Available online at www.sciencedirect.com
ScienceDirect
Atmospheric Environment 41 (2007) 617–628
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

A statistical approach for estimating uncertainty in dispersion modeling: An example of application in southwestern USA

Darko Koračin[a,*], Anna Panorska[b], Vlad Isakov[c], Jawad S. Touma[c], Jenise Swall[c]

[a]Desert Research Institute, 2215 Raggio Parkway, Reno, Nevada 89512, USA
[b]University of Nevada, Reno, NV 89557, USA
[c]NOAA/Atmospheric Sciences Modeling Division, USA[1]

Received 22 March 2006; received in revised form 7 July 2006; accepted 17 August 2006

Abstract

A method based on a statistical approach of estimating uncertainty in simulating the transport and dispersion of atmospheric pollutants is developed using observations and modeling results from a tracer experiment in the complex terrain of the southwestern USA. The method takes into account the compensating nature of the error components by representing all terms, except dispersion error and variance of stochastic processes. Dispersion error and the variance of the stochastic error are estimated using the maximum likelihood estimation technique applied to the equation for the fractional error. Mesoscale Model 5 (MM5) and a Lagrangian random particle dispersion model with three optional turbulence parameterizations were used as a test bed for method application. Modeled concentrations compared well with the measurements (correlation coefficients on the order of 0.8). The effects of changing two structural components (the turbulence parameterization and the model grid vertical resolution) on the magnitude of the dispersion error also were examined. The expected normalized dispersion error appears to be quite large (up to a factor of three) among model runs with various turbulence schemes. Tests with increased vertical resolution of the atmospheric model (MM5) improved most of the dispersion model statistical performance measures, but to a lesser extent compared to selection of a turbulence parameterization. Method results confirm that structural components of the dispersion model, namely turbulence parameterizations, have the most influence on the expected dispersion error.

Keywords: Uncertainty; MM5 simulations; Complex terrain; Turbulence parameterizations; Lagrangian dispersion model

1. Introduction

Air quality models are the primary tools for quantitatively simulating future emission control strategies. Air quality models integrate knowledge of emissions, meteorology, as well as physical and chemical processes affecting pollutant transport and dispersion in the atmosphere. Because air quality models play a central role in design of strategies for reducing local-scale impacts of pollutants, it is critical that comprehensive, systematic uncertainty analysis of these models be undertaken (NRC, 2004).

So far, a unique approach has not been developed for evaluating models and estimating their

*Corresponding author. Tel.: +1 775 674 7091; fax: +1 775 674 7060.
E-mail address: darko.koracin@dri.edu (D. Koračin).
[1]In partnership with the US Environmental Protection Agency, Research Triangle Park, NC 27711, USA.

1352-2310/$ - see front matter
doi:10.1016/j.atmosenv.2006.08.048

uncertainties and errors. The American Society for Testing and Materials (ASTM) has issued a Standard Guide for Statistical Evaluation of Atmospheric Dispersion Model performance (ASTM, 2000). This guide provides information and a series of options for determining the accuracy of model predictions and guidance on testing differences among models. The guide suggests that initial analysis should focus on determining, on average, how close to observations each of the models is and then estimating the significance of differences among models. Although the guide does not recommend a specific course of action, it represents a valuable step toward developing systematic methods of model evaluation and comparison.

There are three major types of uncertainty in simulating the transport and dispersion of gases and particles in the atmosphere: uncertainty due to input data, uncertainty caused by model physics assumptions, and uncertainty due to random variability (turbulence effects) (Sax and Isakov, 2003; Rao, 2005). Uncertainty due to model physics assumptions usually represents a superposition of errors from both atmospheric and dispersion models. As noted by Hanna (1994) and Koračin et al. (2000), even perfect agreement between modeled and measured concentrations of atmospheric pollutants and tracers does not necessarily mean that the atmospheric and dispersion components are behaving properly. As indicated by Weil et al. (1992) among others, understanding this error requires a full understanding of the model physics, model algorithm, and even the actual program code.

The classical approach to estimate uncertainty in model results, usually referred to as the Monte Carlo method, is widely used in different applications (Bergin et al., 1999; Hanna et al., 2001; Sax and Isakov, 2003). In this approach, the value of an unknown variable is estimated using a sample of empirical observations. This approach is computationally extensive, however, especially for complex modeling applications.

In this study, we demonstrate use of a computationally efficient statistical approach for estimating uncertainty in model parameters and inputs for air quality modeling applications. This approach uses the maximum likelihood estimation technique to quantify uncertainty and identify influential sources of uncertainty. We have developed a stochastic model for estimating uncertainty in model results that accounts for the compensating nature of error components. The model fractional error has been represented as a combination of error components due to emissions, measurements, atmospheric and dispersion models (total model error), and stochastic processes. By decomposing the total model error into its atmospheric and dispersion components, we are able to estimate error in the dispersion computation as well as to estimate variance of the stochastic component of the error. The method is illustrated using observations and modeling results from a tracer experiment in the complex terrain of the southwestern USA (Green, 1999; Koračin et al., 2000; Podnar et al., 2002).

2. A stochastic model for estimating uncertainty in dispersion modeling

2.1. Fractional error

We define total fractional error (E_t) of a predicted concentration compared to a measured value at a receptor for each sampling time as

$$E_t = \frac{C_p - C_o}{C_a}, \tag{1}$$

where C_p and C_o are the predicted and observed concentrations, respectively, and C_a is the average concentration of C_p and C_o. Generally, fractional error is a useful measure for model evaluation (see, e.g., Wilson and Flesch, 1993). In contrast to normalization by C_o, we use the average value of C_p and C_o as is done similarly for bias and some other statistical parameters. For a case when both C_p and C_o are zero, the left-hand side of Eq. (1) is taken as zero.

2.2. Error components

First, we postulated that total fractional error at a receptor for each sampling time (E_t) is represented by a function of the following normalized error components:

$$E_t = f(E_e, E_m, E_a, E_d, E_s), \tag{2}$$

where E_e is the emission error, E_m is the measurement error, E_a is the atmospheric model and its input data error, E_d is the dispersion model and its data error, and E_s is the stochastic component of the total fractional error. Here, we assume linear superposition of error components, and that the model and its input data error consist of two major components: atmospheric model and its input data error (E_a) and dispersion model and its input data error (E_d). Koračin et al. (2000) presented a method for evaluating atmospheric models using tracer measurements, which separately computes error associated with the atmospheric model and its data error (E_a).

2.3. Dispersion error estimation

Second, since total fractional error (E_t) is known and all other error components (E_e, E_m, E_a, and E_s) can be estimated, we can compute the dispersion and its data error (E_d) as a residual from Eq. (2). Moreover, since different components of the errors could be compensating for each other, they can take either positive or negative values. Consequently, in the third and final step, the residual (E_d) can be estimated from the equation for the total fractional error as

$$\frac{C_p - C_o}{C_a} = \pm E_e \pm E_m \pm E_a + \text{Residual}\,(E_d) \pm E_s. \tag{3}$$

Note that Residual (E_d) can be either positive or negative depending on the combination of other errors and the requirement to balance total fractional error on the left-hand side. This formulation also allows any possible combination of errors. Even if the total error on the left-hand side is zero, there still may be a combination of non-zero compensating errors on the right-hand side of Eq. (3). The linear form of the expression for the dispersion model and data error is a first simple step in estimating this type of error, as shown by others (e.g., Hanna, 1989; Rao, 2005). Generally, errors could be nonlinearly related in a fairly complex manner; however, further methods development is required to treat this nonlinear interaction.

3. Estimation procedure for the stochastic model

We now turn to statistical formulation of the error model (Eq. (3)). We have data on observed and modeled values of tracer concentrations, $C_{o,ij}$ and $C_{p,ij}$, respectively, for each of the n_1 days and each of n_2 receptors ($i = 1,\ldots,n_1$, *and* $j = 1,\ldots,n_2$). Error term distributions are assumed as follows: $E_{e,ij}$, $E_{m,ij}$, and $E_{s,ij}$ are independent and identically distributed with distributions given by $N(0, \sigma_e^2)$, $N(0, \sigma_m^2)$, and $N(0, \sigma^2)$, respectively. Finally, since atmospheric error can vary by day, for each i we have $E_{a,ij}$ independent and identically distributed $N(0, \sigma_{a,i}^2)$, where $\sigma_{a,i}^2$ are daily variances. We assumed the fractional errors $(C_p - C_o)/C_a$ for a given receptor and a given day are to be sufficiently normally distributed according to $Y_{ij} \sim N(\mu, \sigma^2 + \sigma_i^2)$, where σ^2 is the variance of the stochastic error component and σ_i^2 denotes the variance of the combined daily emission, measurement, and atmospheric error components, that is $\sigma_i^2 = \sigma_e^2 + \sigma_m^2 + \sigma_{a,i}^2$. Distribution of Y_{ij}'s is normal because Y_{ij}'s are sums of independent normal random variables. We can then rewrite Eq. (3) as follows:

$$Y_{ij} = E_{e,ij} + E_{m,ij} + E_{a,ij} + \mu + E_{s,ij}, \quad i = 1,\ldots,n_1, \quad j = 1,\ldots,n_2, \tag{4}$$

where μ is the Residual (E_d) from Eq. (3).

Our objective is to find the maximum likelihood estimators (MLEs) for μ and σ^2. For this, we adopt a well known and widely used statistical technique of maximum likelihood estimation (Casella and Berger, 1990). The procedure for finding MLEs entails finding the maximum of the likelihood function of Y_{ij}'s with respect to μ and σ^2. The likelihood function is the probability density function of the random vector $\mathbf{Y} = (Y_{1,1},\ldots,Y_{n_1,n_2})$, regarded as a function of the parameters μ and σ^2 for a given realization $\mathbf{y}$ of $\mathbf{Y}$. The MLEs are the values for μ and σ^2 which maximize the likelihood of obtaining data $\mathbf{y}$ that was actually observed. Sometimes, the MLE is clearly and simply defined, as is the case for μ and σ^2 for a normal distribution, but is not always the case in more complicated estimation problems.

For demonstration purposes, we assume that error component E_i has a normal distribution denoted by $E_i \sim N(0, \sigma_i^2)$. Selection of a statistical distribution representing the distribution of the error components is not limited to a normal distribution. If the specifics of the error components suggest otherwise, other distributions could be used. In a simple case, when we assume that the atmospheric error E_a does not change from day to day, we can obtain explicit solutions for the MLE problem. These are obtained easily from the known MLEs for the mean and variance for a sample from a normal distribution. The MLE for μ is

$$\hat{\mu} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} y_{ij}. \tag{5}$$

The MLE for σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} (y_{ij} - \hat{\mu})^2 - \left(\sigma_e^2 + \sigma_m^2 + \sigma_a^2\right), \tag{6}$$

where σ_a^2 is an estimate of the common atmospheric error's variance.

This simple formulation allows us to relatively easily obtain an estimate of the residual E_d (μ), given the total fractional error (Y_{ij}) and the variances σ_e^2, σ_m^2, and σ_a^2. One significant advantage is that this method is much less computationally demanding than a Monte Carlo simulation might be. Our formulation relies on assumptions about the distribution of the various components of the total fractional error, in particular, the assumption of normality. The method also can be applied in cases in which the components have other distributional forms, although these formulations may not be as simple or as easily implemented.

4. Application of the stochastic model to observed and simulated concentrations of tracers in complex terrain

In order to illustrate this method, we estimated the uncertainty in modeled results from the Project MOHAVE field program in the complex terrain of the southwestern USA (Green, 1999). We chose this case for several reasons: (1) this is a well studied area, providing a rich observational database, including a tracer study; (2) multiple model results are available and have been evaluated extensively (Koračin and Enger, 1994; Yamada, 2000); and 3) uncertainty in transport has been analyzed already (Koračin et al., 2000). Therefore, this project provides a unique opportunity for demonstrating a methodology for the uncertainty analysis.

The Project MOHAVE field program included comprehensive meteorological, chemical, and tracer measurements. Tracer measurements allow us to simplify the procedure for uncertainty estimates. oPCDH inert tracer gas was released from a 152-m tall stack at a rate proportional to power production (Green, 1999). Daily tracer concentrations collected at eleven receptors from 6 to 12 August 1992 were used for analysis.

Lagrangian particle models are widely used to simulate dispersion for various air quality and emergency response applications (Draxler and Hess, 1997). In this study, we used a Lagrangian random particle (LAP) dispersion model (Koračin et al., 1998, 1999) with meteorological inputs from Mesoscale Model 5 (MM5) (Grell et al., 1994). The LAP model was developed following basic concepts described by Pielke (1984), for example. Concepts of the model structure and applications are shown by Koračin et al. (1998, 1999). Model parameterization includes an option of spatially and temporarily variable (Hanna, 1982) or constant time scales (Gifford, 1995), a drift correction term (Legg and Raupach, 1982), a plume rise algorithm, and three optional turbulence parameterizations (Donaldson, 1973; Mellor and Yamada, 1974; Andrén, 1990). Choosing different turbulence parameterizations, which allows us to estimate uncertainty due to model physics, is an essential part of the model capabilities. Meteorological input to the LAP model includes 3D fields of U, V, and W wind components, and potential temperature simulated by MM5. Details of the MM5 setup are shown in Koračin et al. (2000), but some of the main features are highlighted here.

First, the model was run for the period from 6 to 12 August 1992 in a non-hydrostatic mode with 3 km horizontal resolution. Second, the model domain consisted of 124×91 horizontal grid points and 35 vertical levels. One of the obvious choices to improve prediction of meteorological fields is to increase the vertical resolution of the model grid. Third, besides a baseline simulation that is described by Koračin et al. (2000), we have performed simulations with five additional vertical levels within the first 3 km and with each turbulence scheme. This increased vertical resolution improved meteorological model results and the accuracy of the predicted tracer concentrations.

Fig. 1 shows the model domain including the complexity of the terrain and a top view of the simulated particle distribution for the model runs with all three turbulence parameterization options. The figure also illustrates differences in particle distribution due to the treatment of turbulence transfer with respect to channeling along the Colorado River Valley and particle entrapment in the Grand Canyon. Note that differences among simulated particle distributions are apparent, but statistical tests are needed to quantify the differences (see Sections 5 and 6).

Tables 1 and 2 show daily averages of measured and simulated tracer concentrations at 11 receptors for the period from 6 to 12 August 1992 for the model runs with three optional turbulence parameterizations applied to the baseline simulations (Table 1) and the simulation with increased vertical resolution (Table 2). Figs. 2a and b show a comparison among concentrations as measured and as predicted by the three schemes for a baseline simulation (Fig. 2a) and the runs with increased vertical resolution (Fig. 2b).

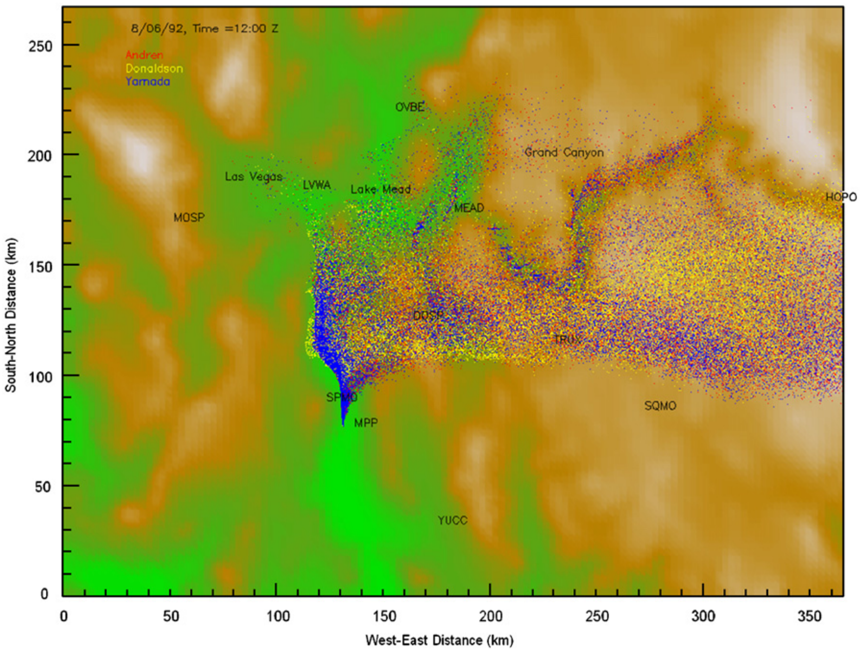


Fig. 1. Top view of simulated particle distribution using three optional turbulence parameterizations (Andrén—red, Donaldson—yellow, and Yamada—blue dots) on 6 August 1992 at 1200 UTC (see text and Tables 1 and 2 for details).

Table 1
Average daily concentrations of tracer ($\mathrm{fl\,l^{-1}}$) for a baseline case measured and simulated by the LAP model with three optional turbulence parameterizations using MM5 results as meteorological input

Station No.		1	2	3	4	5	6	7	8	9	10	11
Measurements	DATE	COCO	DOSP	HOPO	LVWA	MEAD	MOSP	OVBE	SPMO	SQMO	TRUX	YUCC
	6-Aug-92	4.52	1.33	0.025	0.64	1.27	0.23	1.53	2.58	0.04	0.04	0.11
	7-Aug-92	5.5	0.39	0.109	1.02	0.755	0.2	1.39	0.1	0.07	0.06	0.04
	8-Aug-92	2.14	1.29	0.065	1.32	2.095	0.18	1.83	0.04	0.03	0.03	0.03
	9-Aug-92	4.81	1.84	0.23	1.31	1.71	0.16	1	0.12	0.12	0.08	0.08
	10-Aug-92	4.97	1.57	0.035	2.2	0.135	0.8	0.54	0.35	0.05	0.05	0.08
	11-Aug-92	4.2	0.91	0.04	2.74	0.835	1.12	0.87	0.04	0.04	0.06	0.05
	12-Aug-92	3.67	0.33	0.055	2.58	0.51	1.24	0.6	0.59	0.04	0.1	0.13
Andren	DATE	COCO	DOSP	HOPO	LVWA	MEAD	MOSP	OVBE	SPMO	SQMO	TRUX	YUCC
	6-Aug-92	35.743	0.138	0.140	4.506	1.673	0.000	0.620	0.000	0.000	0.004	0.000
	7-Aug-92	46.405	0.082	0.192	4.880	1.398	0.000	4.507	0.000	0.000	0.000	0.000
	8-Aug-92	54.732	0.482	1.490	5.250	5.800	0.000	5.419	0.007	0.000	0.018	0.000
	9-Aug-92	25.900	0.095	0.080	0.010	0.611	0.000	1.843	1.033	0.436	0.332	0.000
	10-Aug-92	70.731	0.187	0.015	6.203	1.228	0.000	2.293	0.112	0.410	0.229	0.000
	11-Aug-92	85.637	0.235	0.054	7.912	3.847	0.000	7.378	1.803	0.068	0.124	0.086
	12-Aug-92	29.670	0.292	0.117	20.169	2.321	0.000	13.397	0.002	0.012	0.007	0.000
Donaldson	DATE	COCO	DOSP	HOPO	LVWA	MEAD	MOSP	OVBE	SPMO	SQMO	TRUX	YUCC
	6-Aug-92	20.577	0.263	0.037	0.841	0.126	0.000	0.055	0.000	0.000	0.109	0.000
	7-Aug-92	27.552	0.243	0.095	0.287	0.289	0.000	1.061	0.000	0.000	0.000	0.000
	8-Aug-92	23.464	0.278	1.184	0.437	0.396	0.000	0.827	0.000	0.000	0.000	0.000
	9-Aug-92	13.586	0.110	0.052	0.000	0.531	0.000	0.105	0.106	0.103	0.234	0.000
	10-Aug-92	25.364	0.199	0.006	1.162	0.074	0.000	0.365	0.065	0.239	0.224	0.000
	11-Aug-92	33.027	0.381	0.003	0.551	0.295	0.002	0.147	0.420	0.070	0.040	0.156
	12-Aug-92	18.066	0.684	0.005	1.298	0.769	0.002	1.233	0.019	0.086	0.027	0.032

Table 1 (*continued*)

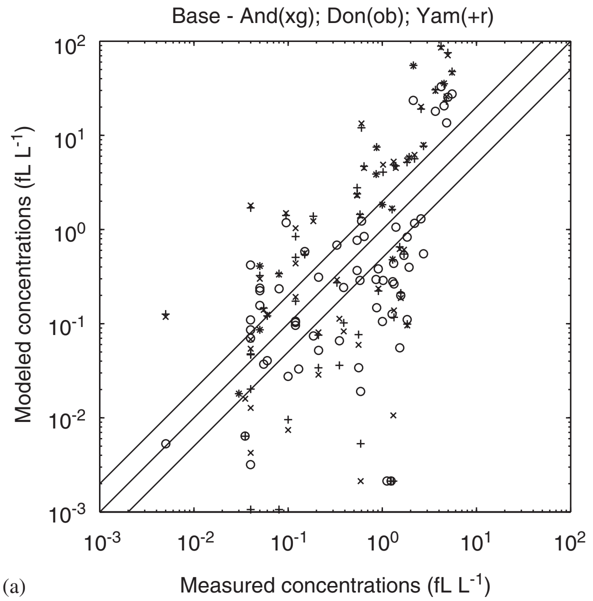
Station No.		1	2	3	4	5	6	7	8	9	10	11
Yamada	DATE	COCO	DOSP	HOPO	LVWA	MEAD	MOSP	OVBE	SPMO	SQMO	TRUX	YUCC
	6-Aug-92	35.300	0.116	0.146	4.700	1.632	0.000	0.654	0.000	0.000	0.001	0.000
	7-Aug-92	47.601	0.102	0.173	4.064	1.447	0.000	4.745	0.000	0.000	0.000	0.000
	8-Aug-92	55.293	0.478	1.411	4.852	5.887	0.000	5.173	0.000	0.000	0.018	0.000
	9-Aug-92	23.286	0.098	0.075	0.002	0.549	0.000	1.834	0.842	0.505	0.340	0.000
	10-Aug-92	75.468	0.211	0.006	5.640	1.384	0.000	2.779	0.036	0.409	0.324	0.001
	11-Aug-92	88.662	0.222	0.047	7.684	3.885	0.000	7.502	1.695	0.046	0.121	0.086
	12-Aug-92	30.453	0.272	0.126	19.044	2.389	0.002	12.025	0.005	0.020	0.009	0.000

Table 2
Average daily concentrations of tracer (fl$\,$l^{-1}) for model runs with increased vertical resolution measured and simulated by the LAP model with three optional turbulence parameterizations using MM5 results as meteorological input

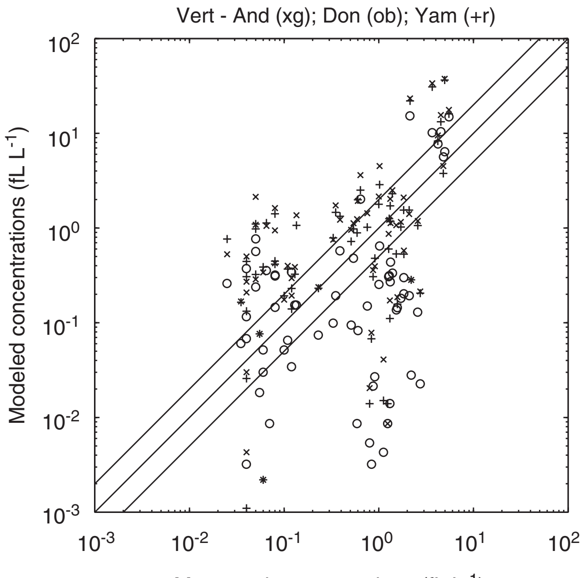
Station No.		1	2	3	4	5	6	7	8	9	10	11
Measurements	DATE	COCO	DOSP	HOPO	LVWA	MEAD	MOSP	OVBE	SPMO	SQMO	TRUX	YUCC
	6-Aug-92	4.52	1.33	0.025	0.64	1.27	0.23	1.53	2.58	0.04	0.04	0.11
	7-Aug-92	5.5	0.39	0.109	1.02	0.755	0.2	1.39	0.1	0.07	0.06	0.04
	8-Aug-92	2.14	1.29	0.065	1.32	2.095	0.18	1.83	0.04	0.03	0.03	0.03
	9-Aug-92	4.81	1.84	0.23	1.31	1.71	0.16	1	0.12	0.12	0.08	0.08
	10-Aug-92	4.97	1.57	0.035	2.2	0.135	0.8	0.54	0.35	0.05	0.05	0.08
	11-Aug-92	4.2	0.91	0.04	2.74	0.835	1.12	0.87	0.04	0.04	0.06	0.05
	12-Aug-92	3.67	0.33	0.055	2.58	0.51	1.24	0.6	0.59	0.04	0.1	0.13
Andren	DATE	COCO	DOSP	HOPO	LVWA	MEAD	MOSP	OVBE	SPMO	SQMO	TRUX	YUCC
	6-Aug-92	15.593	1.126	0.5291	3.6198	0.8679	0	1.071	0	0	0	0
	7-Aug-92	17.652	1.224	0.399	4.5167	1.4314	0	2.506	0	0	0.002	0
	8-Aug-92	23.36	1.221	1.0539	2.0336	1.4077	0	2.1	0.004	0	0	0
	9-Aug-92	4.5178	0.578	0.2377	0.1721	1.1571	0	2.152	0.296	0.194	0.425	0.944
	10-Aug-92	37.613	0.186	0.1678	0.2828	1.3658	0.02	1.125	1.736	1.081	0.287	1.624
	11-Aug-92	9.5055	0.395	0.0301	0.214	0.0785	0.041	0.36	0.501	0.271	0.34	2.132
	12-Aug-92	33.754	0.753	0.0764	1.1883	0.9582	0.009	1.933	1.239	0.209	0.175	0.387
Donaldson	DATE	COCO	DOSP	HOPO	LVWA	MEAD	MOSP	OVBE	SPMO	SQMO	TRUX	YUCC
	6-Aug-92	10.383	0.437	0.2592	2.0142	0.3054	0	0.137	0	0	0	0
	7-Aug-92	14.89	0.574	0.0656	0.6485	0.1506	0	0.336	0	0.009	0.03	0
	8-Aug-92	15.274	0.314	0.357	0.2699	0.1925	0	0.299	0	0	0	0
	9-Aug-92	5.6577	0.202	0.0742	0.014	0.1807	0	0.255	0.343	0.034	0.145	0.319
	10-Aug-92	6.4202	0.144	0.0602	0.028	0.1538	0.005	0.479	0.194	0.565	0.238	0.312
	11-Aug-92	7.7085	0.027	0.0032	0.0226	0.0032	0.004	0.022	0.374	0.068	0.052	0.767
	12-Aug-92	10.139	0.099	0.0183	0.129	0.0946	0.009	0.083	0.009	0.116	0.052	0.154
Yamada	DATE	COCO	DOSP	HOPO	LVWA	MEAD	MOSP	OVBE	SPMO	SQMO	TRUX	YUCC
	6-Aug-92	13.224	1.179	0.7657	2.5111	0.6001	0	0.532	0	0	0	0
	7-Aug-92	16.011	1.33	0.3689	2.867	1.0206	0	2.304	0	0	0.002	0
	8-Aug-92	22.001	1.257	1.1303	1.7142	1.5464	0	1.546	0.001	0	0	0
	9-Aug-92	3.7672	0.53	0.2312	0.1108	1.0238	0	1.783	0.23	0.14	0.416	0.447
	10-Aug-92	36.488	0.157	0.1645	0.2839	1.0679	0.014	0.98	1.463	0.98	0.324	1.412
	11-Aug-92	8.2785	0.479	0.0258	0.2054	0.0678	0.015	0.307	0.444	0.307	0.387	1.103
	12-Aug-92	30.883	0.787	0.0764	1.0604	0.7205	0.014	1.988	0.888	0.132	0.193	0.326

Tables 1 and 2 are used in this study to calculate fractional model error on the left-hand side of Eqs. (1) and (2).

Tracer concentrations predicted by the LAP model agreed fairly well with measured concentrations and yielded a similar order of magnitude. In



(a)



(b)

Fig. 2. Scatter plot of measured versus modeled daily tracer concentrations using three optional turbulence parameterizations (Andrén "x"; Donaldson "o"; and Yamada "+") for a baseline case (a) and the run with increased vertical resolution (b). Values for all receptors are shown for the period 6–12 August 1992.

the case of the baseline run, the model significantly overestimated concentrations at the receptor (COCO) nearest the source. The use of higher vertical resolution efficiently reduced the highest overestimated concentrations at the nearest receptor for all three optional turbulence parameterizations, but overestimated several low concentrations. The former effect reduces some of the statistical error parameters, while the latter effect lowers the correlation coefficient.

To estimate the dispersion error from Eq. (3), errors due to emission, measurements, and atmospheric models must be determined. Then the solution of the system will yield estimates of the mean and standard deviation of the dispersion model error (E_d).

4.1. Emission error

The tracer was released mainly at a rate proportional to the plant power production. Variation in power production, however, could not have been followed exactly by variation in tracer emission. According to Green (1999), the ratio of power production to tracer release rate had a standard deviation of 6.9%. Since there is no precise information on the actual tracer emission and its uncertainty, we assumed that the emission error could be represented by the normal distribution $N(E_m, \sigma_e^2)$ with a zero mean and normalized standard deviation of the emission error (σ_e) of 10%.

4.2. Measurement error

Green (1999) reported an uncertainty in tracer measurements (ΔM) of $\pm 0.052\,\mathrm{fl\,l^{-1}}$. For this field program, the background concentration was estimated as $0.56\,\mathrm{fl\,l^{-1}}$ and subtracted from all measured samples. The reported uncertainty in the background value (ΔC_b) was $\pm 0.06\,\mathrm{fl\,l^{-1}}$. To assume the maximum value of measurement uncertainty due to simultaneous uncertainties in the measurement procedure (ΔM) and estimation as well as variability of the background concentration (C_b), we assumed that the normalized measurement error could be expected in the range ($-E_m$, $+E_m$) with E_m estimated from:

$$E_m = \frac{\Delta M + \Delta C_b}{C_o + C_b} = \frac{0.112}{C_o + C_b}. \tag{7}$$

Since measurement error can be expected in the range [$-E_m$, $+E_m$], the absolute value of E_m can be used as a measure of the standard deviation around the zero mean [$N(0, \sigma_m^2)$].

As shown in Eq. (7), measurement error depends on C_o. In the statistical sense, we would like to find the most probable range for the measurement error. To estimate σ_m^2, we have a choice of values for C_o in

Table 3
Estimates of the success of MM5 wind fields and corresponding atmospheric model error according to the Tracer Potential (TP) method for the base run ("base") and the run with increased vertical resolution ("vert")

Day no.	Day	TP_base (%)	Ea_base	TP_vert (%)	Ea_vert
1	6-Aug-92	4.7	0.95	13.1	0.87
2	7-Aug-92	13.9	0.86	26.0	0.74
3	8-Aug-92	14.7	0.85	24.3	0.76
4	9-Aug-92	6.4	0.94	23.9	0.76
5	10-Aug-92	8.0	0.92	18.5	0.82
6	11-Aug-92	15.8	0.84	28.7	0.71
7	12-Aug-92	12.9	0.87	21.5	0.79

Eq. (7). The choices for C_o range from zero, which will yield the most conservative estimate of σ_m^2, through minimum or average observed concentration to maximum observed concentration, which corresponds to the smallest estimate of σ_m^2. We applied all four choices for C_o to our model, obtaining minimal differences in the MLEs for μ and σ.

4.3. Atmospheric model error

Atmospheric model error was determined via results from the method of evaluating atmospheric models using tracer measurements (Koračin et al., 2000). They used daily averages of tracer concentrations to estimate daily success of four modeled wind fields. The method uses a cost function (*Tracer Potential*, hereafter TP) to quantify the success of a particular atmospheric model in representing tracer transport. A full description of the method is given by Koračin et al. (2000). Table 3 shows estimates of the success (TP) of MM5 atmospheric model runs for the baseline case and the run with increased vertical resolution (present study). The value of the atmospheric model error is then represented by the value of 1-TP and used as input in Eqs. (3) and (4) for the days considered in this study.

Since TP is a measure of the normalized success of the MM5 wind fields, we assumed that atmospheric model error $E_{a,i}$ for a particular day i is equal to the value (1-TP) for that day. Since atmospheric error can be expected in the range $[-E_{a,i}, +E_{a,i}]$, the absolute value of $E_{a,i}$ can be used as a measure of the standard deviation around the zero mean $[N(0, \sigma_{a,i}^{2})]$.

Since emission and measurement errors were assumed constant in time, it is valuable to consider a simplified case with the atmospheric model error constant in time, as well. In that case, σ_a^2 can be taken as a statistic of (1-TP) daily series—such as average, maximum or minimum. Tests with maximum and average (1-TP) as σ_a^2 showed negligible difference in the MLEs for σ and, of course, do not affect the MLE of μ.

4.4. Stochastic component of the error

The stochastic component (E_s) of the fractional total error in Eq. (3) is caused by the stochastic nature of turbulent diffusion in the atmosphere. This error is represented by the normal distribution $[N(0, \sigma^2)]$ with zero mean and variance that is determined jointly with the dispersion error using maximum likelihood estimation.

5. Estimation of the dispersion error using a developed stochastic model

Errors due to emission and measurements are well defined for a given set of observations, while in contrast, atmospheric error depends on selection of how the meteorology is diagnosed or predicted. Variability of success in reproducing atmospheric processes and their implications on the estimation of dispersion error is an important issue that has to be further investigated. As we discussed in the previous sections, one of our choices was to improve vertical resolution of the atmospheric model and estimate the possible improvement in simulating atmospheric fields. Improved vertical resolution generally implies better treatment of vertical transport and mixing that can usually improve model results. Simple reasoning could lead to the conclusion that reduced atmospheric error in the balance equation, such as Eqs. (3) and (4), imposes an increase in dispersion error. Improved treatment of meteorology, however, generally causes better treatment of dispersion processes and reduces model fractional error on the left-hand side of Eqs. (3) and (4). Consequently, although the dispersion estimates are improved, model error can be similar to the previous case that is based on less successfully predicted or diagnosed meteorological fields.

To investigate the hypothesis that the essence of the dispersion error should be more or less quasi-invariant to variations of the error in the atmospheric representation, we calculated dispersion error using the baseline case and also calculated dispersion error using a case of improved MM5

Table 4
Estimates of dispersion model fractional error μ and the standard deviation of stochastic components of the error σ for three model formulations for the base run ("base") and the run with increased vertical resolution ("vert")

Model	Base case		Increased vertical resolution	
	μ	σ	μ	σ
Andren	2.97	6.81	−2.74	6.62
Donaldson	−0.47	2.27	−0.44	2.70
Yamada	−2.92	6.62	2.70	5.55

Values were calculated for all days and all receptors and for daily variable atmospheric model error.

results based on higher vertical resolution. Table 4 shows estimates of dispersion model error and the standard deviation of the stochastic component of the error for two respective MM5 simulations with different vertical resolutions.

Results indicated that both the expected normalized dispersion error and standard deviation of the stochastic error are large and could be up to a factor of three among the runs using different turbulence schemes. Dispersion error appears to be strongly dependent on basic structure of the dispersion model, such as selection of a turbulence scheme. Increased vertical resolution of the atmospheric model efficiently reduces model error for all schemes. Andrén's scheme is essentially a modification of the Yamada scheme and, consequently, yields similar dispersion error and standard deviation of stochastic error estimates. Model results using the Donaldson scheme create approximately five times smaller dispersion error and two times smaller standard deviation of the stochastic error compared to Andrén's and Yamada's schemes. All schemes produce negative values of dispersion error indicating that they generally overestimate measurements.

We used the MM5 baseline case to test several options in treating the atmospheric model error as constant throughout the simulation period as well as being daily variable. As mentioned earlier, dispersion error computation is further complicated by the fact that C_o appears in both the fractional error (left-hand side of Eqs. (3) and (4)) and the measurement error (right-hand side of the same equations). To simplify the computation, we assumed a constant number (average of all observations) for C_o in Eq. (7), but also a zero value for C_o to represent the largest possible measurement error. We also examined the joint effect of variation of atmospheric error and measurement errors on computation of dispersion error and the standard deviation of the stochastic component of the error.

First, we assumed all error components are time independent. In this case, we assumed the largest (conservative) value for the standard deviation of the measurement error (E_m) and assumed that the standard deviation of the atmospheric error (E_a) was the same for all days. Consequently, the standard deviations σ_m and $\sigma_{a,i}$ are constant during the computation procedure. We tested two cases for the choice of the common value of the variance of the atmospheric error: the mean (0.891) and maximum (0.953) of (1-TP). We also explored different values for C_o in the equation for $\sigma_m = (\Delta M + \Delta C_b)/(C_o + C_b)$, namely the most conservative value of 0, then the minimum C_o (0.005), the mean C_o (0.9578), and the maximum C_o (5.5). Different values of C_o produce different values for $\sigma_m = (0.052 + 0.06)/(C_o + 0.56) = 0.112/(C_o + 0.56)$. A selection of various values of C_o and σ_a did not have a significant impact on the value of $\hat{\sigma}$, which varied only up to 1%. We also considered a situation where atmospheric model error is time dependent. Since in this case there is no explicit formula for $\hat{\sigma}$, we had to estimate $\hat{\sigma}$ using a numerical optimization (maximization) process. We chose to use the log-likelihood function. The results indicate a very small difference in the concentration estimates as compared to the previous simple case.

6. Performance measure tests

We computed statistical performance measures to evaluate model results with various turbulence parameterizations and various vertical resolutions of the model grid. We used the same measures reported by Hanna (1989). For each individual model run, the following statistics were computed for modeled and observed concentrations normalized by the emission rate:

- Mean difference between the observed and predicted concentrations.
- Normalized mean square error (NMSE).
- Correlation between observations and model predictions and percent of model-predicted concentrations within a factor of two with respect to observed concentrations.

Table 5
Statistical performance measures for baseline simulations ("base") and simulations with improved vertical resolution ("vert")

Model_run	$\frac{\overline{\text{Obs}}}{\text{Q}}$	$\frac{\overline{\text{Pred}}}{\text{Q}} - \frac{\overline{\text{Obs}}}{\text{Q}}$	NMSE	r	% factor of 2
Don_base	0.43	0.60	16.2	0.79	85.7
And_base	0.43	2.23	41.3	0.78	66.2
Yam_base	0.43	2.25	43.5	0.77	66.2
Don_vert	0.43	0.05	5.2	0.72	87.0
And_vert	0.43	0.69	15.1	0.68	67.5
Yam_vert	0.43	0.57	14.4	0.67	71.4

Values were calculated for a case with daily variable atmospheric model error.

The model runs also were compared in pairs to test any statistically significant differences in their fit to the observed data. Pairing was done in space and time. All tests were two sided and performed at the 5% significance level. Test statistics were computed using a bootstrap technique (Cox and Tikvart, 1990) and following a procedure by Hanna (1989). Bootstrap is a statistical resampling technique for nonparametric inference (see e.g., Efron and Tibshirani, 1993). We used 1000 re-samples. Then, empirical percentiles of the differences were computed which allowed for making decisions. If the interval between the 2.5th and 97.5th empirical percentiles includes zero, we concluded (on the 5% significance level) that the difference between the values of statistic for the two models was not significantly different from zero. Performance measures for all models and results of these comparisons are reported in Table 5. The table shows that model differences in statistical parameters are affected by selection of the turbulence parameterization.

Schemes with similar algorithms (Andrén and Yamada) produced similar statistical parameters compared with measurements. Donaldson's scheme, for example, yielded the highest correlation coefficient, the smallest mean normalized differences between the model and measurements, the smallest NMSE, and the largest number of predicted values that fall within a factor of two of measured values. Higher vertical resolution improved all statistical parameters, except the correlation coefficient. Model runs with higher vertical resolution eliminated over predictions of some of the high concentrations and consequently reduced NMSE (by a factor of 2–3) and the mean normalized differences between the model and measurements, as well as increased the number of predicted concentrations that are within a factor of two of measured concentrations. Model runs with higher vertical resolution over predicted some of the lower concentrations, however, and consequently reduced the correlation coefficient by about 10–14% for all schemes.

Statistical comparison between measurements and various model runs, as well as among the model runs themselves, yielded the following results:

(a) Mean differences between measured and modeled concentrations *are not* significantly different from zero at the 5% significance level. This is true for the run with Donaldson's scheme (improved vertical resolution).
(b) Pairs of models for which the difference in mean differences between the measured and modeled concentrations *is not* significantly different from zero (5% significance level). This is true for the runs with Andrén's scheme (baseline simulation) and Yamada's scheme (baseline simulation).
(c) Pairs of models for which *differences* in mean square error *are not* significantly different from zero (5% significance level). This is true for (1) the run with Donaldson's scheme (improved vertical resolution) and the run with Andrén's scheme (improved vertical resolution); (2) the run with Donaldson's scheme (improved vertical resolution) and the run with Yamada's scheme (improved vertical resolution); and (3) the run with Andrén's scheme (baseline simulation) and the run with Yamada's scheme (baseline simulation).
(d) Models for which the correlation coefficient between measured and predicted concentrations *is not* significantly different from zero. This is true for none (i.e., all correlation coefficients are significantly different from zero).
(e) Pairs of models for which the difference between the correlation coefficients of the measured and predicted concentrations *are not* significantly different from zero. This is true for none (i.e., all pairs of model runs have nonzero difference between their correlations with observations).
(f) Pairs of models for which the difference in "percent is within the factor of two" statistic is not significantly different from zero. This is true for (1) the run with Donaldson's scheme (baseline simulation) and the run with Donaldson's scheme (improved vertical resolution); (2) the run with Andrén's scheme (baseline simulation) and the run with Andrén's (improved vertical

resolution); and (3) the run with Yamada's scheme (baseline simulation) and the run with Yamada's scheme (improved vertical resolution).

In summary, all model runs with various turbulence parameterizations have significant correlation with measurements. The runs with the Donaldson scheme appear to achieve better agreement with measurements and consequent performance measures. The schemes with a similar algorithm (Yamada and Andrén) create similar performance measures and similar dispersion model error. Increased vertical resolution appears to improve performance measures and reduce dispersion model error.

7. Summary and conclusions

We applied a method based on a stochastic approach of estimating model dispersion error to a case study of the propagation of tracers in complex terrain. The method was based on a combination of model fractional error-by-error components due to the emission, measurement, atmospheric and dispersion models, and stochastic processes. Based on a study by Koračin et al. (2000), total model error has been decomposed into separate errors due to atmospheric and dispersion models. Error components due to emission, measurement, and the atmospheric model were determined from measurements and modeling, and, in order to account for the compensating nature of error components, they were represented in the method algorithm in terms of a normal distribution with a specified mean and standard deviation. Error due to stochastic processes was represented in the method algorithm as a normal distribution with zero mean and unknown standard deviation. Dispersion model error and the standard deviation of the error due to stochastic processes then represent unknown numbers in the equation for fractional model error. Our methodology found solutions for dispersion error and the standard deviation of the stochastic error component using the maximum likelihood estimation. Two types of solutions are discussed: a simple one with time-independent error components as well as a complex solution with one of the errors, namely atmospheric error, diurnally variable. We conducted a series of sensitivity tests to investigate the impact of particular error components on the estimate of dispersion error and standard deviation of the stochastic component of the error. Our results show that the method results are robust on variation of the estimates of the known error components and yield stable solutions of the dispersion model error and standard deviation of the stochastic error.

We applied the developed method to observed and modeled tracer concentrations in complex terrain. The daily tracer concentrations for a period of seven days measured at 11 receptors and simulated by a Lagrangian random particle dispersion model that was using MM5 atmospheric simulation results were used to compute all error components. Two major structural components (turbulence parameterization and model vertical grid resolution) of the Lagrangian random particle model were investigated. Although model results achieved noticeably high correlation for all turbulence schemes compared to measurements (from 0.77 to 0.79), the dispersion errors were large (up to a factor of three) among model runs with various turbulence schemes. The range spread in the computed dispersion error also can be seen in the spread of other statistical performance measures shown in Table 5. These results confirm the importance of treating turbulence transfer in dispersion models and their effect on the accuracy of modeled dispersion. We also tested the same method of estimating dispersion model error by improving the vertical resolution of the atmospheric model for the same case. Increased vertical resolution in the atmospheric model generally improved the results from the statistical comparison between the simulated and measured concentrations but did not significantly improve magnitude of the dispersion error. As expected, the study confirms that one of the main structural components of the dispersion model, namely the turbulence scheme, bears much more importance on the magnitude of dispersion model error compared to increased vertical resolution in the atmospheric model.

Acknowledgments

The study was partially supported by DOD-DEPSCoR-ONR Grant N00014-96-1-1235. We thank Dr. Leif Enger of the University of Uppsala, Sweden, Dr. Michael Tjernström of the University of Stockholm, Sweden, and Dr. Julide Kahyaoğlu-Koračin of the Desert Research Institute (DRI), Reno, Nevada for helpful discussions and valuable comments. Mr. Travis McCord and Mr. Domagoj

Podnar (DRI) are acknowledged for assisting in development of the Lagrangian random particle model and performing simulations. The authors thank Mr. Roger Kreidberg of the Desert Research Institute for his thorough editorial assistance.

Disclaimer: The research presented here was performed under a Memorandum of Understanding between the US Environmental Protection Agency (EPA) and the US Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) and under agreement number DW13921548. This work constitutes a contribution to the NOAA Air Quality Program. It does not necessarily reflect agency policies or views.

References

American Society for Testing and Materials, 2000. Standard Guide for Statistical Evaluation of Atmospheric Dispersion Model Performance. Designation: D 6589-00. Available from American Society for Testing and Materials, 100 Barr Harbor Drive, P.O. Box C700, West Conshohocken, PA 19428-2959, 17pp.

Andrén, A., 1990. Evaluation of a turbulence closure scheme suitable for air-pollution applications. Journal of Applied Meteorology 29, 224–239.

Bergin, M.S., Noblet, G.S., Petrini, K., Dhieux, J.R., Milford, J.B., Harley, R.A., 1999. Formal uncertainty analysis of a Lagrangian photochemical air pollution model. Environmental Science and Technology 33, 1116–1126.

Casella, G., Berger, R.L., 1990. Statistical Inference. Duxbury Press, Belmont, CA.

Cox, W.M., Tikvart, J.A., 1990. A statistical procedure for determining the best performing air quality simulation model. Atmospheric Environment 24A, 2387–2395.

Donaldson, C.du.P., 1973. Three-dimensional modeling of the planetary boundary layer. In: Haugen, A. (Ed.), Workshop on Micrometeorology. American Meteorological Society, Boston, MA, pp. 313–390.

Draxler, R.R., Hess, G.D., 1997. Description of the Hysplit_4 modeling system, NOAA Tech Memo ERL ARL-224, December, 24p.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, San Francisco, 436pp.

Gifford, F.A., 1995. Some recent long-range diffusion experiments. Journal of Applied Meteorology 34, 1727–1730.

Green, M.C., 1999. The Project MOHAVE tracer study: study design, data quality, and overview of results. Atmospheric Environment 33, 1955–1968.

Grell, G.A., Dudhia, J., Stauffer, D.R., 1994. A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5). National Center for Atmospheric Research, Techn. Note TN-398, 122pp.

Hanna, S.R., 1982. Applications in air pollution modeling. In: Nieuwstadt, F.T.M., Van Dop, H. (Eds.), Atmospheric Turbulence and Air Pollution Modeling, pp. 275–310.

Hanna, S.R., 1989. Plume dispersion and concentration fluctuations in the atmosphere. In: Cheremisinoff, P.N. (Ed.), Encyclopedia of Environmental Control Technology. Air Pollution Control, vol. 2. Gulf Publishing Company, Houston, Texas, pp. 547–582 (Chapter 14).

Hanna, S.R., 1994. Mesoscale meteorological model evaluation techniques with emphasis on needs of air quality models. In: Pielke, R.A., Pearce, R.P. (Eds.), Mesoscale Modeling of the Atmosphere. Meteorological Monographs, vol. 25(47). American Meteorological Society, pp. 47–58.

Hanna, S.R., Lu, Z., Frey, C., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., Hansen, D.A., 2001. Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. Atmospheric Environment 35, 891–903.

Koračin, D., Enger, L., 1994. A numerical study of boundary layer dynamics in a mountain valley. Part 2. Observed and simulated characteristics of atmospheric stability and local flows. Boundary-Layer Meteorology 69, 249–283.

Koračin, D., Isakov, V., Frye, J., 1998. A Lagrangian particle dispersion model (LAP) applied to transport and dispersion of chemical tracers in complex terrain. Presented at the Tenth Joint Conference on the Applications of Air Pollution Meteorology, Phoenix, AZ, 11–16 January 1998. Paper 5B.5, pp. 227–230.

Koračin, D., Isakov, V., Podnar, D., Frye, J., 1999. Application of a Lagrangian random particle dispersion model to the short-term impact of mobile emissions. Proceedings of the Transport and Air Pollution conference, Graz, Austria, 31 May–2 June 1999.

Koračin, D., Frye, J., Isakov, V., 2000. A method of evaluating atmospheric models using tracer measurements. Journal of Applied Meteorology 39, 201–221.

Legg, B.J., Raupach, M.R., 1982. Markov Chain simulation of particle dispersion in inhomogeneous flows: the mean drift velocity induced by a gradient in Eulerian velocity variance. Boundary-Layer Meteorology 24, 3–13.

Mellor, G.L., Yamada, T., 1974. A hierarchy of turbulence closure models for planetary boundary layers. Journal of Atmospheric Science 31, 1791–1806.

National Research Council (NRC), 2004. Air Quality Management in the United States. The National Academies Press, Washington, DC, 20055.

Pielke, R.A., 1984. Mesoscale Meteorological Modeling. Academic Press, New York, 612pp.

Podnar, D., Koračin, D., Panorska, A., 2002. Application of artificial neural networks to modeling the transport and dispersion of tracers in complex terrain. Atmospheric Environment 36, 561–570.

Rao, K.S., 2005. Uncertainty analysis in atmospheric dispersion modeling. Pure and Applied Geophysics 162, 1893–1917.

Sax, T., Isakov, V., 2003. A case study for assessing uncertainty in local scale regulatory air quality modeling applications. Atmospheric Environment 37, 3481–3489.

Weil, J.C., Sykes, R.I., Venkatram, A., 1992. Evaluating air-quality models: review and outlook. Journal of Applied Meteorology 31, 1121–1145.

Wilson, J.D., Flesch, D.K., 1993. Flow boundaries in random-flight dispersion models: enforcing the well-mixed condition. Journal of Applied Meteorology 32, 1695–1707.

Yamada, T., 2000. Numerical simulations of airflows and tracer transport in the southwestern United States. Journal of Applied Meteorology 39, 399–411.