10. HGSYSTEM VALIDATION

CONTENTS

10. HGSYSTEM VALIDATION

10.1. Objectives of model evaluation exercise

The primary objective of the work described in this section has been to evaluate the new HGSYSTEM version 3.0 package with data from full-scale field experiments. The performance has been compared with the performance of other hazardous gas models for limiting cases such as non-buoyant inert gases and non-reactive dense gases. As a result the typical accuracy's and relative uncertainties of the models can be estimated.

10.2. Evaluations with field data from eight sites

10.3.1. Models to be included

The new HGSYSTEM version 3.0 package has been included in all evaluations. Because nine independent dense gas models (DEGADIS, SLAB, AIRTOX, CHARM, FOCUS, GASTAR, PHAST, TRACE, Britter and McQuaid) had already been evaluated with the field data sets that were used, the performance statistics for HGSYSTEM 3.0 could be directly compared to performance statistics that existed in the files for these nine models (Hanna et al., 1993).

10.3.2. Description of field data sets

The set of field data used for this portion of the evaluations includes the eight experiments used by Hanna et al. (1993) in their evaluation of 14 hazardous gas models. The characteristics of these data sets are summarised in Table 10-1. It is seen that the data include non-buoyant releases (Prairie Grass and Hanford), continuous dense gas releases (Burro, Coyote, Desert Tortoise, Goldfish, Maplin Sands, and part of the Thorney Island tests), and instantaneous dense gas releases (Thorney Island). The three Goldfish trials involved releases of HF (about 4000 kg per trial). There are 41 separate field trials involving dense gases. These data are all stored on Earth Tech's computer files in a so-called Modelers Data Archive (MDA) that has been widely distributed to interested scientists and engineers throughout the world. We do not describe the details of these datasets here, but refer the reader to Section 3 of the Hanna et al. (1993) article, or to Volume II of the Hanna et al. (1991) project report prepared for the U.S. Air Force and the American Petroleum Institute.

The Hanna et al. (1993) model evaluation exercise included the 1990 version (indicated by NOV90 or version 1.0) of HGSYSTEM, which was applied to the eight field data sets listed in Table 10-1.

Table 10-1. Summary of characteristics of the datasets used by Hanna et al. (1993) in their model evaluations

	Burro	Coyote	Desert Tortoise	Goldfish	Hanford Kr^{85} (Continuous)	Maplin Sands	Prairie Grass	Thorney Island (Instantaneous)	Thorney Island (Continuous)
Number of Trials	8	3	4	3	5	4,8	44	9	2
Material	LNG	LNG	NH_3	HF	Kr^{85}	LNG, LPG	SO_2	Freon & N_2	Freon & N_2
Type of Release	Boiling Liquid (dense gas)	Boiling Liquid (dense gas)	2-Phase Jet (dense gas)	2-Phase Jet (dense gas)	Gas (non-buoyant)	Boiling Liquid (dense gas)	Gas Jet (non-buoyant)	Gas (dense gas)	Gas (dense gas)
Total Mass (kg)	10700-17300	6500-12700	10000-36800	3500-3800	11-24*	LNG: 2000-6600 LPG: 1000-3800	23-63	3150-8700	4800
Duration (s)	79-190	65-98	126-381	125-360	598-1191	60-360	600	Instantaneous	460
Surface	Water	Water	Soil	Soil	Soil	Water	Soil	Soil	Soil
Roughness (m)	.0002	.0002	.003	.003	.03	.0003	.006	.005-.018	.01
Stability Class	C-E	C-D	D-E	D	C-E	D	A-F	D-F	E-F
Max. Distance (m)	140-800	300-400	800**	3000	800	400-650	800	500-580	472
Min. Averaging Time (s)	1	1	1	66.6-88.3	38.4	3	Dosage	0.06	30
Max. Averaging Time (s)	40-140	50-90	80-300	66.6-88.3	270-845	3	600	0.06	30
Reference	Koopman et al. 1982	Goldwire et al. 1983	Koopman et al. 1985	Blewitt et al. 1987	Nickola et al. 1970	Puttock et al. 1980	Barad, 1958	McQuaid and Roebuck, 1985	McQuaid and Roebuck, 1985

* Curies, rather than kg, are used as a measure of the amount of this radioactive tracer released

** Concentrations are measured beyond 800 m, but there are not well-instrumented measurement arcs.

Because one new module (HEGABOX for instantaneous sources) has been added to HGSYSTEM 3.0, one module (PLUME) has been superseded by another module (AEROPLUME), and most parts of the model have changed slightly, HGSYSTEM 3.0 has been re-evaluated with the eight sets of field data. The following component modules of HGSYSTEM have been applied to these data in our new model evaluation exercise.

Field Experiment and Source Type	**Applied HGSYSTEM Module**
Burro (evaporation area source of LNG)	HEGADAS-S
Coyote (evaporating area source of LNG)	HEGADAS-S
Desert Tortoise (NH_3 aerosol horizontal jet)	AEROPLUME/HEGADAS-S
Goldfish (HF aerosol horizontal jet)	HFPLUME/HEGADAS-S
Hanford Kr^{85} (trace gas from point)	HEGADAS-S (orifice diameter unknown)
Maplin Sands (evaporating area source of LNG & LPG)	HEGADAS-S
Prairie Grass (trace gas (SO_2) from point)	AEROPLUME/PGPLUME
Thorney Island (instantaneous volume source of Freon & N_2)	HEGABOX/HEGADAS-T
Thorney Island (continuous area source of Freon & N_2)	HEGADAS-S

The existing Modelers' Data Archive (MDA) contained sufficient input data (e.g., mass emission rate, wind speed) to carry out the HGSYSTEM runs described above. The MDA also contained the concentration observations that were necessary for the statistical evaluations.

10.3.3. Model output parameters that were evaluated

Of primary interest in the evaluation is the maximum near-ground-level concentration at each downwind distance; a measure (say the standard deviation) of the plume width and height at each downwind distance; and the geometric characteristics of particular contours of concentration or dosage. For a module (i.e., AEROPLUME) where uniform crosswind and vertical profiles are assumed, the average plume concentration is the same as the maximum centreline concentration.

10.3.4. Statistical model evaluation procedures to be used

The statistical model evaluation software, BOOT, applied in the study described by Hanna et al. (1993), was used. The software has been well-tested in a wide range of studies and is currently in use by a number of groups in the U.S., Europe, and Australia. It involves the use of the relative mean bias, the normalised mean-square-error, the correlation coefficient, and the fraction of predictions within a factor of two of observations. Confidence intervals on these performance measures are generated by bootstrap resampling. Section 4 of the reference describes these procedures in detail.

10.3.5. Standards for accepting or rejecting model performance

Air quality modelers have not yet agreed upon the magnitude of standards for accepting or rejecting model performance. In most cases a model is considered 'acceptable' if most of its predictions are within a factor of two of the observations. However, in the case of dense gas models, the study by Hanna et al. (1993) demonstrated that the performance measures for several models were within a range of acceptability shown in Figure 10-1, which is a reproduction of Figure 1a in Hanna et al. (1993). It is seen that most models fall in a cluster of fair performance, with 0.7 < geometric mean bias < 1.5 and 1.3 < geometric variance < 2.5. Consequently it is expected that, to be acceptable, the performance measures for the new model would at least fall within this same range.

10.3.6. Results of model evaluation at eight field sites

The BOOT model evaluation software produces many tables and figures. Here we have selected a set of figures in which the geometric variance, VG, is plotted versus the geometric mean bias, MG, for each model. These performance measures are calculated from the following formulas:

$$VG = \exp\overline{\left(\ln\left(C_0/C_p\right)\right)^2} \qquad (1)$$

$$MG = \exp\overline{\left(\ln\left(C_0/C_p\right)\right)} \qquad (2)$$

Therefore a 'perfect' model would have VG = MG = 1.0.

Five figures are presented. Figure 10-2a,b,c consists of a set of results for concentration predictions for three groups of data--a) continuous dense gas field data, b) continuous passive gas field data, and c) instantaneous dense gas field data. Figure 10-3a,b is concerned with predictions of plume width for groups a) and b).

Figure sets a) and b) show that there is very little difference between the results for the 'old' (version 1.0) and 'new' (version 3.0) HGSYSTEM models. Both versions overpredict the mean by about 20 to 40% with a geometric variance of about 2, and both versions are within the cloud of the five or six best-performing models. The biggest difference occurs for Figure 10-2c (instantaneous dense gas field data), where the 'old' model did not apply at all, while the 'new' model (with the addition of HEGABOX) now applies and is one of the three best-performing models (along with AIRTOX and the Britter and McQuaid nomograms).

Like the other dense gas models, HGSYSTEM overpredicts the dense gas plume widths by about 50% (see Figure 10-3a) and underpredicts the passive gas plume widths by about 30% (see Figure 10-3b).

It is concluded that the new version 3.0 of HGSYSTEM is among the better performing models, with a typical mean bias of about 20 to 40% and a typical scatter less than a factor of two.

10.3. References

S.R. Hanna, J.C. Chang and D.G. Strimaitis, Hazardous gas model evaluation with field observations. *Atmos. Environ.*, **27A**, 2265-2285, 1993.

S.R. Hanna, D.G. Strimaitis and J.C. Chang, *Hazard Response Modeling Uncertainty (A Quantitative Method) Vol. II, Evaluation of Commonly-Used Hazardous Gas Dispersion Models*, Sigma Research Corp., Concord, MA, 1991.

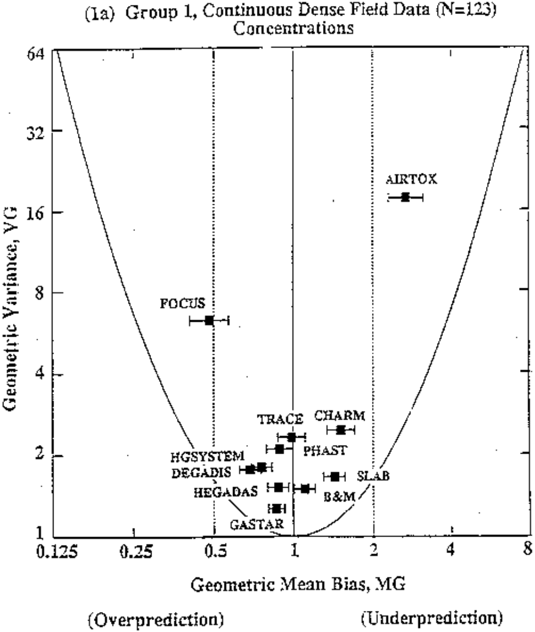

Figure 10-1. Example of presentation of model evaluation results from Hanna et al. (1993).

Group 1. Continuous Dense Field Data (N = 123) Concentrations

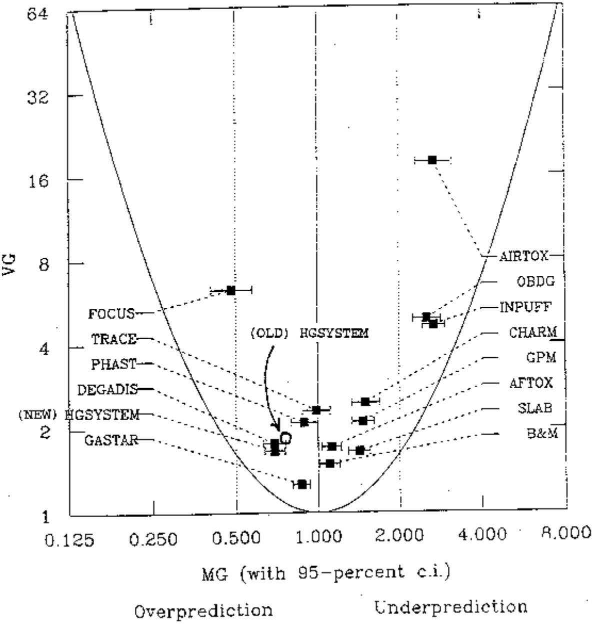

Figure 10-2a. Model performance measures, geometric mean bias MG and geometric variance VG, for maximum plume centreline concentration predictions and observations. 95% confidence intervals on MG are indicated by the horizontal lines. The solid parabola is the 'minimum VG' curve. The vertical dotted lines represent 'factor of two' agreement between mean predictions and observations. Group 1 - Continuous dense gas data sets (Burro, Coyote, Desert Tortoise, Goldfish, Maplin Sands, and Thorney Island), involving a total of 32 trials and 123 points for the shortest available instrument averaging times.

Group 2. Continuous Passive Field Data (N = 222) Concentrations

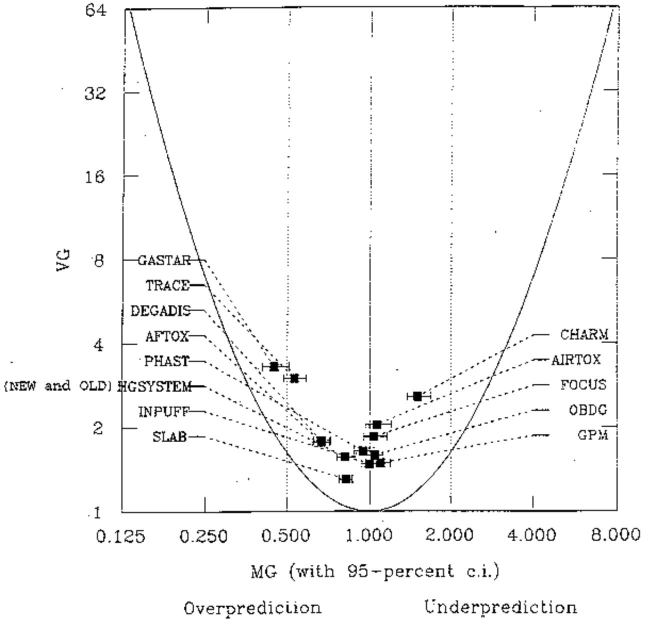

Figure 10-2b. Model performance measures, geometric mean bias MG and geometric variance VG, for maximum plume centreline concentration predictions and observations. 95% confidence intervals on MG are indicated by the horizontal lines. The solid parabola is the 'minimum VG' curve. The vertical dotted lines represent 'factor of two' agreement between mean predictions and observations. Group 2 - Continuous passive gas data sets (Prairie Grass and Hanford), involving a total of 49 trials and 222 points for the shortest available instrument averaging times.

Group 3. Instantaneous Dense Field Data (N = 61) Concentrations

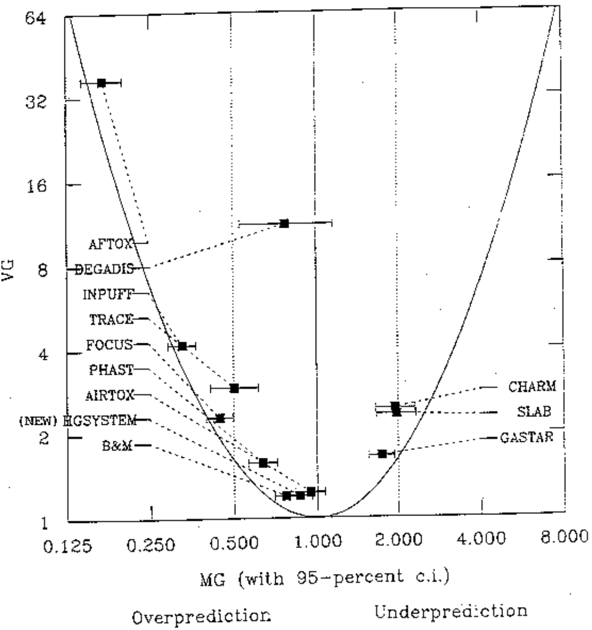

Figure 10-2c. Model performance measures, geometric mean bias MG and geometric variance VG, for maximum plume centreline concentration predictions and observations. 95% confidence intervals on MG are indicated by the horizontal lines. The solid parabola is the 'minimum VG' curve. The vertical dotted lines represent 'factor of two' agreement between mean predictions and observations. Group 3 - Instantaneous dense gas data set (Thorney Island), involving a total of 9 trials and 61 points for the shortest available instrument averaging times.

Group 1. Continuous Dense Field Data (N = 30) Widths

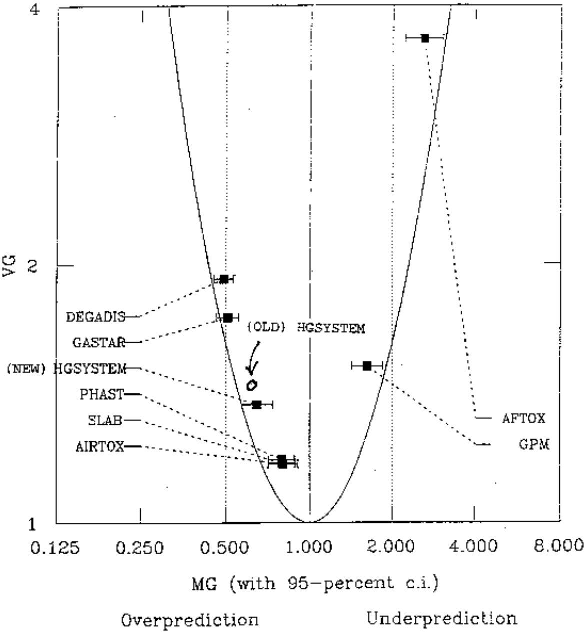

Figure 10-3a. Model performance measures, geometric mean bias MG and geometric variance VG, for plume width predictions and observations. 95% confidence intervals on MG are indicated by the horizontal lines. The solid parabola is the 'minimum VG' curve. The vertical dotted lines represent 'factor of two' agreement between mean predictions and observations. Group 1 - Continuous dense gas data sets (Burro, Coyote, Desert Tortoise, and Goldfish) involving 18 trials and 30 points.

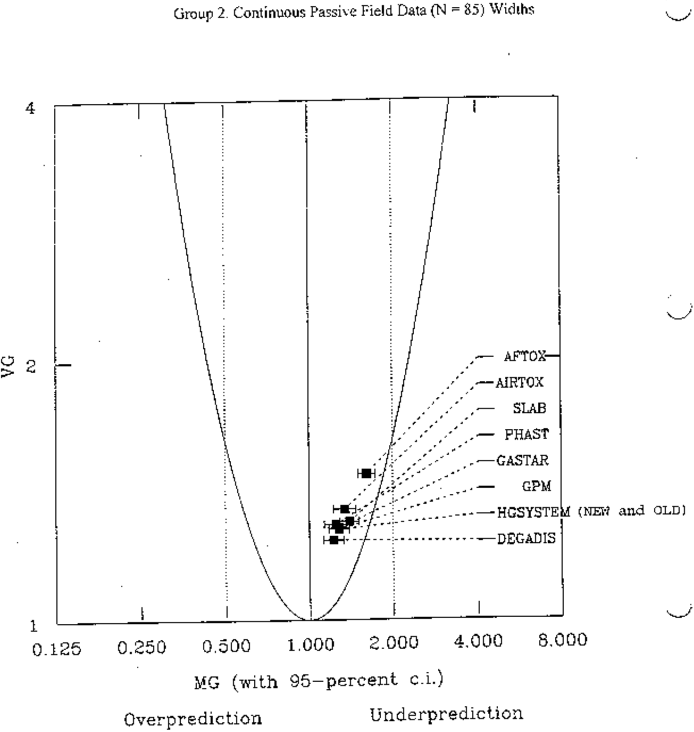

Figure 10-3b. Model performance measures, geometric mean bias MG and geometric variance VG, for plume width predictions and observations. 95% confidence intervals on MG are indicated by the horizontal lines. The solid parabola is the 'minimum VG' curve. The vertical dotted lines represent 'factor of two' agreement between mean predictions and observations. Group 2 - Continuous passive gas data set, (Prairie Grass and Hanford) involving a total of 49 trials and 85 points.