P 452

DETERMINING RELEVANT TURBULENT SCALES AND SOURCE VARIABLES FROM CONTAMINANT CONCENTRATION DATA

Andrew J. Annunzio*, Sue Ellen Haupt, and George S. Young
The Pennsylvania State University, Applied Research Lab and Department of Meteorology

1. Introduction

Various meteorological variables and contaminant source characteristics can be back calculated from contaminant concentration observations. In some atmospheric transport and dispersion (AT&D) applications, such as modeling an accidental or intentional contaminant release, it is necessary to do so in order to mitigate the contaminant threat and predict AT&D downwind. To accurately determine these variables one must ascertain the mean direction of contaminant travel for either instantaneous or continuous releases, and the puff translation speed for an instantaneous release. This issue is addressed in prior work where source characteristics, as well as these atmospheric transport variables are back-calculated from concentration data (Haupt 2005; Allen et al. 2007; Long et al. 2010). There, contaminant plume or puff dispersion parameters were assumed to be known and were determined from Pasquill Gifford stability classes. Here we eliminate that assumption and instead deduce the dispersion parameters from convective scaling parameters (Deardorff 1974), i.e. the length and velocity scales of the boundary layer spanning eddies (BLSE, Deardorff 1970). The former is the depth of the Atmospheric Boundary Layer (ABL), z_i, while the latter is the convective scale velocity, w_*, that is given by

$$w_* = \left(\frac{g}{\theta_v} z_i \overline{w'\theta_v'} \right)^{\frac{1}{3}} \quad (1)$$

where $\frac{g}{\theta_v}\overline{w'\theta_v'}$ is the buoyancy flux. Therefore we also retrieve these scaling variables, which is important in prediction of subsequent contaminant dispersion.

Here we back-calculate not only the convective boundary layer depth and the convective velocity scale, but also the remaining variables required for AT&D modeling of a contaminant release including the source information. Intentional and accidental contaminant releases can occur in the mixed layer or surface layer, and can be either instantaneous or continuous. Therefore, these variables are back-calculated for these four scenarios: surface layer and mixed layer sources, each conducted as both instantaneous and continuous releases. The number of unknown variables depends on the release type. For an instantaneous release, we ascertain seven variables (mean puff transport speed, the direction of contaminant travel, the convective velocity scale, the boundary layer depth, the along-wind and cross-wind source location, and the source strength). For a continuous release, we determine these same variables except for transport speed because this variable is not separable from the source strength in a back calculation from a steady state contaminant concentration field.

In section 2, we describe the technique used to determine the unknown variables. In section 3 we determine the proper sensor domain size and the appropriate dispersion parameters for all four scenarios. Then in section 4, we test the algorithms' ability to back-calculate the unknowns for all four scenarios. Section 5 summarizes the results, discusses the significance of the work, and considers the prospects for future work.

2. Methods

We wish to determine, from a time series of surface contaminant concentration values, the mean depth of an unstable ABL and the convective scale velocity as well as the other variables relevant to AT&D. This back calculation is accomplished via a powerful optimization technique that tunes these variables so as to match observed concentration values to those concentration values calculated by a dispersion model. Thus, the unknowns are determined indirectly through forecasts of surface concentration values as in Haupt (2005), Allen et al. (2007), and Long et al. (2010).

a. Dispersion Model Selection

Because our technique (described in section 2b) requires the creation of a concentration forecast for each of numerous trial estimates of the unknown variables, our dispersion model must be computationally efficient. However, to match the observed concentration values, it is necessary that the dispersion model captures the primary effects of convective boundary layer (CBL) turbulence. The structure of the CBL is relatively well understood from numerical experiments, convective tank experiments, and from atmospheric observations (Deardorff 1972;

*Corresponding author address: Andrew J. Annunzio, The Pennsylvania State University, Applied Research Lab and Department of Meteorology, State College PA, 16804; email: aja199@psu.edu

Kaimal et al. 1976; Panofsky and Dutton 1984). When convection dominates turbulence production (i.e. $\frac{z_i}{L} > 10$ where L in the Monin-Obukhov length), the resulting turbulence statistics scale with the convective velocity scale, w_*, and the boundary layer depth, z_i Deardorff (1970). Thus, under convective conditions, the along-wind and cross-wind turbulence statistics to depend on z_i rather than the height, z, above the surface (Panofsky 1977). The vertical turbulence component, however, is dependent on both z and z_i because of the existence of upper and lower boundaries to the CBL (Wyngaard 1988).

The nature of vertical dispersion depends on the profile of σ_w and the skewness of the vertical velocity probability distribution function (PDF) resulting from an unequal area occupied by updrafts and downdrafts (Wyngaard 1987 and Weil 1990). These phenomena cause non-Gaussian behavior in vertical contaminant dispersion (Wyngaard 1987; Weil 1990) including counter-gradient fluxes at some levels within the CBL. Moreover, they cause the plume centerline or puff centroid to change height with distance downstream of the source (Willis and Deardorff 1975; Lamb 1982). The resulting mean ascent of contaminants for a surface release and mean descent of contaminants for a mixed layer release must be captured in a successful model of CBL dispersion.

In this work we consider instantaneous as well as continuous sources of contaminants. The temporal averaging possibilities differ for these two source types, so we must be careful in choosing the proper domain size for the sensor grid to enable accurate back calculation of the unknown variables. These issues are discussed in section 3.

1) Surface Layer Release

The ensemble averaged contaminant field for a surface release initially spreads laterally and vertically through the surface layer before the BLSEs cause it to meander vertically through the overlying mixed layer (Weil 1988; Willis and Deardorff 1975). Although Gaussian dispersion cannot account for this vertical meander of contaminants, it is shown that a Gaussian model can describe ensemble averaged surface concentration values for a near surface release with some success (Deardorff and Willis 1975; Hadfield 1994). The degree of this success is dependent on the vertical dispersion parameters chosen as well as the non-dimensional distance from the source (see section 3) (Willis and Deardorff 1976). The Gaussian approximation holds best close to the source and for long averaging periods. In this situation, the appropriate dispersion model is a Gaussian with a ground reflection term.

$$C(x,y,z,t) = \frac{Q}{\left((2\pi)U\sigma_y\sigma_z\right)}\exp\left(\frac{-(y-y_o)^2}{2\sigma_y{}^2}\right)\times \left(\exp\left(-\frac{(z-z_o)^2}{2\sigma_z{}^2}\right)+\exp\left(-\frac{(z+z_o)^2}{2\sigma_z{}^2}\right)\right) \quad (2)$$

where x, y, and z span the Euclidean coordinate system, t is time, Q is the release rate, U is the mean wind speed, y_o is the cross-wind source location, z_o is the source height and σ_y and σ_z are the lateral and vertical dispersion parameters respectively. These dispersion parameters are dependent on the downwind distance from the release location.

In contrast, for an instantaneous release in the surface layer, it is not possible to match contaminant AT&D during the initial growth stage because only one release event is available, so one cannot perform an average. Without such averaging, observed contaminant concentration will exhibit eddy-driven meandering and distortions that are not captured by ensemble mean dispersion models such as the Gaussian. In this situation it is necessary to sample a contaminant when it is well mixed in the vertical and back-calculate boundary layer depth by the direct approach. For this situation, we again implement a Gaussian model; however, reflection terms that account for puff entrapment are also included. Assuming that the contaminant puff centroid translates downwind at constant velocity, (2) becomes

$$C(x,y,z,t) = \frac{Q}{\left((2\pi)^{\frac{3}{2}}U\sigma_x\sigma_y\sigma_z\right)}\exp\left(-\frac{(x-(x_o+Ut))^2}{2\sigma_x{}^2}\right)\times \exp\left(\frac{-(y-y_o)^2}{2\sigma_y{}^2}\right)\sum_{n=0}^{N}\exp\left(\frac{-(z-2nz_i-z_o)^2}{2\sigma_z{}^2}\right)+ \exp\left(\frac{-(z+2nz_i-z_o)^2}{2\sigma_z{}^2}\right)+\exp\left(\frac{-(z-2nz_i+z_o)^2}{2\sigma_z{}^2}\right)+ \exp\left(\frac{-(z+2nz_i+z_o)^2}{2\sigma_z{}^2}\right) \quad (3)$$

where U is the puff translation speed.

2) Mixed Layer Release

The average AT&D for a mixed layer release is markedly different than the AT&D of a surface release. Because the average effect of convective elements initially cause contaminants to descend to the surface, predicting surface concentration values via traditional Gaussian dispersion for a mixed layer release will cause significant errors. A passive

contaminant descends to the surface at a velocity proportional to the convective velocity scale, w_*, with this descent resulting in ensemble averaged surface concentrations values near the source location being about 2.9 times greater than predicted by a traditional Gaussian model (Willis and Deardorff 1981; Lamb 1982; Weil 1988; Briggs 1993). To account for this effect, Weil (1988) developed a computationally efficient dispersion model based on the probability density function (PDF) of vertical velocity. Adopting a similar approach to that used to describe the vertical velocity PDF (Misra 1982), vertical contaminant dispersion is described by a superposition of two Gaussians. Assuming that crosswind dispersion is Gaussian, plume dispersion for a mixed layer release is given by Weil (1988, 1997) as

$$C(x,y,z,t)=\frac{Q\Delta t}{2\pi U\sigma_y}\exp\left(-\frac{(y-y_o)^2}{2\sigma_y{}^2}\right)\times$$
$$\sum_{n=-N}^{N}\sum_{j=1}^{2}\frac{\lambda_j}{\sigma_{z,j}}\left(\begin{array}{l}\exp\left(-\frac{(z-2nz_i-z_o-\overline{w_j})^2}{2\sigma_{z,j}{}^2}\right)\\+\exp\left(-\frac{(z+2nz_i+z_o-\overline{w_j})^2}{2\sigma_{z,j}{}^2}\right)\end{array}\right) \quad (4)$$

where λ_j determines the input from the updrafts and downdrafts, $\overline{w_j}$ is the updraft and downdraft speed, and the value of j specifies an updraft or downdraft. The dispersion parameters $\sigma_{z,j}$ depends on $\sigma_{w,j}$, the standard deviation of the vertical velocity, and λ_j, $\overline{w_j}$, and $\sigma_{w,j}$ are found by equating moments of the vertical velocity PDF (Weil 1990; Weil 1997). For an instantaneous mixed layer release, we encounter the averaging problem described above for a surface layer release. Taking a similar approach we again modify (4) by assuming an instantaneous release and that the puff translates downwind at constant velocity to produce

$$C(x,y,z,t)=\frac{Q\Delta t}{2\pi U\sigma_x\sigma_y}\exp\left(-\frac{\left(x-(x_o+Ut)\right)^2}{2\sigma_x{}^2}\right)\times$$
$$\exp\left(-\frac{(y-y_o)^2}{2\sigma_y{}^2}\right)\times \quad (5)$$
$$\sum_{n=-N}^{N}\sum_{j=1}^{2}\frac{\lambda_j}{\sigma_{z,j}}\left(\begin{array}{l}\exp\left(-\frac{(z-2nz_i-z_o-\overline{w_j})^2}{2\sigma_{z,j}{}^2}\right)\\+\exp\left(-\frac{(z+2nz_i+z_o-\overline{w_j})^2}{2\sigma_{z,j}{}^2}\right)\end{array}\right)$$

b. The Hybrid Genetic Algorithm

An optimization technique is required to obtain an accurate estimate of the unknown source and meteorological variables by matching observed and predicted concentration fields. Given the difficulty of this problem, the optimization technique must be robust in the face of a complex solution space. Thus, we use a hybrid genetic algorithm (GA, Haupt and Haupt 2004), a combination of a GA with the Nelder-Meade Downhill Simplex (NMDS) technique. This hybrid method is ideal for this problem because it is a global optimization technique that explores the entire solution space before converging to the best estimate of the unknown variables, and thus, if appropriately configured, avoids falsely converging to a local optimum.

The GA optimization technique mimics the process of natural selection to obtain iteratively improved estimates of the unknown variables (Haupt and Haupt 2004). For a discussion of the GA process, the reader is referred to Haupt and Haupt (2004), Haupt (2005), Allen et al. (2007) and Long et al. (2010).

Our cost function quantifies the fitness of each chromosome by computing the difference between the resulting concentration forecast and the concentration observations. Instead of directly comparing the concentration values, we compare the logarithm of the concentration values so as to transform the Gaussian into a quadratic, thereby enhancing the influence of concentration tails. This transformation enables the algorithm to better match the spread of the contaminant field. To quantify the difference between predictions and observations, the cost function sums the Euclidean distance between the logarithm of the concentration forecast and the logarithm of the concentration observations over all time steps and grid points. This sum is normalized by the logarithm of the concentration observations. Specifically it is,

$$\frac{\sum_{t=1}^{N_t}\left(\sum_{s=1}^{N_s}(\log(C_s+\varepsilon)-\log(R_s+\varepsilon))^2\right)^{\frac{1}{2}}}{\sum_{t=1}^{N_t}(\sum_{s=1}^{N_s}(\log(R_s+\varepsilon))^2)^{\frac{1}{2}}} \quad (6)$$

where ε is a scaling factor added to the predictions and observations to avoid taking the logarithm of zero, C_S is a predicted concentration value at a sensor, R_S represents a concentration value reported by a sensor, N_S represents the number of sensors, and N_t represents the number of time steps. The GA finds an appropriate estimate of the unknown variables by iteratively minimizing this cost function.

After the GA produces an estimate of the unknown variables within the fitness basin of the

global minimum of the cost function, the NMDS algorithm is used to quickly find the bottom of that solution basin. Thus, the GA is used to find the globally optimum solution basin, while NMDS cascades down the gradient of this basin to find the global optimum solution.

c. Identical Twin Experiment

For model development we test our method on synthetic data created in an identical twin experiment, wherein the assimilating model also creates the observations (Daley 1991). Thus, the contaminant concentration observations are created by the dispersion models described above. This formulation allows us to best gauge if our numerical experiments are working properly; however, it may create an artificially simple cost surface. These models may deviate from actual averaged contaminant observations if the averaging is not sufficient, the meteorology is nonstationary, or if the turbulence is not horizontally homogeneous. Therefore, to simulate these effects in a controlled manner, additive clipped Gaussian white noise is applied to the observations as in Long et al. (2009) and Allen et al. (2007). This noise can be used to distort the contaminant puff or plume shape. It does not, however, represent the meandering introduced by atmospheric fluctuations whose length scales are larger than that of the contaminant puff or plume.

3. Domain and Dispersion Parameter Considerations

The size of the sensor grid proves to be a crucial constraint when back calculating the variables relevant to AT&D. In order to accurately determine these variables, the sensor domain size should not only be large enough to capture the contaminant spread, but also extend far enough downwind that BLSE-driven meandering and vertical looping have less impact on the realization's concentration field than do plume spread. This range can be minimized for continuous releases by using temporal averaging. How these factors are accounted for also depends on the contaminant release height as discussed in sections 3.1 and 3.2.

a. Surface Release

1) Continuous Release

A continuous release of contaminants is advantageous for the back calculation problem because we can average surface concentration values in time if the meteorological conditions are steady. The time averaged concentrations then approach the ensemble average if the averaging period is sufficiently large. This averaging will smooth out lateral meandering; however, it will not average out all of the effects of vertical meandering caused by the BLSEs (Lamb 1982; Deardorff and Willis 1975).

This averaging allows us to determine the appropriate sensor domain size for a continuous surface layer release because the ensemble averaged plume grows within the surface layer and follows the Gaussian model, as discussed in section 2.1.1, until the non-dimensional distance $X = 0.5$, where

$$X = \frac{x}{U}\frac{w_*}{z_i} \tag{7}$$

where U is the mean wind speed (Deardorff and Willis 1975; Willis and Deardorff 1976; Hadfield 1994; Dosio et al. 2003). This expression results from scaling time by the large eddy time scale and assuming Taylors' translation hypothesis. Ideally, we would like our domain to terminate at X = 0.5 for all situations; however, this is impractical because the sensor domain then depends on three meteorological variables that are unknown for this problem. Also, it is not viable to adjust the sensor domain in real time applications in response to meteorological conditions. Therefore, the sensor domain is a constant 2 x 2 km grid for continuous, surface release situations. This domain restriction will cause a slight over-prediction of averaged surface concentration values near the edge of the domain for relatively shallow boundary layer depths, but this will not occur for large boundary layer depths.

Deardorff and Willis (1975) developed a vertical dispersion parameter that successfully describes contaminant spread in surface layer when turbulence production is dominated by buoyancy. This σ_z interpolates between Yaglom's free convection limit (Yaglom 1972) and the neutral stability limit, as well as accounting for variations in spread caused by the source height. This parameter is given by

$$\frac{\sigma_z}{z_i} = \left(\left(b' X\right)^2 \frac{{u_*}^2}{{w_*}^2} + 0.25X^3 + 1.8 z_o X^2 \right)^{\frac{1}{2}} \tag{8}$$

where

$$b' = \frac{0.4 + 0.8\left(\frac{w_*}{u_*}\right)\left(\frac{z_o}{X}\right)}{1 + \left(\frac{w_*}{u_*}\right)\left(\frac{z_o}{X}\right)} \tag{9}$$

and u_* is the friction velocity. Because we only search for z_i and w_* in (8), it is assumed that u_* is obtained from wind measurements. The lateral spread parameter does not follow this expression because the BLSEs do not initially cause nonlinear growth of the horizontal contaminant spread. The lateral spread formulation used here is that of Briggs (1993)

$$\frac{\sigma_y}{z_i} = \frac{0.6X}{\left(1+X^2\right)^{\frac{1}{6}}} \qquad (10)$$

which was determined from the Condors experimental data merged with data from other experiments. This σ_y interpolates between the near field linear growth period and the far field $\sigma_y \approx 0.6 z_i X^{\frac{2}{3}}$ growth period (Briggs 1993). It is assumed that the lateral turbulence parameters will not change with release height. Therefore, this expression is used for both the surface release and the mixed layer release.

2) Instantaneous Release

The problem facing the back-calculation technique for an instantaneous release is more difficult because we cannot average the contaminant observations in time. Therefore, we must modify our technique to minimize errors caused by inadequate averaging. In these situations, the domain size must be much larger than that required for the continuous case. This domain enlargement is necessary because horizontal and vertical meandering may disrupt the correspondence between model and observations to even larger values of X if temporal averaging is not used. Thus, we enlarge our square domain such that each side is a distance equivalent to $X = 8$. Increasing the domain size while maintaining the same number of sensors decreases our sensor density. We then determine the unknown variables by matching the lateral contaminant spread and by determining when the contaminant is well mixed in the vertical. The latter effect explicitly provides information on the boundary layer depth from contaminant entrapment. Because our dispersion model includes reflections, the vertical dispersion parameter should not asymptote to a constant value. The entrapment is described by the reflection terms, which captures the vertical homogenization of the contaminant with time. With this enlarged domain size, contaminant dispersion is dominated by mixed layer turbulence. Therefore the vertical turbulence parameter used for the contaminant puff is given by Weil (1997):

$$\frac{\sigma_z}{z_i} = \frac{0.6X}{\left(1+0.7X\right)^{\frac{1}{2}}} \qquad (11)$$

This expression describes the near field linear puff growth and the far field parabolic puff growth. The factor of 0.6 results from assuming $\sigma_w \approx 0.6 w_*$, which is a valid assumption for $0.1 z_i < z < 0.8 z_i$. Although the contaminant release is in the surface layer, dispersion will occur throughout the entire boundary layer, implying that this assumption is valid. Using this formulation, the Gaussian model (3) predicts uniform concentration values near *X = 3*, which is consistent with observations. The information provided by this vertical dispersion is particularly useful in real-time applications, because the existence of vertical boundaries as the surface and z_i ensure that vertical homogenization will eventually occur. In contract, the observed horizontal spread tends to be less than that predicted by (11) if averaging times are inadequate (Briggs 1993).

b. Mixed Layer Release

For a mixed layer release, we implement a dispersion model that accounts for the vertical meandering of the plume centerline or the puff centroid. Therefore our requirements for the domain size and dispersion parameters change slightly.

For a continuous contaminant release in the mixed layer, the initial puff growth occurs in the mixed layer, implying that the vertical turbulence parameter follows an expression similar to (11). Thus, for the continuous release our formulation for σ_z is similar to (11); but is adjusted to apply to either an updraft or downdraft as implemented in equation (4) (Weil 1988; Weil 1997). Because this dispersion model accurately represents the ensemble averaged dispersion for a mixed layer release, no restrictions are placed on the size of the sensor domain. To be consistent with the continuous surface layer release, we use a 2x2 km grid for all boundary layer depths. For the instantaneous release, we take an approach similar to that implemented for the instantaneous release in the surface layer wherein the domain is large to avoid sampling contaminant meandering.

4. Results

We test this back-calculation procedure for boundary layer depths ranging from 500 m to 3000 m. For these boundary layer depths, w_* is also allowed to range through typical atmospheric values, 1.7 ms^{-1} and 2.2 ms^{-1}. For all simulations we maintain a sensor domain as an 8x8 grid; therefore, the sensor density decreases with increasing domain size. The number of time steps, N_t depends on the release type; there is only one concentration field for a continuous release due to time averaging, while we use five fields for an instantaneous release separated by equal times steps. The number of reflection terms, *N,* is five for the situations that require information from entrapment. This number is sufficient to simulate vertical homogenization of the contaminant. The two release heights tested are 1 m for the surface layer release and 100 m for the mixed layer release. The horizontal source location for all scenarios is at the center of the sensor domain, (0,0). Because the GA is initialized with a population of randomized trial solutions, each simulation will take a different path to the global minimum. Thus, we make 50 runs of the method for each boundary layer scenario, and take the median of these runs to estimate a likely result for

a single simulation. We also test the impact of observational and atmospheric noise on the back calculation to determine how much noise the model can handle before the back calculation becomes less successful. The amount of noise added to the concentration data is described by the signal-to-noise ratio (SNR). The SNR values used for this work are infinity, 100, 50, 10, 5, 2, and 1.

For this back calculation method, a low cost function value does not always imply that the optimized variables match the truth. Therefore it is instructive to validate the back calculation with an alternate metric that directly compares the solution to the exact value. The alternate metric used is a skill score that gives an independent assessment of model accuracy and represents the percentage error of the model estimate. It is possible to use skill scores in an identical twin experiment because the variables to be determined are known prior to the back-calculation. The skill score we use for boundary layer depth is given by the error normalized by plausible range

$$S_{z_i} = \begin{cases} \dfrac{z_{i,a} - z_{i,f}}{z_{i,a} - z_{i,l}} : z_{i,a} \geq z_{i,f} \\ \dfrac{z_{i,f} - z_{i,a}}{z_{i,u} - z_{i,f}} : z_{i,a} < z_{i,f} \end{cases} \quad (12)$$

where $z_{i,a}$ is the actual boundary layer depth value, $z_{i,f}$ is the forecast boundary layer depth found by the GA, and $z_{i,l}$ and $z_{i,u}$ are the lower and upper bounds respectively for the GA search range in the back-calculation of boundary layer depth. The skill score computes a percent error; therefore the closer to zero the score, the better the solution. Somewhat arbitrarily, we take a skill score less than 0.1 to indicate a successful back calculation. Similar skill score formulae are applied to the other unknown variables. The total skill of the model is the sum of all the skill scores normalized by the number of unknown variables. Skill scores are used because they provide a means to quantify the combined success of the method across variables with different units.

a. Surface Layer Release

Results for both the continuous and instantaneous release in the surface layer cases are shown in Figure 1; skill scores for the continuous case are presented in Figure 1a. For this source type and these atmospheric conditions, the algorithm accurately estimates the unknown variables when the signal is noiseless, as expected. In fact, the error results for all boundary layer depths is of the order 10^{-8}, and thus excellent agreement with the known solution. As the SNR decreases, we expect the model skill to also decrease because the noise impacts the information content of the observations as seen in Long et al (2010). Figure 1a confirms this: the model error scores increase as noise is added. At $SNR = 5$ the model has an error score less than 0.1, errors increase rapidly, however, with further increases in noise. As SNR decreases from 2 to 1, the information content of the concentration field decreases to the point where the multivariate back calculation is no longer successful.

Table 1 displays the errors and percent errors of just the boundary depth predictions. For $SNR \geq 10$, errors in boundary layer depth are not large and, in this SNR range, the error in boundary layer depth predictions remains below 10%. When SNR is decreased to 5, errors in boundary layer depth remain minor for $z_i \leq 1750$; however, errors become larger (between 10% and 20% error) for deeper boundary layers. These errors increase as the SNR approaches 2, where errors in boundary layer depth become greater than 20%.

For a surface layer instantaneous release, the domain size must be sufficiently large to sample the contaminant when it is vertically homogenized. That occurs near $X = 3$. The back calculation skill scores are shown in Figure 2b. As with the continuous release, the error scores are close to zero when the SNR is infinite, implying that the estimate for each of the unknown variables is close to its true value. As the SNR decreases, the figure shows that the algorithm handles noise better for the instantaneous cases than the continuous cases; skill is not significantly affected by the noise until $SNR = 2$. Even for this noise level, the error remains near 0.1 implying that errors in estimates of the unknown variables are not prohibitively large. It is not until $SNR \approx 1$ that estimates of the unknown variables deviate substantially from the true values, as seen by the large increase in the model skill scores.

Focusing on just boundary layer depth in Table 1, when $SNR \geq 5$, errors are less than 2% for all boundary layer depths tested. Boundary layer depth estimates are still quite good for $SNR = 2$ with the errors being less than 7%. As with the error scores, errors in boundary layer depth become substantial near $SNR \approx 1$, where it grows to between 8% and 36%.

When the concentration values are corrupted by noise that is of the same order of magnitude as the concentration value, the instantaneous release estimates of boundary layer depth are more accurate than those for the continuous release estimates. There are two possibilities for why this occurs. The first possibility is that the instantaneous case has more information available to determine z_i, the extra information coming from contaminant entrapment that is captured by the reflection terms in (2). The contaminant entrapment allows a direct calculation of the mixing depth, and this information along with information from contaminant spread allows accurate determination of z_i. A second possibility is that the temporal dependence of the puff allows multiple

observations of the contaminant field, each corrupted by uncorrelated noise. Thus, the signal can be separated from the noise. An equivalent mathematical effect could be achieved by sampling a continuous release over the same number of time steps and averaging the field. By effectively averaging before noise corruption our test penalized the back calculation method. We undertake the test this way, however, because in a real-world application, an answer is desired in the least time possible.

b. Mixed Layer Release

Figure 2 displays error results for continuous and instantaneous releases in the atmospheric mixed layer; the continuous release errors are presented in Figure 2a. These errors appear somewhat similar to those shown for the continuous release in the surface layer. The errors are small for the noiseless case and increase with lower *SNR* values; the error is less than 0.1 for SNR=5 and is above 0.1 for SNR=2. The information content is no longer sufficient for successful multivariate back calculation once SNR reaches, 1 as seen by the large error scores.

While the pattern in the multivariate back calculation errors are parallel between the continuous mixed layer and surface layer releases, the errors in boundary layer depth deviate somewhat. As shown in Table 2 the errors in boundary layer depth, when noise of the same order as the concentration signal corrupts the observations, are smaller for a mixed layer release, being less than 5% when the $SNR = 5$. Their trend as SNR decreases is similar though, with it becoming difficult to back calculate boundary layer depth when $SNR = 1$. It is interesting that estimates in boundary layer depth for the continuous mixed layer release are improved from the continuous surface layer release. This difference occurs because the rate of plume descent from the mixed layer is a function of boundary layer depth and the convective velocity scale. Therefore, more information is available to back calculate z_i.

As with the surface layer release, Figure 3b error results for the instantaneous release are more successful than for the continuous release. In this case, error estimates remain below 0.1 until the SNR = 2, and these errors do not grow as fast near $SNR = 1$.

Errors of boundary layer depth estimates for the instantaneous mixed layer scenario parallel those for the instantaneous surface layer scenario. As seen in Table 2, errors in boundary layer depth estimates remain less than 2% for all boundary layer depths tested as long as the $SNR \geq 5$. The only difference between the surface and mixed layer release scenarios is that errors in boundary layer depth are not as substantial for a mixed layer release for reasons stated above. Here, percentage errors in boundary layer depth are less than 20% when the $SNR = 1$.

For a mixed layer release, results display more error for a continuous release than the instantaneous release.. This implies that the increased errors reported for continuous contaminant releases are caused by both factors proposed at the end of section 4.1.

5. Conclusions

We have incorporated a back calculation for boundary layer depth, z_i, and the convective velocity scale, w_* into the source term estimation problem. These turbulent scaling variables and other variables relevant to the AT&D problem were back calculated from the surface concentration field via an iterative optimization algorithm and a forward AT&D model. With knowledge of these variables, it is possible to forward model the AT&D problem into the future, allowing authorities to make decisions for source mitigation. By optimizing the source characteristics and relevant atmospheric variables so as to match the contaminant spread in observed and modeled fields, it is possible to estimate these variables skillfully for continuous contaminant releases in the surface layer and the mixed layer. Error estimates of boundary layer depth increase, however, when noise of the same order of magnitude as the concentration signal corrupts the signal. In contrast, for the instantaneous release case we extract information from entrapment as well as spread of the contaminant field, thereby achieving a better estimate of the variables sought.

For source term estimation it is essential to accurately determine the source characteristics of a contaminant and relevant meteorological variables in a timely manner. Unfortunately, the averaging times required for the validity of many dispersion models are large in the convective boundary layer because the ratio of the length and velocity scales of the BLSEs is large. Therefore, dispersion models may not accurately describe the contaminant transport and dispersion in real time applications. We simulate this problem in our back-calculation method by adding noise to that data, using this noise to mimic atmospheric fluctuations and sensor errors. Although this noise cannot account for vertical meandering, at low SNR the noise is of the same order as the signal, causing the observations to deviate considerably from the dispersion model predictions.

The effect of vertical and horizontal meandering also has implications for sensor siting. The continuous source cases are very practical for this back calculation approach because the sensor siting grid need only have an adequate density, the domain size relative to boundary layer depth being much less important. This is possible because we can average the concentration data in time and match the horizontal and vertical contaminant spread to determine boundary layer depth before entrapment

provides additional information on this variable. Such averaging does, however, impose a response time constraint that can limit the timeliness of subsequent contaminant forecasts, a critical issue in the event of a harmful release.

While back calculations for an instantaneous release can be done on small domains using information from only the mean lateral spread of the contaminant, information from vertical spread must also be exploited with instantaneous sources. This extra information is needed because the effects of meandering cannot be averaged away. Rather we require a domain large enough to encompass the eventual vertical homogenization of the contaminant field. Thus, the domain should extend downwind of the source for at least the distance traversed by the contaminant puff in four turnover times of the BLSEs. In real time applications it is not practical to change the domain in response to diurnal and synoptic variations in this scale. Therefore, it is sensible to consider a 15x15 km sensor domain size and a higher sensor density for the instantaneous cases. This ensures that the sensor domain is large enough to sample the contaminant homogenization even in high-wind, deep boundary layer situations.

Here we have exploited surface contaminant concentration observations to find source characteristics as well as the meteorological variables relevant to AT&D. The dispersion models used were basic and computationally efficient. In future work, we may include a more sophisticated model like the Second order Closure Integrated PUFF model (SCIPUFF); however, the same domain considerations will still be necessary for instantaneous releases because this is also an ensemble averaged model. This method can also be applied to other passive tracers, such as pollen, in the atmospheric boundary layer.

6. Acknowledgements

The authors would like to thank the Defense Threat Reduction Agency who supported this research under grant numbers W911NF-06-C-0162 and DTRA01-03-D-0010-0012 and the Educational and Foundational funding supplied by the Applied Research Lab at The Pennsylvania State University. Also, the authors thank John Wyngaard, Kerrie Long, Luna Rodriguez and Yuki Kuroki for their help and suggestions. Lastly, the authors thank John Hannon and Chris Kiley who monitored the project

7. References

Allen C.T., G.S. Young, and S.E. Haupt, 2007: Improving Pollutant source characterization by better estimating wind direction with a genetic algorithm. Atmos. Environ. **41**, 2283-2289.

Briggs G.A., 1993: Final Results from the CONDORS Experiment. *Bound.-Layer Meteor.* **62**, 315-328

Daley R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, UK 457 pp.

Dosio A., J. Vila-Guerau De Arellano, A.A.M. Holstag, and P.J.H. Builthes, 2003: Dispersion of a Passive Tracer in Buoyancy- and Sear-Driven Boundary Layers. *J. Appl. Meteor.,* **42**, 1116-1130.

Deardorff, J.W., 1970: Convective velocity and temperature scales for the unstable planetary boundary layer and Rayleigh convection. *J. Atmos. Sci.*, **27**, 1211-1213.

Deardorff JW, (1972) Numerical Investigation of Neutral and Unstable Planetary Boundary Layers. *J. Atmos. Sci*., **29,** 91-115.

Deardorff J.W., and G.E. Willis, 1975: A parameterization of Diffusion into the mixed layer, *J. Appl. Meteor.,* **14**, 1451-1458.

Hadfield M.G., 1994: Passive Scalar Diffusion From Surface Sources in the Convective Boundary Layer, *Bound.-Layer Meteor,.* **69**, 417-448

Kaimal J.C., J.C. Wyngaard, D.A. Haugen, O.R. Cote, Y. Izumi, S.J. Caughey, and C.J. Readings CJ 1976: Turbulence Structure in the Convective Boundary Layer. *J. Atmos. Sci.,* **33**, 2152 – 2169.

Haupt R.L. and S.E. Haupt, 2004: *Practical Genetic Algorithms, second ed with CD*. Wiley, New york, NY 255 pp.

Haupt S.E., 2005: A Demonstration of Coupled Receptor/Dispersion Modeling with a Genetic Algorithm. *Atmos. Environ*., **39**, 1781-1788

Hsu S.A., 2003: Nowcasting mixing height and ventilation factor for rapid atmospheric dispersion estimates on land. Nat. Wea. Dig.

Lamb R.G., 1982: Diffusion into the Convective Boundary Layer, in F.T.M. Nieustadt and H. van Dop (eds.), *Atmospheric Turbulence and Air Pollution Modelling*, D. Reidel, Dordrecht.

Long K.J., S.E. Haupt, and G.S. Young, 2010: Assessing Sensitivity of Source Term Estimation. *Submitted to Atmos. Environ., accepted*

Panofsky, H.A. and J.A. Dutton, 1984: *Atmospheric Turbulence: Models and Methods for Engineering Applications*. Wiley, New York, 437 pp.

Panofsky H.A., H. Tennekes, D.H. Lenschow, and J.C. Wyngaard, 1977: The characteristics of turbulent velocity components in the surface layer under convective conditions. *Bound.-Layer Meteor*. **11**, 355-361.

Stull R.B., 1988: An *Introduction to Boundary Layer Meteorology.* Springer 680 pp.

Weil J.C., 1988: Dispersion in the convective boundary layer. *Lectures on Air Pollution Modeling*, Venkatram A., and Wyngaard JC., Eds., Amer. Meteor. Soc., 167-227.

Weil J.C., 1990: A diagnosis of the Asymmetry in Top-Down and Bottom-Up Diffusion Using a Lagrangian Stochastic Model, *J. Atmos. Sci,* **47**, 501-515

Weil J.C., L.A. Corio, and L.A. Brower, 1997: A PDF Dispersion Model for Buoyant Plumes in the Convective Boundary Layer, J. Atmos. Sci. 36. 982-1003

Willis GE and J.C. Deardorff, 1976: A laboratory model of Diffusion into the Convective Planetary Boundary Layer, *Quart. J. Roy. Meteor. Soc*., **102**, 427-445

Willis G.E. and J.C. Deardorff, 1981: A laboratory study of dispersion from a source in the middle of the convective mixed layer, *Atmos. Environ*. **15**, 109-117.

Wyngaard J.C., 1987: A physical mechanism for the asymmetry in top-down and bottom up scalar diffusion. *J. Atmos. Sci.*, **44**, 1083-1087

Wyngaard J.C., 1988: Structure of the PBL. Lectures on Air Pollution Modeling, Venkatram A., and Wyngaard JC., Eds., *Amer. Meteor. Soc.*, 9-61.

Yaglom A.M., 1972: Turbulent Diffusion in the Surface Layer of the Atmosphere Izv. *Atmos. Ocean Phys*., **8**, 333-340.

8. Figures and Tables

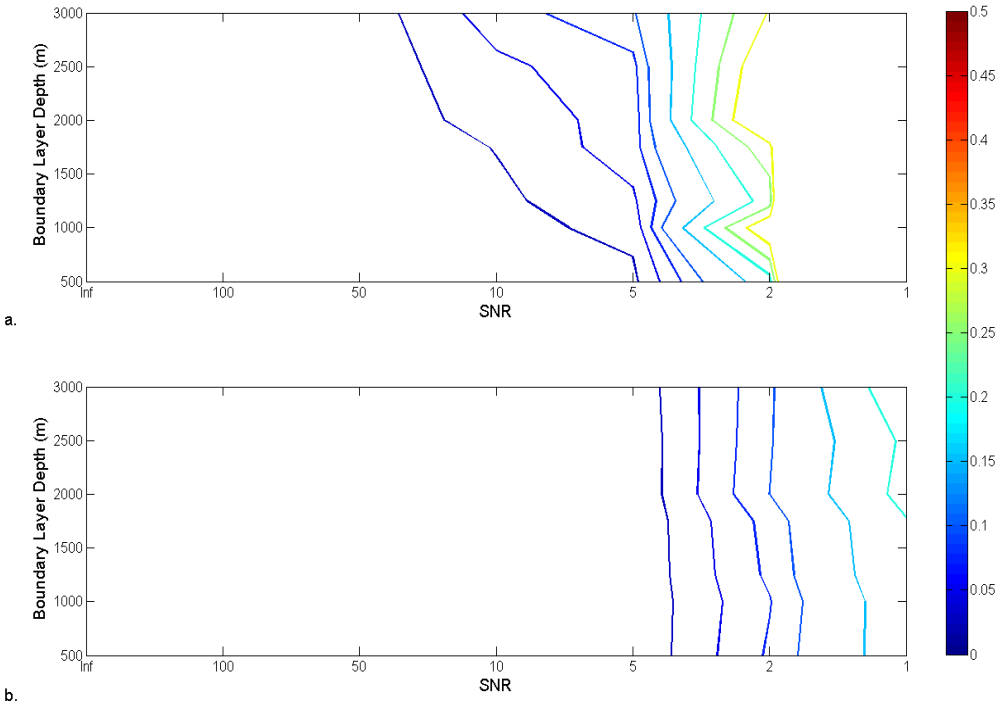

Figure 1 Contoured Skill Scores for Surface Layer Simulations. Figure a is for a continuous release and Figure b is for an instantaneous release.

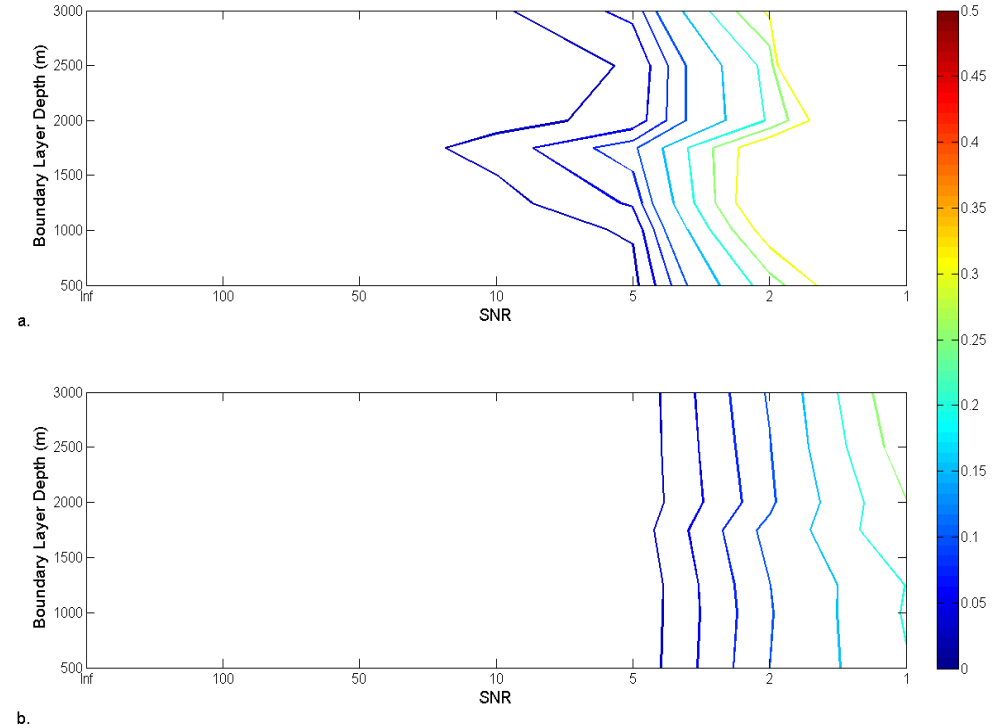

Figure 2 Contoured Skill Scores for Mixed Layer Simulations. Figure a is for a continuous release and Figure b is for an instantaneous release.

Table 1: Boundary layer depth errors for instantaneous and surface releases in the surface layer

Surface Layer Release															
SNR	**Inf**		**100**		**50**		**10**		**5**		**2**		**1**		
Boundary Layer Depth	**Error (m)**	**Error (%)**	**Error (m)**	**Error (%)**	**Error (m)**	**Error (%)**	**Error (m)**	**Error (%)**	**Error (m)**	**Error (%)**	**Error (m)**	**Error (%)**	**Error (m)**	**Error (%)**	
500	0.0	0.0	0.60	0.12	0.69	0.14	4.51	0.90	4.05	0.81	4.10	0.82	435.45	87.09	**Continuous Release**
1000	0.0	0.0	3.84	0.38	5.81	0.58	2.22	0.22	14.04	1.40	178.66	11.20	765.79	76.58	
1250	0.0	0.0	2.7	0.22	2.40	0.19	42.26	3.38	103.7	8.26	37.77	3.02	1090.6	87.25	
1750	0.0	0.0	6.99	0.40	11.39	0.65	111.26	6.36	120.68	6.90	225.51	12.89	1651.8	94.38	
2000	0.0	0.0	6.83	0.34	17.69	0.88	41.36	2.07	210.23	10.51	106.78	5.34	838.38	41.92	
2500	0.0	0.0	50.03	2.00	31.72	1.27	211.84	8.47	448.88	22.44	346.81	13.87	1004.5	40.18	
3000	0.0	0.0	4.66	0.16	20.75	0.69	40.47	1.35	381.00	12.7	679.07	22.66	2765.3	92.18	
500	0.00	0.00	0.03	0.01	0.05	0.01	7.76	1.55	0.11	0.02	4.55	0.20	60.45	12.09	**Instantaneous Release**
1000	0.00	0.00	0.02	0.00	0.04	0.00	18.10	1.81	3.07	0.03	65.52	6.56	123.78	12.38	
1250	0.00	0.00	0.06	0.00	0.05	0.00	24.54	1.96	3.02	0.03	79.73	6.37	282.32	22.59	
1750	0.00	0.00	0.07	0.00	0.1	0.01	32.28	1.84	0.12	0.01	72.17	4.12	186.00	10.63	
2000	0.00	0.00	0.15	0.01	0.03	0.00	37.92	1.90	3.02	0.15	44.83	2.24	716.00	35.80	
2500	0.00	0.00	0.57	0.02	0.35	0.01	30.30	1.21	2.05	0.08	74.27	2.97	204.00	8.16	
3000	0.00	0.00	0.26	0.01	0.015	0.00	2.51	0.08	6.25	0.21	22.93	0.76	300.00	10.0	

Table 2: Boundary layer depth errors for mixed layer release

Mixed Layer Release															
SNR	**Inf**		**100**		**50**		**10**		**5**		**2**		**1**		
Boundary Layer Depth	**Value (m)**	**Error (%)**	**Value (m)**	**Error (%)**	**Value (m)**	**Error (%)**	**Value (m)**	**Error (%)**	**Value (m)**	**Error (%)**	**Value (m)**	**Error (%)**	**Value (m)**	**Error (%)**	
500	0.00	0.00	0.04	0.00	0.0	0.00	6.29	1.25	9.47	1.89	95.65	19.13	251	50.2	**Continuous Release**
1000	0.00	0.00	0.21	0.0	0.11	0.00	8.60	0.86	18.04	1.80	35.17	3.51	695	69.65	
1250	0.00	0.00	0,96	0.0	1.51	0.00	11.93	0.95	12.91	1.03	50.42	4.07	764	61.12	
1750	0.00	0.00	0.90	0.0	5.65	0.30	5.86	0.33	7.65	0.13	296	16.91	391	22.34	
2000	0.00	0.00	4.56	0.0	3.37	0.16	4.34	0.22	54.00	2.70	426	21.30	1524	76.2	
2500	0.00	0.00	6.05	0.0	2.47	0.01	7.18	0.29	16.14	0.65	264	10.56	637	25.48	
3000	0.00	0.00	1.31	0.0	5.65	0.18	21.7	0.72	94.73	3.15	311	10.37	2225	74.16	
500	0.0	0.00	0.04	0.08	0.06	0.01	0.28	0.00	0.30	0.06	10.05	0.20	84.55	16.80	**Instantaneous Release**
1000	0.0	0.00	0.02	0.00	0.18	0.02	0.50	0.00	1.30	0.13	31.88	3.10	4.05	0.40	
1250	0.0	0.00	0.10	0.00	0.43	0.03	1.25	0.10	0.82	0.01	42.28	3.30	154.06	12.32	
1750	0.0	0.00	0.11	0.00	2.15	0.12	16.57	0.95	34.70	1.90	10.65	0.60	110.05	6.28	
2000	0.0	0.00	13.50	0.68	0.29	0.02	1.98	0.10	0.90	0.05	57.45	2.80	397.08	19.00	
2500	0.0	0.00	16.88	0.68	0.45	0.02	0.68	0.03	6.85	0.40	46.70	1.80	30.10	1.20	
3000	0.0	0.00	19.50	0.65	0.28	0.01	0.78	0.03	3.62	0.12	10.80	0.36	274.43	9.10	