SPECIAL
NOAA/EPA Golden Jubilee
ISSUE

Probabilistic Characterization of Atmospheric Transport and Diffusion

JOHN S. IRWIN

John S. Irwin and Associates, Raleigh, North Carolina

WILLIAM B. PETERSEN AND STEVEN C. HOWARD

NOAA Atmospheric Sciences Modeling Division, Research Triangle Park, North Carolina*

(Manuscript received 9 January 2006, in final form 21 September 2006)

ABSTRACT

The observed scatter of observations about air quality model predictions stems from a combination of naturally occurring stochastic variations that are impossible for any model to simulate explicitly and variations arising from limitations in knowledge and from imperfect input data. In this paper, historical tracer experiments of atmospheric dispersion were analyzed to develop algorithms to characterize the observed stochastic variability in the ground-level crosswind concentration profile. The algorithms were incorporated into a Lagrangian puff model ("INPUFF") so that the consequences of variability in the dispersion could be simulated using Monte Carlo methods. The variability in the plume trajectory was investigated in a preliminary sense by tracking the divergence in trajectories from releases adjacent to the actual release location. The variability in the near-centerline concentration values not described by the Gaussian crosswind profile was determined to be on the order of a factor of 2. The variability in the trajectory was determined as likely to be larger than the plume width, even with local wind observations for use in characterizing the transport. Two examples are provided to illustrate how estimates of variability 1) can provide useful information to inform decisions for emergency response and 2) can provide a basis for sound statistical designs for model performance assessments.

1. Introduction

Implicit in current air quality models, and in the discussion of this paper, is the assumption that the temporal and spatial variations in observed hourly concentration values can be envisioned as being partly deterministic and partly stochastic. For specified boundary conditions, the deterministic part of the concentration variations in time and space are the ensemble-average hourly concentrations to be seen at each location in the modeling domain. What we observe at any given location and time represents an individual realization from a population of possible outcomes, which "scatter" in some random fashion about the true ensemble average. Current models attempt to simulate the ensemble averages, but uncertainties arise that are due to limitations in our understanding of atmospheric processes and imperfect input data (e.g., meteorological conditions, emissions, terrain, buildings, and land use). Thus, the observed scatter of observations about model predictions is a combination of naturally occurring stochastic variations that are impossible for any model to ever explicitly simulate and variations ("uncertainties") arising from limitations in our knowledge and imperfect input data. The concept that models cannot predict what is actually observed is not new (Ramsdell and Hinds 1971; Venkatram 1983; Sykes et al. 1984; Hanna 1984; Dabberdt and Miller 2000); however, more discussion appears to be needed on the communication of the magnitude of stochastic variations to decision makers (e.g., the effect such variations have on decision making) and on the importance of stochastic variations in developing meaningful model evaluations (e.g., the statistical sampling problem posed for discerning "skill" in the midst of large stochastic variations).

We analyzed historical tracer dispersion experiments

* NOAA Atmospheric Sciences Modeling Division is in partnership with the U.S. Environmental Protection Agency.

Corresponding author address: John S. Irwin, 1900 Pony Run Road, Raleigh, NC 27615.
E-mail: jsirwinetal@nc.rr.com

DOI: 10.1175/JAM2515.1

© 2007 American Meteorological Society

to provide quantitative estimates of two sources of variability in atmospheric transport and diffusion, and we provide two examples of how a quantitative estimate of the variability of outcomes can have practical usefulness. The variability is attributed to 1) unresolved (diffusion) variability not currently characterized by the model parameterizations and 2) wind field (trajectory) variability. There are limits to what can be gleaned from these tracer experiments, as they were conducted for purposes other than investigations of concentration variability. However, they do provide results for larger downwind transport distances and more complex settings than the concentration fluctuation laboratory and field studies conducted to date. It is beyond the scope of this discussion to provide mechanistic explanations for the nature and causes of variability, other than to recognize, as discussed above, that it has many sources and is an active area of research (e.g., Jones and Thomson 2006; Cassiani et al. 2005a,b; Luhar and Sawford 2005a,b; Weil et al. 2002).

2. Characterizations of unresolved variability

a. Defining what is "unresolved"

For atmospheric transport and dispersion, regardless of the sophistication of the model employed, we have concluded that we cannot reproduce exactly what is observed at a given time and location; however, we can hope to predict the average characteristics of the concentration distribution (e.g., the variance of the distribution of outcomes) seen at given locations. The problem of predicting the onset, duration, and intensity of a precipitation event (which involves the transport and dispersion of moisture in the atmosphere) is routinely viewed and the model output is cast in a probabilistic manner. As another analogy, consider the problem of predicting the outcomes from a series of tosses of a pair of dice. We cannot predict exactly the sequence of individual outcomes in a series of tosses, but we can predict the distribution of outcomes and their respective probabilities of occurrence, including the mean, variance, and other moments of the distribution. One difference between modeling atmospheric transport and diffusion and dice tossing is that the physics of dice tossing is well known, whereas we have much to learn regarding atmospheric transport and diffusion processes. The point is, though, even with more complete characterizations of atmospheric transport and diffusion physics in our models, we cannot reproduce exactly what is observed at a given time and location, and hence, the prediction of atmospheric transport and diffusion (as with most environmental and meteorological processes) is best characterized in a probabilistic sense.

The characterization of the unresolved variance, often referred to as "inherent uncertainty" (e.g., Venkatram 1983; Fox 1984) from field data, ultimately involves either an analysis of residuals about predictions or an analysis of residuals about observed averages. In either case, it is necessary to understand the differences to be seen between what is simulated by a model and that which is observed. The observed concentrations C_o can be envisioned as (adapted from Venkatram 1988)

$$C_o = C_o(\alpha, \beta) = \overline{C_o}(\alpha) + c(\Delta c) + c''(\alpha, \beta), \qquad (1)$$

where α are the model input parameters, β are the variables needed to describe the unresolved transport and dispersion processes, the overbar represents an average over all possible values of β for the specified set of model input parameters α, $c(\Delta c)$ represents the effects of measurement uncertainty of the concentration values, and $c''(\alpha, \beta)$ represents our ignorance of β (unresolved deterministic variations and stochastic fluctuations). Because $\overline{C_o}(\alpha)$ is an average over all β, it is only a function of α. The modeled concentrations C_m can be envisioned as

$$C_m = [\overline{C_o}(\alpha) + f(\alpha)] + d(\alpha) + [c''(\alpha, \beta) + g(\Delta c'')], \qquad (2)$$

where $f(\alpha)$ is the deterministic error in the estimate of $\overline{C_o}(\alpha)$ and $d(\Delta\alpha)$ represents the effects of uncertainty in specifying the model inputs. The next two terms in (2) are not present in most current operational atmospheric transport and diffusion models, because they represent an attempt to estimate the unresolved variability, $c''(\alpha, \beta)$, and any deterministic error, $g(\Delta c'')$, in the estimate of $c''(\alpha, \beta)$. Although we have explained (1) and (2) in terms of observed and estimated concentration values, anything that can be observed (or derived from observations) and is predicted by a model (e.g., plume rise, deposition, lateral dispersion) can be substituted for "concentration."

The sources for variance in (1) and (2) are different. This means that if we attempt to estimate $c''(\alpha, \beta)$ from an analysis of residuals involving $(C_o - C_m)$ or C_o/C_m, then the estimate of $c''(\alpha, \beta)$ will be inflated by any deterministic error in the model's estimate of $\overline{C_o}(\alpha)$, $f(\alpha)$, measurement uncertainties $c(\Delta c)$, and uncertainty in specifying the model inputs $d(\Delta\alpha)$. The variance due to $d(\Delta\alpha)$ (input uncertainty) has been estimated to be on the order of the magnitude of the ensemble averages (see Irwin et al. 1987), which is similar in magnitude to estimates of the variance due to natural variability (see Hanna 1993). This suggests that the analysis of residuals involving $(C_o - C_m)$ or C_o/C_m may be problematic. As a further incentive to investigate $c''(\alpha, \beta)$ empirically, Weil et al. (2002) noted that laboratory studies of

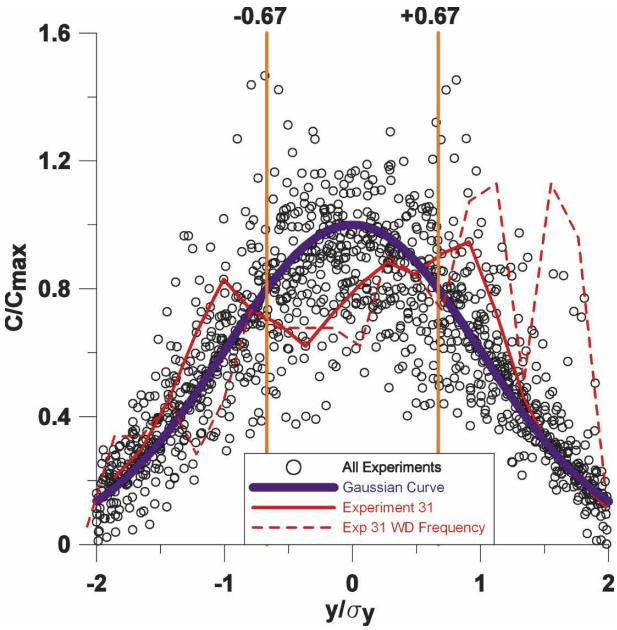

FIG. 1. Illustration of the natural variability that is not characterized in the crosswind of a dispersing puff or plume by a Gaussian puff or plume model (American Society for Testing and Materials 2005). Results depicted are for the 50-m downwind arc of the PG experiment. Solid red line depicts $C/C_{\max}$ values for expt 31. Dashed red line depicts normalized wind direction frequencies of occurrence for comparison with actual lateral concentration profile seen for expt 31.

concentration fluctuations suggest $c''(\alpha, \beta)$ decreases as transport distance increases, whereas empirical investigations sometimes suggest otherwise (see discussion of Weil et al. 2002, their Fig. 18). Given this, we have chosen to explore the effects of unresolved variations by a direct analysis of observations, in essence by looking at the variance about observed averages. This is not without problems, as uncertainties in empirically determining the observed averages and the variance will lead to the overestimation of the magnitude of $c''(\alpha, \beta)$. However, this does allow our estimates $c''(\alpha, \beta)$ to be model-independent.

b. *Crosswind concentration profile variations*

The widely used Gaussian approximation for characterizing the crosswind distribution of mass of a dispersing plume as it is carried downwind provides an ensemble-average view of what is really seen in the world. Figure 1 illustrates the observed concentration during Project Prairie Grass (PG; near O'Neill, Nebraska), averaged over 10 min, measured by near-surface sampling along a 50-m arc downwind of a near-surface point source release of sulfur dioxide. For each 10-min experiment or realization, the crosswind receptor positions y relative to the observed center of mass along the arc have been divided by σ_y, which is the second moment of the lateral concentration distribution along the arc for that experiment, and the observed 10-min concentration values have been divided by $C_{\max} = C^Y/[\sigma_y(2\pi)^{1/2}]$, where C^Y is the crosswind integrated concentration along the arc.

If one looks at the visual impression given by all the individual experiments plotted in Fig. 1, the crosswind concentration profile is seen to follow a Gaussian shape on average, which is what is predicted by all atmospheric transport and diffusion models, regardless of their sophistication. However, as illustrated by the results shown for experiment 31, the actual crosswind profile may not necessarily be Gaussian for any single experiment (e.g., the results for experiment 31 in Fig. 1). The PG and Round Hill (near South Dartmouth, Massachusetts) experiments are unique in that they provide histograms of wind direction frequencies of occurrence during the tracer release periods. We anticipated that if one actually had such information, it might improve the prediction of lateral crosswind concentration profile (see red dashed line in Fig. 1), which is the normalized wind frequency such that it represents the lateral crosswind profile that would be predicted utilizing the reported wind direction frequencies of occurrence during the tracer sampling interval. When we correlated the wind direction frequencies with the lateral concentration profiles, we found that there was some improvement in several cases, but overall a Gaussian profile was seen to be equally effective in characterizing the lateral profile. This suggests that summary statistics of the wind direction variations for the tracer release are interesting, but insufficient in themselves to improve predictions of crosswind concentration profiles. It would appear that an actual time history of wind direction fluctuations at each location downwind is needed, to provide a better prediction of the lateral concentration profile than that provided by a Gaussian assumption. These results confirm our belief that regardless of the sophistication of the model and its inputs, operational dispersion models cannot predict the observed variation in ground-level concentration values at specific locations and times, except in a statistical sense.

Previous investigations of the concentration fluctuations about the Gaussian profile have determined that they can be well characterized by using clipped normal distributions near the release and using lognormal distribution farther downwind. From historical tracer dispersion experiments, with averaging times of no less than 10 min and typically average concentrations of 30 min or longer, Irwin and Lee (1996) found that a lognormal distribution seemed to well characterize the centerline concentration fluctuations. To extend the re-

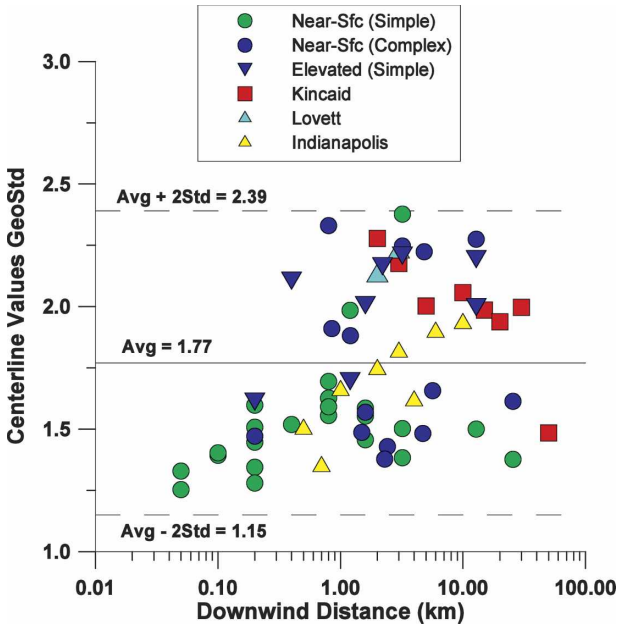

FIG. 2. Summary of GeoStd determined for near-centerline concentration fluctuations about a Gaussian fit for the crosswind concentration distribution.

sults presented by Irwin and Lee (1996), an analysis was conducted of 13 tracer dispersion experiments, each of which had intensive near-ground level sampling along crosswind arcs at various distances downwind of the release.[1] Our purpose is to investigate the unresolved variability that a Gaussian lateral profile does not characterize; hence, we have not removed observations from consideration that might have experienced nonsteady-state conditions (e.g., more than one mode along one or more crosswind arcs). For instance, what we have called Green Glow I are the 16 "steady-state" experiments selected for analysis by Fuquay et al. (1964), and Green Glow II are the 10 experiments that were deemed by them to have nonsteady conditions. This means that nonsteady conditions were identified to have occurred in about 38% of the Green Glow experiments (near Richland, Washington). A Gaussian fit (as described above) was computed for each release at each arc, and statistics were computed for all $c/C_{\max}$ ratio values for "centerline" receptors ($y < |0.67\sigma_y|$). The definition of the position of centerline receptors follows the American Society for Testing and Materials (2005). Results were tabulated only for arcs having at least 50 ratio values for analysis. In these analyses, and those that follow, the following expressions were used for determining the geometric average (GeoAvg) and the geometric standard deviation (GeoStd):

$$\mathrm{GeoAvg} = \exp\left(\sum_{i=1}^{N} c_i^*/N\right) \quad \text{and} \tag{3}$$

$$\mathrm{GeoStd} = \exp\left\{\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}[c_i^* - \ln(\mathrm{GeoAvg})]^2}\right\}, \tag{4}$$

where $c_i^* = \ln(c_i/C_{\max})$, c_i are the centerline concentration values, and N is the number of values.

Figure 2 depicts the results obtained, summarized into six groups. Tables 1–4 provide the statistics for each experiment. The "near-surface simple" group consists of PG, Round Hill, Hanford-30, Green Glow I, and Hanford-67 (near Richland, Washington) and involves releases at or below 2 m in nearly flat terrain with steady-state meteorological conditions. The "near-surface complex" group is Green Glow II, Ocean Breeze (Cape Canaveral, Florida), and Dry Gulch (Vandenberg Air Force Base, California) and involves releases at or below 2 m in complex nonsteady meteo-

TABLE 1. Summary statistics for normalized centerline concentration fluctuations ($c/C_{\max}$). The results listed are for the near-surface releases (simple), where X is the downwind distance, N is the number of centerline values, Avg is the average, std dev is the standard deviation, GeoAvg is the geometric average, GeoStd is the geometric std dev, z_0 is surface roughness length, and duration is the concentration averaging time.

X (km)	N	Avg	Std dev	GeoAvg	GeoStd
PG (z_0 = 0.6 cm, duration = 10 min)					
0.05	406	0.96	0.20	0.93	1.25
0.10	366	0.93	0.30	0.89	1.39
0.20	329	0.92	0.35	0.85	1.51
0.40	288	0.93	0.34	0.86	1.52
0.80	506	0.90	0.40	0.81	1.63
Round Hill I (z_0 = 10 cm, duration = 10 min)					
0.05	171	0.92	0.26	0.89	1.33
0.10	155	0.93	0.29	0.88	1.40
0.20	134	0.95	0.34	0.89	1.45
Hanford 30 (z_0 = 3 cm, duration = 20–75 min)					
0.20	118	0.97	0.28	0.93	1.34
0.80	109	0.99	0.47	0.88	1.69
1.60	198	0.92	0.36	0.84	1.59
3.20	202	0.98	0.60	0.81	2.38
Green Glow I (z_0 = 3 cm, duration = 30 min)					
0.20	74	0.94	0.23	0.91	1.28
0.80	66	0.96	0.37	0.88	1.56
1.60	107	0.93	0.40	0.85	1.55
3.20	96	0.93	0.36	0.86	1.50
12.80	104	1.08	0.53	0.99	1.50
25.60	246	0.96	0.28	0.92	1.38

[1] Space limitations preclude providing complete descriptions of all of the experiments. References and descriptions for all the classical tracer experiments discussed in this paper are provided in Irwin (1983, 1984). The Kincaid, Lovett, and Indianapolis experiments are described in Environmental Protection Agency (2003).

TABLE 2. Same as in Table 1 but for near-surface releases (complex).

X (km)	N	Avg	Std dev	GeoAvg	GeoStd
Hanford-67 (z_0 = 3 cm, duration = 30 min)					
0.20	101	0.88	0.38	0.80	1.60
0.80	87	0.89	0.35	0.81	1.59
1.20	55	0.93	0.61	0.74	1.99
1.60	109	0.86	0.30	0.81	1.46
3.20	103	0.84	0.25	0.81	1.38
Green Glow II (z_0 = 3 cm, duration = 30 min)					
0.20	62	0.98	0.39	0.90	1.47
0.80	59	1.00	0.74	0.75	2.33
1.60	67	1.07	0.50	0.98	1.57
3.20	103	1.20	1.06	0.88	2.25
12.80	118	1.35	1.50	0.99	2.28
25.60	109	1.06	0.57	0.94	1.61
Ocean Breeze (z_0 = 5–15 cm, duration = 30 min)					
1.21	240	0.91	0.40	0.80	1.88
2.41	216	0.93	0.27	0.88	1.43
4.83	101	1.02	0.94	0.85	2.22
Dry Gulch (z_0 = 5–15 cm, duration = 30 min)					
0.85	232	0.98	0.88	0.85	1.91
1.50	204	0.98	0.35	0.91	1.49
2.30	201	0.93	0.27	0.89	1.38
4.72	335	1.00	0.39	0.93	1.48
5.67	355	0.91	0.47	0.82	1.66

TABLE 4. Same as in Table 1 but for elevated releases (complex).

X (km)	N	Avg	Std dev	GeoAvg	GeoStd
Kincaid (z_0 = 5–14 cm, duration = 60 min)					
2.00	122	1.00	1.14	1.01	2.28
3.00	235	0.88	0.85	0.96	2.18
5.00	277	0.93	1.10	1.00	2.00
10.00	128	1.29	1.63	1.12	2.06
15.00	67	0.91	0.96	1.04	1.99
20.00	71	0.90	1.05	1.02	1.94
30.00	67	1.12	2.17	1.61	2.00
50.00	63	0.91	0.47	0.88	1.48
Lovett (z_0 = 15–30 cm, duration = 60 min)					
2.00	604	1.07	1.01	1.01	2.12
3.00	23 656	0.81	1.10	0.83	2.22
Indianapolis (z_0 = 30 cm, duration = 60 min)					
0.50	52	0.98	0.62	0.96	1.50
0.70	103	0.96	0.41	0.96	1.35
1.00	374	0.98	0.62	0.92	1.66
2.00	244	1.07	0.92	1.01	1.74
3.00	208	1.11	0.88	1.00	1.81
4.00	159	1.14	0.61	1.04	1.62
6.00	187	1.22	0.96	1.04	1.90
10.00	99	1.21	1.10	1.03	1.93

rological conditions. The "elevated simple" group is Hanford-67 and Hanford-64 and involves elevated releases mostly at 26 and 56 m, with a few at 111 m over nearly flat terrain. The last "elevated complex" group (Kincaid; Lovett; and Indianapolis, Indiana) involved tracers injected into the exhaust gases of operating electric power generation plants. The stacks were 187, 145, and 87 m in height for Kincaid, Lovett, and Indianapolis, respectively. Kincaid is located in rural, relatively flat terrain near Springfield, Illinois. Lovett is located in rural, complex terrain near Stony Point, New York. The Indianapolis release and initial sampling arcs were in the suburbs, and the final sampling arcs were in city center.

As mentioned before, determining the statistical properties of concentration values empirically does involve some uncertainty. In an effort to minimize these effects, we only provide statistics for arcs having at least 50 values for analysis. The uncertainty in GeoStd values increases as sample size decreases and as the GeoStd value increases. For a sample size of 50, the uncertainty in the GeoStd increases from about 3% to 20% as the value of the GeoStd increases from about 1.2 to 3. For a sample size of 500, the uncertainty in the GeoStd increases from about 1% to 6% as the value of the GeoStd increases from about 1.2 to 3.

We did not see any definite correlation in the variation statistics for the centerline concentration fluctuations as a function of release height, surface roughness length, or averaging time, but we have provided the detailed results in Tables 1–4 for inspection. For those near-surface releases in nearly flat terrain with steady-state meteorological conditions, the GeoStd of the near-centerline concentration values is about 1.5 for all downwind distances. The average GeoStd for all the results depicted is 1.8 with a variability (standard deviation) of 0.62. A GeoStd equal to 1.8 means that 95% of the centerline c/C_{max} ratio values are within about a factor of 3 of C_{max}. There is a suggestion in looking at the results shown in Fig. 2 that the GeoStd of the near-centerline concentration values may increase as transport distance increases, at least out to 1 km. However, this may be a false impression, because the results shown for transport less than 1 km involve the simplest

TABLE 3. Same as in Table 1 but for elevated releases (simple).

X (km)	N	Avg	Std dev	GeoAvg	GeoStd
Hanford-67 (z_0 = 3 cm, duration = 30 min)					
0.20	95	0.98	0.52	0.89	1.61
0.40	122	1.03	0.87	0.83	2.11
1.20	206	0.92	0.47	0.81	1.70
2.20	150	0.96	0.70	0.76	2.17
12.80	152	1.13	0.76	0.93	2.00
Hanford-64 (z_0 = 3 cm, duration = 30 min)					
1.60	75	0.94	0.53	0.77	2.01
3.20	108	1.05	0.59	0.83	2.21
12.80	77	0.89	0.69	0.68	2.19

of situations. If we had data for transport of less than 1 km involving more complex settings, our perception of the results shown in Fig. 2 might change.

The averaging times of concentration values shown in Fig. 2 range from 10 min (Round Hill and PG) to 1 h (Kincaid and Indianapolis), but most of the data have averaging times of 30 min or longer. The Round Hill and PG experiments have nearly identical experimental designs. The 29 Round Hill I releases were conducted in 1955 and in essence tested the setup for the more extensive PG experiments conducted in 1956. An objective of the 10 Round Hill II releases (conducted in 1957 after the completion of PG) was to investigate the effects of averaging time on atmospheric transport and diffusion. Sulfur dioxide concentrations were sampled along 3 arcs (50, 100, and 200 m), and the release height and sampling height was 1.5 m. Samples were taken for the first 30 s and the first 3 min of each 10-min sample. The average ratio over all arcs of C_{max}(0.5 min)/C_{max}(10 min) is 1.66 and of C_{max}(3 min)/C_{max}(10 min) is 1.42, which as expected shows that the maximum concentration decreases as averaging time increases, with all other factors being equal. Also as expected, the lateral dispersion seems to increase as averaging times increase. The GeoStd of the near-centerline concentration values about C_{max} was nearly invariant ranging from 1.28 to 1.23 to 1.27, in going from 0.5- to 3- to 10-min samples. Because the GeoStd is a measure of the relative scatter about the C_{max} and because C_{max} was seen to increase as averaging time decreases, this suggests that the actual variability may increase as averaging time decreases.

The investigation of whether the variability of the near-centerline concentration values is stability-dependent requires sufficient data to stratify the values along an arc into various "atmospheric stability" groups. The only experiments having sufficient data for this purpose are PG and Kincaid, which are very different situations, because PG involves near-surface releases in nearly ideal circumstances and Kincaid involves a highly buoyant release from a tall stack. Figure 3 shows the variation of the GeoStd as a function of atmospheric stability (where L is the Monin–Obukhov length). We have shown only the results for three arcs from each experiment, because they are illustrative of what is found at the other arcs from each experiment. The decrease in the GeoStd as we go from unstable ($1/L < 0$) to stable ($1/L > 0$) conditions is statistically significant for PG but is not statistically significant for Kincaid. For PG, convective eddies may be responsible for the larger values of the GeoStd during very unstable conditions. However, another possible explanation may be that the 10-min averaging time is too short for sampling adequately a dispersing plume during convectively unstable conditions (where convective eddies may be on the order of 1 km in depth and thus may take longer than 10 min to move across an arc). With 1-h samples for PG, we might see different results.

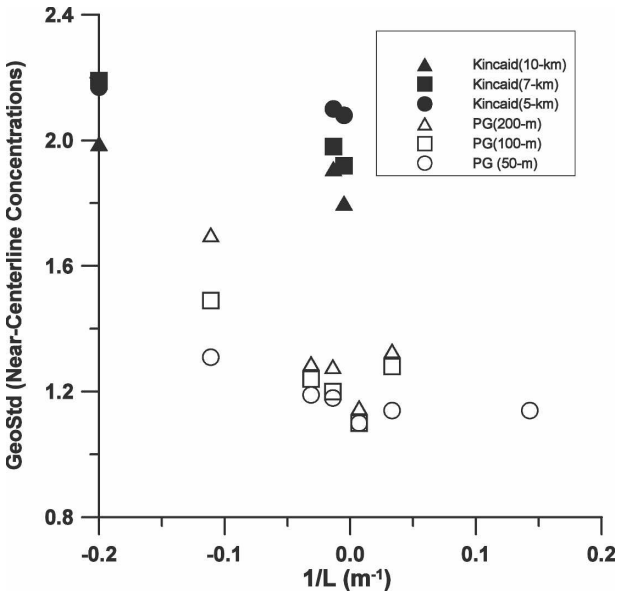

FIG. 3. The variation of the GeoStd of near-centerline concentration values as a function of atmospheric stability ($1/L$), where L is the Monin–Obukhov length. Results are shown for 3 arcs from PG (50, 100, and 200 m) and for 3 arcs from Kincaid (5, 7, and 10 km).

We conclude that the variability of the near-centerline concentration values not resolved by the assumption that the lateral profile is of a Gaussian shape can be approximated as having a lognormal distribution with a GeoStd on the order of 1.5–2.5, depending on the complexity of the situation being characterized (e.g., dispersion over flat uncomplicated rural terrain versus an urban city). Such variations are sufficiently large to explain why factor of 2 differences in concentration values are observed between what is measured (an event or individual realization) and what is predicted as the average concentration (first moment). For instance, we can expect peak-to-mean ratios of centerline concentration values greater than 2 to occur 15% of the time, and peak-to-mean ratios of centerline concentration values greater than 5 to occur about 1% of the time. We have limited evidence from one experiment with averaging times ranging from 0.5 to 10 min that the GeoStd of the near-centerline concentration values (which is a relative measure of variability) may not be strongly affected by variation in the averaging time, and we have limited evidence that the GeoStd may be largest during convectively unstable conditions.

c. Dispersion parameter variability

Irwin (1984) calculated the bias in the dispersion parameter (σ_y and σ_z) estimates and observed that the bias varied from one site to the next. Irwin (1984) also calculated the random errors about the systematic bias at each site. To explore further these uncertainties, an analysis was conducted of the 26 tracer field experiments listed and discussed in Irwin (1983). For each experiment we 1) computed the average and geometric mean of ratio P/O, where P is the predicted and O is the observed growth rate of the dispersion (i.e., the increase in the lateral or vertical dispersion in going from one arc to the next downwind arc), and 2) computed the standard deviation and geometric standard deviation of the P/O ratio values. We limited the analysis to transport distances of less than 5 km. For the current analysis, Model 3 as described in Irwin (1983) was used for the predictions, because it had been found to have the best overall performance of those tested in Irwin (1983). Table 5 summarizes the results obtained from the analysis described. A lognormal distribution was seen to be a reasonable characterization for all of the random error distributions, even though a normal distribution is seen to be indicated at 10 experiment sites (see notations in Table 5). We looked to see if the variability in the growth rates had a distance dependence or release height dependence, but such was not seen.

Assuming that the random biases and random errors in the dispersion parameter growth rates come from independent lognormal distributions, we can model the variability in the growth rates of the dispersion parameters as $\Delta\sigma_{y,z} = b_{y,z} r_{y,z} \Delta\sigma^o_{y,z}$, where the subscripts y and z respectively refer to the lateral and vertical dispersion, b and r are random bias and error factors, $\Delta\sigma^o_{y,z}$ is the model's estimate of the increase in the dispersion parameter, and $\Delta\sigma_{y,z}$ is the simulated increase including the effects of variability. We can use the Table 5 results to characterize the distributions of b and r. We can characterize the 26 biases as a lognormal distribution with a GeoStd of 1.48 (e.g., $b_{y,z}$), and we can characterize the 26 GeoStd values by their average, 2.02 (e.g., $r_{y,z}$). Note, a lognormal distribution with a GeoStd of 1.5 means 90% of the values are within a factor of 2.

For now, our focus is on whether growth rate variations would greatly affect the variability of near-centerline concentration values. To investigate this, we simulated the variability in the growth rate of the dispersion coupled with the variability in the lateral profile for non-Gaussian effects with a modified version of a Lagrangian puff model ("INPUFF"; Petersen and Lavdas 1986). Note that any model that simulates dispersion in a Lagrangian sense would have served our purpose. In our implementation, we made multiple runs of INPUFF, and for each run we randomly picked separate and independent values for b_y and b_z, selecting them from a lognormal distribution. Each time the meteorological conditions were updated, which in our case was hourly, we randomly picked separate and independent values for r_y and r_z from a lognormal distribution with a GeoStd. Given b_y, b_z, r_y, and r_z, which were assumed to be spatially invariant over the entire modeling domain, we then adjusted the predicted vertical and lateral growth rates for each puff. The assumption that b_y, b_z, r_y, and r_z are the same everywhere in the modeling domain is admittedly simplistic and can be modified in the future when we have some basis to do so.

It was seen that variability in the vertical and hori-

TABLE 5. Summary of comparison of Irwin's (1983) Model-3 predictions of the growth rate of vertical and lateral dispersion with field data from 26 sites. Asterisk denotes cases in which a normal distribution best characterizes the random errors but for which we also found that a lognormal distribution fits nearly as well. Established in 1949, the National Reactor Testing Station (NRTS) is between Arco and Idaho Falls, ID.

Expt site	No.	Bias	GeoStd
Elevated lateral dispersion sites			
Hanford(64) 56 m	11	1.29	3.05*
Hanford(67) 56 m	46	0.99	2.66*
Hanford(67) 26 m	158	1.22	2.13
NRTS	80	1.10	1.22
Karlsruhe, Germany	26	2.44	2.55
Hanford	48	1.07	1.82*
Suffield, AB, Canada	80	1.07	1.63
Elevated vertical dispersion sites			
Ågesta, Sweden	21	0.82	1.66*
Karlsruhe	58	1.12	2.24
Hanford	13	0.40	2.01*
NRTS	80	0.62	1.36*
Near-surface lateral dispersion sites			
Mount Iron, Vandenberg AFB, CA	49	1.29	2.02
NRTS-B	31	0.98	2.24*
NRTS-A	66	0.71	1.67*
Hanford 30	83	1.19	1.75
Green Glow	44	1.30	1.67
PG	251	0.81	2.02
Dry Gulch	98	0.62	2.03
Ocean Breeze	101	0.59	1.69
Round Hill II	47	0.75	1.62
Round Hill I	20	0.54	1.63*
Hanford(67) 2 m	64	1.00	2.30
Near-surface vertical dispersion sites			
NRTS-B	74	1.69	3.03
NRTS-A	25	1.70	2.79*
PG	154	0.92	1.92
Round Hill I	32	1.16	1.77

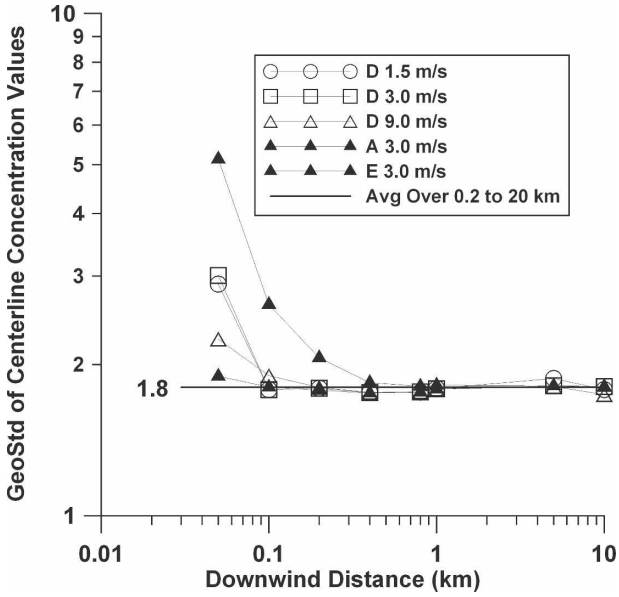

FIG. 4. The GeoStd determined for the centerline concentration values for several idealized situations having steady-state conditions for 200 h. In these simulations variability was simulated in the lateral concentration profile and in the growth rates of the dispersion. In the legend is presented the Pasquill stability category and the horizontal wind speed.

zontal puff growth rates mostly affected concentration values for locations on the edges of the puffs, where concentration values are small, but the relative variability is high considering the difference between a zero concentration and a value not quite zero. It was also seen that the variability in the growth rates affected centerline concentration values in the near field when the puffs are small and the growth rates are at a maximum (see Fig. 4). Once puffs attain some size, the centerline fluctuations were seen to result primarily from the fluctuations imposed on the lateral concentration profile.

d. Puff trajectory variability

Three studies will be summarized here that provide strong evidence that the assumption that the local wind field is homogeneous over some broad area is suspect. In the first study, Finkelstein et al. (1986) compared 6 wind sensors located on 10-m masts, located about 25 m west of the National Oceanic and Atmospheric Administration (NOAA) Boulder Atmospheric Observatory. The towers were approximately 5 m apart. Twenty-minute wind observations for a 7-h daytime period and 7-h nighttime period during 9 September 1982 were compared. Small biases in the wind speeds on the order of 0.4 m s^{-1} and in the wind directions on the order of 2.7° were seen between the instruments. The standard deviation of the differences in wind speed ranged from 0.30 to 0.35 m s^{-1}, and the standard deviation of the differences in wind direction ranged from 4.0° to 5.2°. In the second study, Lockhart and Irwin (1980) compared differences in wind observations made at 25 sites with separation distances of 3–100 km. Eight of the wind sensors were at 10-m height while the other 17 were at a 30-m height. Comparisons were made of the hourly wind observations taken during 1976, where the data recovery was better than 90%. Biases were generally less than 1 m s^{-1} and 5° between the 25 stations. The standard deviation in the differences in wind speed ranged from 0.7–1.0 m s^{-1} at 3-km separation to 1.4–1.7 m s^{-1} at 90-km separation. Extrapolation of the results provides an estimate of 0.47 m s^{-1} for the standard deviation at a separation of 1 km, which is close to that seen by Finkelstein et al. (1986). The standard deviation in the differences in wind direction ranged from 17°–25° at 3-km separation to 37°–45° at 90-km separation. Extrapolation of the results provides an estimate of 15° for the standard deviation at a separation of 1 km, which is somewhat larger than that seen by Finkelstein et al. (1986). In the third study, Hanna and Yang (2001) compared winds predicted for a 9-day period by the Regional Atmospheric Modeling System and fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) for the 12-km modeling domain covering the eastern United States. For these comparisons, the root-mean-square difference between what was predicted and what was observed was less than 1.9 m s^{-1} for wind speed and on the order of 60° for wind direction. They also compared MM5 predictions for a 4-day period for the 4-km grid in the central California SARMAP domain (SARMAP is a complicated acronym denoting the SJVAQS/AUSPEX Regional Model Adaptation Project modeling and data analysis program, where SJVAQS/AUSPEX is the San Joaquin Valley Air Quality Study with Atmospheric Utilities Signatures, Predictions, and Experiments field management program) where four-dimensional data assimilation was employed. For these comparisons, the root-mean-square difference between predictions and observations was 2.5 m s^{-1} for the wind speed and 66° for the wind direction. It is admitted that these results are anecdotal, but they do suggest that local wind observations (and accordingly the local transport of gases and particulate in the atmosphere) can differ significantly from that suggested by the mesoscale and synoptic wind patterns.

Irwin and Smith (1984) using Green Glow tracer results and Weil et al. (1992) using Kincaid tracer results concluded that the microscale transport direction of a

plume can be defined to be about 25% of the overall width of the plumes, which typically are on the order of 20° in width. Irwin and Hanna (2005) analyzed the following list of experiments for which there was sufficient information to compare directly the transport wind direction (as indicated by a wind direction sensor near the release) with the actual transport directions (as indicated by the location of the center of mass of the tracer at each downwind arc): PG, Green Glow, Hanford 30, Hanford 64, and Hanford 67. The Hanford 64 study involved releases at 26 m, and the Hanford 67 study involved releases at 2, 26, and 56 m. The variability in the transport direction could be characterized as having a Gaussian distribution with a standard deviation of approximately 4.5°, which is in good accord with the differences seen in local wind observations by Finkelstein et al. (1986). The transport direction variability is substantially less for the tracer experiments than might be inferred from Hanna and Yang's (2001) MM5 comparisons. So Irwin and Hanna (2005) concluded that the results determined for these tracer experiments represent a lower bound on the wind speed and transport variability.

Previous investigations of the variability of "ensembles" of trajectories (e.g., Stunder 1996; Stohl 1998; Harris et al. 2005) have shown that the dispersion of the trajectories within an ensemble increases with distance and the rate of dispersion is dependent on the synoptic situation. To extend these results to examine the variability of the trajectories relative to the width of the dispersing plume, we conducted an investigation using approximately thirty 24-h Eta forecasts (yeardays 182–212, 2005), which are publicly available and have a horizontal grid size of 12 km. We selected four locations along the eastern United States (New York, Washington, Atlanta, and Miami). The puff trajectories were simulated using INPUFF, which cannot treat the effects of variations in the winds as a function of height. A puff was released at the start of the 0000 UTC forecast from each of the 8 cells surrounding each central location plus 1 from the central location and tracked for the 24 h of the forecast.

An analysis of the 0000 UTC 10-m and first-level (midpoint is at approximately 75 m) winds at the 9 grids suggests that for this period in 2005, the standard deviation of the differences in wind direction between these adjacent grids at both levels is less than 6° and 1 m s^{-1}. These differences are similar to what Finkelstein et al. (1986) found for 10-m masts located within 5–30 m of one another, and these differences are much less than those determined by Lockhart and Irwin (1980) for the St. Louis metropolitan area. We cannot confirm but we anticipate that use of such mesoscale wind analyses will underrepresent the possible variation in the trajectory outcomes, since only mesoscale and larger-scale variations are represented in such wind fields.

At the end of each hour, the median separation of the puffs from the central puff was determined as well as the central puff's σ_y. Trajectories were generated using the 10-m winds and the winds for the first layer (midlevel of which was at 75 m) at each location. At all four locations (see Fig. 5), the separation of the puff trajectories for both sets of winds was seen to be greater than the puff's lateral dimensions ($4\sigma_y$) at least out to 100 km.

3. Examples and implications

a. Estimates for emergency response applications

The variability in the transport is likely larger than the entire plume or puff width. So, for emergency response planning, what is needed is a forecast of the trajectory variability plus an estimate of the probability distribution of the centerline concentration values as a function of transport downwind distance (Dabberdt and Miller 2000). For now, since ensemble mesoscale meteorological forecasts are not routinely available, the trajectory variability could be characterized using the technique we applied to generate nine trajectories, and collocating the results from the central release point. This would generate a figure similar to Fig. 6. The probability distribution of centerline concentration values can be generated by estimating the centerline concentration values, and then applying a lognormal distribution having a geometric mean of 1.0 and a geometric standard deviation of 2.0 to each value. This would generate a figure similar to Fig. 7. For planning exercises, summarizing the consequences of a release using illustrations like Fig. 6 and Fig. 7 would provide decision makers a sense of the variability to be seen in the transport and also a sense of at what distance concentration values of importance might exist. Combining the results of Figs. 6, 7 with a prescribed concentration value of concern provides a forecast of the downwind area where emergency services may be required.

The example provided is meant to illustrate how the probabilistic nature of short-term concentration values might be displayed for use in emergency assessment applications. The empirical estimate of the variability of the concentration values used in this example is but one approach available. There are operational models that provide predictions of the first and second moment of the concentration values. One such model is "SCIPUFF" (Sykes et al. 1998), a Lagrangian transport and diffusion puff model for atmospheric dispersion ap-

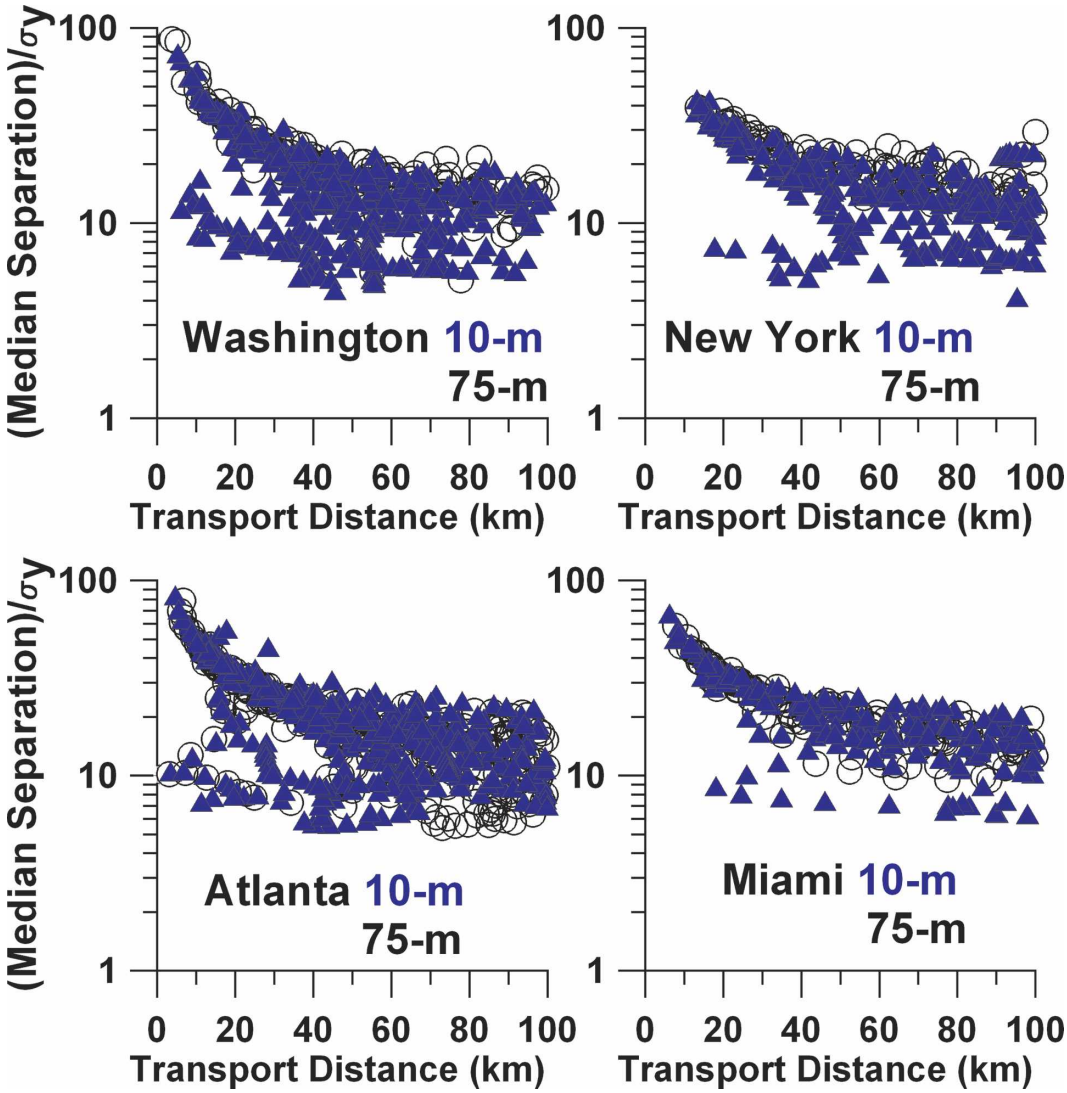

FIG. 5. Median separation of nine puffs at the end of each hour divided by the central puff's lateral dispersion (σ_y) for Washington D.C., New York, Atlanta, and Miami. Results are for INPUFF simulations conducted using Eta 12-km 0000 UTC forecasts for yeardays 182–212, 2005. Results for 10-m winds are shown with triangles and results using the first-layer winds (midlevel is at about 75 m) are shown with open circles.

plications. The closure model used in SCIPUFF has been applied on local scales up to a 50-km range (Sykes et al. 1988) and also on continental scales up to a 3000-km range (Sykes et al. 1993). The intent of this example was not to recommend a specific approach, but to show how the stochastic nature of estimated concentration centerline values can be conveyed (Fig. 6), to show how the uncertainty in plume trajectory can be conveyed (Fig. 7), and to stimulate discussion on other approaches. The point being that we believe it is time for the probabilistic nature of the problem to be openly discussed not only in the research community but with those conducting operational assessments of air quality impacts and with decision makers.

b. Implications for model evaluation studies

The variability in the trajectory is likely larger than the entire plume or puff width, which will preclude a pairing in time and space of model and observation results for the statistical evaluation of dispersion model performance. A better approach would be to separately characterize the differences seen in the transport and the differences to be seen in the centerline concentration values (once transport uncertainties are mitigated). The American Society for Testing and Materials (ASTM) guide D 6589 (American Society for Testing and Materials 2005) outlines a procedure for conducting a statistical comparison of centerline concentration

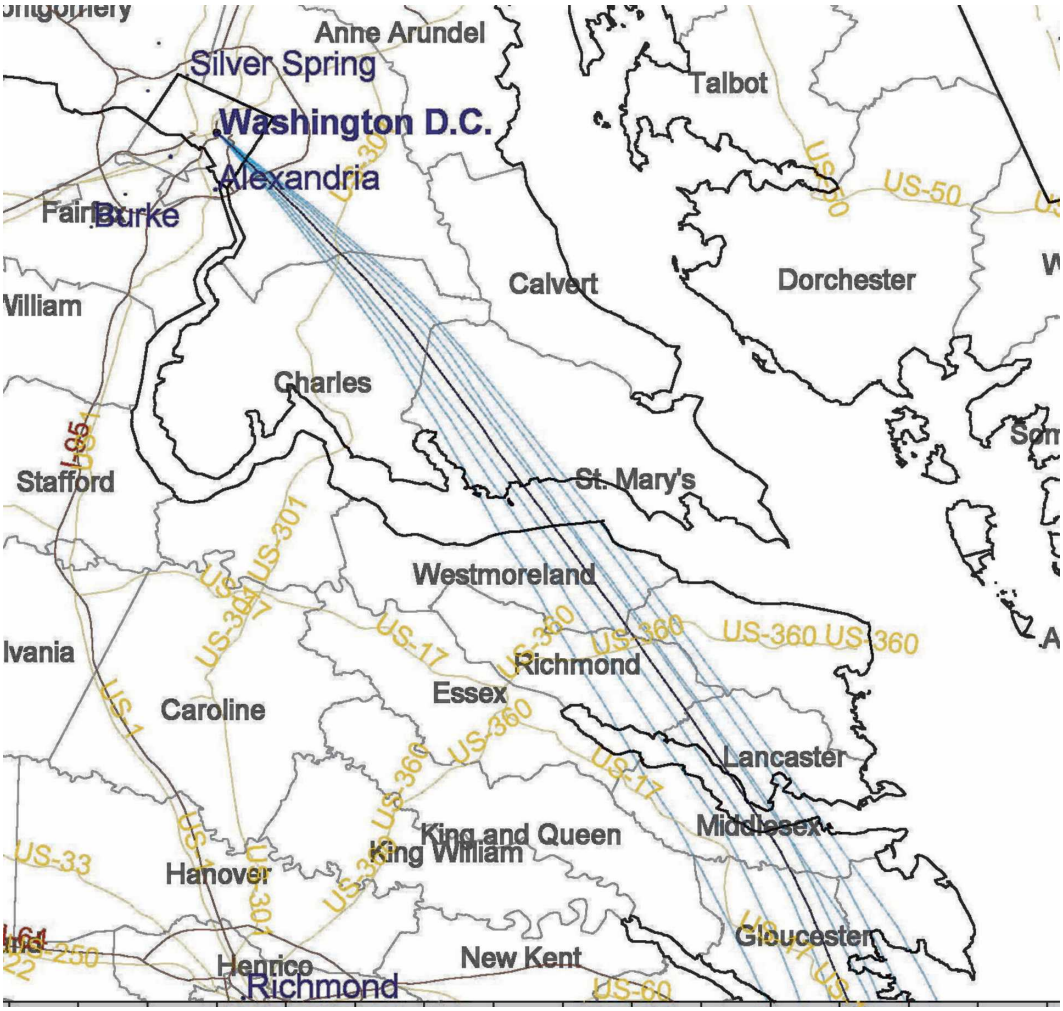

FIG. 6. Trajectory forecast for nine puffs released from Washington D.C. Simulated trajectories have been adjusted to appear as if they all were released from the central release point. Simulations were generated using INPUFF and the Eta-12-km 0000 UTC 19 May 2005 forecast 10-m winds. The trajectory for the central release point is shown as a thick red line.

values, where an allowance is made for the fact that models are predicting the ensemble average (first moment) of the centerline concentration value whereas the observations are individual realizations from a population of possible outcomes. As part of the procedure, experiments having similar dispersive conditions are grouped together (imperfect ensembles) in order to determine an average observed and estimated centerline concentration at each downwind distance having sufficient observations [this is an extension of an idea discussed in the last paragraph of Venkatram (1979), regarding model evaluation].

Part of the problem of forming groups of data then becomes how many observations are needed in a group to have an average that is sufficiently certain that it will provide a means for discerning whether differences in performance by alternative models is statistically significant. This is analogous to deciding how many tosses are needed of a pair of dice to define the distribution of outcomes sufficiently that one can determine whether the dice are "fair." The uncertainty in determining the average centerline concentration $C_{\max}$ as a function of the number of samples NS can be expressed as, $\mathrm{Std}(C_{\max})/C_{\max} = [\exp(\ln^2\mathrm{GeoStd}) - 1]^{1/2}/(\mathrm{NS})^{1/2}$. If the GeoStd in the centerline concentration values was 2 and we had 30 "near centerline" values for the computation of the GeoAvg ($C_{\max}$ in this case), then $\mathrm{Std}(C_{\max})/C_{\max}$ would be approximately 14%. Attempting to achieve 30 samples for analysis places demands on the spacing of the receptors (i.e., the closer the receptors, the greater the number of near-centerline values) and on the number of experiments con-

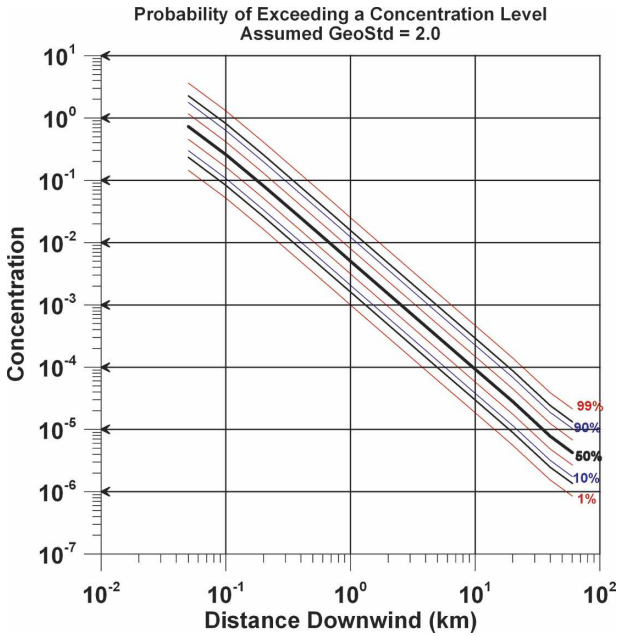

FIG. 7. Illustration of how the variations in the centerline concentration values could be displayed as a probability of occurrence.

ducted. For PG, the lateral dispersion of the plume (σ_y) decreases from about 20° for unstable conditions to about 4° for stable conditions for the 50-m arc and decreases from about 10° for unstable conditions to about 1.5° for stable conditions for the 800-m arc. If we define centerline receptors as before in section 2b ($y < |0.67\sigma_y|$) and if we desire to have at least 5 nonzero concentration values for analysis along the 50-m arc, then for stable conditions we would need a receptor spacing on the order of 3°, and this would only provide one centerline sample per experiment. For analysis of 5 nonzero concentration values along the 800-m arc for stable conditions, we would need a receptor spacing on the order of 1°. In the actual PG experiment, the receptor spacing was 2° for the 50-, 100-, 200-, and 400-m arcs and was decreased to 1° spacing for the 800-m arc.

It would be wrong to portray ASTM D 6589 as arguing only for the comparison of group averages. Group averages are discussed in ASTM D 6589 as "an example" of how one might attempt to compute ensemble averages for comparison with model predictions. The stochastic variation in short-term near-centerline concentration values for inert species is large, on the order of a factor of 2–4 as compared with its average value (Fig. 1). The direct comparison of measured short-term maximum concentration values with model predictions is analogous to attempting to determine if a pair of dice are fair by comparing the individual outcomes from a series of tosses with seven (the average of all possible outcomes). At best, air quality models (be they puff or plume dispersion models, or grid-based chemical-transport grid models) will only be capable of predicting the properties of a distribution (first moment, second moment) of concentration values as measured at a given location; to be able to predict what is directly measured would be analogous to being able to predict the exact precipitation amount for a specified hour and location (or the sequence of outcomes to be seen in a series of tosses of a pair of dice). The basic message of ASTM D 6589 is that model evaluations that directly compare observations with predictions and do not prove that the unresolved stochastic variability can be assumed to be negligible are suspect. ASTM D 6589 is promoting a revision in how one evaluates model performance, from naively believing that models predict what is directly measured at given locations from one hour to the next (which is impossible) to understanding that at best models can only predict the distributional properties of concentration values measured at given locations. One can mitigate the effects of stochastic variations through the use of long-term time averages or large-scale spatial averages (which are forms of "groups"), but it is anticipated that assessing model performance for the shorter averaging times and smaller spatial scales will require testing how well a model predicts the distributional properties of data either determined from analyses of data directly grouped together or determined through time series and spatial-scale analyses. The example illustrates that knowledge of the stochastic variability in the concentration values that the model is to be challenged to predict provides a basis for designing meaningful model evaluation studies, and also provides a basis for judging the adequacy of existing datasets for the purpose of conducting statistical assessments for transport and diffusion model performance.

4. Summary

Our analyses confirm the longstanding finding that the crosswind concentration profile of a dispersing plume on average is well characterized as having a Gaussian shape. However, the variability not described by a Gaussian crosswind profile is substantial, on the order of a factor of 2 for centerline concentration values, and the variability in the centerline concentration values is seen in field data out to distances of at least tens of kilometers. The variability in the trajectory of a dispersing plume will be dependent on the actual syn-

optic situation, but likely is larger than the plume width, even with local wind observations for use in characterizing the transport. The variability in the plume trajectory was investigated by tracking the divergence in trajectories from releases adjacent to the actual release location. The analysis provided preliminary results that in the future can be refined using meteorological ensembles as will be generated by the Weather Research and Forecasting meteorological model.

In this paper, two examples are given to illustrate how estimates of variability 1) can provide useful information to inform decisions for emergency response and 2) can provide a basis for sound statistical designs for model performance assessments. For emergency response, it is suggested that an effort be made to convey to decision makers that the uncertainty in the trajectory is likely larger than the width of the dispersing plume and that centerline concentration estimates can vary by at least a factor of 2—about what is predicted to be seen on average (and this assumes no uncertainty in the release characterization). For model evaluation, it is suggested that one should account for the unresolved variability in the statistical design, otherwise one may well be asking the model to replicate random variations that are unresolved by the deterministic physics in the model, and this is especially important when the unresolved variations are large in comparison with the observed concentration values.

Although the focus of our discussion has been on atmospheric transport and diffusion modeling, we believe it would be worthwhile to objectively characterize the unresolved variability in all forms of atmospheric models, so that 1) research can be focused on providing predictions that describe all possible outcomes and 2) research can be conducted to develop meaningful model evaluation strategies. This will provide a firmer basis for decisions and will promote the development of model evaluation metrics and procedures that test a model's ability to characterize the moments of the distribution of possible outcomes.

Acknowledgments. The research presented here was performed under the Memorandum of Understanding between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Commerce's National Oceanic and Atmospheric Administration and under Agreement DW13921548; J. S. Irwin collaborated in the research described here under Contract NOAA Order EA133R04SE1097. This work constitutes a contribution to the NOAA Air Quality Program. Although it has been reviewed by EPA and NOAA and approved for publication, it does not necessarily reflect their policies or views.

REFERENCES

American Society for Testing and Materials, 2005: Standard guide for statistical evaluation of atmospheric dispersion model performance (D 6589). American Society for Testing and Materials, 17 pp. [Available from 100 Barr Harbor Drive, P.O. Box C700, West Conshohocken, PA 19428 and online at http://www.astm.org.]

Cassiani, M., P. Franzese, and U. Giostra, 2005a: Application of a mixing model to dispersion in atmospheric turbulence. Part I: Development of the model and application to homogeneous turbulence and neutral boundary layer. *Atmos. Environ.,* **39,** 1457–1469.

——, ——, and ——, 2005b: A PDF micromixing model of dispersion for atmospheric flow. Part II: Application to convective boundary layer. *Atmos. Environ.,* **39,** 1471–1479.

Dabberdt, W. F., and E. Miller, 2000: Uncertainty, ensembles, and air quality dispersion modeling: Applications and challenges. *Atmos. Environ.,* **34,** 4667–4673.

Environmental Protection Agency, 2003: AERMOD: Latest features and evaluation results. Tech Rep. EPA-454/R-03-003, 42 pp. [Available from Office of Air Quality Planning and Standards, Emissions Monitoring and Analysis Division, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711 and online at http://www.epa.gov/scram001/7thconf/aermod/aermod_mep.pdf.]

Finkelstein, P. L., J. C. Kaimal, J. E. Gaynor, M. E. Graves, and T. J. Lockhart, 1986: Comparison of wind monitoring systems. Part I: In situ sensors. *J. Atmos. Oceanic Technol.,* **3,** 583–593.

Fox, D. G., 1984: Uncertainty in air quality modeling. *Bull. Amer. Meteor. Soc.,* **65,** 27–36.

Fuquay, J., C. L. Simpson, and W. T. Hinds, 1964: Prediction of environmental exposures from sources near the ground based on Hanford experimental data. *J. Appl. Meteor.,* **3,** 761–770.

Hanna, S. R., 1984: Concentration fluctuations in a smoke plume. *Atmos. Environ.,* **18,** 1091–1106.

——, 1993: Uncertainties in air quality model predictions. *Bound.-Layer Meteor.,* **62,** 3–20.

——, and R. Yang, 2001: Evaluations of mesoscale model's simulations of near-surface winds, temperature gradients, and mixing depths. *J. Appl. Meteor.,* **40,** 1095–1104.

Harris, J. M., R. R. Draxler, and S. J. Oltmans, 2005: Trajectory model sensitivity to differences in input data and vertical transport method. *J. Geophys. Res.,* **110,** D14109, doi:10.1029/2004JD005750.

Irwin, J. S., 1983: Estimating plume dispersion—A comparison of several sigma schemes. *J. Climate Appl. Meteor.,* **22,** 92–114.

——, 1984: Site-to-site variation in performance of dispersion parameter estimation schemes. *Air Pollution Modelling and Its Application III,* C. De Wispelaere, Ed., Plenum, 605–617.

——, and M. E. Smith, 1984: Potentially useful additions to the rural model performance evaluation. *Bull. Amer. Meteor. Soc.,* **65,** 559–568.

——, and R. F. Lee, 1996: Comparative evaluation of two air quality models: Within-regime evaluation statistic. *Int. J. Environ. Pollut.,* **8,** 346–355.

——, and S. R. Hanna, 2005: Characterizing uncertainty in plume dispersion models. *Int. J. Environ. Pollut.,* **25,** 6–24.

——, S. T. Rao, W. B. Petersen, and D. B. Turner, 1987: Relating error bounds for maximum concentration predictions to diffusion meteorology uncertainty. *Atmos. Environ.,* **21,** 1927–1937.

Jones, A. R., and D. J. Thomson, 2006: Simulation of time series of concentration fluctuations in atmospheric dispersion using a correlation-distortion technique. *Bound.-Layer Meteor.,* **118,** 25–54.

Lockhart, T. J., and J. S. Irwin, 1980: Methods for calculating the "representativeness" of data. *Proc. Symp. on Intermediate Range Atmospheric Transport Processes and Technology Assessment,* Gatlinburg, TN, 169–176. [NTIS CONF-801064.]

Luhar, A. K., and B. L. Sawford, 2005a: Micromixing modelling of concentration fluctuations in inhomogeneous turbulence in the convective boundary layer. *Bound.-Layer Meteor.,* **114,** 1–30.

——, and ——, 2005b: Micromixing modelling of mean and fluctuating scalar fields in the convective boundary layer. *Atmos. Environ.,* **39,** 6673–6685.

Petersen, W. B., and L. G. Lavdas, 1986: INPUFF 2.0–A multiple source Gaussian puff dispersion algorithm user's guide. Atmospheric Sciences Research Laboratory, U.S. Environmental Protection Agency, EPA/600-8-86-024, 105 pp.

Ramsdell, J. V., and W. T. Hinds, 1971: Concentration fluctuations and peak-to-mean concentration rations in plumes from a ground-level continuous point source. *Atmos. Environ.,* **5,** 483–495.

Stohl, A., 1998: Computation, accuracy and applications of trajectories–A review and bibliography. *Atmos. Environ.,* **32,** 947–966.

Stunder, B. J. B., 1996: An assessment of the quality of forecast trajectories. *J. Appl. Meteor.,* **35,** 1319–1331.

Sykes, R. I., W. S. Lewellen, and S. F. Parker, 1984: A turbulent transport model for concentration fluctuations and fluxes. *J. Fluid Mech.,* **139,** 193–218.

——, ——, ——, and D. S. Henn, 1988: A hierarchy of dynamic plume models incorporating uncertainty. Vol. 4: Second-order closure integrated puff. Electric Power Research Institute (EPRI) EA-6095, Project 1616-28, 180 pp. [Available from R. I. Sykes, ARAP/Titan, 50 Washington Rd., P.O. Box 2229, Princeton, NJ 08543-2229.]

——, S. F. Parker, D. S. Henn, and W. S. Lewellen, 1993: Numerical simulation of ANATEX tracer data using a turbulence closure model for long-range dispersion. *J. Appl. Meteor.,* **32,** 929–947.

——, ——, ——, C. P. Cerasoli, and L. P. Santos, 1998: PC-SCIPUFF version 1.2PD, technical documentation. ARAP Rep. 718, 180 pp. [Available from Titan Research & Technology Division, ARAP Group, Titan Corporation, P.O. Box 2229, Princeton, NJ 08543-2229 and online at http://www.titan.com/products-services/336/download_scipuff.html.]

Venkatram, A., 1979: The expected deviation of observed concentrations from predicted ensemble means. *Atmos. Environ.,* **13,** 1547–1549.

——, 1983: Uncertainty in predictions from air quality models. *Bound.-Layer Meteor.,* **27,** 186–196.

——, 1988: Topics in applied modeling. *Lectures on Air Pollution Modeling,* A. Venkatram and J. C. Wyngaard, Eds., Amer. Meteor. Soc., 267–324.

Weil, J. C., R. I. Sykes, and A. Venkatram, 1992: Evaluating air quality models: Review and outlook. *J. Appl. Meteor.,* **31,** 1121–1145.

——, W. H. Snyder, R. E. Lawson Jr., and M. S. Shipman, 2002: Experiments on buoyant plume dispersion in a laboratory convection tank. *Bound.-Layer Meteor.,* **102,** 367–414.