AERMOD: A Dispersion Model for Industrial Source Applications. Part II: Model Performance against 17 Field Study Databases

STEVEN G. PERRY,* ALAN J. CIMORELLI,+ ROBERT J. PAINE,# ROGER W. BRODE,@ JEFFREY C. WEIL,&
AKULA VENKATRAM,** ROBERT B. WILSON,++ RUSSELL F. LEE,@@ AND WARREN D. PETERS&&

*Air Resources Laboratory, NOAA, and National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina
+U.S. Environmental Protection Agency Region 3, Philadelphia, Pennsylvania
#ENSR International, Westford, Massachusetts
@MACTEC Federal Programs, Inc., Durham, North Carolina
&Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado
**College of Engineering, University of California, Riverside, Riverside, California
++U.S. Environmental Protection Agency Region 10, Seattle, Washington
@@Charlotte, North Carolina
&&OAQPS, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina

(Manuscript received 21 January 2004, in final form 26 October 2004)

ABSTRACT

The performance of the American Meteorological Society (AMS) and U.S. Environmental Protection Agency (EPA) Regulatory Model (AERMOD) Improvement Committee's applied air dispersion model against 17 field study databases is described. AERMOD is a steady-state plume model with significant improvements over commonly applied regulatory models. The databases are characterized, and the performance measures are described. Emphasis is placed on statistics that demonstrate the model's abilities to reproduce the upper end of the concentration distribution. This is most important for applied regulatory modeling. The field measurements are characterized by flat and complex terrain, urban and rural conditions, and elevated and surface releases with and without building wake effects. As is indicated by comparisons of modeled and observed concentration distributions, with few exceptions AERMOD's performance is superior to that of the other applied models tested. This is the second of two articles, with the first describing the model formulations.

1. Introduction

In 1991, the U.S. Environmental Protection Agency (EPA) in conjunction with the American Meteorological Society (AMS) formed the AMS and EPA Regulatory Model (AERMOD) Improvement Committee (AERMIC) with the expressed purpose of incorporating the current understanding of the planetary boundary layer (PBL) into a state-of-the-art applied dispersion model, AERMOD.

AERMIC's work clearly has benefited from the model development activities worldwide over the past few decades, especially in the parameterization of mean winds and PBL turbulence, dispersion in the CBL, the treatment of plume/terrain interactions, plume–building interactions, and urban dispersion.

AERMOD (Cimorelli et al. 2003) is a steady-state plume model aimed at short-range (up to 50 km) dispersion from stationary industrial-type sources—the same scenarios that are currently handled by the EPA's Industrial Source Complex Short-Term model (ISCST3) (U.S. Environmental Protection Agency 1995). The meteorological conditions are assumed to be steady during the modeling period (typically 1 h) and horizontally homogeneous. Vertical variations in the PBL, however, are incorporated into the model's predictions. For flow in complex terrain AERMOD incorporates the concept of a dividing streamline (Snyder et al. 1985). The model considers the influence of building wakes on plume rise and dispersion using the algorithms of the Plume Rise Model Enhancements (PRIME) model (Schulman et al. 2000). In urban areas, AERMOD accounts for the dispersive nature of the "convective like" boundary layer that forms during nighttime conditions by enhancing the turbulence resulting from urban heat flux (Oke 1978, 1982).

This paper is the second of two describing the newly developed AERMOD modeling system. Cimorelli et al. (2005, hereinafter Part I) describe the model formula-

---

Corresponding author address: Steven G. Perry, U.S. Environmental Protection Agency, MD-81, Research Triangle Park, NC 27711.
E-mail: perry.steven@epa.gov

TABLE 1. Description of field studies (without building wake effects).

Database	Description of field study
Prairie Grass (SO_2)	Very flat, rural (Nebraska); nonbuoyant single-point source; 0.46-m release; 44 data hours; SO_2 samplers in arcs out to 800 m; 16-m meteorological tower (wind, turbulence, and temperature data); Barad (1958) and Haugen (1959).
Kincaid (SF_6)	Flat, rural (Illinois), highly buoyant single source; tall stack release (187 m); 375 data hours; SF_6 samplers in arcs out to 50 km; 100-m tower (wind, turbulence, and temperature); Liu and Moore (1984) and Bowne et al. (1983).
Indianapolis (SF_6)	Flat, urban (Indiana), highly buoyant release (84 m); 170 data hours; SF_6 samplers in arcs out to 12 km; Urban tower (94 m); 10-m suburban and rural towers (wind, turbulence, and temperature); Murray and Bowne (1988).
Kincaid (SO_2)	Flat, rural (Illinois), highly buoyant single source; tall stack release (187 m); 4614 data hours; 30 samplers out to 20 km; 100-m tower (wind, turbulence, and temperature); Liu and Moore (1984) and Bowne et al. (1983).
Lovett (SO_2)	Hilly, rural (New York), highly buoyant release (145 m); 12 monitors out to 3 km; 1 yr of data; 100-m (wind, turbulence, and temperature); Paumier et al. (1992).
Baldwin (SO_2)	Flat, rural (Illinois); three highly buoyant stacks (184 m); 10 fixed samplers out to 10 km; 1 yr of data; 100-m (wind and temperature data); Hanna and Chang (1993).
Clifty Creek (SO_2)	Moderately hilly, rural (Indiana); three highly buoyant stacks (each 208 m); six fixed samplers out to 15 km; 1 yr of data; 60-m tower on nearby plateau, 115 m above stack base (wind and temperature data).
Martins Creek (SO_2)	Hilly, rural (Pennsylvania); multiple highly buoyant releases (122–183 m); 1 yr of data; seven fixed samplers out to 8 km; 10-m tower plus sodar (wind, turbulence, and temperature data).
Westvaco (SO_2)	Hilly, rural (Maryland); highly buoyant stack (183 m); 11 fixed samplers out to 3 km; 1 yr of data; two 30-m towers; 100-m tower (wind, turbulence, and temperature data); Strimaitis et al. (1987).
Tracy (SF_6)	Mountainous, rural (Nevada); moderately buoyant stack (91 m); 128 h of data; SF_6 samplers out to 8 km; 150-m tower (wind, turbulence and temperature data); tethersonde temperatures; acoustic sounder; DiCristofaro et al. (1985).

tion, while this paper provides an overview of the model's performance against the concentration observations at 17 field study databases. The studies include sites with flat and complex terrain, urban and rural conditions, and elevated and surface releases with and without building wake effects. The evaluation measures are focused on those that are relevant to regulatory applications, that is, emphasis on ability of the model to simulate the upper end of the concentration distributions. AERMOD estimates have been compared with those of other applied models, including ISCST3 (U.S. Environmental Protection Agency 1995), the Hybrid Plume Dispersion Model (HPDM) (Hanna and Paine 1989), the Rough Terrain Diffusion Model (RTDM) (Paine and Egan 1987), and the Complex Terrain Dispersion Model Plus Algorithms for Unstable Situations (CTDMPLUS) (Perry 1992).

2. Model evaluation field studies

Of the 17 databases that were considered, 10 were designed to collect data for overall model performance where building wakes were not an issue, while the remaining 7 were specifically focused on building influences. The studies are summarized in Tables 1 and 2. Maps of the various sites can be found in Paine et al. (1998, 2003). The first five databases listed in Table 1 were used during the AERMOD development process to identify major problems with the model algorithms but generally were not used to set empirical parameters to improve the model results. An exception is found with the use of the Prairie Grass Experiment data to specifically develop the formulation for the lateral dispersion parameter. The remaining five databases were independently applied to the developed model code. The first four building wake databases in Table 2 were each subdivided to provide data for both model development and evaluation of the PRIME building downwash algorithms. However, all of the data from the seven building wake databases were used in the performance results described in this paper.

3. Performance measures

Although the model evaluation examined the quality of the predictions relative to the model physics, the results reported here are focused primarily on answering the questions: how well does AERMOD predict the high-end, ground-level concentrations that are generally used to assess compliance with air quality regulations; and is AERMOD's performance distinguishably better than that of other applied models for this purpose? To answer these questions the analyses of the model's performance utilized all of the relevant input data that are available with each dataset. The performances of earlier versions of AERMOD were examined with reductions in the number of height levels in the measured profiles of PBL variables (e.g., wind, temperature). Paine (2003) found that the performance of the model tended to degrade as more and more levels of data were removed from the analysis. In general, the model predictions tended toward higher concentrations and, thus, provided more conservative results when compared to the observed concentrations.

In the absence of model formulation errors and stochastic variations, the major reasons for deviations between model estimates and observations are errors in the model inputs, and the concentration observations

TABLE 2. Description of field studies (with building wake emphasis).

Database	Description of field study
Bowline Power Plant (SO_2)	Rural, locally flat (New York); buoyant twin stacks; 87-m release; dominant building height = 65 m; SO_2 samplers at 250 and 850 m; 100-m meteorological tower (wind, temperature); full year of data; Schulman and Hanna (1986).
Millstone Nuclear Plant (SF_6, CF_3Br)	Rural (coastal Connecticut) with terrain variation < 10 m; nonbuoyant releases at 48 and 29 m; building height = 45 m; sampling on arcs from 350 to 1500 m; 43-m meteorological tower (wind and temperature; mostly high winds and onshore flow); Bowers and Anderson (1981).
Duane Arnold Energy Center (SF_6)	Rural (Iowa); with terrain variations up to 30 m; rooftop (nonbuoyant) releases at 46 and 24 m plus ground-level releases; samplers on arcs at 300 and 1000 m; 50-m meteorological tower (mostly light wind, convective conditions); Thullier and Mancuso (1980).
Alaska North Slope (SF_6)	Isolated, very flat (Prudhoe Bay area, Alaska); buoyant release at 39 m; building height of 34 m; samplers in seven arcs from 50 to 3000 m; 33-m meteorological tower (wind, temperature, and velocity variance; primarily stable to very stable conditions); Guenther et al. (1989) and Guenther and Lamb (1990).
American Gas Association Study (SF_6)	Rural, flat (Texas and Kansas); highly buoyant releases varied from about 1 to 2.5 times building heights; sampler arrays from 50 to 200 m; 10-m meteorological tower; Engineering Science (1980).
EOCR Study (SF_6)	Rural (Idaho); terrain variations < 10 m; nonbuoyant releases at 30 m, 25 m, and near ground level; building height = 25 m; sampling arcs at seven distances from 50 to 1600 m; wide range of stabilities and wind speeds; many stable; Start et al. (1981).
Lee Power Plant (wind tunnel study)	Rural simulation, flat (wind tunnel); buoyant release at 1.5*H* (*H* is model building height); sampling at arcs of 150–900 m (full scale); neutral and stable conditions; stack Froude numbers varied; wind directions varied; Melbourne and Taylor (1994).

themselves. An individual model prediction will most likely differ from the corresponding observation because the model cannot include all of the variables that affect the observation at a particular time and location. It is the experience of model developers (e.g., Weil 1992 and Liu and Moore 1984) that wind direction uncertainties alone can, and do, cause disappointing results from what otherwise may be well-performing dispersion models. However, a model that is based on appropriate characterizations of the important physical processes should be able to reproduce the distribution of observations as long as the range of model inputs is similar to that of the observations (Venkatram et al. 2001). A model with the ability to adequately predict the distribution of concentrations provides information for regulatory questions, such as what is the probability that a certain concentration is exceeded?

Concentration distributions can be readily assessed with quantile–quantile (Q–Q) plots (Chambers et al. 1983) that are created by ranking the predicted and observed concentrations and then pairing by rank. Specifically, a good model will have a slope in this plot similar to that of the 1:1 line and, specifically for regulatory applications, will have values in the important upper end of the distribution near to those of the measurements. Paumier et al. (1992) demonstrated the usefulness of Q–Q plots in characterizing the performance of CTDMPLUS. Venkatram et al. (2001) argue for the use of Q–Q plots for evaluating regulatory models.

Furthermore, the ability of AERMOD to assess the high end of the concentration distribution is examined with the robust highest concentration (RHC) statistic (Cox and Tikvart 1990). The RHC represents a smoothed estimate of the highest concentrations based on an exponential fit to the upper end of the concentration distribution:

$$\mathrm{RHC} = \chi\{n\} + (\chi - \chi\{n\}) \ln\left(\frac{3n-1}{2}\right), \tag{1}$$

where n is the number of values used to characterize the upper end of the concentration distribution, χ is the average of the $n - 1$ largest values, and $\chi\{n\}$ is the nth largest value; $n = 26$ (suggested by Cox and Tikvart 1990) for most comparisons reported here. RHC is a preferred statistic because it yields a representative high-end estimate while mitigating the undue influence of individual unusual events. In summary, for regulatory applications, a good model would produce a concentration distribution parallel to the slope of the measured distribution and produce high-end concentrations (RHCs) that are similar to that of the observations.

Other methods for comparing model performance could have been applied in this analysis [e.g., the recently approved American Society for Testing and Materials (ASTM) methodology for comparing the performance of dispersion models (Irwin et al. 2003)], but the focus here is on the estimates of the high end of the concentration distributions.

4. Model performance results

Of the 17 databases considered, seven emphasized near-field concentrations resulting from building wake effects. Four of the no-wake studies involve short-term, intensive measurements with extensive sampler arrays, while six include long-term, continuous sampling at more limited locations. In the intensive studies, where experimental periods are rarely continuous, only 1-h averages are considered. With the long-term studies, results are also reported for 3-h, 24-h, and annual averages.

a. *Model comparisons with data from intensive studies (no building wakes)*

These studies involve a nonbuoyant, surface release in very flat terrain (Prairie Grass, Nebraska), an elevated buoyant release in flat terrain (Kincaid, Illinois; SF_6), an elevated buoyant release in a midsized urban area (Indianapolis, Indiana), and an elevated, weakly buoyant release in mountainous terrain (Tracy, Illinois). For these data, observations and predictions correspond to maximum concentrations on each arc of samplers to minimize the effect of wind direction uncertainties. The RHC and Q–Q plots for each model and field study are developed from the ranked and paired distributions of observations and predictions. Table 3 summarizes the ratio of modeled to observed RHC values for AERMOD, ISCST3, and CTDMPLUS. CTDMPLUS was designed for applications in complex terrain and is, therefore, compared with AERMOD for Tracy, which is a mountainous location.

The RHC ratios reveal generally good performance for AERMOD at all four sites. The model shows a tendency to underpredict the higher concentrations for the flat-terrain, rural sites (Prairie Grass and Kincaid) and a tendency to overpredict at the urban and mountainous sites. In all cases, AERMOD shows improvement over ISCST3, which, except for Kincaid, tends to overpredict (particularly in complex terrain). For Tracy, AERMOD is found to be unbiased, while CTDMPLUS shows underprediction. This is particularly interesting because Tracy was one of the primary databases used in the original development of CTDMPLUS.

TABLE 3. Ratio of modeled to observed robust highest concentrations—intensive studies.

Database	AERMOD (1 h)	ISCST3 (1 h)	CTDMPLUS (1 h)
Prairie Grass	0.87	1.50	—
Kincaid SF6	0.77	0.68	—
Indianapolis	1.18	1.30	—
Tracy	1.07	2.81	0.77

Figure 1 shows the Q–Q plots that are relevant to these models and the intensive databases. The slope of the ISCST3 concentration distribution compares well to that of the Prairie Grass data (Fig. 1a) to a great degree because ISCST3 dispersion is based on the Pasquill–Gifford–Turner (PGT) curves (Pasquill 1961; Gifford 1961). However, the modeled values tend to remain consistently above the observations (RHC ratio of 1.5). AERMOD (with an RHC ratio of 0.87) also shows a concentration distribution that matches observations well, suggesting that both models are capable of simulating near-field dispersion for a near-surface release. In contrast, for an elevated plume in flat terrain (Fig. 1b; Kincaid SF_6 data), both models have distributions that drop below the observations for the lower concentrations. Because short-term maximum concentrations from elevated plumes are generally associated with convective conditions, this dropoff is related primarily to stable conditions. This is borne out upon comparing the Kincaid convective and stable residual plots (not shown here), a pairing of modeled and predicted concentrations in time and downwind distance (Paine et al. 1998). Figure 1b also displays a distinction between the two models' performance in the upper end of the distribution with AERMOD tracking the observations noticeably better. This stems from the fact that AERMOD's plume formulation is non-Gaussian in convective conditions, in keeping with field and laboratory observations (Willis and Deardorff 1981; Baerentsen and Berkowicz 1984; Weil et al. 1997).

The Indianapolis database provides a test of model performance in a moderately sized urban area with a population of approximately 700 000. Based on the RHC ratios in Table 3, both ISCST3 and AERMOD perform well. For this single stack situated near the downtown business district, the Q–Q plot (Fig. 1c) shows AERMOD's superiority in tracking the full distribution of observed concentrations; suggesting an added ability to predict longer averaging times. Although not shown here, an examination of the residuals (Paine et al. 1998) found that for convective conditions, during which most of the highest measured concentrations were observed, AERMOD performed well at all downwind distances. For stable conditions, the residuals show that the model underpredicted the very small concentrations occurring within a kilometer of the stack but performed better for the more distant, higher concentrations.

Of the intensive tracer studies, the most notable difference in the performance of AERMOD and ISCST3 appears with the Tracy data. Tracy, a tall buoyant stack in mountainous terrain, is equipped with a very high quality meteorological and tracer sampling network (Fig. 1d). The Tracy data were collected during predominately stable conditions. The concept that the plume in stable conditions is influenced by the flow in the layers separated by the dividing streamline has been shown to be integral to plume modeling in complex terrain. This concept is central to the formulations in both AERMOD and CTDMPLUS but is not considered by ISCST3. The distributions of AERMOD and CTDMPLUS are well matched throughout the range of the Tracy data. ISCST3 estimates exhibit an approximate factor-of-3 overprediction.

b. *Model comparisons with data from long-term, continuous studies (no building wake)*

The six long-term field studies provide data for both individual and multiple elevated, buoyant stack releases. The study sites are located in predominantly rural areas of flat to complex terrain (with topography in some cases extending above stack height). Receptor fields were more limited for these studies, such that

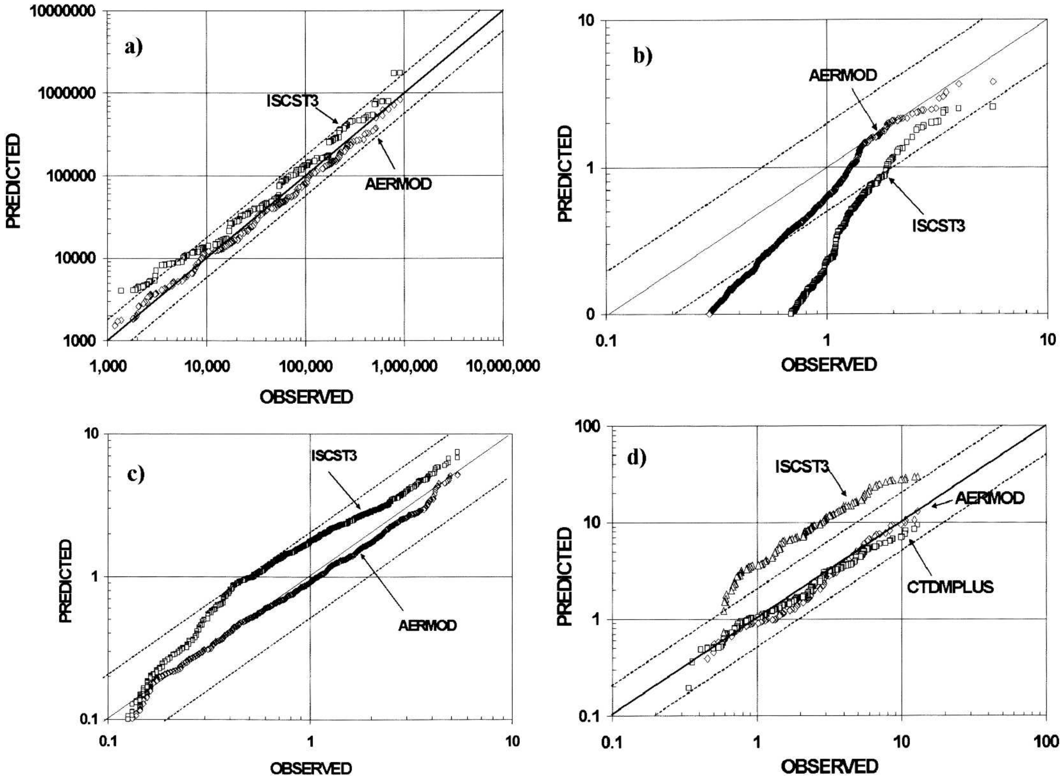

FIG. 1. Quantile–quantile plot of model-predicted vs observed 1-h-averaged concentrations (μg m^{-3}) for (a) Prairie Grass data, (b) Kincaid SF$_6$ data, (c) Indianapolis data, and (d) the complex terrain case of Tracy Power Plant. Solid line indicates a one-to-one correspondence in the concentration distributions; the dashed lines indicate factor-of-2 over- and underestimates.

well-defined arcs as a function of downwind distance did not exist. Therefore, the distributions of modeled and observed concentrations are based on the 1-h maximum concentrations throughout the network.

Table 4 shows RHC ratios for the two studies in relatively flat terrain (Kincaid SO$_2$ and Baldwin, Illinois). AERMOD does very well in capturing the upper end of the distribution for both the 3- and 24-h averaging period. This is confirmed by the Q–Q plots of Figs. 2 and 3. ISCST3 displays a considerable underprediction at Kincaid and an overprediction at Baldwin for the upper end of the 3-h-averaged concentrations. Although the two studies share similar stack heights and parameters, Kincaid involves a single stack, while Baldwin has three that are separated over a distance of 100 m. Baldwin also has many fewer sampler locations than Kincaid. The better performance shown by AERMOD is not unexpected in these comparisons because the shorter averaging times are dominated by near-field impacts that are controlled by convective conditions for which AERMOD's formulation is superior. Additionally, the HPDM model (with a convective formulation not unlike that of AERMOD) predictions matched the Baldwin observations exceptionally well for all of the averaging periods (Table 4 and Fig. 3).

Clifty Creek, Indiana, is an interesting database because this highly buoyant source, with nearly collocated stacks, has near-field monitors at elevations of about 0.5 times the stack height in elevation. Both AERMOD and ISCST3 perform well for 3- and 24-h averages (Table 4 and Fig. 4). HPDM shows a tendency to overpredict the RHC ratios particularly for the 24-h average but has a distribution that otherwise "parallels" the observations well. The hilly terrain setting at Clifty Creek provided the only full-year database for which ISCST3 showed a very good performance.

The three complex terrain databases (Lovett, New York; Martin's Creek, Pennsylvania; and Westvaco, Maryland) all contain samplers at elevations above the stack top and at locations generally between 2 and 8 km from the stacks. Lovett and Westvaco each involve a single stack while Martin's Creek has emissions from three identical stacks with horizontal separations on the order of 100–200 m. All stacks in these studies are tall (145–189 m), and emissions are highly buoyant.

AERMOD performed well for all three complex terrain studies with RCH ratios of 1.0–1.65 for the 3-h and 24-h averages (Table 4 and Figs. 5, 6, and 7). This is a very satisfying result given the complexity of the terrain, the source configurations, and the relatively few sampling locations in these studies. When examined in

TABLE 4. Ratio of modeled to observed robust highest concentrations—continuous studies.

Database	Time avg	AERMOD	ISCST3	CTDMPLUS	HPDM	RTDM
Kincaid SO_2	3 h	1.02	0.56	—	—	—
	24 h	0.97	0.45	—	—	—
	Annual	0.31	0.14	—	—	—
Baldwin	3 h	1.35	1.48	—	1.06	—
	24 h	1.04	1.13	—	1.02	—
	Annual	1.00	0.63	—	1.15	—
Clifty Creek	3 h	1.26	0.98	—	1.33	—
	24 h	0.73	0.67	—	1.46	—
	Annual	0.55	0.31	—	0.96	—
Lovett	3 h	1.00	8.20	2.37	—	—
	24 h	1.00	9.11	2.01	—	—
	Annual	0.79	7.49	1.34	—	—
Martins Creek	3 h	1.06	7.25	4.80	—	3.33
	24 h	1.65	8.88	5.56	—	3.56
	Annual	0.76	3.37	2.19	—	1.32
Westvaco	3 h	1.08	11.00	2.14	—	—
	24 h	1.14	8.74	1.54	—	—
	Annual	1.65	10.33	0.93	—	—

conjunction with the Tracy results, it is clear that AERMOD is very capable of estimating the important regulatory concentrations for situations in which the plume is either impinging directly on the terrain or significantly interacting with the distorted flow near and above the terrain. This reflects well on the robust nature of the model.

The remaining three models tested in complex terrain were all consistently high in their estimates. For short-term averages, CTDMPLUS's and RTDM's RHC ratios ranged over a factor of 2–5 too high, and modeled estimates remained overpredictive over much of the distributions. Despite the fact that CTDMPLUS accounts for much of the flow details, AERMOD was found to be superior in determining the distribution of concentrations (particularly the important high end). ISCST3 demonstrates consistently poor performance for all three studies, overpredicting the high concentrations by on the order of a factor of 10!

All of the Q–Q plots reflect a dropoff in the modeled distributions for the low concentrations. This has obvious implications for the annual average estimates. As Table 4 suggests, both ISCST3 and AERMOD have some problems with the annual average estimates. For flat and simple terrain, the models underpredict the observed annual averages. For complex terrain, ISCST3 overpredicts observations by a factor of 3–10. The use of SO_2 as a tracer and the detection limits of the samplers contribute to these poor comparisons. When determining background, residual concentrations from previous periods interfere with its true estimate. Also, SO_2 monitors typically have a detection limit on the order of 16 $\mu g\ m^{-3}$. Concentrations below this are set to one-half of the limit even though actual concentrations may be much less or even zero. These uncertainties, in combination with a great many small sampled concentrations throughout extended periods, reflect on the reliability of the long-term averages. Peak short-term concentrations are much less impacted because background estimates and instrument limitations are a much smaller percentage of the total concentration. In contrast, CTDMPLUS and RTDM do not underestimate the annual averages, perhaps due somewhat to their overestimates for the short-term averages. HPDM does very well for the annual averages for the two databases with which it was compared. For the applications for which HPDM was designed (particularly, tall-stack, simple-terrain databases) it performs like a state-of-the-science model. Additional evaluations of HPDM can be found in Hanna and Chang (1993).

c. *Model comparisons with databases where building wakes are important*

Of the seven building wake studies (Table 2), Millstone Nuclear Plant in Waterford, Connecticut, Duane Arnold Energy Center (DAEC) in Cedar Rapids, Iowa, Alaska North Slope, Alaska, and (one-half of) the Bowline Power Plant in Haverstraw, New York, databases were used to some degree in the development of the PRIME algorithms that have been most recently implemented within the AERMOD framework. A summary of the AERMOD and ISCST3 performances against these building wake studies with RHC as the indicator is shown in Table 5. A more detailed discussion, including comparisons using Q–Q plots, can be found in Paine et al. (2003).

Looking first at the Bowline Power Plant study (the only full-year, continuous building wake database), there is little difference in the two models. Both show a modest overprediction for the upper end of the concentration distribution. However, the short-term estimates (3 h), which are most dominated by downwash conditions, are nearly unbiased. The increased overprediction of RHCs for the longer averaging times suggests that the models may be finding a higher incidence of

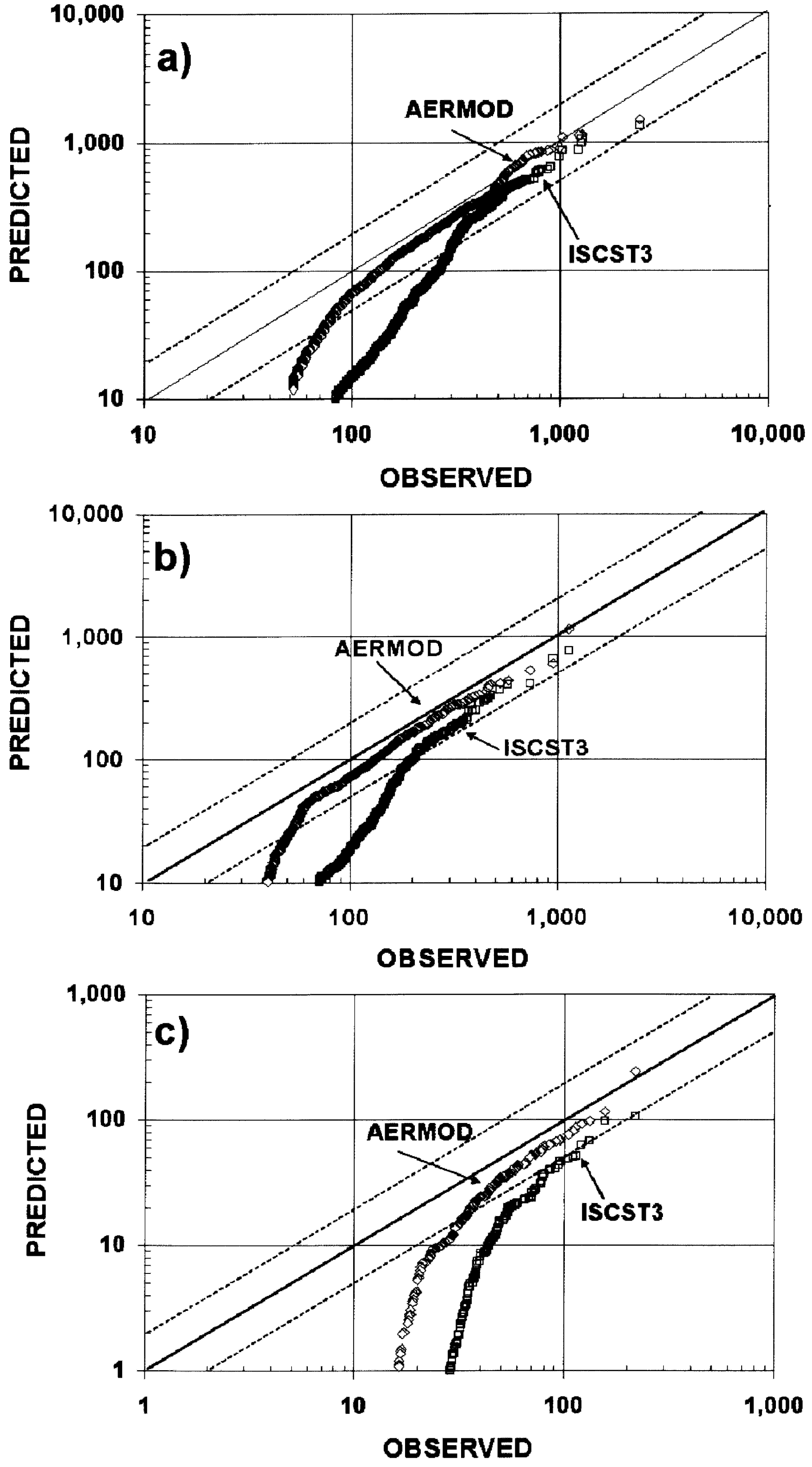

FIG. 2. Quantile–quantile plot of model-predicted vs observed concentrations (μg m^{-3}) for the Kincaid SO_2 database for (a) 1-, (b) 3-, and (c) 24-h averages. Dashed and solid lines mean the same as in Fig. 1.

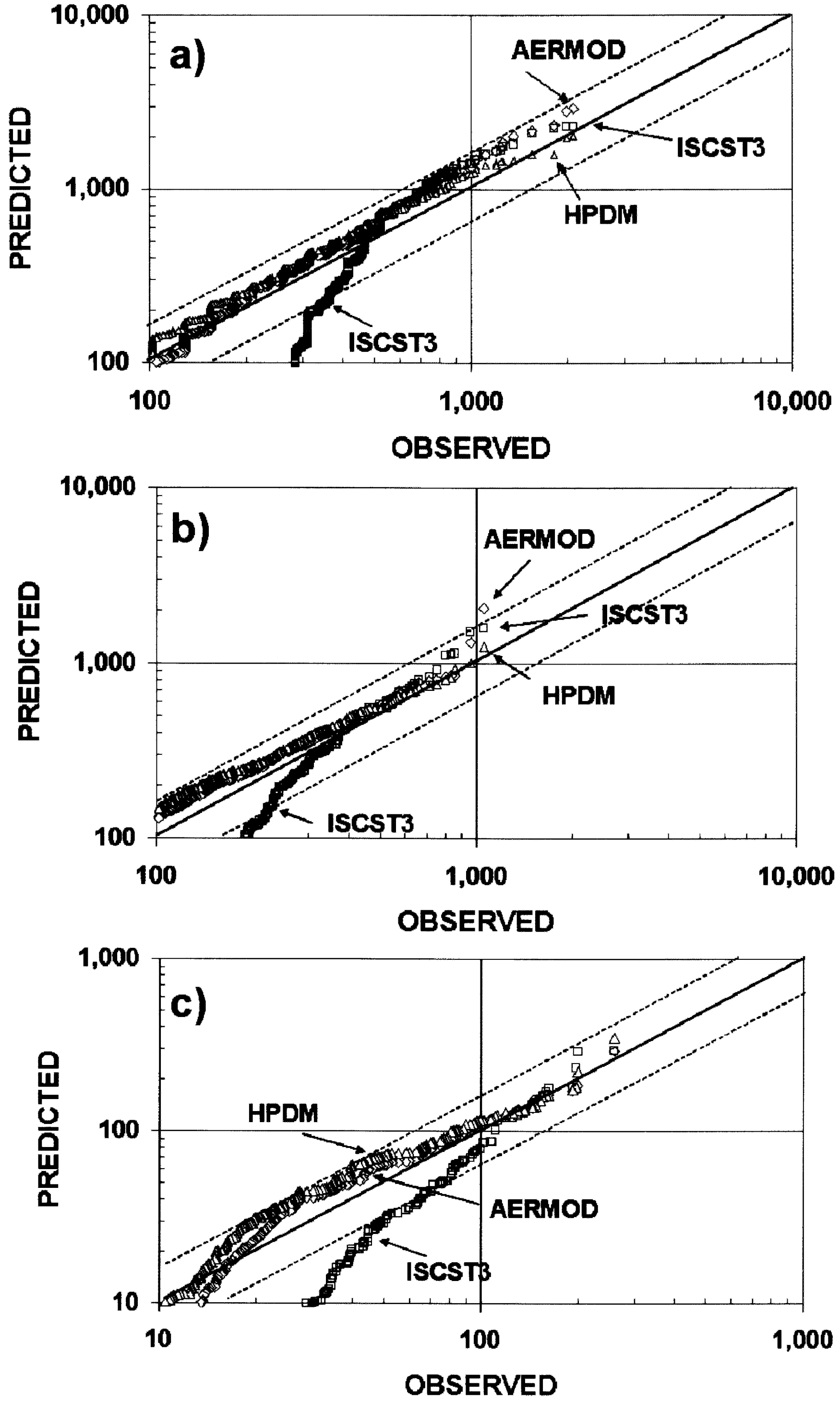

FIG. 3. Quantile–quantile plot of model-predicted vs observed concentrations (μg m^{-3}) for the Baldwin SO_2 database for (a) 1-, (b) 3-, and (c) 24-h averages. Dashed and solid lines mean the same as in Fig. 1.

downwashed plumes that are represented in the measurements.

The Millstone Nuclear Plant facility provides an opportunity to examine the sensitivity of the model estimates to plume/cavity geometry, because its nonbuoyant plumes are released at the height of the building and at one-half of that height. With release of material in the cavity (29-m release), AERMOD is overpredict-

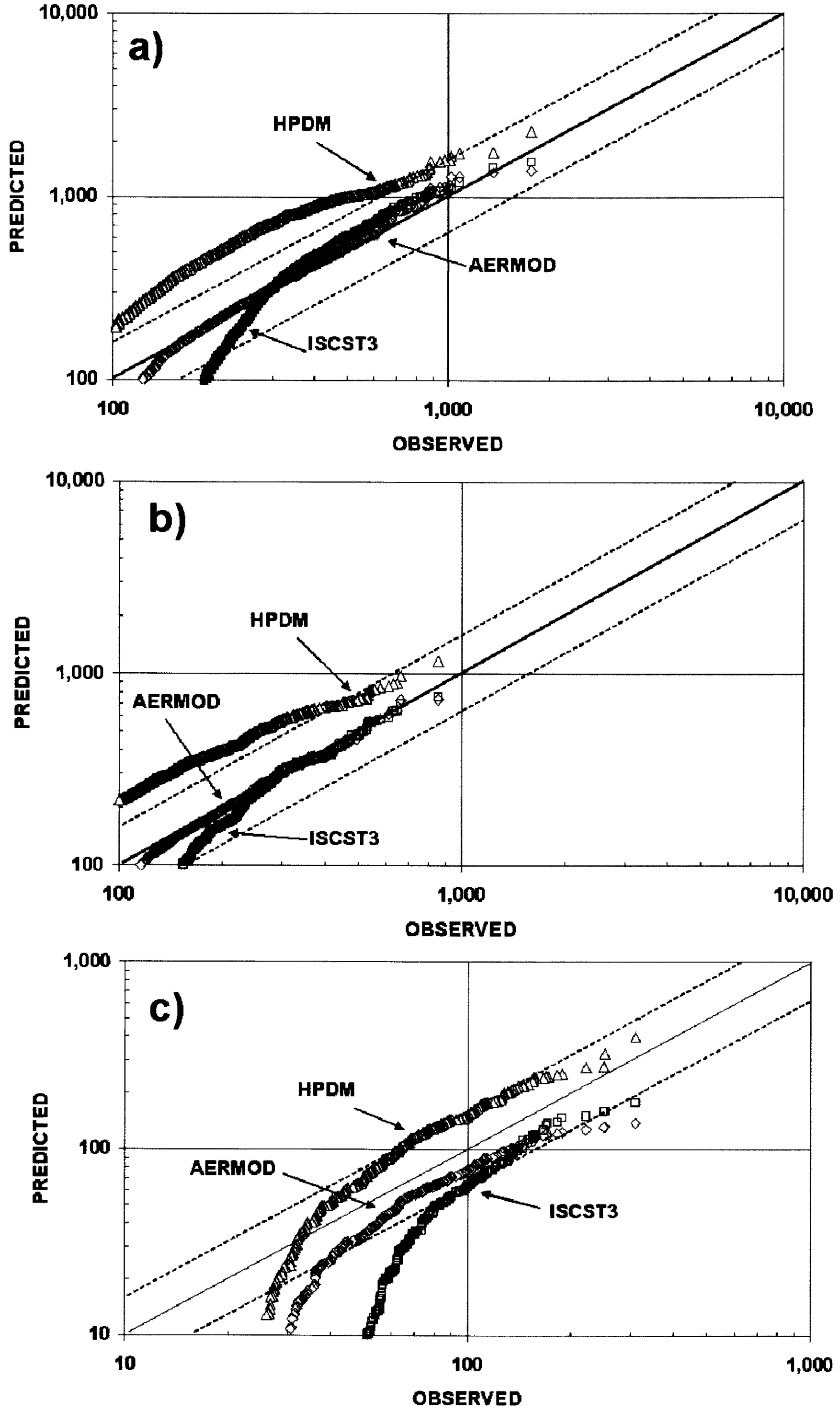

FIG. 4. Quantile–quantile plot of model-predicted vs observed concentrations (μg m^{-3}) for the Clifty Creek SO_2 database for (a) 1-, (b) 3-, and (c) 24-h averages. Dashed and solid lines mean the same as in Fig. 1.

ing the RHC. The plume is assumed to be well mixed in the cavity. However, when the plume is released near the boundary of the cavity or wake (near or slightly above the building top), the specification of those boundaries apparently becomes critical to the determination of peak ground-level concentrations. With an underprediction of over a factor of 2 for the near-building-height release, it appears that the models may

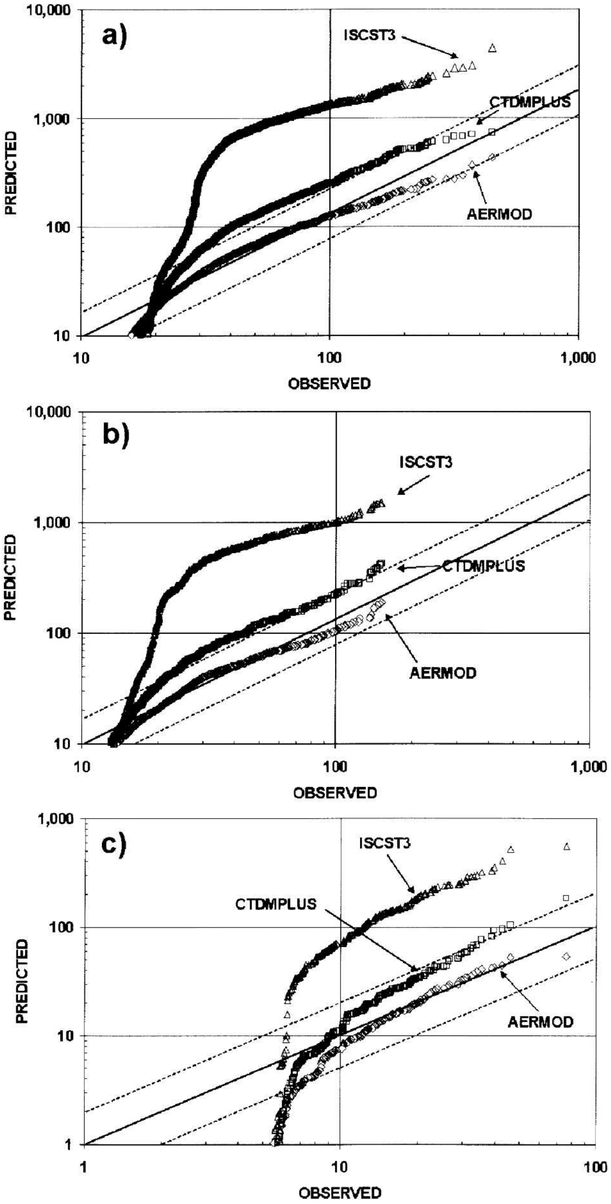

FIG. 5. Quantile–quantile plot of model-predicted vs observed concentrations (μg m^{-3}) for the Lovett SO_2 database for (a) 1-, (b) 3-, and (c) 24-h averages. This is a complex terrain site. Dashed and solid lines mean the same as in Fig. 1.

not be specifying the plume/cavity relationship particularly well.

The remainder of the databases yields comparisons with the model estimates that considerably emphasize the sensitivity of the models to the geometry and meteorological characteristics of a particular site. For example, the Duane Arnold and Experimental Organic Cooled Reactor (EOCR), Idaho Falls, Idaho, studies both involve nonbuoyant, near-building-height (or below) releases, yet the Duane Arnold study was dominated by convective conditions and EOCR by stable conditions. As Table 5 shows, the models consistently underpredict at Duane Arnold and equally overpredict for EOCR.

A more compounded contrast can be found with the Duane Arnold and the Alaska North Slope studies. For the nonbuoyant plumes at Duane Arnold, the conditions were mostly convective, while in Alaska the plume was buoyant and conditions were very stable. Again, at Duane Arnold AERMOD underpredicts (particularly for the lower releases), and yet the model shows little bias for the Alaska cases. The sensitivity of the model estimates to the specification of the plume/cavity/wake geometry is suggested.

The Alaska North Slope results (with very stable conditions) also contrast with the stable wind-tunnel results for the Lee Power Plant (Pelzer, South Carolina) case. These studies are similar in plume/wake geometry with the difference being that, in some cases, the wind-tunnel study had releases much higher than the building. Despite the fact that the meteorological conditions and geometry were comparable in these two studies, the model is unbiased in its estimates for Alaska and yet overpredicts by more than a factor of 2 for the stable wind-tunnel comparisons.

In general, it is not surprising that these results are complex for modeling regimes that are themselves complex. It seems that a combination of meteorological conditions (obviously affecting plume rise) and the building geometry and, thus, specification of wake and cavity dimensions highly influences these high-end concentration predictions represented by the RHC ratios in Table 5. One summary observation after reviewing all the results (Table 5) is that AERMOD is generally capable of capturing the important regulatory concentrations within a factor of 2 or better. Further comparisons (particularly with the wind-tunnel data) may help to articulate the specific algorithms in the model for which sensitivity is greatest.

d. AERMOD performance for area and volume sources

Because all 17 studies previously discussed in this paper involved point source releases, one particular study in the literature that examines the model's performance for area and volume source types is worth discussing here. A report commissioned by the American Petroleum Institute (Hanna et al. 1999) was summarized in a paper by Hanna et al. (2000) in which they examined the performance of AERMOD against the Kincaid, Indianapolis, and Lovett databases, already discussed here, and two additional studies—a nonbuoyant release within a refinery complex (OPTEX data) and a nonbuoyant release from area and volume source configurations in an open grassy area (Duke Forest, North Carolina, data). Both of these studies, described in some detail by Hanna et al. (1999), involved multiple release points that are generally near the surface. The OPTEX data were intended to simulate point, area,

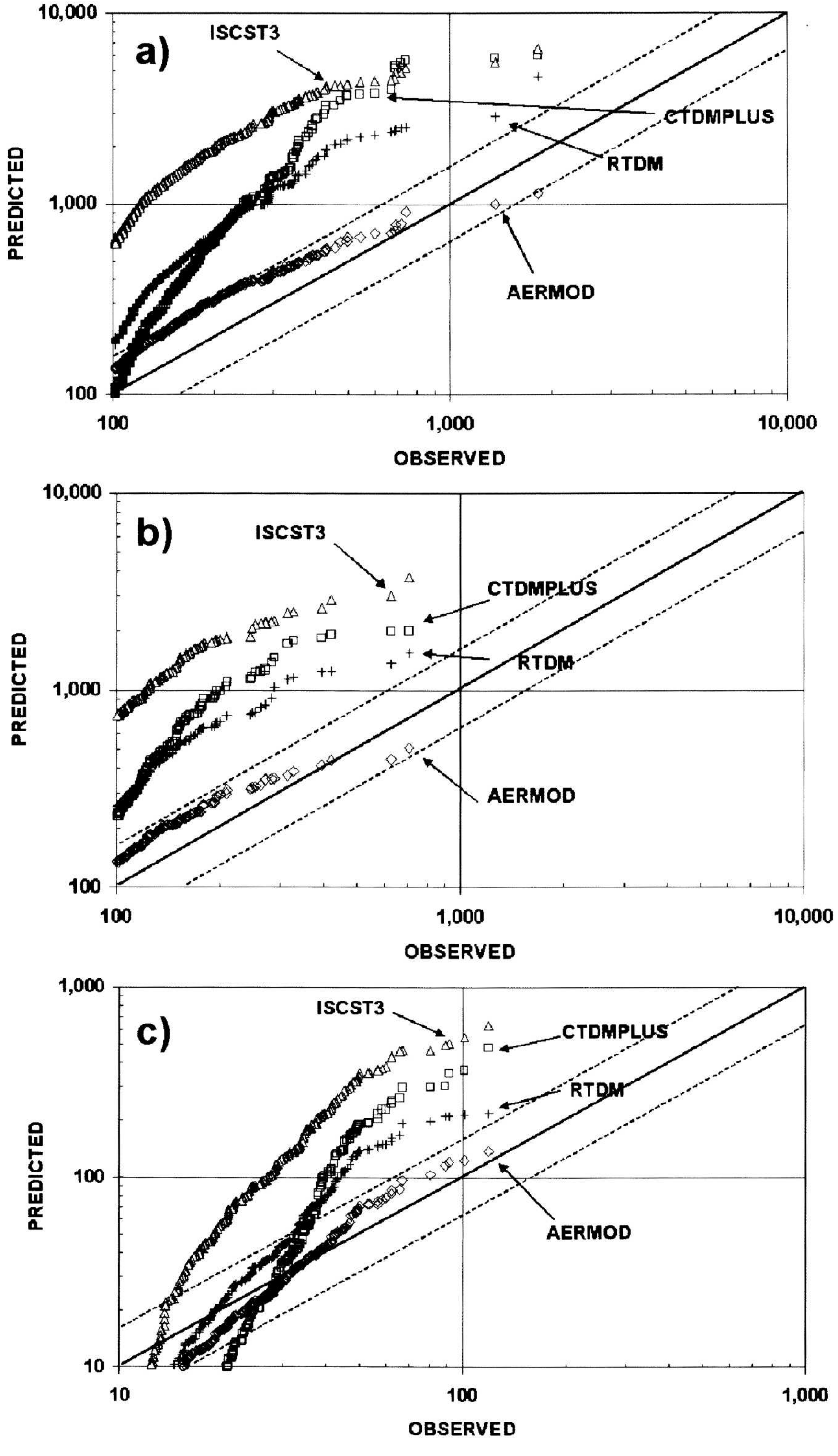

FIG. 6. Quantile–quantile plot of model-predicted vs observed concentrations (μg m^{-3}) for the Martin's Creek SO_2 database for (a) 1-, (b) 3-, and (c) 24-h averages. This is a complex terrain site. Dashed and solid lines mean the same as in Fig. 1.

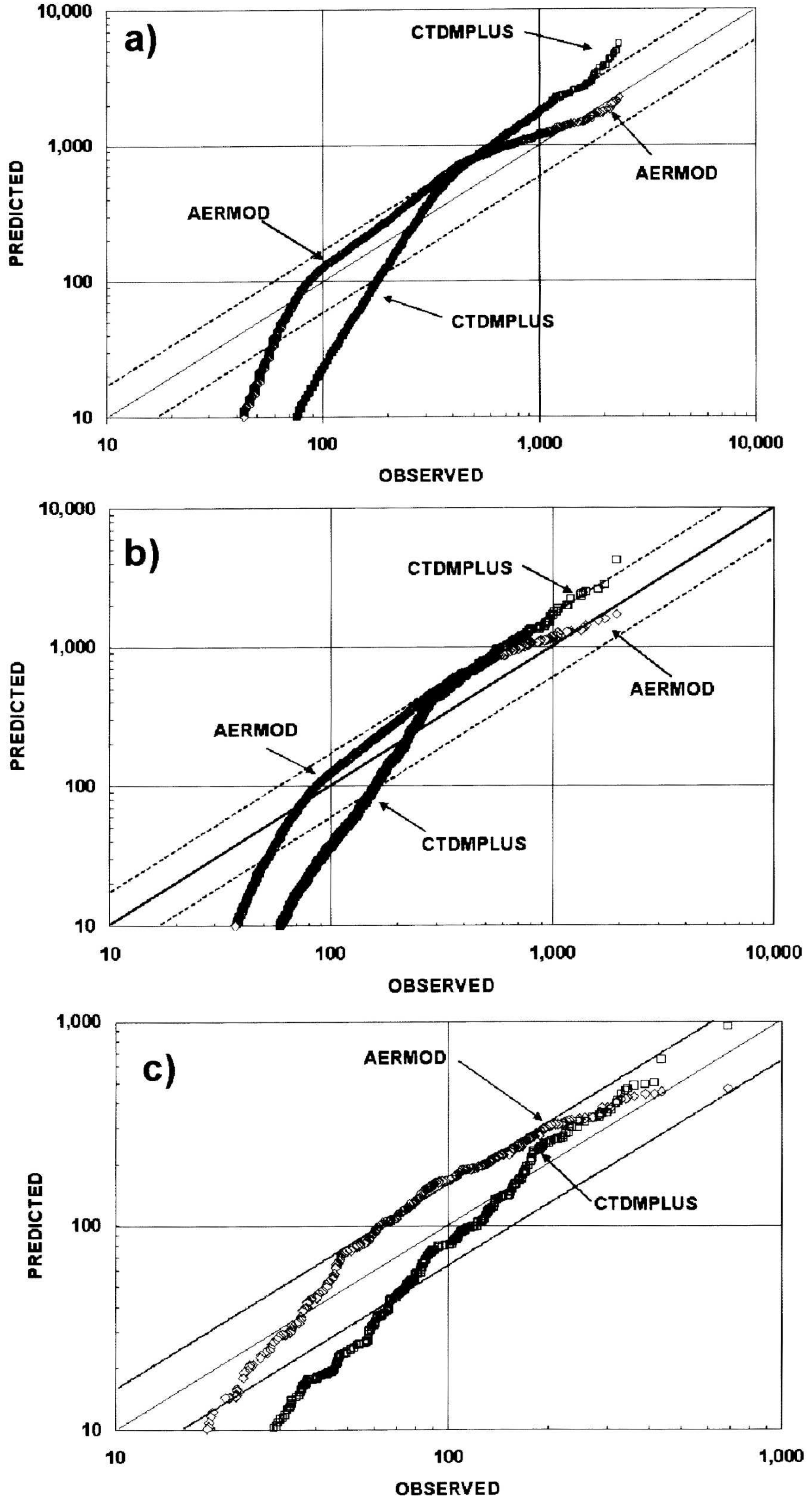

FIG. 7. Quantile–quantile plot of model-predicted vs observed concentrations (μg m^{-3}) for the Westvaco SO$_2$ database for (a) 1-, (b) 3-, and (c) 24-h averages. This is a complex terrain site. Dashed and solid lines mean the same as in Fig. 1.

TABLE 5. Ratio of modeled to observe robust highest concentrations—downwash studies.

Database	Time avg	AERMOD	ISCST3-PRIME
Bowline Point	3 h	1.14	1.23
	24 h	1.43	1.42
	Annual	1.50	1.35
Alaska North Slope	1 h	1.06	1.49
Duane Arnold	1 h (1-m release)	0.51	0.38
	1 h (24-m release)	0.25	0.29
	1 h (46-m release)	0.69	0.76
Millstone Nuclear Power Plant	1 h (29-m release)	1.32	1.42
	1 h (46-m release)	0.44	0.41
American Gas Association	1 h	0.92	0.76
EOCR	1 h	1.72	1.69
Lee Power Plant	1 h (stable)	2.50	2.11

line, and volume source types within a refinery complex. The experiments were separated into what was called a "matrix source" configuration and a "tank-farm source" configuration. The matrix source involved up to nine point sources (with heights of 2, 15, and 40 m) that were placed among the refinery piping. The tank source involved releases from 1.5 m above the surface in the vicinity of storage tanks. At the Duke Forest site, arrays of point sources were arranged to simulate area and volume sources with release heights from about 3 to 8 m. Unlike the previously discussed databases for which the sources were generally elevated point releases, these databases challenge the model's ability to simulate the dispersion of area and volume sources near the surface.

AERMOD had difficulty simulating the dispersion from the OPTEX tank source, perhaps because the model was applied using old building downwash algorithms. The geometric mean ratio of AERMOD to monitored concentration was 2.47. The model has been modified, since these comparisons, to include the downwash algorithms of the PRIME model as discussed earlier. For the OPTEX matrix source and the Duke Forest comparisons, AERMOD was found to perform very well with geometric means of 1.02 and 1.42 respectively. It should be noted that the Hanna et al. (2000) geometric means are computed for concentrations paired in time (a much more demanding comparison) and give no special emphasis for the higher concentrations, as was discussed earlier with the RHC calculations. Additionally, AERMOD was found to provide predictions for all three datasets that were within a factor of 2 of the observations over 70% of the time.

5. Summary and discussion

The formulations of the newly developed AERMOD steady-state plume dispersion model are described in the companion to this paper (Part I). The model was designed to fill the niche of applications currently filled by ISCST3 where comparisons with air quality standards are important. Because of the inherent uncertainties in individual model simulations, plume models, such as AERMOD, find their greatest potential for success in simulating the overall distributions of concentrations related to a wide variety of modeling conditions. The new modeling system's concentration distributions have been compared with those of 16 separate tracer (field) studies and one laboratory wind tunnel study. The primary purpose of this evaluation study is to identify the scenarios for which AERMOD shows good performance and those for which it may be lacking. Additionally, there is a desire to examine the value obtained in these improved formulations by comparing AERMOD with some currently used regulatory models with these varied databases.

Ten databases avoided the complications of building wakes to challenge the remainder of the model algorithms. The other seven studies provided a focused look at cases in which building downwash was dominant. Of the 10 nonwake studies, AERMOD found its greatest overall success in reproducing the concentration distributions for buoyant, tall-stack releases in moderate to complex topography (Lovett, Martin's Creek, Westvaco, and Tracy). This is believed to be due to AERMOD's characterization and utilization of the vertical structure of the boundary layer in combination with its implementation of the dividing streamline concept for flow in complex terrain.

Similarly, for tall, buoyant stacks in flat terrain (Kincaid, Baldwin), AERMOD performs well in reproducing the upper end of the concentration distribution. This success is most likely related to the improved algorithms for convective conditions. The bi-Gaussian vertical concentration distribution in AERMOD is based upon years of model development following the observations and calculations of laboratory, field, and numerical studies over the past 30 yr. These formulations result in a much more appropriate treatment of elevated plume material in convective conditions. In contrast, the model is still somewhat challenged in reproducing some of the lower concentration values, particularly in stable conditions, as suggested by the annual average comparisons.

The model also performed well in the only urban database in the study. The Indianapolis data were utilized to some extent in the model development (specifically, in the formulation of the urban mixing height). The authors believe that it is not unreasonable to expect AERMOD's formulation to translate well to other urban areas because the urban formulation (Part I) is based on meteorological observations in a variety of urban areas. Obviously, evaluation in other urban areas is desirable.

The comparison of AERMOD with the measurements of seven building wake studies provided a very interesting variety of sometimes contrasting results.

This served to highlight the sensitivity of dispersion to local meteorological behavior and the geometry of the building wakes and cavities. Overall, the model found the representative high-end concentrations (i.e., RHC) within a factor of 2 or better. Although it seems rather obvious, the results here strongly suggest that specification of the cavity extent and plume material height and spread (near the building) is critical to appropriately simulating the downwash effect.

AERMOD (Part I) represents many formulation improvements over commonly applied regulatory models such as ISCST3. In model-to-model comparisons, AERMOD's performance is clearly superior to that of ISCST3. Models such as HPDM and CTDMPLUS perform similarly to AERMOD in the selected circumstances for which these models were designed. This is not surprising because many of the formulations of AERMOD are based, to some extent, on earlier work by others in developing these and other models.

Acknowledgments. We thank the many scientists who provided evaluation databases and conducted peer reviews and beta testing throughout the period of the AERMOD development. This paper has been reviewed in accordance with the U.S. Environmental Protection Agency's peer review and administrative review policies for approval for presentation and publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

REFERENCES

Baerentsen, J. H., and R. Berkowicz, 1984: Monte Carlo simulation of plume dispersion in the convective boundary layer. *Atmos. Environ.,* **18,** 701–712.

Barad, M. L., 1958: Project Prairie Grass, A Field Program in Diffusion. Vols. I and II, Geophysical Research Papers No. 59, Air Force Cambridge Research Center Rep. AFCRC-TR-58-235, 439 pp.

Bowers, J., and A. J. Anderson, 1981: An evaluation study for the Industrial Source Complex (ISC) dispersion model. Environmental Protection Agency Office of Air Quality Planning and Standards Rep. EPA-450/4-81-002, 32 pp.

Bowne, N. E., R. J. Londergan, D. R. Murray, and H. S. Borenstein, 1983: Overview, results, and conclusions for the EPRI Plume Model Validation and Development Project: Plains site. Electric Power Research Institute Rep. EA-3074, Project 1616-1, 234 pp.

Chambers, J. M., W. S. Cleveland, B. Kleiner, and J. A. Tukey, 1983: Comparing data distributions. *Graphical Methods for Data Analysis (Bell Laboratories),* Wadsworth International Group and Duxbury Press, 47–73.

Cimorelli, A. J., S. G. Perry, A. Venkatram, J. C. Weil, R. J. Paine, R. B. Wilson, R. F. Lee, and W. D. Peters, 2003: AERMOD description of model formulation. U.S. Environmental Protection Agency Rep. EPA 454/R-03-002d, 85 pp.

——, and Coauthors, 2005: AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization. *J. Appl. Meteor.,* **44,** 682–693.

Cox, W., and J. Tikvart, 1990: A statistical procedure for determining the best performing air quality simulation model. *Atmos. Environ.,* **24A,** 2387–2395.

DiCristofaro, D. C., D. G. Strimaitis, B. R. Green, R. J. Yamartino, A. Venkatram, D. A. Gooden, T. F. Lavery, and B. A. Egan, 1985: EPA complex terrain model development: Fifth milestone report—1985. U.S. Environmental Protection Agency Rep. EPA-600/3-85-069, 277 pp.

Engineering Science, 1980: Field validation of atmospheric dispersion models for natural gas compression stations. Prepared for the American Gas Association Rep. PR-133, 152 pp.

Gifford, F. A., 1961: Uses of routine meteorological observations for estimating atmospheric dispersion. *Nucl. Saf.,* **2,** 47–51.

Guenther, A., and R. G. Lamb, 1990: Atmospheric dispersion in the Arctic: Winter-time boundary layer measurements. *Bound.-Layer Meteor.,* **49,** 339–366.

——, ——, and E. Allwine, 1989: Building wake dispersion at an Arctic industrial site: Field tracer observations and plume model evaluations. *Atmos. Environ.,* **24A,** 2329–2347.

Hanna, S. R., and R. J. Paine, 1989: Hybrid Plume Dispersion Model (HPDM) development and evaluation. *J. Appl. Meteor.,* **28,** 206–224.

——, and J. S. Chang, 1993: Hybrid Plume Dispersion Model (HPDM), improvements and testing at three field sites. *Atmos. Environ.,* **27A,** 1491–1508.

——, B. A. Egan, J. Purdum, and J. Wagler, 1999: Evaluation of ISC3, AERMOD, and ADMS dispersion models with observations from five field sites. Hanna Consultants Rep. P020DR, 10 pp.

——, ——, ——, and ——, 2000: Comparison of AERMOD, ISC3, and ADMS model performance with five field data sets. Preprints, *AWMA Annual Meeting,* Salt Lake City, UT, Air and Waste Management Association, 1–14.

Haugen, D. A., 1959: Project Prairie Grass, A field program in diffusion. Vol. III, Geophysical Research Paper 59, Air Force Cambridge Research Center Rep. AFCRC-TR-58-235, 439 pp.

Irwin, J. S., D. Carruthers, J. Stocker, and J. Paumier, 2003: Application of ASTM D6589 to evaluate dispersion model performance. *Int. J. Environ. Pollut.,* **20,** 4–10.

Liu, M. K., and G. E. Moore, 1984: Diagnostic validation of plume models at a plains site. Electric Power Research Institute Rep. EA-3077, Research Project 1616-9, 400 pp.

Melbourne, W. H., and J. Taylor, 1994: Wind tunnel studies of plume dispersion from the Lee Power Plant Power Station. Electric Power Research Institute Rep. TR-135274, 277 pp.

Murray, D. R., and N. E. Bowne, 1988: Urban power plant plume studies. Electric Power Research Institute Rep. EA-5468, Research Project 2736-1, 236 pp.

Oke, T. R., 1978: *Boundary Layer Climates.* John Wiley and Sons, 372 pp.

——, 1982: The energetic basis of the urban heat island. *Quart. J. Roy. Meteor. Soc.,* **108,** 1–24.

Paine, R. J., and B. A. Egan, 1987: User's guide to the Rough Terrain Diffusion Model (RTDM)—Rev. 3.20. ENSR ERT Doc. PD-535-585, 260 pp.

——, and Coauthors, cited 1998: Model evaluation results for AERMOD. Environmental Protection Agency. [Available online at http://www.epa.gov/scram001/.]

——, and Coauthors, 2003: AERMOD: Latest features and evaluation results. Preprints, *96th Annual Meeting Air and Waste Management Association,* San Diego, CA, Air and Waste Management Association, 1–35.

Pasquill, F., 1961: The estimation of the dispersion of windborne material. *Meteor. Mag.,* **90,** 33–49.

Paumier, J. O., S. G. Perry, and D. J. Burns, 1992: CTDMPLUS: A dispersion model for sources near complex topography. Part II: Performance characteristics. *J. Appl. Meteor.,* **31,** 646–660.

Perry, S. G., 1992: CTDMPLUS: A dispersion model for sources in complex topography. Part I: Technical formulations. *J. Appl. Meteor.,* **31,** 633–645.

Schulman, L. L., and S. R. Hanna, 1986: Evaluation of downwash modifications to the Industrial Source Complex Model. *J. Air Pollut. Control Assoc.,* **36,** 258–264.

——, D. G. Strimaitis, and J. S. Scire, 2000: Development and evaluation of the PRIME plume rise and building downwash model. *J. Air Waste Manage. Assoc.,* **50,** 378–390.

Snyder, W. H., R. S. Thompson, R. E. Eskridge, R. E. Lawson, I. P. Castro, J. T. Lee, J. C. R. Hunt, and Y. Ogawa, 1985: The structure of the strongly stratified flow over hills: Dividing streamline concept. *J. Fluid Mech.,* **152,** 249–288.

Start, G. E., N. F. Hukari, J. F. Sagendorf, and J. H. Cate, 1981: EOCR building wake effects on atmospheric diffusion. National Oceanic and Atmospheric Administration Rep. NUREG/CR-1395, 220 pp.

Strimaitis, D. G., R. J. Paine, B. A. Egan, and R. J. Yamartino, 1987: EPA complex terrain model development: Final report. U.S. Environmental Protection Agency Rep. EPA/600/3-88/006, 486 pp.

Thullier, R. H., and R. L. Mancuso, 1980: Building effects on effluent dispersion from roof vents at nuclear power plants. Electric Power Research Institute Rep. NP-1380, Research Project 1073-1, 240 pp.

U.S. Environmental Protection Agency, 1995: User instructions. Vol. I, User's Guide for the Industrial Source Complex (ISC3) Dispersion Models (revised). U.S. Environmental Protection Agency, EPA-454/b-95-003a.

Venkatram, A., and Coauthors, 2001: A complex terrain dispersion model for regulatory applications. *Atmos. Environ.,* **35,** 4211–4221.

Weil, J. C., 1992: Updating the ISC model through AERMIC. Preprints, *85th Annual Meeting of Air and Waste Management Association,* Kansas City, MO, Air and Waste Management Association, 1–14.

——, L. A. Corio, and R. P. Brower, 1997: A PDF dispersion model for buoyant plumes in the convective boundary layer. *J. Appl. Meteor.,* **36,** 982–1003.

Willis, G. E., and J. W. Deardorff, 1981: A laboratory study of dispersion in the middle of the convectively mixed layer. *Atmos. Environ.,* **15,** 109–117.