Journal of Hazardous Materials, 26 (1991) 127–158
Elsevier Science Publishers B.V., Amsterdam

Evaluation of fourteen hazardous gas models with ammonia and hydrogen fluoride field data

Steven R. Hanna, David G. Strimaitis and Joseph C. Chang

Sigma Research Corporation, 234 Littleton Rd., Suite 2E, Westford, MA 01886, (USA)

(Received March 20, 1990; accepted in revised form July 2, 1990)

Abstract

Fourteen hazardous gas models are evaluated using data from the Desert Tortoise ammonia (NH_3) and Goldfish hydrogen fluoride (HF) field experiments , which involved horizontal releases of aerosol jets. Seven experiments are available for analysis, with data at three downwind monitoring arcs, at distances ranging from 100 m to about 3000 m. The models include eight publicly available models, (ADAM, AFTOX, ALOHA, Britter and McQuaid, DEGADIS, HEGADAS, OB/DG, and SLAB) and six proprietary models (CHARM, EAHAP, PHAST, SAFETI, TRACE, and WHAZAN). In addition, the methods of initializing the ALOHA, CHARM, DEGADIS, and HEGADAS models were modified to account for initial dilution in an aerosol jet and these revised predictions were included in the evaluation. About one-half of the models yield relatively good performance in their predictions of maximum concentrations on monitoring arcs, with relative mean biases of $\pm 30\%$ or less and root mean square error (rmse) values that are about 40% to 60% of the mean. It is interesting that the simple Britter and McQuaid model performs just as well as some of the more sophisticated models, indicating that the simple model has captured the essence of the plume thermodynamics. Because this data set is not large, no significant differences can be shown (at the 95% confidence level) among the better models. This analysis will be expanded in the future to include other field datasets (e.g., Thorney Island and Maplin Sands).

1. Objectives of study

In the event that any hazardous chemical is accidentally released into the atmosphere, there is a need to estimate the concentration downwind of the release. The overall objective of this study is to evaluate several available hazardous gas models with field data. This paper describes some preliminary results based on two field data sets.

The U.S. EPA has recommended a set of dispersion models (the UNAMAP models) for application to continuous releases of neutrally or positively buoyant gases, but has not yet made any recommendations of models for application to transient releases of hazardous chemicals, which are often negatively buoyant. Only 15 years ago, no transport and dispersion models existed that were

0304-3894/91/$03.50

relevant to this problem and no comprehensive field data sets were available. Spurred by increased public concern, several government agencies and industries have sponsored model development and field programs over the past decade, so that there are now several dozen models and over 100 hourly periods of field data from numerous sites available [1]. Of the total number of models, about seven or eight of the publicly available models and an equal number of the proprietary models are in common use. But no one has a good idea of whether any model is better than any other model, or what the level of uncertainty is in a typical application.

Some preliminary evaluations of hazardous gas models have been conducted, including one sponsored by the EPA [2], one sponsored by the U.S. Air Force [3], and others carried out by model developers [4,5]. These studies demonstrate that many of the more recent state-of-the-art models are able to reproduce the plume centerline concentrations in the field tests at downwind distances of a few hundred meters within a factor of about two, although the model predictions tend to diverge at distances less than about 100 m from the source and greater than about 1.0 km from the source.

The primary objective of the preliminary study reported in this paper is to evaluate several of these models using the Desert Tortoise ammonia (NH_3) [6] and the Goldfish hydrogen fluoride (HF) [7] field data sets. Uncertainties in the models have been assessed from the results of this evaluation and from sensitivity tests with the models. Further evaluations of the model uncertainties will be carried out in the future, expanding the model evaluation data base and including analysis of data input uncertainties and stochastic fluctuations.

As shall be discussed in Section 2, hazardous gas models that are publicly available, that are in wide use, and that are of interest to the American Petroleum Institute (API) and the U.S. Air Force were acquired and tested. All of these codes either originally ran on an IBM PC or have been converted to run on a PC. Because some of the models (e.g., AFTOX) are not applicable to jet releases and most do not handle aerosols, it has been necessary to devise methods for initializing them so that they can simulate the Desert Tortoise and Goldfish field tests, both of which involve aerosol jets.

The Desert Tortoise and Goldfish field data were sent to us on magnetic tape and floppy disk. A modelers' data base was set up in the form of an electronic spreadsheet, including all input data required to run the models and all observations needed to test the models (see Section 3). The model evaluation software described by Hanna [8] was applied to the observations from the two field tests and the predictions from the following hazardous gas models: publicly available models: ADAM, AFTOX 3.1., ALOHA, Britter and McQuaid, DEGADIS 2.1., HEGADAS, OB/DG, SLAB; proprietary models: CHARM, EAHAP, PHAST, SAFETI, TRACE, WHAZAN.

We ran the publicly available models and the proprietary CHARM model using the modelers' data base produced and presented in Section 3. Engineers of

API ran the proprietary EAHAP, PHAST, SAFETI, TRACE, and WHAZAN models using the same data base, and have provided the results to us. Section 4 gives the details of the model performance measures that were applied and provides tabular and graphical results of the model evaluation exercise.

2. Summary of models to be evaluated and their input requirements

The hazardous gas models evaluated in this paper are representative of those likely to be applied in the early 1990 time-frame by industries and by government agencies to estimate the impact of accidental releases of hazardous gases and aerosols to the atmosphere. Typical examples of these releases would include the rupture of a one-ton chlorine cylinder, the failure of a sour gas pipeline, or the spill of ammonia from a storage tank [9]. The following sections briefly describe the publicly available and proprietary models that are evaluated. More details on the models and the methods of initializing them are given in the report [10] and in the model user's guides.

2.1 Publicly available models

For the purposes of this study, a hazardous gas model is considered to be "publicly available" if the model description (user's guide) and model source code are available (at cost or less) to anyone who requests them. From our point of view, it is preferable that the code be available on IBM-PC compatible floppy disk, and that a test case be included. Hanna and Drivas [1] list dozens of publicly available models, but some of them have been abandoned, some have been superseded, and some are very similar to other models. The following eight hazardous gas models have been chosen for this evaluation.

2.1.1 ADAM model

The U.S. Air Force recently sponsored the development of the ADAM model [11], which is intended to be a supplement for the OB/DG and the AFTOX models. The new model can treat dense gases and the thermodynamics of aerosols. It has an advantage over many of the other publicly available models in that it contains methods to calculate the emission rates for evaporating spills or momentum jets. For example, given the pressure, temperature, and mass content of a pressurized tank, as well as the geometry of the opening and the ambient air temperature and pressure, the model will calculate the characteristics of the emission (e.g., mass flux, temperature, jet velocity, and fraction of aerosol). Because the model can be applied to only six chemicals (H_2S, SO_2, Cl_2, NH_3, N_2O_4, and hydrazine or N_2H_4), it can be tested with the Desert Tortoise (NH_3) field data but not with the Goldfish (HF) field data. A version of the ADAM model that can be applied to HF emissions is currently being developed.

2.1.2 AFTOX 3.1 model

The AFTOX 3.1 model was developed by the U.S. Air Force [12,13] as a replacement for the OB/DG model. AFTOX 3.1 is very similar to the Shell SPILLS model [14], which contains an evaporative emissions source algorithm for liquid spills and a Gaussian plume dispersion model for downwind transport and dispersion. The AFTOX 3.1 model approximates the plume as a sequence of Gaussian puffs. The model does not treat jet sources or dense gas effects.

2.1.3 ALOHA Gaussian plume model

The ALOHA transport and dispersion model [15] is part of a comprehensive hazard assessment software package called CAMEO, that was developed and distributed by the National Oceanic and Atmospheric Administration (NOAA). The current version of the model does not account for dense gas effects, and reduces to the standard U.S. EPA Gaussian plume formulation [16] for most applications.

2.1.4 Britter and McQuaid model

The Britter and McQuaid (B&M) model is not a computer code, but is a set of simple equations and nomograms suggested in the "Workbook on the Dispersion of Dense Gases" by Britter and McQuaid [17]. The authors fit some simple curves to available field and laboratory data, making sure that the curves agree asymptotically with known relations for passive diffusion. Because the Desert Tortoise and Goldfish data were not used in the derivation of the model, they represent an independent test of the model.

The B&M model requires that source input conditions be specified (initial excess density and volume flux), and can treat either instantaneous or continuous sources. It assumes that surface roughness is unimportant and that atmospheric stability is neutral (i.e., Pasquill class D).

2.1.5 HEGADAS model

The Shell HEGADAS model was originally a proprietary model [18] but a special version has been recently released for public distribution and is available from the U.S. National Technical Information Service (NTIS). The HEGADAS model was used as a basis for the development of the DEGADIS model. A user's guide for HEGADAS is available [19]. The code treats only area or volume sources, does not handle jets, and requires input of all source emission parameters. It can handle dense gases. This code has been applied to the Desert Tortoise NH_3 field test.

An extension of the HEGADAS model to account for the thermodynamics of HF and the near-field jet (HFSYSTEM) is being developed but has not yet been released. Concentration predictions for the Goldfish HF field test have been supplied to us by the model developers for this preliminary exercise.

2.1.6 DEGADIS 2.1 model

The DEGADIS 2.1 model was originally developed for application to evaporative emissions from liquid spills, but has recently been modified to account for vertical jet emissions [20]. The U.S. Coast Guard sponsored its initial development [4], and it is now in wide use. The DEGADIS 2.1 model is the most complicated of the publicly available models analyzed in this project, since it includes algorithms accounting for variable emission rates, dense gas thermodynamics at the source, and cross-wind concentration distributions. Source emission rates are not calculated but must be provided as input to the model. the DEGADIS code in common use requires a VAX computer, although a test version of the code that will run on an IBM-PC (or compatibles) was received.

2.1.7 OB/DG model

The Ocean Breeze/Dry Gulch (OB/DG) model [21] was developed for use in support of the rocket fuel handling operations at Cape Canaveral and Vandenberg. Dispersion data were collected at those two sites (Cape Canaveral, Florida is the site of the Ocean Breeze; Vandenberg AFB, California is the site of the Dry Gulch experiment) and at the Prairie Grass, Kansas, site during the 1950s and 1960s. These data were used to develop a purely empirical correlation known as the OB/DG model, which gives the concentration as a function of Q, x, σ_θ, and ΔT. Note that the standard deviation of wind direction fluctuations σ_θ is in angular degrees, and ΔT is defined as the temperature difference (°F) between the 54 ft and 6 ft levels on a tower. The model is limited to source scenarios where there is a continuous release from a ground-level source of inert gas with no initial buoyancy.

2.1.8 SLAB model

The SLAB model was originally developed by Ermak and Chan [22] at Lawrence Livermore National Laboratory (LLNL) for application to dense gases that are emitted from liquid spills. The code considers the concentration integrated over a cross-section perpendicular to the plume centerline, and calculates the downwind variation of this integrated concentration. The version that is used in the current project is dated November, 1989, and includes algorithms for a horizontal jet. The source conditions must be input to the model.

Because this model was originally a research tool that was part of several LLNL internal projects, until recently a user's guide was not written and little attention was paid to making the code "user friendly." However, the SLAB model code has been distributed throughout the world and is in frequent use because of its relative simplicity and because of the fact that it agrees fairly well with observations at dense gas field tests. During the past year, LLNL devoted considerable effort towards improving the model and writing a user's guide [23].

2.2 Proprietary models

For the purposes of this study, proprietary models are defined as those that are either not available to the public or must be purchased from vendors at a cost greatly exceeding the expenses associated with copying and distributing the user's guide and code. In addition, the source code for these models is generally not available for inspection. Because these models simulate a wide range of release characteristics and they are in use at many industrial sites, they have been included in this model evaluation exercise. We ran the CHARM model, and under an agreement with one of the sponsors of this research, engineers in oil companies ran the TRACE, EAHAP, PHAST, SAFETI, and WHAZAN models, given the model input data bases listed in Section 3.

2.2.1 CHARM model

The CHARM model is a Gaussian puff model [24] that treats all releases as a series of puffs, regardless of whether the release is continuous or instantaneous. The puff characteristics can be calculated internally by means of several release-rate modules, or they can be specified by the user. The evaporative emission rate from liquid spills is obtained from the Shell SPILLS model [14]. The growth of heavier than air puffs is modeled using the approach of Eidsvik [25].

2.2.2 EAHAP model

The EAHAP model has its roots in earlier models for hazards and risk assessments of fire and radiation effects [26]. It includes algorithms for two-phase jet releases, and it incorporates the HEGADAS and DEGADIS algorithms for dense-gas dispersion. Unlike DEGADIS, however, the internal emissions algorithms provide a complete specification of the characteristics of the source.

2.2.3 TRACE model

The TRACE model uses release and dispersion algorithms similar to those in Dupont's SAFER system [27]. It contains a complete set of algorithms for modeling two-phase dense-gas releases, including elevated releases and jets subject to flashing and aerosol formation. These releases may be instantaneous, steady-state, or transient, and may result in either neutral, or buoyant clouds (heavier or lighter than air).

2.2.4 PHAST/SAFETI/WHAZAN models

These three models were developed by Technica, Inc. [28–30]. The least complex of the three is WHAZAN. The SAFETI package is a more comprehensive quantitative risk assessment system that runs on a VAX computer, while PHAST is a PC-based product that contains many of the dispersion modeling algorithms of SAFETI. Each of these models contains dense-gas dispersion algorithms and a two-phase, momentum jet algorithm, so the user need not devise

methods for simulating this type of release. The models accept both instantaneous and "prolonged" release modes.

2.3 Overview of model input requirements and capabilities

The models evaluated here have considerable variation in their capabilities and input requirements. Some models simulate all aspects of a complex release typified by the Desert Tortoise and Goldfish experiments, including aerosols, entrainment processes associated with momentum jets, variable averaging times, detailed meteorological data, and site roughness. Others contain no modules that explicitly simulate aerosols, or dense-gas effects for that matter. Many of these attributes are identified in the descriptions of the individual models in the references, but are summarized here in order to highlight differences among the models which influence how each is applied to the various source scenarios. The attributes are given in Table 1, and some explanations of the terminology and the assumptions in the models are given below. The table contains only a representative subset of all the attributes that could have been listed; e.g., questions of ground heat transfer and chemical reactions are not addressed here, and the reader should consult the references for more details.

2.3.1 Surface roughness input requirements

The surface roughness at the site is generally characterized by the roughness length scale, z_0, expressed in meters. Models that allow the user to input a specific value of the roughness length are identified with a "✓" in Table 1. Note that 11 of the 14 models accept a roughness length in this way. Of these, the EAHAP and the ALOHA models restrict the selection of a roughness length to particular choices which correspond to general descriptive classes. The EAHAP model permits several classes. The ALOHA model recognizes two general types of surface roughnesses, one for urban sites and one for open, rural sites. Although the roughness of the surface is not the only factor in choosing between the two sets, it is the only way in which surface roughness can enter the dispersion calculation in the ALOHA model. The CHARM model, the OB/DG model, and the Britter and McQuaid [17] Workbook model (B&M) do not allow the user to account for variations in surface roughness.

Three of the models (PHAST, SAFETI, and WHAZAN) make use of a surface roughness parameter (SRP), defined as

$$SRP = 0.4/\ln(10/z_0) \tag{1}$$

This parameter arises from boundary layer wind profile formulas in a well-mixed (i.e., neutral stability) boundary layer, where the wind speed, u, at a height of 10 m is given by the formula:

$$u = u_*/SRP = (u_*/0.4)\ \ln\ (10/z_0) \tag{2}$$

TABLE 1

Attributes of models[a]

Attribute	Model													
	ADAM	AFTOX	ALOHA	B&M	CHARM	DEGADIS	EAHAP	HEGADAS	OB/DG	PHAST	SAFETI	SLAB	TRACE	WHAZAN
Surface roughness (z_0)	✓	✓ ($z \geq 0.005$ m)	Rural or urban	–	–	✓	✓	✓	–	✓	✓	✓	✓	✓ ($z_0 \geq 0.01$ m)
Avg. Time (min)	✓	✓	10(rural) 60(urban)	–	10	✓	10	✓	–	10	10	✓	✓	10
Height of wind (m)	✓	–	–	10	–	✓	10	✓	–	–	–	✓	✓	–
Receptor height (m)	0	✓	0	0	✓	0	1	0	0	0	0	✓	✓	0
Release:														
Continuous	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Instantaneous	✓	✓	–	✓	✓	✓	✓	–	–	✓	✓	✓	✓	✓
Variable rate	–	–	–	–	✓	✓	✓	✓	–	–	–	–	✓	–
Aerosols	✓	–	–	–*	✓	–*	✓	–*	–	✓	✓	✓	✓	✓
Flash fraction	✓*	–	–	–*	✓*	–*	✓	–*	–	✓	✓	–*	✓	✓
Jet mixing	✓	–	–	–*	–	–	✓	–	–	✓	✓	✓	✓	✓
Relative humidity	✓	–	–	–	✓	✓	✓	✓	–	✓	✓	✓	✓	✓
Ambient pressure	✓	–	–	–	✓	✓	–	–	–	✓	–	–	–	–

[a]Explanations of terms are given in Section 2.3. A check (✓) indicates that the model accounts for variation in that attribute. A dash indicates that the model does not account for variations in that attribute. A number indicates the values assumed by the model. An asterisk indicates that we accounted for this attribute in our model initialization assumptions.

where u_* is referred to as the friction velocity. As can be seen, z_0 and *SRP* have a one-to-one relation, although the use of *SRP* implies that wind speeds will be observed at 10 m above the surface.

2.3.2 Averaging time input requirements

Six of the models in Table 1 allow the averaging time for concentration estimates to be specified directly (indicated by check-marks). The HEGADAS model provides guidance on how to alter the rate of growth in the function for lateral dispersion, σ_y, to simulate variations in averaging time and therefore requires involvement of the user. The rest of the models make no provision for adjustments to the concentrations to account for various averaging times. These models appear to make use of dispersion relations that are commonly considered representative of averaging times on the order of 10 minutes, and so are identified by a "10" in the table.

2.3.3 Wind speed measurement height input requirements

Most of the models do not request information on the height at which winds are measured, or else specifically identify the height expected by the model (usually 2 m or 10 m). Those models that do request the height of the measurement are identified by check-marks in Table 1. The B&M and EAHAP models are assigned a "10" in the table, indicating that if the wind observations are made at a height of 2 m, they should be scaled to a height of 10 m using a power-law relation or a boundary layer equation (e.g., eqn. 2) before running this model. This adjustment obviously adds a degree of uncertainty to the predictions.

2.3.4 Receptor height input requirements

Only five of the models in Table 1 provide an option for specifying the height at which concentrations are calculated (identified by check-marks). All the rest except the EAHAP model provide concentration estimates at the surface. (We note that the DEGADIS model will provide lateral distance information to two specified concentrations at an elevation specified by the user, but "centerline" concentrations are provided only at the surface.) The EAHAP model provides concentrations at an elevation of 1 m above the surface. These differences can produce an unfortunate discrepancy in concentration predictions at various receptor heights among the models, since concentrations in shallow dense gas plumes near the source can exhibit large gradients between the ground surface and a height of 1 m. In the case of the Desert Tortoise and Goldfish field experiments, the closest monitoring arcs were at least 100 m from the source, and therefore the modeled vertical gradients would not be as large as they would be closer to the source.

2.3.5 Release properties input requirements

The first three release properties listed in Table 1 indicate the flexibility of the models in handling scenarios in which the material is released over a finite period of time (i.e., transient releases). All of the models except the OB/DG, ALOHA, and HEGADAS models provide a choice between a steady-state limit (continuous release) and a single, instantaneous limit. Most of these models are also able to accept a steady release that has a finite duration. Models for this type of release typically make use of continuous release algorithms at short times of travel (compared to the duration of the release), and either instantaneous release algorithms or more complex algorithms for long times of travel. The transition between the two types of algorithms may be accomplished internally, or guidance may be given as to when the instantaneous mode should be used. Another, more complicated, group of releases is characterized by a finite but unsteady rate of release (i.e., variable rate). Approximately half of the models in the table cannot treat this type of release. They can simulate a variable release by an approximation in which the release rate is adjusted such that the total mass of the release and the period of the release are conserved, but it is assumed that the rate of release is steady with time.

The remaining three items in the "release properties" section of Table 1 characterize the degree to which the models are able to handle a two-phase release typified by the Desert Tortoise and Goldfish tests. Many of the models are able to account for the initial two-phase mixture (aerosols) explicitly by either having the user specify the presence of aerosols, or by having the model simulate the release and apportion the liquid phase between a liquid pool on the ground and aerosols in the cloud. Other models may simulate the effect of aerosols on the density of the mixture when properly initialized by the user, and those models are denoted by an asterisk in Table 1.

The DEGADIS model contains a procedure that simulates the effects of the aerosols on density as a function of the mole fraction of the material in the diffusing cloud. This procedure mimics (to a greater or lesser extent depending on the details provided by the user) the evaporation of aerosols as the cloud mixes with the ambient atmosphere. As a result, its results are more similar to that class of models that are identified by check-marks in the table.

The amount of liquid that immediately flashes to vapor is computed by some of the models. We have chosen to fix this value in all of the models that we ran in order to facilitate intercomparisons. We made no such demand on those engineers who supplied us with results from the other models. Most of the models were run with similar values. A notable exception is the PHAST model, in which the fraction of HF vapor flashed during the Goldfish tests was calculated to be an order of magnitude less than the fraction that we had calculated.

Another prominent feature of the Desert Tortoise and Goldfish tests is the initial mixing of the cloud due to the jet-like nature of the release which coincides with the rapid expansion of the cloud as a portion of the liquid flashes to

vapor. Many of the models do not account for this momentum jet at all. Those that do are indicated by the check-marks in Table 1. Without this or an alternate method for entraining air in the process of evaporating the aerosol, a model is likely to overestimate concentrations at receptors closest to the point of release.

Some of the models have the capability to treat heat exchanges with the underlying surface. This effect was thought to be minor compared to the thermodynamic effects of the aerosol evaporation and the entrained air. To adequately parameterize the ground heat flux, the thermal conductivity of the ground would have to be included as an input parameter.

2.3.6 Relative humidity/ambient pressure input requirements

The last two items listed in Table 1 are relative humidity and ambient pressure. Both have a bearing on the relative density of the cloud. Those models that include water vapor (change of phase with temperature, chemical interactions with the gas cloud) will certainly require the relative humidity, and may exhibit a greater sensitivity to these data than those that do not include water vapor.

2.4 Approximate methods of initializing turbulent, two-phase jets

Although many of the models listed in Table 1 contain explicit treatments of two-phase jets, several do not treat either horizontal jets, or two-phase releases, or both. Videotapes of the Desert Tortoise and Goldfish field tests indicate that the initial dispersion is dominated by the influence of the aerosol jet out to distances of 100 m or more. The aerosol jet can be approximately accounted for in a simple manner by assuming that air is entrained at the source in sufficient quantities to provide the heat required to evaporate all of the aerosol, while maintaining a cloud temperature equal to the boiling-point temperature. This procedure is recommended by Britter and McQuaid [17] for application of their model to ammonia releases. Although the actual sequence of events surrounding the development of the vapor/aerosol cloud is more complicated (e.g., the complete energy balance formula would include the effects of water vapor condensation, heat gained from the ground, and several other terms), this method of initialization roughly maintains the overall energy balance while accounting for the entrainment of a substantial volume of air. The end result of this initialization procedure is a single-phase mixture of vapor and air at a uniform temperature equal to the boiling-point temperature at ambient pressure.

To determine if the inclusion of initial dilution might lead to better model performance for those models do not contain aerosol jet formulations, we have modified the initialization procedures for the ALOHA, DEGADIS, HEGADAS, and CHARM models in this manner, and the results are indicated in later tables and figures by adding the suffix "id" to the names of these models. The amount of

air that is needed to supply the heat to completely evaporate the aerosol is calculated as follows. Assume that all of the heat needed to evaporate the aerosol comes from entrained air. The heat energy given up (per unit mass) by the entrained air, as its temperature drops by ΔT, equals $c_{pa}\Delta T$, where c_{pa} is the specific heat of air at constant pressure. It is assumed that this heat energy is used to evaporate the liquid, which has a latent heat of vaporization of L_l. Then the ratio of mass of air entrained to mass of liquid evaporated equals $L_l/c_{pa}\Delta T$.

Finally, the mass-relationship is converted to a molar relationship, and it is recognized that the air is entrained to evaporate that fraction of the contaminant that did not flash. Let f equal the fraction flashed, M_a equal the molecular weight of air, and M_c equal the molecular weight of the contaminant; then the ratio of the volume of entrained air to the volume of unflashed liquid equals $(L_l/c_{pa}\Delta T)(M_c/M_a)(1-f)$. The volume of entrained air is then determined and added to the volume of the gaseous (flashed) contaminant in order to estimate the initial dilution of the plume.

3. Description of field experiments and tabulation of database

This section contains details of the Desert Tortoise (NH_3) and Goldfish (HF) experiments, including maps of instrument locations and tabulation of model input data.

3.1 Desert Tortoise (NH_3) experiments

The Lawrence Livermore National Laboratory (LLNL) conducted four large-scale (15–60 m^3) ammonia spill tests at its Frenchman Flat, Nevada, field site during the summer and fall of 1983 [6,31]. The characteristics of these tests are summarized in Table 2. It is seen that the meteorological conditions were quite similar during all four tests (moderate wind speeds and nearly neutral stabilities). The pressurized liquid NH_3 was released from a spill pipe pointing downwind at a height of 0.79 m above the ground. The liquid jet flashed as it exited the pipe and its pressure decreased, resulting in about 18% of the

TABLE 2

Test summary for Desert Tortoise NH_3 spills [6,31]

Test	Date (1983)	Spill rate (m^3/min)	Spill duration (s)	Mean wind speed (m/s)	Atmospheric stability class (Pasquill)
1	24 Aug.	7.0	128	7.4	D
2	29 Aug.	10.3	255	5.7	D
3	1 Sept.	11.7	166	7.4	D
4	6 Sept.	9.5	380	4.5	E

liquid changing phase to become a gas. The remaining 82% of the NH_3-jet remained as a liquid, which was broken up into an aerosol by the turbulence inside the jet. Very little of this liquid was observed to deposit on the ground.

3.1.1 Instrumentation

Measurements of the ammonia concentration in the atmosphere were obtained by means of three types of detectors: Lawrence Livermore National Laboratory (LLNL) IR gas sensors, Mine Safety Appliances nondispersive infrared (NDIR) sensors, and International Sensor Technology (IST) solid-state gas sensors. The accuracy of the concentration measurements is thought to be $\pm 20\%$. Figure 1 shows the relative positions of the major sets of instruments at the site. The primary sampling rows are located 100 m and 800 m downwind of the point of release. Concentration data were obtained at heights of 1.0, 2.5, and 6 m on the 100 m arc and at heights of 1.0, 3.5, and 8.5 m on the

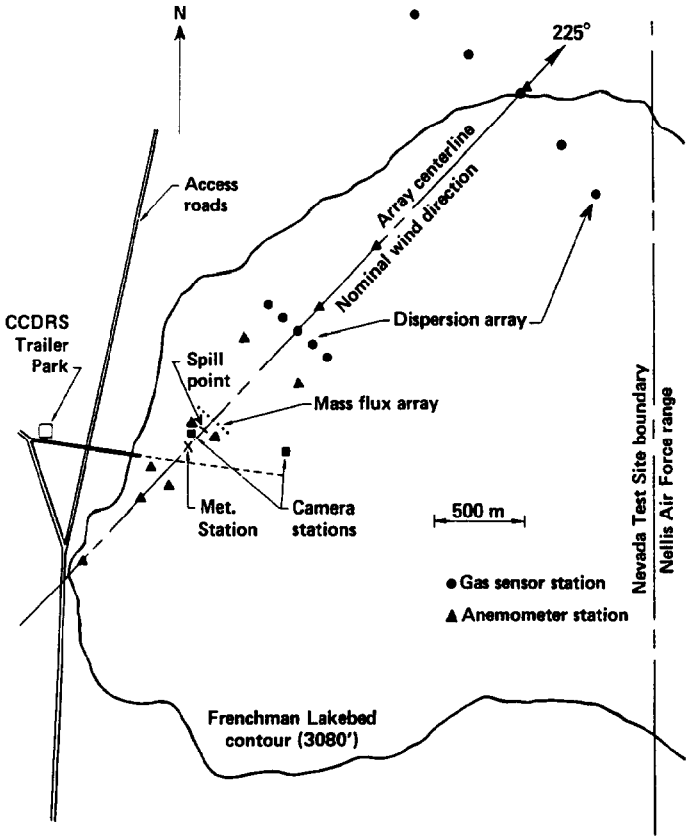

Fig. 1. Diagnostic instrument array for Desert Tortoise and Eagle series experiments [6].

800 m arc. In addition, eight portable IST ammonia vapor samplers were placed at ground-level between 1.4 km and 5.5 km downwind of the release. No information on vertical distribution of NH_3 concentration was available from these more distant arcs, and, because of the wide spacing of the instruments, there is little assurance that the concentrations obtained are representative of centerline values.

Numerous wind and temperature measurements were made at the site. The locations of some of these instruments are shown in Fig. 1. Eleven cup-and-vane anemometers were located at a height of 2 m at various positions within the test array in order to define the wind field for planning the releases and subsequent calculation of plume trajectories. A 20-m tall meteorological tower was located 50 m upwind of the point of release (see point M in Fig. 1), with temperature measured at four levels and wind speed and turbulence measured by bivane anemometers at three levels. A 10-m tall tower was located 100 m downwind of the release, with bivane anemometers at three heights, and thermocouples at two. Additional temperature measurements were made at three heights on each of the sampler masts located 100 m and 800 m downwind of the release. Soil temperature was measured 3, 5, and 9 m downwind of the point of release, and heat fluxes in the ground were measured by heat flux gauges located just below the surface 3 m downwind of the release, at both meteorological towers, and at the two sampler masts on either side of the meteorological tower 100 m downwind of the release.

Wind and temperature data obtained from the tower located 50 m upwind of the spill site were analyzed [16,31] to calculate the friction velocity (u_*) and the Monin–Obukhov length (L). A surface roughness length of 0.003 m was assumed. However, the boundary layer during these tests may not be adequately characterized by these derived values, since water covered portions of the test area during Tests 1, 2, and 3. Several other measures of turbulence were observed on the central meteorological tower, such as the standard deviations of wind speed and wind direction fluctuations, σ_u and σ_θ.

In each of these tests, the following basic properties of ammonia are assumed: molecular weight 17.03 g/mol, normal boiling point temperature 239.7 K, latent heat of vaporization 1.37 MJ/kg, heat capacity for vapor 2190.0 J/kg K, heat capacity of liquid 4490.0 J/kg K, density of liquid 682.8 kg/m^3 at normal boiling point.

3.1.2 Modeler's database

The listing of data for applying dispersion models to the four Desert Tortoise tests is given in Table 3, which contains information about the release and meteorological conditions [6]. All releases were in the horizontal, downwind direction, from a pipe at an elevation of 0.79 m. Several aspects of the data in Table 3 are explained below.

The spill rate, and the exit temperature and pressure represent average val-

TABLE 3

Modelers' data base for Desert Tortoise ammonia experiments

1. Description				
Test number	1	2	3	4
Date	8/24/83	8/29/11:20	9/1/83	9/6/83
Time	16:37 PDT	PDT	15:37 PDT	18:15 PD
2. Release conditions				
Exit pressure (atm) (avg.)	10.00	11.02	11.23	11.64
Exit temperature (K)	294.7	293.3	295.3	297.3
Nozzle diameter (m)	0.081	0.0945	0.0945	0.0945
Spill rate (kg/s)	79.7	111.5	130	96.7
Spill duration (s)	126	255	166	381
3. Site conditions				
Ambient pressure (atm)	0.897	0.898	0.895	0.891
Rel. humidity (%)	13.2	17.5	14.8	21.3
Air temperature @ 2.5 m (K)	302.4	303.9	306.9	306.0
Soil temperature (K)	304.8	303.8	304.8	304.0
Wind speed @ 2 m (m/s) (3-min avg. over 11 sites)	7.4	5.8	7.4	4.5
σ_u (m/s)	1.2	0.7	1	–
σ_θ @ 2 m (deg)	5.7	7.5	8.3	5
Friction velocity, u_*, (m/s)	0.44	0.34	0.45	0.27
Monin-Obukhov length, L, (m)	93	95	571	45
Cloud cover (%)	1	4	70	1
Pasquill stability class	D	D	D	E
T (16 m) – T (2 m) (°C)	0.87	0.46	0.13	0.9
4. Peak concentrations (ppm)				
Averaging time (s)	80	160	120	300
100 m arc	50000	83200	76900	57300
800 m arc	8800	10800	7099	15400
Other arcs	328(3500m) 101(5500m)	5000(1400m)	693(1400m)	3890(2800m)

ues over that period of the release which was considered steady. These periods correspond to the following times, where "0" seconds is the time at which the signal was given to start the release: Test 1 60–140 s, Test 2 50–200 s, Test 3 100–170 s, Test 4 100–300 s. The "site conditions" are composed of meteorological data averaged over the first three minutes of the test, and derived boundary layer parameters. Note that the wind speed is the average speed over measurements made at 2 m above the ground at 11 sites during the 3-min averaging period.

Concentrations are available at many sites in the array. The distances of these measurements are tabulated, along with an averaging time. This averaging time was selected to encompass the period over which the properties of the gas cloud at 100 m and 800 m downwind of the spill are "steady." That is, the period does not include the leading or the trailing tails of the cloud.

3.2 Goldfish database

Amoco Oil Company and the LLNL conducted six anhydrous HF spill tests during the summer of 1986 at the Department of Energy (DOE) Liquefied Gaseous Fuels Facility (LGF) at Frenchman Flat, part of the Nevada Test Site. This is the same site used for the Desert Tortoise ammonia field experiment. The only publicly available discussion of the experiment is in a set of papers by Blewitt et al. [7,32,33] and Chan et al. [34]. These papers discuss the experiment and compare the predictions of the FEM3, SLAB, and DEGADIS 1.4 models with observed concentrations for the first three of the tests. The other three tests involved the study of water sprays and are not so useful for dispersion model evaluation. General characteristics of the three tests that were analyzed are given in Table 4. As with the Desert Tortoise tests, these tests were all conducted during similar meteorological conditions (moderate wind speeds and nearly neutral stabilities).

The pressurized liquid HF was released from a spill pipe pointing horizontally downwind at a height of 1 m above the ground. An electric heater maintained the temperature of the HF in the pipe at approximately 40 °C. The liquid jet flashed as it exited the pipe into the atmosphere at ambient pressure, resulting in about 15% of the liquid changing phase to become a gas. The remaining 85% of the HF-jet remained as a liquid, which was broken up into an aerosol by the turbulence inside the jet. No liquid was observed to deposit on the ground, and no liquid was collected in the liquid collection pad during any of the tests.

Hydrogen fluoride was sampled along three rows which were set up normal to the expected wind direction at distances of 300 m, 1000 m, and 3000 m from the point of the release. Although the site of the HF tests was the same as the site of the NH_3 tests (see Fig. 1), the downwind distances to the monitoring arcs are slightly different.

The amount of water vapor in the ambient air during Test 3 was artificially increased by means of a 4.9 MW (500 hp) steam boiler/water injection systems and a shallow pond 250 m long and 600 m wide. The purpose of this test

TABLE 4

Test summary for Goldfish HF spills [6]

Test	Date (1986)	Spill rate (m^3/min)	Spill duration (s)	Mean wind speed (m/s)	Atmospheric stability class (Pasquill)
1	Aug. 1	1.78	125	5.6	D
2	Aug. 14	0.66	360	4.2	D
3	Aug. 20	0.65	360	5.4	D

was to investigate the effect of humidity on the development of the cloud. However, the moisture source did not raise the relative humidity in this desert environment above 20%.

3.2.1 Instrumentation

Measurements of the hydrogen fluoride concentration in the atmosphere were obtained by means of integrated filter samplers (IFS), located on arcs at downwind distances of 300 m and 1000 m. At each position, HF concentrations were measured at elevations of 1, 3, and 8 m.

Numerous wind and temperature measurements were made throughout the sampling array. Eighteen cup-and-vane anemometers were located at a height of 2 m at various positions within the test array in order to define the wind field for planning the releases and subsequent calculation of plume trajectories. In addition, three towers or masts provided profiles of wind speed and direction and turbulence in the flow upwind of the point of release. Temperature measurements were obtained from a mast and from an array of thermocouples located between 20 m and 200 m downwind of the point of release, and also along the sampling row at 1000 m downwind of the point of release. These thermocouples were located at a height of 1 m above the ground.

3.2.2 Modeler's data archive

The archive of data for applying dispersion models to the three Goldfish tests is listed in Table 5. In addition, the following basic information about hydrogen fluoride can be obtained from chemistry textbooks: molecular weight 20.01 g/mol, normal boiling point temperature 292.7 K, latent heat of vaporization 3.73 10^5 J/kg, heat capacity for vapor 1450.0 J/kg K, heat capacity for liquid 2528.0 J/kg K, and density of liquid 987.0 kg/m^3 at normal boiling point.

Much of the material in Table 5 is taken from a tabulation prepared by Blewitt et al. [32]. The "site conditions" are composed of meteorological data averaged over fifteen minutes during the test. The Monin–Obukhov length is set to a large value on the basis of the estimated stability class (neutral). The friction velocity is calculated from the observed mean wind speed, and the estimate of the roughness length for the site (0.003 m). Because soil temperature data are not available, we have set soil temperature equal to the air temperature. Note that the wind speed is the average speed over measurements made at 2 m above the ground at 15 locations at the site. The largest concentration measured within each monitoring arc is listed at the end of the table.

The stated molecular weight assumes that all hydrogen fluoride is present as a monomer. In fact, several different oligomers are present (e.g., HF, H_2F_2, H_6F_6, etc.) in the initial plume, and subsequent chemical changes may be important if a model were to attempt to accurately simulate the initial thermodynamics [35]. Of the models that are being evaluated, only the HEGADAS-HF

TABLE 5

Modelers' data base for Goldfish hydrogen fluoride experiments

1. Description			
Test number	1	2	3
Date	8/1/86	8/14/86	8/20/86
Time	18:15 PDT	18:15 PDT	18:15 PDT
2. Release conditions			
Exit pressure (atm) (avg.)	6.80	7.35	7.48
Exit temperature (K)	313.2	311.2	312.2
Nozzle diameter (m)	0.0419	0.0242	0.0242
Spill rate (kg/s)	27.67	10.46	10.27
Spill duration (s)	125	360	360
3. Site conditions			
Ambient pressure (atm)	0.893	0.889	0.894
Rel. humidity (%)	4.9	10.7	17.7
Air temperature @ 2.5 m (K)	310.2	309.2	307.2
Soil temperature (K)	310.2 (=air T)	309.2	307.2
Windspeed @ 2 m (m/s) (3-min avg. over 18 sites)	5.6	4.2	5.4
σ_θ @ 2 m (deg.)	10.7	14.9	10.7
Friction velocity, u_*, (m/s)	0.34[a]	0.26[a]	0.33[a]
Monin-Obukhov length, L, (m)	9999	9999	9999
Pasquill stability class	D	D	D
T (16 m) − T (2 m) (°C)	0.527	0.0276	0[b]
4. Concentration data (ppm)			
Averaging time (s)	66.6	66.6	66.6 (x=0.3,1 km), 88.3 (x=3 km)
300 m arc	25473	19396	18596
1000 m arc	3098	2392	2492
3000 m arc	411	–	224

[a]Friction velocity, u_*, is estimated from u, z_0, and assumed value of L.
[b]This temperature difference was observed to be 1.35 °C, but is assumed to be 0 °C because the observed value was affected by evaporation from an artificial pond constructed just upwind of the tower.

model (HFSYSTEM) accounts for the variation of molecular weight of hydrogen fluoride.

4. Description of model evaluation procedures

The predictions of the hazardous gas models described in Section 2 were evaluated using the data sets from the two field experiments described in Section 3. Statistical tests were applied to the sets of observations and model predictions in order to determine whether there were significant differences between the model predictions and the observations and between the predictions of the various models. The following subsections describe the performance measures and statistical tests that have been applied, as well as some of the methods used to evaluate the sensitivity of the models.

4.1 Performance measures

The performance measures that are most relevant to evaluation of hazardous gas models have been considered by several researchers [35,36,3]. Because of the potentially serious toxic effects of hazardous gases, both the maximum concentration over some averaging period within the cloud and the lateral dimension of the cloud should be correctly simulated by the model. Only the results for the maximum concentrations are discussed in this paper. The results for the cloud dimensions were less conclusive [10] and will be discussed in a separate paper. The appropriate averaging periods in our analysis would be the values listed in Tables 3 and 5, ranging from 66 seconds to 300 seconds.

Once a table of observed concentrations (C_o) and predicted concentrations (C_p) is available, it can be used to calculate the three performance measures that we recommend [8,37]. These performance measures are the fractional bias, *FB*, normalized mean square error (NMSE), and fraction within a factor of two (FAC2) defined by:

$$FB = (\bar{C}_o - \bar{C}_p)/(0.5(\bar{C}_o + \bar{C}_p)) \tag{3}$$

$$\text{NMSE} = \overline{(C_o - C_p)^2}/\bar{C}_o\bar{C}_p \tag{4}$$

$$\text{FAC2} = \text{Fraction of } C_p \text{ within a factor of 2 of } C_o \tag{5}$$

where an overbar indicates an average over all points in the data set. The fractional bias, *FB*, is also used by the EPA in their model evaluation exercises [39].

4.2 Estimation of confidence intervals

In order to estimate confidence intervals, this model evaluation exercise employs a statistical computer code that has been applied in several other model evaluation exercises [8,38]. The code uses a blocked bootstrap or jackknife resampling method to estimate whether the calculated performance measures *FB*, NMSE, and FAC2 defined above are significantly different from zero for each model and whether the differences in these measures between pairs of models are significantly different from zero. In this manner the goodness of each model can be determined, as well as whether the predictions of the "best" model are significantly different from the predictions of the other models.

The bootstrap of jackknife resampling methods allow the standard deviation, σ, of any performance measure to be estimated, from which confidence limits can be calculated by the student-t procedure:

$$95\% \text{ confidence limits} = \text{mean} \pm t_{95}\sigma(n/(n-1))^{1/2} \tag{6}$$

where tables in which the student-t parameter, t_{95}, is given as a function of degrees of freedom, $n-1$, can be found in most statistics textbooks [40]. This estimate of the degrees of freedom assumes that the n data points are truly independent of each other. In most applications, there is likely to be some

correlation among the data (e.g., the concentration observed during one hour is likely to be highly correlated with the concentration observed during the next hour), and the actual number of degrees of freedom may be less than $(n-1)$. Consequently the 95% confidence limits will be broader. It is recommended that the degrees of freedom be set equal to the quantity $(n-1)$ minus the number of pairs of data points that are highly correlated with each other. For example, the degrees of freedom for the four Desert Tortoise runs would equal 3, since the runs were made on different days.

As an example of how these confidence limits might be applied, suppose the fractional bias, *FB*, of the SLAB model for some experiment is 0.2 and *FB* of the ADAM model for the same data is 0.1. Is the ADAM model a significant improvement over the SLAB model? In this case the bootstrap method might be applied to the difference in *FB* between the two models, $DFB = FB(\text{SLAB}) - FB(\text{ADAM})$. The statistical question is phrased in the following way: Can we say with 95% confidence that the difference *DFB* is not zero? Suppose the standard deviation of DFB is found to equal 0.1. If there are $n=20$ independent data points, then $t_{95}=2.09$, and from eqn. (6) the 95% confidence limits will be $0.10 \pm (2.09)(0.1)(20/19)^{1/2}$, i.e. ranging from -0.11 to 0.31. Because these confidence limits overlap zero, we cannot say with 95% confidence that the difference *DFB* is not zero. We would conclude that the ADAM model is not a significant improvement over the SLAB model in this contrived example. In general, it is difficult to show 95% significant differences between air quality models unless there are large quantitative differences in the model predictions (factor of two or greater) or the size of the data set is large ($n=100$ or greater). Unfortunately, n is quite small for the data sets that are studied in this report ($n=4$ for Desert Tortoise and $n=3$ for Goldfish).

The application of the jackknife resampling procedure requires that the data be separated into blocks characterized by similar concentrations or meteorological conditions. There are two sets of field experiments (Desert Tortoise and Goldfish), and three monitoring arcs in any field test. In this model evaluation exercise, the data set from each arc distance for either the Desert Tortoise or Goldfish experiment has been considered to be a single block. Thus there are six blocks, with each Desert Tortoise arc-block containing four data points and each Goldfish arc-block containing three data points (one point for each field test), at most.

5. Results of model evaluation

The model evaluation procedures described in Section 4 were applied to hazardous gas model predictions generated using the Desert Tortoise and Goldfish input data presented in the tables in Section 3. The predicted and observed peak concentrations on each monitoring arc for each of the Desert Tortoise and Goldfish tests are given in Table 6. Note that there are predictions from

14 basic models included in this table, plus alternate predictions for the ALOHA, CHARM, DEGADIS, and HEGADAS models employing an initial dilution (i.d.) assumption.

5.1 Qualitative analysis of model performance

Many insights can be gleaned from an analysis of the numbers in Table 6 prior to application of the statistical performance evaluation software. For example, Fig. 2 contains plots of observed concentrations, C_o, as a function of downwind distance for the Desert Tortoise 4 and Goldfish 1 tests, respectively. Predictions of concentrations, C_p, by the ALOHA, CHARM, DEGADIS, and HEGADAS models are also plotted on the graphs. The ALOHA model was chosen for presentation in Fig. 2 because it is a Gaussian model and it typically underpredicts C by a factor of about two. However, the ALOHA model (as well as the other three models on the figure) overpredicts the concentration in the Desert Tortoise 100 m arc by an order of magnitude. The CHARM model is seen to consistently overpredict all concentrations on the figure. The HEGADAS and DEGADIS models are similar to each other in structure and are both seen to provide fairly good agreement with the concentration observations at all distances except the Desert Tortoise 100 m arc. The Desert Tortoise 100 m arc is unique because the excess density in the plume was still relatively large at that distance and the effect of the initial air entrainment were still significant.

The ALOHA model's overprediction of concentrations at the 100 m distance at Desert Tortoise is not much different from similar overpredictions by the AFTOX, CHARM, DEGADIS 2.1, EAHAP, WHAZAN, and HEGADAS models. For Desert Tortoise 4, the ALOHA and DEGADIS model predictions are nearly equal.

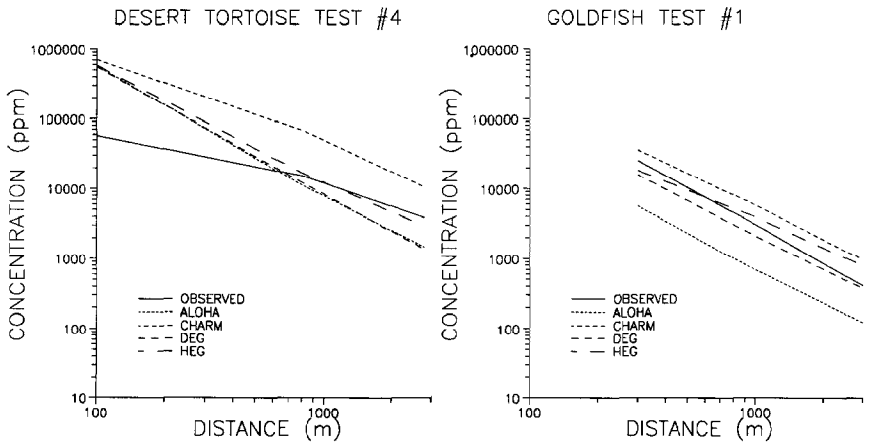

Fig. 2. Graphs of observed and predicted concentrations as a function of downwind distance for Desert Tortoise test 4 (left) and Goldfish test 1 (right). Each figure contains predictions of the ALOHA, CHARM, DEGADIS, and HEGADAS models.

TABLE 6

Observed and predicted peak concentrations, C (ppm), at each downwind monitoring arc for four Desert Tortoise tests and three Goldfish tests

Test	Arc	Comparison of peak concentration (ppm)								
		OBS. @ 1 m	ADAM @ 0 m	AFTOX 3.1 @ 1 m	ALOHA @ 0 m	B&M2 @ 0 m	CHARM @ 1 m	DEG 2.1 @ 0 m	EAHAP @ 1 m	HEGADAS @ 0 m
DT1	100 m	50000	29300	235328	112100	65000	346300	259000	154100	287024
	800 m	8800	2630	6943	2504	9100	18760	7480	9300	10663
	3500 m	328	555	716	247	300	2009	713	900	1499
	5500 m	101	378	367	130	121	705	295	500	732
DT2	100 m	83200	25400	371670	207300	71000	488000	446000	163800	439104
	800 m	10800	3650	10924	4633	15200	30760	11300	10100	14958
	1400 m	5000[a]	1430	4366	1845	2800	15170	4290	3500	6457
DT3	100 m	76900	27100	291726	185500	68000	479800	387000	170500	425566
	800 m	7099[b]	3920	8002	4146	13700	29140	11200	12100	13132
	2800 m	693	815	1026	566	600	5026	1440	1500	1880
DT4	100 m	57300	27300	612505	579500	70000	704600	555000	165100	586589
	800 m	15400	4510	21456	11270	14000	68840	12300	8600	16877
	2800 m	3890	875	3218	1487	900	10640	1380	1300	2929
GF1	300 m	25473	–	14625	5655	31500	36090	15900	31600	18372
	1000 m	3098	–	1816	691	2100	5830	2120	2700	3944
	3000 m	411	–	335	124	315	986	375	270	804
GF2	300 m	19396	–	3129	2841	21000	19520	7990	13500	11100
	1000 m	2392	–	357	347	2100	2995	1090	1300	2128
	3000 m	96[c]	–	49	62	280	392	193	130	393
GF3	300 m	18596	–	5952	2156	17500	16360	6990	15800	13004
	1000 m	2492	–	746	263	2100	2389	1010	1600	1943
	3000 m	224	–	135	47	280	361	131	110	300

[a]Sampler was saturated.
[b]During the averaging period (190 s–310 s), the plume is outside the network almost half of the time.
[c]The plume centerline was not seen by the monitor.

None of these models adequately accounts for the initial conditions associated with the aerosol jet. For example, the ALOHA model uses the Gaussian plume formulation which reduces to the following equation for a ground level source and a receptor on the plume centerline:

$$C = Q/(\pi u \sigma_y \sigma_z) \tag{7}$$

where Q is the initial mass emission rate (g/s) and C is in units of g/m^3. Because the model assumes empirical formulas in which both σ_y and σ_z approach zero as x approaches zero, the predicted concentration becomes unrealistically large near the source. One solution to this problem would be to use the following interpolation formula:

$$C = Q/(q_0 + \pi u \sigma_y \sigma_z) \tag{8}$$

OB/DG @ 0 m	PHAST @ 0 m	SAFETI @ 0 m	SLAB @ 1 m	TRACE @ 1 m	WHAZAN @ 0 m	ALOHA (id) @ 0 m	CHARM (id) @ 0 m	DEG (id) @ 0 m	HEG (id) @ 0 m
484853	44500	40080	155000	44819	233697	35660	61810	70200	79939
8408	7760	9200	5370	7859	6207	2128	4361	5630	5821
473	457	559	453	630	355	239	576	450	711
196	201	220	213	284	156	128	293	163	401
492377	48165	47030	202000	50057	287734	49770	74240	84000	97245
8536	11092	11360	8230	13120	9979	3727	5473	9280	9876
2866	5096	6861	3180	4814	3446	1635	2709	3480	3732
440688	53057	51530	206000	51432	277867	47790	81460	86600	95600
7640	11438	12480	7130	12454	9822	3383	6317	9130	7938
664	978	1522	829	1438	917	537	1233	1140	999
613303	41422	40480	198000	59855	320155	72770	77420	83700	107141
10633	9654	9378	9630	19296	12618	7732	6431	10900	11545
924	1704	2872	1260	2244	2006	1358	1254	1280	1497
13232	25966	24510	9990	14055	17318	4205	11620	13500	12763
1265	1830	3340	1230	2270	1688	635	2499	2340	1977
148	288	166	180	351	176	121	543	321	357
3286	13788	14330	5150	7487	8373	2285	6981	7420	7319
314	896	1521	587	1214	775	326	1352	1240	1048
37	111	133	83	190	85	61	253	141	167
3110	9615	12920	4240	6174	6935	1762	6439	6390	5752
297	738	1057	457	956	529	249	1247	1090	929
35	78	90	62	138	61	46	240	99	109

where q_0 is the initial volume flux (m^3/s), thus assuring that C approaches the initial concentration (Q/q_0) at the release point.

Another reason for some of the concentration overpredictions at the 100 m distance at Desert Tortoise is the failure of models such as the ALOHA model to account for the thermodynamic effects of the aerosol. Britter and McQuaid [17] suggest that this effect can be included by assuming that the cloud is initially diluted by a volume of air that would thoroughly evaporate the aerosol (see the derivation in Section 2.4). The notation "id" after some of the models in Table 6 indicates that this "initial dilution" has been accounted for. This method has also been included in the Britter and McQuaid model applications. It is obvious in the table that the "id" assumption greatly improves the models' performance on the closest arc. However, in some cases the "id" model performance deteriorates at distances farther downwind because of the modification to the plume's internal thermodynamics that come about as a result of the initial dilution.

Figure 3 contains plots of the ratios of predicted to observed concentration, C_p/C_o, for all models in Table 6 as a function of downwind distance for all the Desert Tortoise and Goldfish tests. These data illustrate the range of predictions of all the models and their relation to the line of "perfect agreement" (i.e., C_p/C_o equal to one). The distribution of C_p/C_o is fairly consistent at all

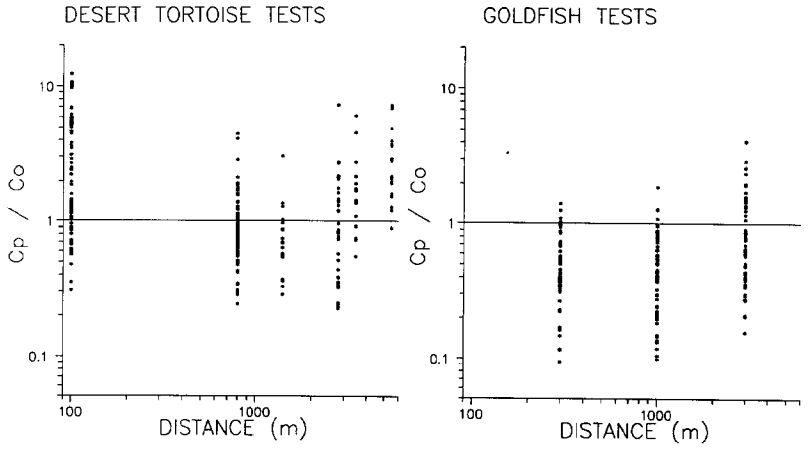

Fig. 3. Graph (bilogarithmic) of predicted and observed concentration ratios (C_p/C_o) as a function of downwind distance, for combined Desert Tortoise tests (left) and Goldfish tests (right). Each point represents a single model prediction for a single test.

distances for both tests, with a typical range of about one to one and one-half orders of magnitude (i.e., a factor of about 10 to 30 or 40). However, the median C_p/C_o value at the closest distance shows a disagreement between the two field experiments, with a value of about 2 (i.e., overprediction by a factor of 2) for the Desert Tortoise tests and 0.4 (i.e., underprediction by a factor of 2.5) for the Goldfish tests. The median C_p/C_o is closer to unity at moderate distances for both tests. The apparent overpredictions at the farthest two distances for the Desert Tortoise tests is probably due to the fact that the plume centerline may not have been measured. Furthermore, steady-state conditions may not have been achieved at these distances due to the relatively short release duration time. The differences in model performance between the two field programs may be due to the fact that the emission rate is an order of magnitude less in the Goldfish field experiments. Consequently the ability of the models to treat dense gases was tested more severely for the Desert Tortoise field experiments. Furthermore, the initial dilution calculation is less certain for HF since its boiling point is only a few degrees less than the temperature of the air.

The qualitative behavior of the model predictions in Table 6 can be explained based on study of the physical assumptions in the various models:

ADAM is capable of making predictions only for the Desert Tortoise NH_3 runs. The model accounts for the initial jet and the NH_3 thermodynamics, but tends to underpredict concentrations on the 100 and 800 m arcs by a factor of two or three. Its entrainment rate has been "tuned" with the Desert Tortoise data, so that its predicted plume width matches the observed widths during Test 4 [11]. It is clear that further tuning of model parameters should be carried out in order to increase the predicted concentrations.

AFTOX neglects dense gas effects and behaves more like a standard Gaussian plume formula. It overpredicts the NH_3 concentrations on the 100 m arc by a factor of 4 or 5 because it has no mechanism for forcing the concentration to approach C_o as x approaches zero. The AFTOX model matches the NH_3 concentrations at $x \geq 800$ m fairly well. On the other hand, it underpredicts the HF concentrations by a factor of 2 to 8 at all distances. Although the AFTOX model does not account for the initial jet or dense gas slumping, its predicted (Gaussian) plume widths are, by coincidence, only slightly low at all distances. Apparently the AFTOX σ_y predictions for stability class D are nearly equal to the σ_y values for a momentum jet released in a horizontal direction near the ground.

ALOHA is another standard Gaussian model, and tends to underpredict concentrations by factors of 2 to 5. However, it overpredicts the NH_3 data on the 100 m arc for the same reasons as the AFTOX model (i.e., failure to force C to approach C_o as x approaches 0). When initial dilution is accounted for (ALOHA-id), the 100 m arc predictions are improved, but now are about 30% low.

B&M shows little bias at all downwind distances. The good agreement at $x = 100$ m has been produced by the addition of algorithms that account for initial dilution. There is an ambiguity in the B&M model predictions at the 800 of 1000 m distances, which are in the interpolation zone between the ranges of applicability of the continuous plume equations (valid at shorter distances) and the instantaneous puff equations (valid at larger distances). We generally chose the continuous release solution in this range because it was more conservative by a factor of two or three.

CHARM consistently overpredicts concentrations by a factor of 5 to 10. The model developers state that this tendency is deliberate, so that the model will be conservative [24].

DEGADIS accounts for plume thermodynamics and performs fairly well at the $x \geq 800$ m arcs of the Desert Tortoise experiment, but underpredicts concentrations by about a factor of two in the Goldfish experiment. Because it lacks a formula for a horizontal jet it simulates these experiments as a pool evaporation source and, consequently, overpredicts by a factor of 5 on the Desert Tortoise 100 m arc. When initial dilution is accounted for, the DEGADIS-id model is seen to match the 100 m arc data fairly well.

EAHAP and HEGADAS model predictions are seen to be highly correlated with the DEGADIS model predictions, because the EAHAP and DEGADIS models are based on the dense gas slumping equations contained in the HEGADAS model. However, the HEGADAS model predictions are greater than the DEGADIS model predictions by about 30% to a factor of 2, probably because of slight differences in entrainment assumptions. The HEGADAS model tends to overpredict the NH_3 data but the predictions of the HF-specific HEGADAS model (HFSYSTEM) are fairly close to the HF data. In all cases these models overpredict on the NH_3 100 m arc, but it is seen that this problem can be eliminated by accounting for initial dilution.

OB/DG is an empirical formula based on applying multiple linear regression techniques to the results of inert tracer experiments. Because it does not allow C to approach C_o as x approaches 0, it, like most other models, overpredicts concentrations by a factor of 5 to 10 on the Desert Tortoise NH_3 100 m arc. However, its performance improves at $x \geq 800$ m, where the dense gas influences become unimportant. It tends to underpredict all Goldfish HF data by a factor of two to eight.

The PHAST, SAFETI, and WHAZAN models are all in the family of proprietary models distributed by a single consulting firm. The WHAZAN model is a screening model and clearly does not include sufficient jet-induced mixing, as evidenced by its factor of 4 overpredictions on the Desert Tortoise 100 m arc. On the other hand, the PHAST and SAFETI models, which account for initial jets and plume thermodynamics, produce more realistic results on the 100 m arc, as well as on the $x \geq 800$ m arcs at Desert Tortoise. The models tend to slightly underpredict the Goldfish HF concentration data.

SLAB has recently been revised to include initial jet effects, but is still seen to overpredict the Desert Tortoise concentrations on the 100 m arc by a factor of three to four. The SLAB model then underpredicts by a factor of 2 to 4 at all other arcs and at the Goldfish HF experiment.

TRACE is a proprietary model that includes many of the important physical and thermodynamic effects. Its concentration predictions are good at all arcs of the Desert Tortoise NH_3 experiment, but it tends to underpredict by about a factor of two at the Goldfish HF experiments.

5.2 Results of quantitative performance evaluation

Table 7 contains listings of the fractional bias (FB), normalized mean square error (NMSE), and fraction within a factor of two (FAC2) for each model for the predicted concentrations. Results are given for each downwind distance, and a weighted average is given at the bottom of each group, where the weighting factor is proportional to the number of tests. These quantitative results confirm what was suggested in the qualitative analysis discussed in the last subsection. For example, the ALOHA model overpredicts at the closest distance in the Desert Tortoise experiments ($FB = -1.21$). Also, some models (e.g., AFTOX) have as low a weighted average fractional bias, FB, as other models (e.g., EAHAP), but their variability, as reflected by the normalized mean square error, NMSE, is much greater. Based on the weighted average results in Table 7, the following models have the best performance (seven given for each performance measure) for the concentration data:

FB: AFTOX, B&M, DEGADIS, EAHAP, HEGADAS, SAFETI, TRACE, WHAZAN

NMSE: B&M, EAHAP, PHAST, SAFETI, TRACE, DEGADIS(id), HEGADAS(id)

FAC2: B&M, HEGADAS, PHAST, SAFETI, TRACE, CHARM(id), DEGADIS(id)

Note that the B&M, TRACE and SAFETI models are the only ones that appear on all lists.

TABLE 7

Fractional bias (*FB*), normalized mean square error (NMSE), and fraction within a factor of two (FAC2) for predicted centerline concentrations on a given downwind arc for each site for each model. Weighted averages for each model are also given

Test Arc	ADAM	AFTOX v 3.1	ALOHA v 4.2	B & M	CHARM v 5.0	DEGADIS v 2.1	EAHAP	HEGADAS	OB/DG	PHAST	SAFETI	SLAB	TRACE	WHAZAN	ALOHA id	CHARM id	DEGADIS id	HEGADAS id
Fractional bias (FB) of centerline concentrations predicted by various models versus observations																		
DT 100m	0.84	−1.40	−1.21	−0.02	−1.53	−1.44	−0.84	−1.47	−1.54	0.35	0.40	−0.96	0.26	−1.23	0.26	−0.10	−0.19	−0.35
DT 800m	0.96	−0.12	0.61	−0.21	−1.11	−0.00	0.05	−0.20	0.18	0.05	−0.01	0.32	−0.22	0.09	0.85	0.60	0.19	0.18
DT > 800m	0.85	0.03	0.80	0.72	−1.08	0.21	0.26	−0.30	0.65	0.17	−0.18	0.51	0.06	0.37	0.88	0.49	0.42	0.31
GF 300m	N/A[a]	0.91	1.43	−0.10	−0.13	0.69	0.04	0.40	1.06	0.25	0.20	1.06	0.78	0.64	1.54	0.87	0.80	0.84
GF 1000m	N/A	0.93	1.44	0.24	−0.34	0.62	0.35	−0.00	1.24	0.79	0.30	1.11	0.57	0.91	1.47	0.44	0.52	0.68
GF 3000m	N/A	0.30	1.15	0.07	−0.72	0.23	0.50	−0.54	1.11	0.54	0.85	0.90	0.26	0.91	1.17	−0.21	0.41	0.31
Weighted average	0.88	0.01	0.59	0.15	−0.90	−0.02	0.02	−0.40	0.33	0.32	0.18	0.40	0.24	0.18	0.96	0.38	0.33	0.29
Normalized mean square error (NMSE) of centerline concentrations predicted by various models versus observations																		
DT 100m	0.98	4.67	4.19	0.03	6.17	4.72	0.86	5.03	5.88	0.17	0.19	1.21	0.13	2.47	0.18	0.03	0.06	0.15
DT 800m	1.41	0.08	0.44	0.12	2.47	0.06	0.17	0.10	0.08	0.13	0.15	0.16	0.09	0.06	0.94	0.54	0.10	0.07
DT > 800m	2.71	0.06	1.84	1.46	2.54	0.46	0.65	0.23	1.30	0.29	0.22	0.86	0.18	0.44	2.28	1.04	0.72	0.52
GF 300m	N/A	1.08	4.16	0.03	0.08	0.55	0.06	0.17	1.56	0.11	0.05	1.58	0.73	0.47	5.89	0.93	0.75	0.86
GF 1000m	N/A	1.14	4.32	0.07	0.26	0.43	0.14	0.05	2.50	0.75	0.18	1.80	0.38	1.06	4.77	0.22	0.31	0.52
GF 3000m	N/A	0.09	2.09	0.07	0.82	0.06	0.27	0.46	1.81	0.31	0.96	1.04	0.07	1.09	2.18	0.07	0.18	0.11
Weighted average	1.78	1.24	2.73	0.40	2.38	1.17	0.41	1.11	2.20	0.28	0.24	1.05	0.25	0.91	2.49	0.53	0.37	0.37
Fraction within a factor of 2 (FAC2) of centerline concentrations predicted by various models versus observations																		
DT 100m	0.25	0.00	0.00	1.00	0.00	0.00	0.25	0.00	0.00	1.00	1.00	0.00	1.00	0.00	1.00	1.00	1.00	1.00
DT 800m	0.25	1.00	0.50	1.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.25	0.50	1.00	1.00
DT > 800m	0.40	0.60	0.60	0.80	0.00	0.20	0.20	0.40	0.80	0.80	0.60	0.60	0.60	1.00	0.60	0.60	0.80	0.40
GF 300m	N/A	0.33	0.00	1.00	1.00	0.33	1.00	1.00	0.33	1.00	1.00	0.00	0.33	0.33	0.00	0.00	0.33	0.33
GF 1000m	N/A	0.33	0.00	1.00	1.00	0.33	1.00	1.00	0.00	0.33	0.67	0.00	0.67	0.33	0.00	1.00	0.67	0.33
GF 3000m	N/A	1.00	0.00	1.00	0.50	1.00	0.50	1.00	0.00	0.50	0.00	0.00	1.00	0.00	0.00	1.00	0.50	0.50
Weighted average	0.31	0.52	0.24	0.95	0.33	0.43	0.62	0.67	0.43	0.81	0.76	0.33	0.76	0.52	0.38	0.67	0.76	0.62

[a]N/A mean not available.

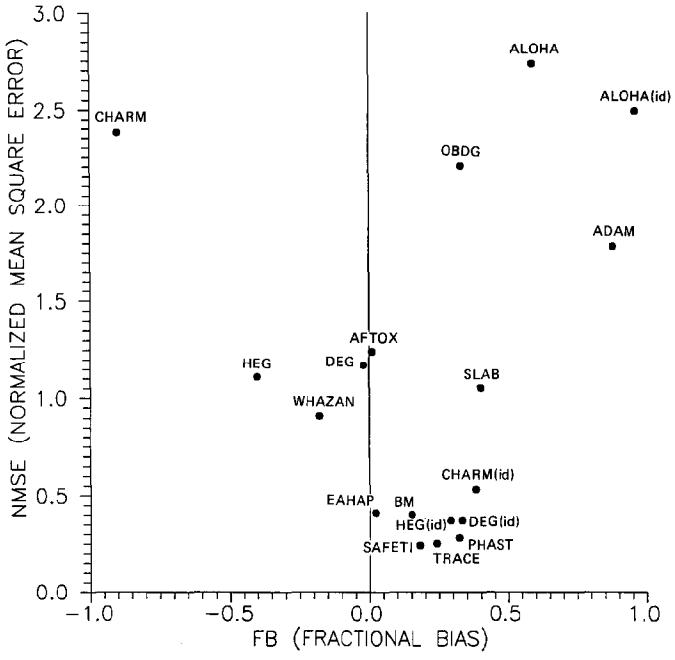

Fig. 4. Weighted average fractional bias, *FB*, and normalized mean square error, NMSE, for concentration predictions for each model (from Table 7), for Desert Tortoise and Goldfish data. The "id" suffix indicates that the initial dilution assumption has been applied to that model. Note that if a model tends to overpredict, its *FB* will be less than 0.

The jackknife resampling method [8] was applied to the *FB* and NMSE data in order to generate 95% confidence limits, using the blocking procedure that separates the data in Table 7. The following results were obtained regarding significant differences:

- Models with fractional bias, *FB*, for C_p not significantly different from zero: AFTOX, B&M, DEGADIS, EAHAP. Of these models, none of the model pairs show a significant difference.
- Of the models (B&M, EAHAP, PHAST, SAFETI, TRACE) with the lowest normalized mean square error, NMSE, for C_p, none of the model pairs shows a significant difference (The models with the initial dilution assumption are not included in this comparison)

These results suggest that with only seven field tests, it is very difficult to discern significant differences between models. In any given Desert Tortoise or Goldfish scenario, the models produce predictions that cover one to one and one-half orders of magnitude. No single model is a clear improvement. One surprising result is that the simple B&M model does very well at predicting maximum concentrations at all distances, with weighted *FB*, NMSE, and FAC2 of 0.15, 0.40 and 0.95, respectively. The B&M model has evidently accounted

for the most important aspects of the plume thermodynamics. The best models yield typical errors of $\pm 50\%$ in individual predictions.

In an attempt to demonstrate the differences among the models, the weighted fractional bias, *FB*, and normalized mean square error, NMSE, for each model are plotted in Fig. 4. A perfect model would be represented on this figure by a point at zero *FB* and NMSE. The data for concentration predictions in Fig. 4 suggest that there is a cluster of eight models (B&M, EAHAP, PHAST, SAFETI, CHARM-id, HEGADAS-id, DEGADIS-id, and TRACE) with relatively good performance. These models have $0 \leq FB \leq 0.4$ and $0.25 \leq \text{NMSE} \leq 0.50$.

6. Conclusions

This preliminary hazardous gas model evaluation exercise has employed data from only seven field tests, resulting in relatively broad 95% confidence intervals that do not permit any firm conclusions regarding which single model performs "best." In later phases of this work, the evaluations will be extended to approximately 50 to 100 additional field tests, which will result in a narrowing of the confidence intervals. At this stage, we can say only that there is a group of about 8 models with reasonable performance (i.e., mean biases in the range of 0% to 40% and typical uncertainties equal to about 50 to 70% of the mean). These models include the B&M, EAHAP, PHAST, SAFETI, and TRACE models, as well as three models that have been corrected for initial dilution (CHARM-id, DEGADIS-id, and HEGADAS-id). The initial dilution effects are important for the Desert Tortoise and Goldfish field tests, which both involve aerosol jets. It may be that stochastic processes in the atmosphere will prevent us from reducing the uncertainty much below about 50%. Consequently a simple model that accounts for the basic plume thermodynamics (e.g., the B&M model) is seen to perform as well as more complex models.

Some of the models exhibit relatively poor performance. For example, the CHARM model consistently overpredicts by a factor of about 2 to 10. The ALOHA and OB/DG models have uncertainties that are equal to about 2 times the mean observed value. Models such as AFTOX and DEGADIS have little average bias, but it is seen that this is the result of the cancellation of overpredictions on some arcs by underpredictions on other arcs. As a result, the uncertainties of the predictions by these models are greater than the mean value.

The statistical results from this model evaluation exercise would be more valid if the field data were independent. However, several of the model developers acknowledge that the Desert Tortoise and Goldfish data have been used in the derivation of parameters in their models, and many of the models have been previously tested with these data. Unfortunately, there are so few available full scale hazardous gas dispersion experiments that it is impossible to find a truly independent dataset.

The model evaluation exercise described here is being expanded to include

more field experiments (Maplin Sands, Thorney Island, Porton Down, Burro, Coyote, and some non-buoyant tracer experiments), which should reduce the range of the confidence limits on the model performance measures.

Acknowledgements

This work was sponsored by the American Petroleum Institute and the U.S. Air Force. We appreciate the guidance of Mr. Kenneth Steinberg of Exxon Research and Engineering Company, Mr. Howard Feldman of the API, and Captain Michael Moss of the U.S. Air Force Engineering and Services Center. Some of the proprietary hazardous gas model runs were made by Mr. Gib Jersey of Mobil Research and Development Corporation, Mr. Lee Gilmer and George Partridge of the Texaco Port Arthur Research Laboratories, Mr. Bruce Davis of Exxon Chemical Company, Mr. William Schikora of Chevron Corporation, and Ms. Becky Locher of Exxon Research and Engineering Company.

References

1 S.R. Hanna and P.J. Drivas, Guidelines for vapor cloud dispersion modeling, CCPS/AIChE, 345 East 47th St., New York, NY 10017, 1987.
2 D.E. Layland, D.J. McNaughton and P.M. Bodner, Assessing the reliability of dispersion models for hazardous materials releases, Proc. 5th Joint Conf. on Applic. of Air Poll. Meteorol., AMS, Boston, MA, 1986.
3 S.R. Hanna, T. Messier and L.L. Schulman, Hazard response modeling uncertainty (a quantitative method), ESL-TR-88-16, AFESC, Tyndall AFB, FL 32403, 1988.
4 J.A. Havens and T.O. Spicer, Development of an atmospheric dispersion model for heavier-than- air gas mixtures, Report No. CG-D-22-85 for the U.S. Coast Guard by the Chem. Eng. Dept., University of Arkansas, Fayetteville, AR 72701, Three Volumes, 1985.
5 D.W. Heinold, R.J. Paine, K.C. Walker and D.G. Smith, Evaluation of the AIRTOX dispersion algorithms using data from heavy gas field experiments, Proc. 5th Conf. on Applic. of Air Poll. Meteorol., AMS, Boston, MA, 1986.
6 H.C. Goldwire, T.G. McRae, G.W. Johnson, D.L. Hipple, R.P. Koopman, J.W. McClure, L.K. Morris and R.T. Cederwall, Desert Tortoise series data report: 1983 pressurized ammonia spills, UCID-20562, Lawrence Livermore National Laboratory, Livermore, CA, 1985.
7 D.N. Blewitt, J.F. Yohn and D.L. Ermak, An evaluation of SLAB and DEGADIS heavy gas dispersion models using the HF spill test data, Proc. Int. Conf. on Vapor Cloud Modeling, 1987, AIChE, New York, NY, pp. 56-80.
8 S.R. Hanna, Confidence limits for air quality models, as estimated by bootstrap and jackknife resampling methods, Atmos. Environ., 23 (1989) 1385-1398.

9 S.R. Hanna and D.G. Strimaitis, Workbook of test cases for source characterization and dispersion models for vapor clouds, CCPS/AIChE, 345 East 47th St., New York, NY 10017, 1988.

10 S.R. Hanna, D.G. Strimaitis and J.C. Chang, Results of hazard response model evaluation using Desert Tortoise (NH_3) and Goldfish (HF) databases. Final Report (2 volumes), prepared for API and USAF by Sigma Research Corp., Westford, MA, 1990.

11 P.K. Raj and J.A. Morris, Source characterization and heavy gas dispersion models for reactive chemicals, AFGL-TR-88-0003(I), Hanscom AFB, MA 01730, 1987.

12 B.A. Kunkel, Development of an atmospheric diffusion model for toxic chemical releases, AFGL-TR-85-0338, AFGL, Hanscom AFB, MA 01731, 1985.

13 B.A. Kunkel, User's Guide for the Air Force Toxic Chemical Dispersion Model (AFTOX), AFGL-TR-88-0009. AFGL, Hanscom AFB, MA 01731, 1988, 57 pp.

14 M.T. Fleischer, SPILLS, An evaporation/air dispersion model for chemical spills on land. Shell Development Center, Westhollow Research Center, P.O. Box 1380, Houston, TX 77001, 1980.

15 National Oceanic and Atmospheric Administration, ALOHA — Areal locations of hazardous atmospheres. Version 4.0 User's Manual, Hazardous Material Response Branch, NOAA, Seattle, WA 98115, 1988.

16 D.B. Turner, Workbook of atmospheric dispersion estimates. U.S. Dept. of HEW, Public Health Service, Cincinnati, OH, 1967.

17 R.E. Britter and J. McQuaid, Workbook on the Dispersion of Dense Gases, HSE Contract Research Report No. 17/1988, Health and Safety Executive, Sheffield, 1988, 129 pp.

18 G.W. Colenbrander, A mathematical model for the transient behavior of dense gas vapor clouds, Proc. 3rd Int. Symp. on Loss Prevention and Safety Promotion in the Process Industries, Basel, 1980.

19 H.W.M. Witlox, User's Guide for the HEGADAS Heavy Gas Dispersion Program, Shell Research Ltd., 1988, 22 pp.

20 J.A. Havens, A dispersion model for elevated dense gas jet chemical releases, Univ. of Arkansas, Fayetteville, AR 72701, 1987.

21 J.V. Nou, The ocean breeze and dry gulch diffusion programs, AFCRL, Hanscom AFB, MA, 1963.

22 D.L. Ermak and S.T. Chan, A study of heavy gas effects on the atmospheric dispersion of dense gases, Paper VI.6. Proc. of 15th Int. Tech. Meeting on Air Poll. Modeling and Its Applic., NATO/CCMS, 1985.

23 D.L. Ermak, User's Manual for SLAB: An atmospheric dispersion model for denser-than-air releases (Draft, November 1989 Version), Lawrence Livermore National Laboratory, Livermore, CA 94550, 1989.

24 W. Balentine and W. Eltgroth, Validation of a hazardous spill model using N_2O_4 and LNG spill data. Presented at the 78th Ann. Meeting of the Air Pollut. Control Assoc., 1985, Paper No. 85-25B.1.

25 K.J. Eidsvik, A model for heavy gas dispersion in the atmosphere, Atmos. Environ., 14 (1980) 69–77.

26 J.B. Cornwell, EAHAP User Manual, Energy Analysts, Inc., 2001 Priestley Ave., Norman, OK 73070, 1987.

27 E.D. Chikhliwala, M. Oliver and A.N. Shinde, Comparison of an accidental release modeling technique with recent spill test data, Air Pollut. Control Assoc. Annual Meeting Paper, 1987, 87-102.6.

28 Technica, PHAST User Manual, Technica Ltd, 7/12 Tavistock Square, London WCIH 9LT, 1989.

29 Technica, SAFETI User Manual, Technica Ltd, 7/12 Tavistock Square, London WCIH 9LT, 1988.

30 Technica, WHAZAN computer system for process plant hazard analysis. Technica Ltd., Lynton House 7/12, Tavistock Square London WCIH 9LT, 1988.
31 D.L. Ermak, R. Chapman, C. Goldwire, J. Govela and H.C. Rodean, 1988: Heavy gas dispersion test summary report, Report for AFESC, Tyndall AFB, FL, 1988.
32 D.N. Blewitt, J.F. Yohn, R.P. Koopman and T.C. Brown, Conduct of anhydrous hydrofluoric acid spill experiments. Proc. Int. Conf. on Vapor Cloud Modeling, AIChE, New York, NY, 1987, pp. 1–38.
33 D.N. Blewitt, J.F. Yohn, R.P. Koopman, T.C. Brown and W.J. Hague, Effectiveness of water sprays on mitigating hydrofluoric acid releases, Proc. Int. Conf. on Vapor Cloud Modeling, 1987, AIChE, New York, NY, 1987, pp. 155–180.
34 S.T. Chan, H.C. Rodean and D.N. Blewitt, FEM3 modeling of ammonia and hydrofluoric acid dispersion, Proc. Int. Conf. on Vapor Cloud Modeling, 1987, AIChE, New York, NY, 1987, pp. 116–154.
35 P.N. Clough, D.R. Grist and C.J. Wheatley, The mixing of anhydrous hydrogen fluoride with moist air, Proc. Int. Conf. on Vapor Cloud Modeling, 1987, AIChE, New York, NY, 1987, pp. 39–55.
36 D.L. Ermak and M.H. Merry, A methodology for evaluating heavy gas dispersion models, Final Report for AFESC, Tyndall AFB, FL, 1988.
37 A. Mercer, Methods of validation of heavy gas dispersion models: a review, IMA Conf. on Stably Strat. Flows and Dense Gas Dispersion, The Inst. Math. and its Appl., Southend-on-Sea, 1987.
38 S.R. Hanna, Air quality model evaluation and uncertainty, J. Air Pollut. Control Assoc., 38 (1988) 406–412.
39 U.S. Environmental Protection Agency, Interim procedures for evaluating air quality models (revised), EPA-450/4-84-023, OAQPS, U.S. EPA, Research Triangle Park, NC, 1984.
40 H.A. Panofsky and G.W. Brier, Some applications of statistics to meteorology, The Pennsylvania State University, University Park, PA, 1968, 224 pp.