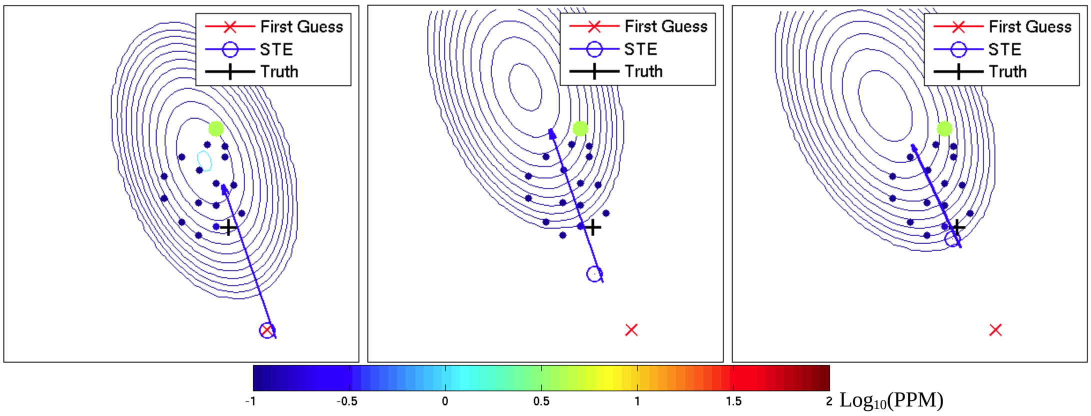

Figure 4: Dispersion results at t = 192s from: first guess (left), estimated source without wind adjustment (center) estimated source with wind adjustment (right). The blue arrow is proportional to wind speed and points to wind direction. The dots show the sensors that responded above the 10^{-17} threshold.

STE ALGORITHM OPERATIONAL IMPLIMENTATION

The algorithm described above has been implemented into a proto-type which has been integrated into a demonstration version of DTRAs Hazard Prediction and Analysis Capability (HPAC) modelling tool. HPAC is widely used by the United States Department of Defense (DoD) and civilian emergency management groups to provide rapid response T&D simulation solutions. This proto-type implementation demonstrates the entire process in the STE from loading the meteorological and CB observations into HPAC, to the two source term estimation components, to the final depiction of the plume hazard based on the observational data provided. The demonstration works within the framework of the graphical user interface (GUI) that the HPAC user community is familiar with, and can utilize observational data formatted in the US DoD Nuclear, Biological, and Chemical (NBC) messaging formats. Preliminary results suggest that the algorithm is capable of providing source parameters that are significant improvements over the standard operational practice of developing model source terms based on the sensor observation locations and measurements. The performance validation of this algorithm is an ongoing effort and we anticipate that quantitative results over a broader range of cases will be completed in the coming year.

CONCLUSIONS AND FUTURE WORK

The STE algorithm is a step towards providing a fully automated capability to quickly characterize the source parameters of an outdoor CB dispersion event and produce a T&D solution, which closely matches the observations. While numerous approaches to this problem exist, each with their own strengths and weaknesses, this approach attempts to addresses this problem in an operational environment where computational resources are limited and a timely solution is critical. An early demonstration of this capability in an operational tool, in this case HPAC, illustrates that it adds an important capability to the tool set available to the emergency management and first responder community. Future work will progress in a number of areas ranging from:

1. Continuing qualitative and quantitative performance validation;
2. Hardening of the software and ensuring the existing algorithm works robustly for a broader set of scenarios;
3. Extension of the algorithm to support urban STE requirements.

REFERENCES

Bieringer, P. E., J. Weil, J. Hurst, G. Bieberbach, A. Wyszogrodzki, R. Sheu, M. Raines, and T. Warner, 2008: A framework for developing synthetic chemical and biological agent release data sets for use in virtual test and evaluation. *Chemical and Biological Defense Physical Science and Technology Conference*, 17-21, Nov., 2008, New Orleans,

Courtier P., Thepaut, J.-N. and Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc*., **120**, 1367-1388

Delle Monache, L., Lundquist, J., Kosovic, B., Johannesson, G., Dyer, K., Aines, R., Chow, F., Belles, R., Hanley, W., Larsen, S., Loosmore, G., Nitao, J., Sugiyama, G., Vogt, P., 2008. Bayesian inference and Markov Chain Monte Carlo to reconstruct a contaminant source at continental scale. *Journ. Appl. Meteor. Climatol.*, **47**, 2600-2613.

Donaldson, C., 1973: Atmospheric turbulence and the dispersal of atmospheric pollutants. *AMS Workshop on Micrometeorology*, 313-390.

Hannan, J. R., H. E. Fuelberg, A. M. Thompson, G. Bieberbach Jr., R. D. Knabb, Y. Kondo, B. E. Anderson, E. V. Browell, G. L. Gregory, G. W. Sachse, and H. B. Singh. 2000. Atmospheric chemical transport based on high-resolution model-derived winds: A case study. *J. Geophys. Res.* **105**, 3807-3820.

Lewellen, W.S., 1977: Use of invariant modeling. *Handbook of Turbulence*, 237-280.

Sykes, R.I., S.F. Parker, D.S. Henn, and B. Chowdhury, 2008: SCIPUFF Version 2.4 Technical Documentation.