Carissimo B., S.F. Jagger, N.C. Daish , A. Halford, S. Selmer-Olsen, K. Riikonen, J.M. Perroux, J. Würtz, J. Bartzis, N.J. Duijm, K. Ham, M. Schatzmann, and R. Hall, "The SMEDIS Database and Validation Exercise", *International Journal of Environment and Pollution*, **16**, 1-6, 2001.

The SMEDIS Database and Validation Exercise

B. Carissimo
Electricité de France, Chatou, France
S.F. Jagger
Health & Safety Laboratory, Buxton, UK
N.C. Daish
CERC Ltd., Cambridge, UK
A. Halford
BG Technology
S. Selmer-Olsen
Det Norske Veritas
K. Riikonen
Finnish Met. Institute
J.M. Perroux
Gaz de France
J. Würtz
EC Joint Research Centre, Ispra
J. Bartzis
NCSR "DEMOKRITOS"
N.J. Duijm
Risø National Laboratory
K. Ham
TNO Institute of Environmental Affairs
M. Schatzmann
Met. Institute, University of Hamburg
R. Hall
WS Atkins Safety and Reliability

Abstract

SMEDIS is an ongoing research project funded by the European Union under the Environment and Climate Research Programme for the period 1996-1999. The project is coordinated by the Health and Safety Executive (HSE, UK) with two other main partners: Cambridge Environmental Research Consultants (CERC, UK) and Electricité de France (EDF, F). Ten additional partners from across Europe are also participants in the project.

The main objective of the project is to develop a methodology for the evaluation of dense gasatmospheric dispersion models used in the study of accidental releases of explosive or toxic materials. This evaluation is composed of a scientific assessment of each model, together with a validation by comparison with available experimental data.

This paper describes more specifically the data set database constructed and the validation performed by the participants involved in the project. Preliminary results indicate that the restriction to arcwise

concentrations lead to an optimistic view of model performance when complex effects are present and that, in general, statistical performance is better for more sophisticated models.

1. Introduction

SMEDIS is an ongoing EU research project funded under the Environment and Climate RTD Programme for the period 1996-1999. Its main objective is to develop a methodology for the *scientific evaluation* of dense gas dispersion (DGD) models - i.e. taking into account not only validation but also scientific assessment or review of each model - and to test this methodology by actually carrying out the scientific evaluation of a large number of DGD models currently available in Europe. The project is focusing on situations in which complex effects such as aerosols, topography and obstacles are important as well as "simple" situations.

The project is coordinated by the Health and Safety Executive (HSE, UK) with two other main partners, Cambridge Environmental Research Consultants (CERC, UK) and Electricité de France (EDF, F). There are also ten associated partners participating in the project: BG Technology (BG, UK), Det Norske Veritas (DNV, NO), Finnish Meteorological Institute (FMI, FI), Gaz de France (GDF, F), Joint Research Center, Ispra (JRC, I), National Center for Scientific Research "Demokritos" (NCSR, GR), Risø National Laboratory (DK), TNO Institute of Environmental Science (TNO, NL), University of Hamburg (UH, D) and WS Atkins (WSA, UK). Additionally, a number of external sponsors contribute both financially and technically to the project.

The origin of this project dates back to previous model intercomparison work carried out in the framework of the cooperation agreement between UK AEA and France CEA (Brighton *et al.*, 1994) and from more recent activities from a subgroup of the Model Evaluation Group (MEG), the Heavy Gas Dispersion Expert Group (HGDEG). This group has adapted the general guidelines of the MEG by drawing up a list of heavy gas dispersion models, identifying data sets, making further development of the protocol more specific to the type of model and conducting an informal open exercise to test this protocol. This open exercise comprised the distribution of a limited number of data sets to interested participants followed by an attempt to carry out a statistical analysis of returned model predictions. Details and experience of this limited exercise are given by Cole & Wicks (1995) and the work of the HGDEG is described by Duijm *et al.* (1997).

In the US there has also been considerable activity in this area during the 1980's and 90's. Hanna *et al.* (1993) carried out a seminal validation study on a number of commonly used dense gas dispersion models, focusing on flat unobstructed terrain.

In practice industrial sites consist of complex, obstructed terrain, and failure of pressurised storage and transportation vessels will commonly lead to a dense gas cloud containing liquid droplets, i.e. aerosols. Such complex effects - aerosol releases, terrain effects and obstacles - are therefore crucial elements in a real site and accident scenario.

SMEDIS is the first project to combine scientific assessment with validation against observed data, including these complex effects, and to apply its procedure to a large number of dense gas dispersion models (DGD) - the majority in use across Europe, as shown in Table 1. Furthermore, the goal of SMEDIS is to encourage continual model improvement - rather than to rank a set of models at one instant in time - by leaving in place a protocol and archived database of test cases which can be used by all DGD developers and users in the future.

Daish *et al.* (1999) have given a general description of the entire project. In this paper, we describe more specifically the construction of the validation data set database, how the validation phase was conducted and the results obtained up to this point.

Model	Developer
Screening tools	
Britter-McQuaid Workbook	HSE & CUED, UK
VDI Guideline 3783 Part 2	Met. Institute, U. Hamburg, GERMANY
Integral models	
AERCLOUD	Finnish Met. Institute, FINLAND
DEGADIS	US Coastguard, US EPA and GRI, USA
DRIFT	AEA Technology, UK
EOLE	Gaz de France, FRANCE
ESCAPE	Finnish Met. Institute, FINLAND
GASTAR	CERC Ltd., UK
GReAT	Risø National Laboratory, DENMARK
HAGAR	BG Technology, UK
HGSystem	Shell Research, UK
OHRAT/Multi-Stage	Det Norske Veritas, UK/Norway
PHAST/UDM	Det Norske Veritas, UK/USA
SLUMP	WS Atkins Safety & Reliability, UK
WHAZAN/HVYCLD	Det Norske Veritas, UK/USA
Shallow-layer models	
DISPLAY-1	EC Joint Research Centre, Ispra, ITALY
DISPLAY-2	EC Joint Research Centre, Ispra, ITALY
SLAB	Lawrence Livermore Natl. Lab., USA
SLAM	Risø National Laboratory, DENMARK
TWODEE	HSE/HSL, UK
CFD models	
ADREA-HF	NCSR “DEMOKRITOS”, GREECE
CFX	AEA Technology, UK
COBRA	Mantis Numerics Ltd., UK
FLACS	Christian Michelsen Research, NORWAY
FLUENT	FLUENT, UK
KAMELEON FireEx 98	SINTEF, NORWAY
MERCURE	Electricité de France, FRANCE
STAR-CD	Computational Dynamics Ltd., UK

Table 1 Models participating in SMEDIS. In the above table: HSE = Health & Safety Executive, UK; HSL = Health & Safety Laboratory, UK; CUED = Cambridge University Engineering Department, UK. The VDI model uses a worst case approach and has not been included in the statistical analysis of results.

2. Database description

2.1 Data selection

Identification of suitable data sets for the model validation was carried out with input from all participants. First a preliminary list of over 40 data sets and corresponding references was prepared by collecting information from partners. Based on this preliminary list, a detailed questionnaire was sent to all partners to obtain information about the data sets, including previous use of the data for model validation, availability of data and an opinion of the user on the data set quality in several areas, e.g. source specification and concentration measurements.

The replies were then analysed to produce a priority list of data sets which had been previously used for the validation of a range of models, including integral and CFD models, and were judged of sufficient quality by the participants. There turned out to be relatively few high quality data sets available for validation in these complex situations. These are presented in Table 2.1 A provision was made to include more recent data sets, which had not already used for validation, as the project progressed, including data sets for which near-field effects are important (BMT and ETH-Z datasets). These datasets, processed within the project, were added to the REDIPHEM database.

Identifier	**Scale**	**Material**	**Source type**	**No. tests**	**Complex effects**
Burro	field	LNG	pool	8	fast aerosol evaporation
Desert Tortoise	field	Ammonia	jet	4	aerosol
FLADIS-Risø	field	Ammonia	jet	16	aerosol
BA-Hamburg	wind tunnel	Sulphur hexafluoride	continuous instantaneous	146	obstacles, slopes
BA-Propane	field	Propane	jet/cyclone	51	aerosol, fences
BA-TNO	wind tunnel	Sulphur Hexafluoride	continuous instantaneous	13	fence
Thorney Island	field	Freon	instantaneous	30	fence, building
EMU-Enflo	wind tunnel	Krypton	continuous	2	buildings, real site

Table 2.1 Data set groups selected based on questionnaires returned by all participants.

There is a mixture of the well-established experimental programmes used by Hanna *et al.* (1993) and more recent EC-funded programmes. Most of the data sets from both categories can be found in the REDIPHEM database (Nielsen and Ott, 1996); see also Britter (1998) for a review of the European experimental programmes.

Both field trials and wind tunnel experiments were allowed, although preference was given to field trials where these were available. However, it became clear that the exclusion of wind tunnel tests would be unduly restrictive to the range of scenarios considered and so a mixture of field trials and wind tunnel experiments was used in the final selection.

2.2 Scenario classification

It is clearly preferable to test any model over as wide a range of conditions as possible, i.e. for a broad range of scenarios. In order to delimit the possible range of scenarios, a scheme was devised to classify them. This scheme is based on four main characteristics of a scenario, as shown in Table 2.2.

Main code	Main characteristic	Values	Value code
S	source type	continuous	c
		instantaneous	i
C	complex effect	aerosol	a
		complex terrain	t
		obstacle	o
		congestion/confinement	c
		none	n
D	gas density	weak	w
		strong	s
A	atmospheric conditions	neutral/unstable	n
		stable	s
		low/no wind	l

Table 2.2 Classification scheme for dense gas release scenarios. The code letters can be used to denote a particular scenario, e.g. ScCnDwAn refers to a scenario with a continuous source, no complex effects, weak dense gas effects and neutral atmospheric conditions.

The complex effect “congestion/confinement” and atmospheric conditions “no/low wind” were added later to the basic formulation in order to highlight situations where near-field effects are particularly important, e.g. because of very low wind or where there is interaction with obstacles near the source. After a search for suitable data sets, two groups of wind tunnel tests were identified, namely continuous dense releases inside an offshore module and instantaneous dense releases on a slope in calm conditions (Daish *et al.* 1999).

This gives a total of 45 combinations, or cells in a four-dimensional matrix, and so ideally there would be at least one high quality data set per combination. In fact, based on the short list given in Table 2.1 there were data sets to fill most but not all of the cells in the four-dimensional array, as described in Daish *et al.* (1999). There are few instantaneous releases showing aerosol effects and no releases with terrain (slopes) for stable atmospheric conditions. There is heavy reliance on the wind tunnel experiments designated “BA Hamburg” which were necessarily carried out under neutral conditions. Nevertheless, there is a sufficient spread of conditions for which data sets of adequate quality can be found to carry out effective validation

2.3 Specific cases

The final stage in the selection of cases to simulate is to choose one or more cases from as many of the complete cells as possible. Taking into account resource constraints, it was decided to limit the exercise to a total of 30 cases, divided into three main batches: some details are given in Table 2.3. The number of cases a particular model is required to simulate depends on the type of model: thus screening tools and integral models are required to simulate all 30 cases, shallow layer models approximately 50% of these and CFD models approximately 20% of the total. For all model types the cases are specified, so that all models of a given type simulate the same (or a subset of the same) cases.

Data set group	Test name	Code	Rationale	C/I	F/W	A	T	O	ST	I	SL	3d
Prairie Grass	PG8	ScCnDwAn	Simple, well-used data set	C	F				o	o	o	
	PG17	ScCnDwAn	As Test 17, but with unstable atmosphere	C	F				o	o		
Desert Tortoise	DT1	ScCaDsAn	Aerosol and stronger dense gas effects	C	F	x			o	o	o	
	DT2	ScCaDsAn	Repeat of DT1 with different flow rate	C	F	x			o	o		
BA Propane	EEC360	ScCnDsAn	Strong density effects	C	F	x			o	o		
	EEC361	ScCoDsAn	As EEC360, but with one obstacle (fence)	C	F	x		x	o	o		
	EEC362	ScCoDsAn	As EEC360, but with two obstacles (fences)	C	F	x		x	o	o		
	EEC550	ScCnDsAn	Strong density effects	C	F	x			o	o	o	o
	EEC551	ScCoDsAn	As EEC550 but with one obstacle (fence)	C	F	x		x	o	o	o	
	EEC560	ScCnDsAn	Strong density effects	C	F	x			o	o		
	EEC561	ScCoDsAn	As EEC560, but with obstacle (porous fence)	C	F	x		x	o	o	o	
	EEC170	ScCnDsAn	Strong density effects	C	F	x			o	o		
	EEC171	ScCoDsAn	As EEC170, but with obstacle (circular fence)	C	F	x		x	o	o	o	o
	LAT49	SiCaDsAn	Instantaneous release with aerosol effects	I	F	x			o	o		
BA Hamburg	DAT638	SiCtDsAn	Instantaneous source with steep slope	I	W		x		o	o	o	o
	DAT648	SiCtDsAn	As DAT638 but slope less steep	I	W		x		o	o		
	DAT231	SiCoDsAn	Instantaneous release with wall parallel to wind	I	W			x	o	o		
	DAT647	ScCtDsAn	Continuous release on slope	C	W		x		o	o	o	
	DAT458	SiCoDsAn	Instantaneous release with canyon	I	W			x	o	o	o	
	049101	SiCoDsAn	Instantaneous release and near-field array of obstacles	I	W			x	o	o		
	129034	ScCoDsAn	Continuous release with near-field obstacles	C	W			x	o	o	o	
BA TNO	TUV11	ScCnDsAn	Reference continuous release without obstacle	C	W				o	o		
	TUV13	ScCoDsAn	As TUV11 but with obstacle (oblique fence)	C	W			x	o	o		
EMU ENFLO	EMUDJ	ScCt/oDsAn	Dense release in complex terrain with obstacles (buildings)	C	W		x	x	o	o	o	o
	EMUNJ	ScCt/oDwAn	As EMUDJ but with weak density effects	C	W		x	x	o	o		
FLADIS Risø	FLADIS16	ScCaDwAs	Aerosol effects and stable atmosphere	C	F	x			o	o	o	
	FLADIS24	ScCaDwAn	Aerosol effects and neutral atmosphere	C	F	x			o	o		
	FLADIS9	ScCaDsAn	Dense release with aerosol effects	C	F	x			o	o		o
Thorney Island	TI08	SiCnDsAn	Instantaneous release with strong density effects	I	F				o	o	o	
	TI21	SiCoDsAn	Instantaneous release with strong density effects and fence	I	F			x	o	o	o	o

Table 2.3 Specific cases used in SMEDIS validation exercise. Additional cases will also be selected for the near field cases (BMT and ETH-Z).

KEY: C = continuous release; I = instantaneous release. F = field trial; W = wind tunnel experiment. An "x" indicates aerosol effects (A), terrain effects (T) and obstacles (O) important. Fences are linear unless otherwise stated. An "o" indicates the case should be attempted by screening tools (ST), integral models (I), shallow-layer models (SL) and CFD models (3d).

2.4 Validation procedure

Finally, we consider how the validation comparisons between observations and predictions are made. This comprises two aspects, namely defining the parameters used in the comparisons - physical parameters and statistical parameters - and setting up and running the models.

(a) Physical comparison parameters

These are the physical quantities, either directly measured or derived from direct measurements. Table 2.4 summarises the physical parameters being considered. There are two main divisions:

- first, distinction is made between continuous and instantaneous releases, where the parameters are broadly based on concentration and dose, respectively
- secondly, there are both pointwise and arcwise comparisons. In the former case, observations and predictions are compared over a given set of points, e.g. the dose at a given set of points; while in the latter case the comparisons are made between sets of arcs, e.g. the maximum concentration across an arc for each of a set of arcs. For a given case (trial) some or all of the arcwise comparisons may not be possible, e.g. when the arcs contain insufficient numbers of points, are not well-defined or when the presence of an obstacle makes the definition of an arc difficult.

Although the arcwise comparisons are often of most practical use, the above difficulty in complex situations has led us to include the pointwise comparisons as well, so as to give credit to models which provide spatial information on the concentration field (e.g. in situations where the cloud is distorted by the presence of obstacles and/or terrain). The cloud width is calculated using moments of the concentration distribution across the arc, while the arrival and departure times of the cloud are defined as the times at which 10% and 90% of the dose, respectively, has been recorded.

(b) Statistical comparison parameters

For a given physical parameter ψ, we need procedures and parameters for comparing the observed values $\{(\psi_o)_i\}_{i=1}^{N}$ and the predicted values $\{(\psi_p)_i\}_{i=1}^{N}$ (where the index i runs over the set of N points or N arcs) to give a measure of the overall agreement between the two sets of numbers. Many so-called statistical comparison parameters have been devised for this purpose (see, for example, Duijm *et al.*, 1996), each with their own merits and limitations, but in general the parameters most commonly adopted (including here) come in pairs, one member of the pair giving a measure of bias in the predictions - do they represent consistent over/under-prediction of the observed values ? - while the other member gives a measure of the spread in the predictions - is there wide scatter in the predicted values compared with the observed values ?

Table 2.5 shows the statistical comparison parameters chosen in SMEDIS. In the table the angle brackets <…> denote an average over N observed/predicted pairs, while the notation $N_{a<\zeta<b}$ denotes the number of members of the set of N values of ζ which lie between a and b. The main parameters are the Mean Relative Bias, MRB, and the Mean Relative Square Error, MRSE, together with the Factor of n with n = 2 and 5. The Geometric Mean and Geometric Variance, MG and VG, respectively, have also been included for comparison with previous studies, although they are not the primary comparison parameters in SMEDIS. All the parameters used are based on concentration ratios, and the MRB/MRSE pair has preferred status on the basis of previous work (Duijm *et al.*, 1996).

APPROACH	RELEASE TYPE	
	Continuous	**Instantaneous**
pointwise comparisons	a) Time-averaged concentration at sensor position ***x***	a) Dose *D* at sensor position ***x*** b) Cloud arrival time at ***x*** c) Cloud departure time at ***x***
arcwise comparisons	b) Maximum concentration across arc at a given radius *x* c) Width of cloud across arc at (*x*, *z*)	d) Maximum dose across arc at a given radius *x* and height *z* e) Maximum concentration across arc at (*x*, *z*) f) Time for maximum concentration across arc at (*x*, *z*) g) Cloud width across arc at (*x*, *z*) based on dose h) Cloud arrival time across arc at (*x*, *z*) i) Cloud departure time across arc at (*x*, *z*

Table 2.4 Physical comparison parameters. A precise definition of each parameter is given in the protocol.

Name	Definition	Advantages	Disadvantages
Mean Relative Bias	$\mathrm{MRB} = \left\langle \frac{\psi_o - \psi_p}{\frac{1}{2}(\psi_p + \psi_o)} \right\rangle$	• Accepts zero values. • Less sensitive than other measures to minimum thresholds. • Symmetric for under/over- prediction.	• Allows differences between models with ψ_o/ψ_p up to ~10 to become apparent, but not so outside this range.
Mean Relative Square Error	$MRSE = \left\langle \frac{(\psi_p - \psi_o)^2}{\frac{1}{4}(\psi_p + \psi_o)^2} \right\rangle$	• More transparent than VG (see below) in allowing standard deviation of predictions to be obtained.	
Factor of n	$\mathrm{FACn} = \frac{\mathrm{N}_{1/n<\psi_o/\psi_p<n}}{\mathrm{N}}$	• Robust, consistent, easy to understand	
Geometric mean bias	$MG = \exp\left\langle \log_e\left(\frac{\psi_o}{\psi_p}\right) \right\rangle$	• Log_e(MG) symmetric about zero in under/over-prediction.	• Cannot accept zero values.
Geometric mean variance	$VG = \exp\left\langle \left[\log_e\left(\frac{\psi_o}{\psi_p}\right)\right]^2 \right\rangle$	• Variance measure related to MG	• Cannot accept zero values.

Table 2.5 Statistical comparison parameters used in SMEDIS. The MRB and MRSE are the principal parameters used together with FA2 and FA5. MG and VG are included for comparison with previous validation exercises which have used these.

Thresholds have been applied to the parameters being compared to allow for cases where either the predicted or observed concentration is zero. In practice, this only applies to the pointwise concentration or dose comparisons, since all other values are non-zero, and is only required by MG and VG unless both observed and predicted values are zero. The effect of thresholds is still being actively assessed, although a working value of 10^{-3} units for concentration (and for dose $10^{-3} t_d$, where t_d is the duration of the test) is being used initially. (It would be preferable to base thresholds on sensor sensitivity, but such information is not available in all cases.)

(c) Other aspects of validation procedure

In general the model proponent is expected to carry out as many of the allocated runs as possible (see Table 2.3). They are expected to document the setting up of their model, especially if there is a deviation from the "normal" method, such as a change to the recommended values of parameters.

In many cases models will not have the capability to handle specific features, in particular aerosol clouds, sloping ground or obstacles. In these circumstances it was agreed that users on a voluntary basis could attempt to simulate all cases assigned to their model type even if their model was missing the necessary capability. Conclusions would be drawn from the results on the effect of using a model beyond its limits of applicability (when validation results are shown, any such mismatch will be clearly indicated).

SMEDIS also allows an element of model development/improvement during the project, since in many cases the model proponent is the same as the model developer. For this reason the validation exercise provides users with the observations (i.e. the "answers") to assist the improvement process. However, the users are required to document any changes made to the model and provide this as additional input to the scientific assessment of their model. They should also carry out all runs with the same version of the model.

2.5 Distribution of data sets

The data for each case have been circulated in the form of Excel workbooks : there is one workbook per case, containing the data describing the test set-up (to be used in setting up models) together with the concentration measurements (and others, for example temperature if available) against which the model predictions are to be compared. A common format is adopted for each case. The data have been derived from the REDIPHEM database (Nielsen and Ott, 1996) wherever possible. A second workbook is provided for users to return their results.

In the case of wind tunnel experiments, the data workbook contains equivalent "full-scale" data as well, representing (where relevant) the specific full-scale analogue associated with the test. It was agreed that models would be run to simulate this scaled-up version, on the basis that models (other than the general purpose CFD models) were developed to simulate such problems.

The SMEDIS database of test cases will be made available to all DGD model developers and users for future use.

2.6 Implementation of validation procedure

Once a given model has been run for its allotted cases, the results workbooks are sent electronically to the two partners responsible for processing the results (EDF and HSL) where the statistical parameters are calculated for all cases and models. The results of the validation exercise are also added to complete the "model evaluation report" resulting from the scientific assessment.

3. Results

The entire exercise has led to the collection of over 300 model results returned (each containing up to 130 data points).

A preliminary analysis has been performed on a limited set of these results (roughly one-third). The following discussion is based on this preliminary analysis.

Although all statistical parameters have been computed for all models and all datasets, we have formed groups to summarize the results. First we have formed groups of models for the four model categories (see Table 1) and then groups of datasets comprising similar complex effects (including none). These groups are presented in Table 3.1, indicating the number of results processed in each category[1].

	Type_effect			
Model_Types	1-no effect	2-obstacle	3-aerosols	4-terrain
1-WorkBook	2	2	4	0
2- Integral	18	17	42	9
3- Shallow-Layer	4	5	8	4
4- CFD-3D	0	7	7	6

Table 3.1 : Number of model results processed in each category.

		Type_effect			
Statistical Parameter	Model_Types	1-no effect	2-obstacle	3-aerosols	4-terrain
ln(MG)	1-WorkBook	-0,88	-0,73	1,06	
	2- Integral	-0,01	0,01	0,18	1,07
	3- Shallow-Layer	-0,49	0,75	0,35	1,79
	4- CFD-3D		-0,01	-0,27	0,50
ln(VG)	1-WorkBook	1,46	0,82	10,20	
	2- Integral	0,39	0,52	2,34	1,74
	3- Shallow-Layer	1,04	1,04	1,51	7,32
	4- CFD-3D		0,14	0,35	0,40
MRB	1-WorkBook	-0,72	-0,66	0,14	
	2- Integral	-0,00	0,01	0,03	0,88
	3- Shallow-Layer	-0,39	0,65	0,31	0,94
	4- CFD-3D		-0,01	-0,25	0,47
MRSE	1-WorkBook	0,98	0,65	1,39	
	2- Integral	0,33	0,37	0,61	1,08
	3- Shallow-Layer	0,66	0,74	1,01	1,31
	4- CFD-3D		0,13	0,30	0,35

Table 3.2 : Results of statistical analysis of **arcwise** comparison between model results and experimental values. Model and datasets have been grouped into four categories.

Table 3.2 and Fig. 1 present the results of the arcwise comparison. This part can be compared to previous studies that have similarly used the maximum arcwise comparison. For the "FAC2" statistics (Fig. 1), we can first notice the weakness of integral models with complex terrain and shallow layer model with aerosols. All other results are above the 40% which have been reported elsewhere (for example in Duijm *et al.* 1996, and other papers in this conference). We can also note

[1] Note that the term "Workbook" is used here since there is only one model - the Britter-McQuaid Workbook - included in the analysis of results (see also caption to Table 1).

a general improvement of the FAC2 with increasing complexity of the models (more than 70% within a factor of two for CFD).

For the two measures of bias (ln(MG) and MRB, Table 3.2) there is no clear over- or under-prediction if obstacles or aerosols are present. However with no complex effects, the models over-predict whereas with complex terrain they under-predict. Based on these numbers alone, the best model type is the integral model excepted for terrain effects.

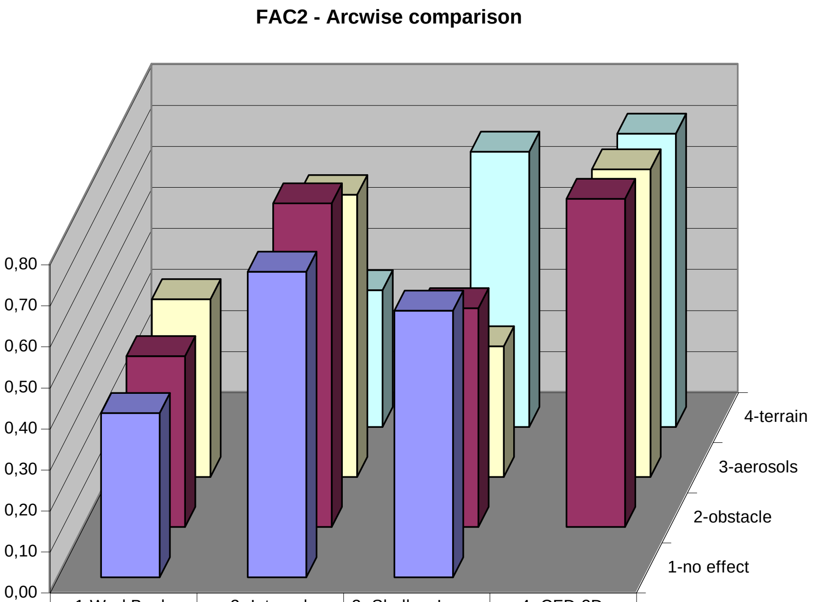
Finally for the two measures of spread (ln(VG) and MRSE) we again see the improvements with model complexity excepted for the shallow layer models.

		Type_effect			
Stat. Parameter	Model_Types	1-no effect	2-obstacle	3-aerosols	4-terrain
ln(MG) =	1-WorkBook	-1,49	0,39	0,54	
	2- Integral	0,45	0,42	0,15	0,87
	3- Shallow-Layer	1,10	1,19	0,77	1,81
	4- CFD-3D		0,90	0,96	0,44
ln(VG)=	1-WorkBook	3,94	8,20	13,85	
	2- Integral	6,43	7,55	9,42	1,56
	3- Shallow-Layer	13,24	6,42	5,48	7,73
	4- CFD-3D		5,53	7,57	0,43
MRB =	1-WorkBook	-0,99	-0,09	0,16	
	2- Integral	0,08	0,11	0,02	0,71
	3- Shallow-Layer	0,16	0,64	0,43	0,92
	4- CFD-3D		0,47	0,37	0,41
MRSE =	1-WorkBook	1,58	1,60	1,91	
	2- Integral	1,40	1,64	1,68	0,97
	3- Shallow-Layer	1,47	1,60	1,45	1,35
	4- CFD-3D		1,08	1,25	0,36

Table 3.3 : Results of statistical analysis of **pointwise** comparison between model results and experimental values. Model and datasets have been grouped into four categories.

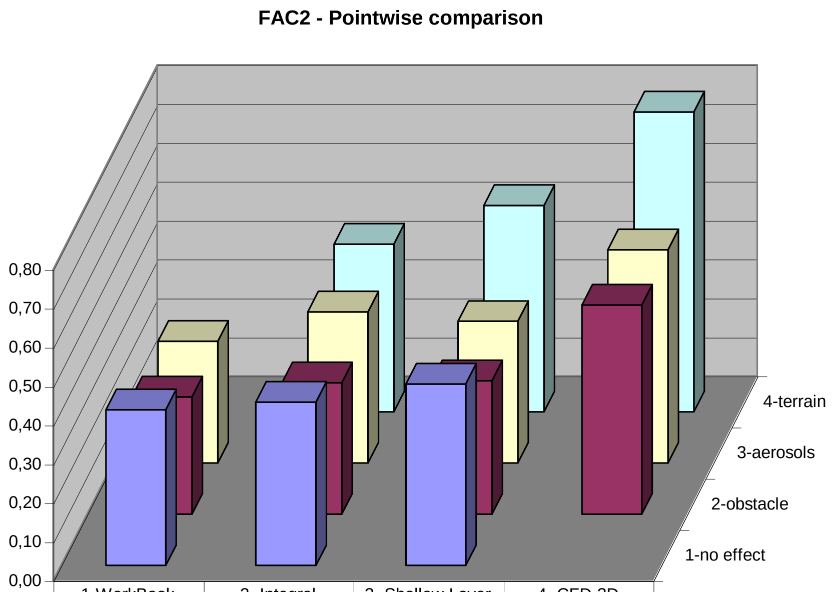
Table 3.3 and Fig. 2 present the statistical results obtained for the pointwise comparison. It is important to note that this comparison involves a much larger number of points as all the measurements on each arc are included and there are some additional measurement point that are sometimes not included in any arc.

We can first notice that the pointwise comparisons show a global decrease in the statistical performance measure compared with the arcwise values, indicating that all models are better at predicting centerline maximum concentration than the general cloud shape. Putting aside this global decrease in performance, all comments made from the arcwise analysis can be carried over to the pointwise statistics.



	1-WorkBook	2- Integral	3- Shallow-Layer	4- CFD-3D
1-no effect	0,40	0,74	0,65	
2-obstacle	0,42	0,79	0,53	0,89
3-aerosols	0,43	0,69	0,32	0,75
4-terrain		0,33	0,67	0,71

Figure 1 : Fraction of model results within a factor of 2 of experimental results ("FAC2") for the arcwise comparison



	1-WorkBook	2- Integral	3- Shallow-Layer	4- CFD-3D
1-no effect	0,40	0,42	0,47	
2-obstacle	0,30	0,34	0,34	0,54
3-aerosols	0,31	0,39	0,36	0,55
4-terrain		0,43	0,53	0,77

Figure 2 : Fraction of model results within a factor of 2 of experimental results ("FAC2") for the pointwise comparison

4. Conclusion

In this paper we have described the database of experimental results and its use for the model validation performed in the SMEDIS project. This validation was carried out according to the general protocol designed to specify all aspects of model evaluation, including validation but also scientific model assessment.

In the course of constructing the database, we found that there are significant gaps in the experimental data available for dense gas release problems with complex effects, in particular if wind tunnel experiments are excluded. Instantaneous releases featuring complex effects are particularly sparsely studied, with very few examples available for aerosol releases. In addition, stable atmospheric conditions are not well-represented (wind tunnels are not able easily to simulate such cases). Situations where near-field effects dominate (congested/confined releases, low/no wind conditions) are an area in need of further experiments, in particular for flammable clouds.

Information on sensor accuracy and data uncertainty is not always available: this information can have an important role in defining “acceptable” agreement with model predictions, and is necessary to define a threshold to be used for certain statistical measures.

In total, over 300 sets of model results have been returned by the participants (with some having over 100 data points). From these we have carried out a preliminary analysis of approximately one third of the total for the results presented here.

The pointwise statistical measures of model performance are globally lower than for the arcwise comparison, indicating that all models are better at predicting centerline maximum concentration than the general cloud shape. Based on the Fraction within a factor of 2 (FAC2), the Geometric Variance (ln(VG)) and the Mean Relative Square Error (MRSE) we clearly see a general improvement of model performance with increasing complexity, that has been quantified by these statistical measures. However it should be remembered that some results are based on smaller samples size as given by Table 3.1.

Acknowledgements

The authors would like to gratefully acknowledge the financial support of the European Commission (Contract ENV4-CT96-0245) and of sponsoring organisations.

References

Brighton P. *et al.*, 1994 : Comparison of heavy gas dispersion models for instantaneous releases. *J. Hazardous Materials*, Vol.36, pp 193-208

Britter, R.E. 1992 : The evaluation of technical models used for major-accident on hazard installations. *EUR 14774.*

Britter, R.E. 1998 : Recent research on the dispersion of hazardous materials. *EUR 18198.*

Cole, S.T. and Wicks, P.J. , 1995 : ‘Model Evaluation Group Seminar - The Evaluation of Models of Heavy Gas Dispersion, Mol, Belgium, November 1994’, EUR 16146.

Daish N., R.E. Britter, P.F. Linden, S. Jagger and B. Carissimo, 1999 : SMEDIS : Scientific Model Evaluation Techniques applied to Dense Gas Dispersion Models in Complex Situations. Paper presented at the CCPS/AIChE Conference on Modelling Hazardous Releases (San Francisco, Sep/Oct 99)

Duijm, N.J., Carissimo, B., Mercer, A., Bartholome, C. and Giesbrecht, H. , 1997 : ‘Development and test of an evaluation protocol for heavy gas dispersion models’, *J. Hazardous Materials*, Vol. 56, pp. 273-285.

Duijm, N.J., Ott, S. and Nielsen, M. 1996 : An evaluation of validation procedures and test parameters for dense gas dispersion models. *J. Loss Prev. Process Ind.* **9**, 323-338.

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

Nielsen, M. and Ott, S. 1996 : A collection of data from dense gas experiments. *RISO-R-845(EN)* http://www.risoe.dk/vea-atu/densegas/rediphem.htm