Available online at www.sciencedirect.com
SCIENCE DIRECT®
Atmospheric Environment 40 (2006) 674–685
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

A dispersion modelling system SILAM and its evaluation against ETEX data

M. Sofiev[a,*], P. Siljamo[a], I. Valkama[a], M. Ilvonen[b], J. Kukkonen[a]

[a]*Finnish Meteorological Institute, Erik Palmenin Aukio 1, Fin-00561, Helsinki, Finland*
[b]*Technical Research Centre of Finland, P.O. Box 1600, FIN-02044, Helsinki, Finland*

Received 17 November 2004; received in revised form 27 September 2005; accepted 28 September 2005

Abstract

This paper presents the SILAM dispersion modelling system that has been developed for solving various forward and inverse dispersion problems. The current operational version is based on a Lagrangian dispersion model that applies an iterative advection algorithm and a Monte Carlo random-walk diffusion representation. The system can utilize meteorological data from either the HIRLAM or ECMWF numerical weather prediction models. We present an evaluation of SILAM against the data of the European Tracer Experiment (ETEX). The model showed an overall time correlation coefficient of 0.6 (over 150 stations), with specific values for the two ETEX measurement arcs of 0.75 and 0.74, respectively. The number of well-reproduced observation sites are 55, 37, and 40—for a Figure of Merit in Time of >0.2, a correlation coefficient of >0.7, and mean observed and modelled values being within a factor of 2, respectively. We have also investigated the sensitivity of the model to the meteorological input data and model setup. The most important factors with regard to the model performance were (i) the selection of the meteorological input data set and (ii) the method used for the atmospheric boundary layer height estimation. The study allowed selection of the optimum setup for the operational model configuration. We also tried to find explanations for the successes and failures of the specific methodologies in order to facilitate broader conclusions on their applicability in emergency dispersion modelling.

Keywords: Dispersion modelling; Model validation; Emergency preparedness

1. Introduction

A growing variety of emergency situations, their complexity, and the limited time available for the decision-making have resulted in an intensifying use of operational dispersion models as the key source of supporting information. Such models are often formulated in Lagrangian terms (Zannetti, 1992), since this mechanism facilitates handling of point emission sources and naturally includes information on the parcel age, which is useful for radioactivity-related computations. Advection then comprises the wind-driven motion of the centers of masses of the air parcels (Lagrangian particles). Turbulent diffusion is simulated by a random relocation of the Lagrangian particles; this diffusion can be parameterized in various ways ranging from a simple well-mixed-layer approach to sophisticated and computationally expensive solutions of the Langevin equations (Thomson, 1987; Rodean, 1996; Stohl

*Corresponding author. Tel.: +358 9 1929 3151; fax: +358 9 9129 3146.
E-mail address: mikhail.sofiev@fmi.fi (M. Sofiev).

1352-2310/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2005.09.069

and Thomson, 1999). The characteristic computed time period for a Lagrangian model does not exceed a couple of weeks, with a spatial coverage of up to a continent. Examples of such models are SNAP (Saltbones et al., 1996), NAME (Maryon et al., 1991), and FLEXPART (Stohl et al., 1998). There are also operational models utilizing Eulerian advection mechanisms—either stand-alone or combined with Lagrangian near-source pre-computation, e.g., MATCH (Robertson et al., 1999) or DREAM (Brandt et al., 2000).

There are several available data sets that can be utilized for the evaluation of emergency models. Local-scale dispersion may be verified by the model validation kit using its own database (Olesen, 1995). Regional-to-continental-scale data sets have been collected after some past accidents, such as the Chernobyl catastrophe, and the Algeciras release (e.g. Pobanz et al., 1999). Probably the best data set was collected during the European tracer experiment ETEX (Graziani et al., 1998; ETEX, 1998). It included two planned releases with controlled conditions and extensive measurement campaigns over most of Europe and succeeded in following the plume evolution during several days after the release (at least in Experiment 1).

The SILAM model (version 3.5) discussed below is formulated in a similar way to mainstream Lagrangian Monte-Carlo random-walk models. The system also includes original solutions, which allow its utilization in source delineation tasks (inverse dispersion problems), in analysis of the dispersion of non-radioactive species and in long-term simulations covering up to several years. The system has been evaluated against several emergency data sets and in international exercises and model inter-comparison projects (Pöllanen et al., 1997; Graziani et al., 2000), ETEX (Graziani et al., 1998), and ENSEMBLE (Galmarini et al., 2004b).

The objective of the current study was to validate v.3.5 of SILAM system against the ETEX-1 data set and to analyze its sensitivity to variations in the input data and the model setup. Since SILAM incorporates several methodologies for meteorological data treatment and dispersion computation, a specific goal was to select the best configuration for operational use and quantify its performance. Finally, we aimed at an evaluation of the methodologies and their combinations in order to highlight their compatibility, strengths, and limitations for regional dispersion simulations.

2. Overview of the SILAM modelling system

Here, we shortly present SILAM and highlight its similarities and differences from other emergency models. A more detailed presentation of the model and its results can be found at http://silam.fmi.fi.

2.1. Basic equations

SILAM can solve both forward and adjoint (often called "inverse") dispersion problems. The forward dispersion equation can be written in the following form:

$$L\varphi = E, \quad L = \frac{\partial}{\partial t} + \frac{\partial}{\partial x_i}(u_i) - \frac{\partial}{\partial x_i}\mu_{ii}\frac{\partial}{\partial x_i} + \xi, \tag{1}$$

where L is a differential operator, φ is the concentration of the pollutant, t is time, E is an emission term, x_i, $i = 1, \ldots, 3$ denote the three spatial axes, u_i are the components of the wind velocity along these axes, μ_{ii} is the turbulent diffusion coefficient in air, and ξ represents all sink processes. The boundary conditions are

$$\begin{aligned} &\varphi(t=0) = 0 \\ &\mu_{33}\frac{\partial\varphi}{\partial x_3}\bigg|_{x_3=h_1} = v_d\varphi(h_1), \quad \frac{\partial\varphi}{\partial x_3}\bigg|_{x_3=H} = 0, \\ &\frac{\partial\varphi}{\partial x_i}\bigg|_{(x_1,x_2)\in\partial\Omega} = 0, \quad i = 1,2, \end{aligned} \tag{2}$$

where the heights h_1 and H are the lower and upper boundaries of the transport domain, Ω is a horizontal computation area with border $\partial\Omega$, and v_d is the dry deposition velocity at height h_1.

The corresponding adjoint dispersion equation can be written in the following form (e.g., Marchuk, 1982):

$$\begin{aligned} &L^*\varphi^* = S, \\ &L^* = -\frac{\partial}{\partial t} - \frac{\partial}{\partial x_i}(u_i) - \frac{\partial}{\partial x_i}\mu_{ii}\frac{\partial}{\partial x_i} + \xi, \\ &\varphi^*(t=t_{end}) = 0, \quad \mu_{33}\frac{\partial\varphi^*}{\partial x_3}\bigg|_{x_3=h_1} = v_d\varphi^*(h_1), \\ &\frac{\partial\varphi^*}{\partial x_3}\bigg|_{x_3=H} = 0, \quad \frac{\partial\varphi^*}{\partial x_i}\bigg|_{(x_1,x_2)\in\partial\Omega} = 0. \end{aligned} \tag{3}$$

Here t_{end} is the end time of the simulations; φ^* is the sensitivity distribution that arises from the sensitivity source function S. The specific form of S depends on the problem. For example, it is proportional to the population density for an assessment of population exposure and comprises

a sum of Dirac δ-functions for model-measurement comparison or a map of concentration-to-exposure conversion factors in the general case. The adjoint equation is directed backwards in time, so the sensitivity distribution φ^* starts from the receptor with sensitivity S at $t = t_{\text{end}}$ and covers the areas where the sources E affect the receptor.

The SILAM system is capable of solving both dispersion problems—the forward equations (1) and (2) with unknown φ and the adjoint equation (3) with unknown φ^* (Sofiev and Atlaskin, 2004). Following the standard way of solving these equations, SILAM involves time- and process-based first-order splits. The second-order Strang split brings about only a minor improvement in the final results at the price of a significant loss of efficiency.

2.2. Input data

In contrast to many other models, SILAM can directly utilize the meteorological data from several NWP models and dynamically adjust pre-processing routines in accordance with the availability and completeness of the input variables. Data sets used operationally originate from the HIRLAM (Unden, 2002) and European Centre on Medium-Range Weather Forecast (ECMWF, http://www.ecmwf.int) models. A particularly difficult task for the meteodata pre-processing—estimating the boundary-layer height (Maryon and Buckland, 1994; Seibert et al., 2000; Sørensen, 1998)—is approached via four different methods: (i) the dry parcel method (Troen and Mahrt, 1986), (ii) the critical Richardson number method (e.g. Nieuwstadt and Van Dop, 1982, Chapter 1), (iii) a constant ABL, and (iv) the ABL height from the NWP input (if available). The user can select the method for the specific run or allow the model to choose automatically between (i) and (ii), whatever estimate is higher.

The emission term in SILAM v.3.5 consists of a list of sources, which can be points, areas or a nuclear explosion(s). In the latter case, the mushroom-shaped cloud is parameterized following Directions (1994).

2.3. Transport and deposition computations

The SILAM dynamics computation is based on a precise (and resource-consuming) iterative Lagrangian advection scheme (representing the second terms in the right-hand side of the second equation in Eqs. (1) and (3)) based on Eerola (1990), combined with a 4D interpolation of all input data to the actual positions of the Lagrangian particles. The random-walk mechanism (the third terms on the right-hand side of the second equation in Eqs. (1) and (3)) is less resource-consuming and based on a well-mixed ABL assumption after the SNAP model (Saltbones et al., 1996), which is a simple but well-working variation of the methodology of Thomson (1987) and Rodean (1996).

The sink terms in Eqs. (1) and (3) include dry and wet deposition. Dry deposition is computed using the standard resistance analogy (Hicks et al., 1987; Lindfors et al., 1993), with the gravitational settling velocity being found from Stoke's law. For all aerosols, the surface resistance $r_s = 0$, and for coarse particles the diffusive velocity is set to 0 due to the laminar resistance $r_b \to \infty$. Scavenging with precipitation is also computed in a standard manner using the scavenging coefficient Λ that depends on the precipitation rate R: $\Lambda = \eta R^{\nu}$. The values ν and η are specific for the types of precipitation and scavenging (in-cloud or sub-cloud). Numerical values are primarily based on the Chernobyl accident data (Horn et al., 1987; Smith and Clark, 1989; Jylhä, 1991).

2.4. Computed species and output parameters

Currently, the list of species handled by the SILAM system includes radioactive nuclides, size-segregated aerosols, natural allergenic species, and risk probabilities. The aerosol size spectrum can be described via both the bin and modal approaches.

Compared to other emergency models, the SILAM radiation dose assessment module contains an extensive database of 496 nuclides, 80 dose pathways (dependent on the source and receptor locations and type of radiation), and 23 human target organs. It allows computation of the decay chains, environmental removal (migration) of nuclides after deposition, and external and internal exposure doses. External dose pathways include direct irradiation from cloud and fallout (gamma and beta); internal pathways include inhaled radionuclides. The dose rate conversion factors for the external exposure of human organs to photons and electrons are based on Kocher (1980, 1983), while the factors for intakes of nuclides are based on NRPB data (Phipps et al., 1991a) and those of the Oak Ridge National Laboratory (RSIC, 1981, 1988, 1989). The radioactive decay chains are treated via a three-nuclide scheme.

To our knowledge, SILAM is the only operational model in Europe that computes probabilities of the territories to be affected by releases of unknown or unlisted materials. For such computations, the cloud of Lagrangian particles is considered an ensemble of volumes of air stochastically picked from the plume and dispersed following either the forward (1) or adjoint (3) equations. The probability densities for the affected area or contaminated air volume are obtained as integrals over the ensemble providing the 3D volume- and 2D area-of-risk fields, respectively.

3. Evaluation of the SILAM model against the ETEX-1 data

The ETEX experiment 1 (Graziani et al., 1998; http://rem.jrc.cec.eu.int/etex/) was conducted during the period 23–26 October 1994. The release of a passive non-depositing tracer was started at $t_0 = 16:00$ on 23 October 1994 from a point source in Western France and lasted for 11 h and 50 min. The resulting plume was followed over Europe by 168 stations that provided three-hourly mean tracer concentrations near the surface for 3 days. A detailed description of the meteorological situation during the whole experiment can be found in ETEX (1998). The high temporal resolution of the observations and the large number of stations allowed the selection of two arcs of affected sites, one for the beginning of the development of the tracer cloud and the other for the second day of the experiment (Fig. 1).

For comparison of SILAM results with the above data, we used both time- and space-related statistical measures: (i) the temporal correlation coefficient and figures of merit as measures of the plume time evolution; (ii) the absolute and relative deviations that reflect the size of the plume and distribution of mass within it; and (iii) the root-mean-square error as an overall characteristic of the model performance. The maximum and cumulative concentrations, and the plume arrival/departure times, though presented below, are of lesser value for the model evaluation because of unstable nature of these parameters (Mosca et al., 1998; ETEX, 1998).

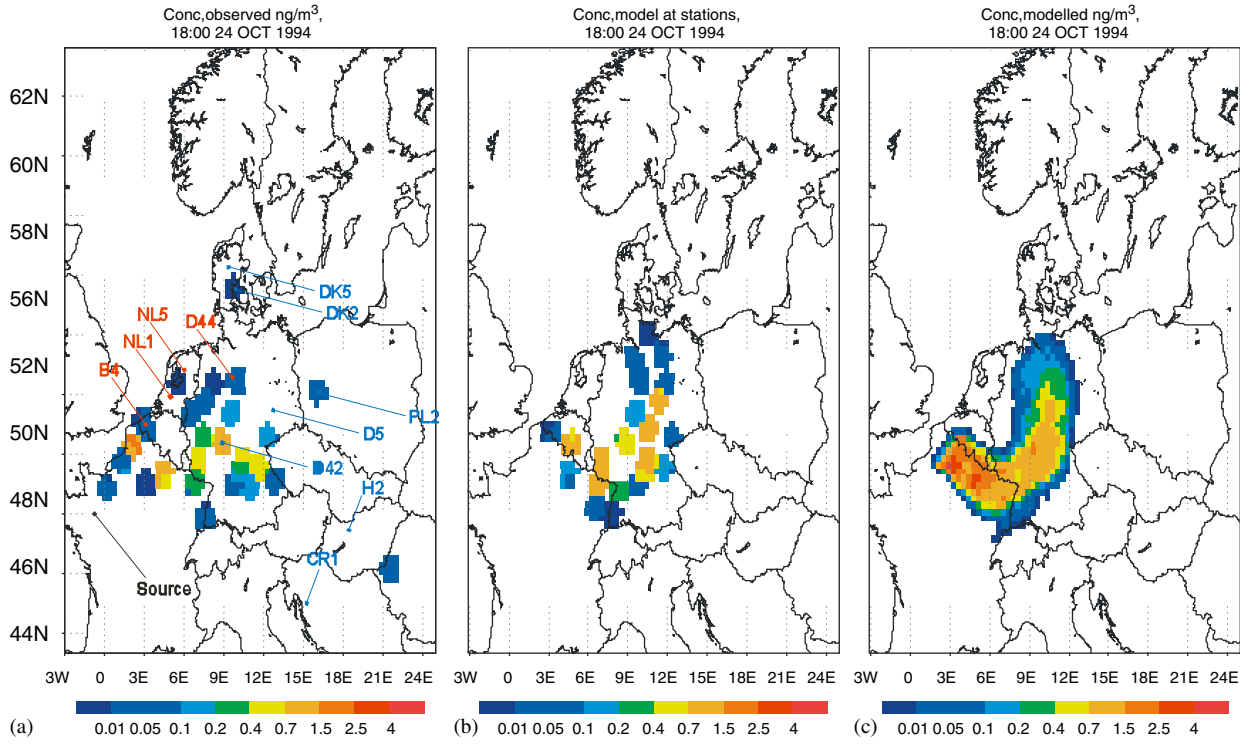

Fig. 1. Concentrations (ng m^{-3}) averaged over the period from $T_0 + 23$ to T_0+26 h after the release at time T_0. The left-hand panel shows the observations, the middle panel shows the model data projected to the station locations, and the right-hand panel shows the model fields. *Note*: sites in ETEX arcs 1 and 2 are marked by red and blue pointers, respectively.

3.1. Model configurations for the sensitivity run

A list of the model parameters varied during the study and the input data are presented in Table 1. In total, six meteorological data sets were used for the runs (Table 1). We utilized the data of the Finnish Meteorological Institute (FMI) HIRLAM model (two versions) and the ECMWF model T213. The older HIRLAM 2.6 and ECMWF T213 were operational in 1994, while the new HIRLAM v.5.2.1 (operational in 2004) was re-run for the ETEX period. The temporal resolution of the HIRLAM 2.6 and ECMWF T213 models was 6 h, while the HIRLAM 5.2.1 produced forecast fields every hour. To make the three data sets temporarily comparable, an additional data set was created from HIRLAM 5.2.1 by taking each sixth hourly forecast field. In order to study the impact of time resolution and spin-up of the meteorological model after the data assimilation at 00:00, 06:00, 12:00, and 18:00 UTC, an intermediate (third) data set was created from every third hourly output of the HIRLAM 5.2.1 model.

The accuracy of HIRLAM in terrestrial regions can be considered good, but discrepancies from the surface observations can occur in cases of stable stratification (Savijärvi and Kauhanen, 2001). Pirazzini et al. (2002) also found that the HIRLAM model tends to underpredict inversions in coastal and sea areas. The latter problem was seemingly present in this study too (see the ABL height discussion below).

The default fully mixed boundary-layer random-walk scheme was used in almost all runs, while the fixed-length approach was applied once with a constant ABL height (the simplest case in which the vertical profiles and diffusion are independent of the actual conditions).

The computation results were analyzed twice: the near-surface concentration was computed by counting the Lagrangian particles from the surface up to 150 and 550 m. Such an averaging addresses the known problem of spatial non-uniformity of the output fields. This problem originates from the limited spatial representativeness of a single Lagrangian particle and exists in all Lagrangian models. However, its impact on the final results can be reduced by spatial averaging or other means, such as a Lagrangian puff approach, various

Table 1
Meteorological data sets used as input for the SILAM model

Meteorological input data					
Abbreviation	Model	Grid resolution	Vertical levels	Time step	Forecast length
H5_fc0	FMI HIRLAM 5.2.1	0.25°	31	6 h	0 h (analysis)
H5_fc1	FMI HIRLAM 5.2.1	0.25°	31	1 h	From 1 to 6 h
H5_fc3	FMI HIRLAM 5.2.1	0.25°	31	3 h	3 and 6 h
H5_fc6	FMI HIRLAM 5.2.1	0.25°	31	6 h	6 h
H2	FMI-HIRLAM 2.6	0.50°	31	6 h	6 h
EC	ECMWF(T213)	0.50° (MARS interpolation)	31	6 h	6 or 12 h

ABL height assessments	
parcel	Dry parcel method of the ABL height assessment
Ri	Critical Richardson number method
comb	A combination of parcel and Ri methods (the higher value is taken)
const	A constant ABL set at 850 hPa
hirABL	HIRLAM's own ABL height value (HIRLAM 5 only, computed from the profile of the turbulence kinetic energy)

Random-walk type	
—	Default well-mixed ABL along the vertical, fixed-length horizontal
simple	A combination of the const ABL and fixed-length random-walk displacement

Model time step	
dt5/dt15/dt30	5/15/30 min time step

Vertical averaging of the model output	
150/550	Vertical averaging from the surface up to 150 m/550 m

Note: MARS is an automatic extraction and interpolation system at ECMWF that provides user-requested features of the meteorological fields. Native resolution of T213 fields is ~75 km.

horizontal and/or vertical averaging kernels, etc. In this study, the horizontal resolution was always kept at approximately 25 km, while the sensitivity to vertical averaging was investigated.

In the following sections, each run is referred to by its parameters in accordance with the above abbreviations in Table 1. For example, the notation H5_hirABL_fc3_dt5_150 denotes a run with HIRLAM 5 input data, native HIRLAM ABL height, a 3-h forecast time step, a 5-min SILAM internal time step, and vertical averaging over the lowest 150 m.

3.2. Results of model evaluation

The statistics were computed for three comparison data sets: time statistics for the two arcs, and space and time statistics for all stations that reported more than five observations (150 sites). For qualitative analysis, the observed and simulated shapes of the tracer cloud at two different times are shown in Figs. 1(a)–(c) and Figs. 2(a)–(c). The middle panels in Figs. 1 and 2 are constructed from the model fields (obtained from the operational model setup described further) projected to the station locations and then drawn with the same correlation radius as the observations in the left-hand panels. The middle panels are therefore directly comparable with both observations (left-hand panels a) and the model fields (right-hand panels c).

A consideration of the patterns leads to three general conclusions. Firstly, the model patterns are close to those obtained from the measurements. Both the length of the cloud and its shape are reproduced correctly, as well as the orientation in an east–west direction after 1 day and in north–south direction with two peaks of concentration showing up after 2 days of dispersion. Secondly, the simulated cloud is somewhat more condensed than the observed distribution. The model tends to miss the low-concentration parts of the cloud, collecting the mass into correctly placed but over-emphasized peaks. Thirdly, the model plumes in the right-hand panel do not look exactly like those in the middle, despite the fact that they are both built from the same model data. For example, the peaks at +1 day as well as the northern maxima and a high-concentration band to the southern peak along the

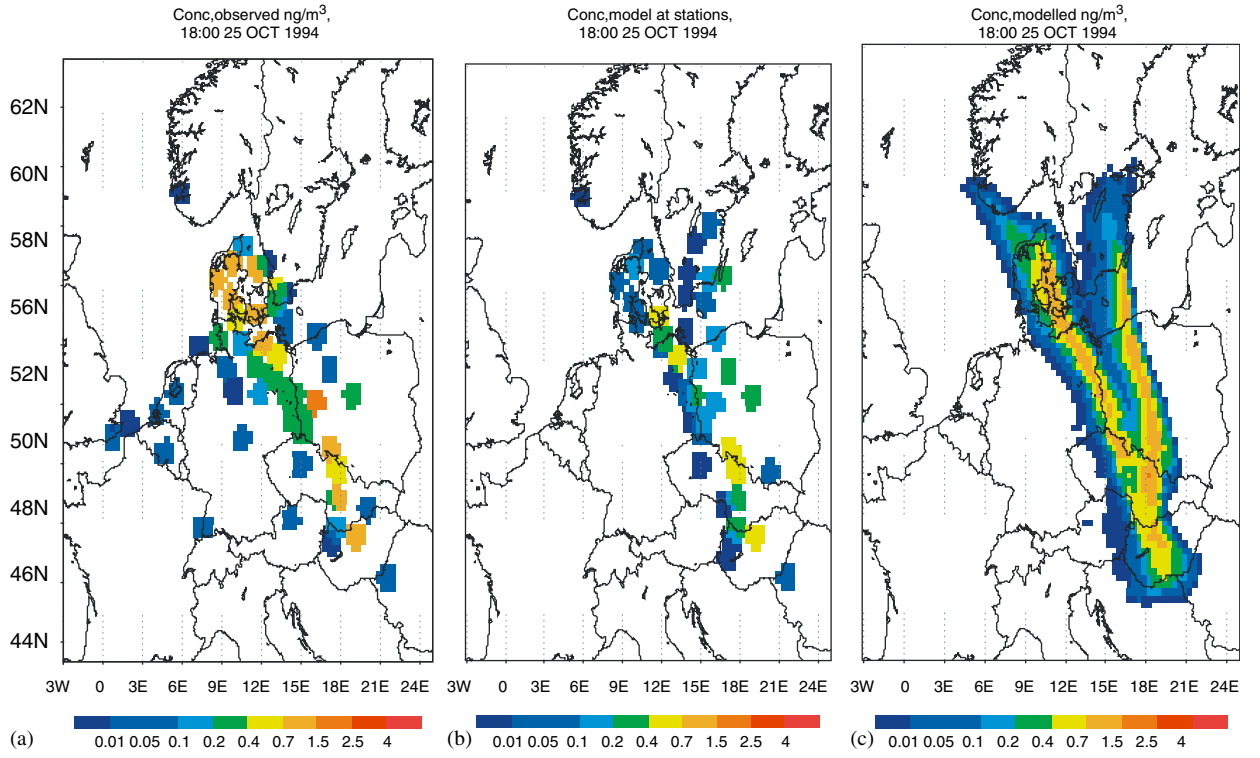

Fig. 2. Concentrations (ng m^{-3}) averaged over the period from T_0+47 to T_0 + 50 h after the release at time T_0. The left-hand panel shows the observations, the middle panel shows the model data projected to the station locations, and the right-hand panel shows the model fields.

eastern border of Germany at +2 days are all but lost. As a result, if the middle-panel patterns of Figs. 1 and 2 were to be interpolated to continuous fields, they would correlate quite poorly with the original model patterns. This shows a limitation of the ETEX data set: there is still insufficient density in the station network, which resulted in a partial loss of the details of the plume development.

The quantitative results of the SILAM verification are presented in Tables 2 and 3 and Fig. 3. The

Table 2
Time-related statistics for the H5_hirABL_fc3_dt5_550 run for the ETEX arcs (see Fig. 1 for locations)

Stations ETEX (1998)	$\bar{C}$ (ng m^{-3})	$\bar{C}-\bar{O}$ (ng m^{-3})	R_t	RelD	FMT	RMSET (ng m^{-3})	dT_{arr} (h)	dT_{dep} (h)	$(C_{cum}-O_{cum})/O_{cum}$	$(C_{max}-O_{max})/O_{max}$
B5	0.06	−0.04	0.87	−0.05	0.44	0.10	−3	−3	−0.56	−0.28
D44	0.14	0.10	0.88	0.04	0.52	0.29	0	−3	0.67	1.44
NL1	0.22	−0.16	0.36	0.16	0.12	0.50	−6	−6	−0.72	−0.72
NL5	0.218	−0.217	0.88	−0.56	0.0001	0.56	12	−15	−1.00	−1.00
Av.Arch 1	0.16	−0.08	0.75	−0.10	0.27	0.36	0.75	−6.75	−0.40	−0.14
CR3	0.11	0.12	0.88	−0.07	0.42	0.32	0	−6	1.06	1.37
D5	0.27	−0.13	0.63	−0.01	0.34	0.29	−3	−3	−0.48	−0.29
D42	0.44	−0.33	0.68	−0.20	0.23	0.61	−3	−6	−0.75	−0.50
DK2	0.34	−0.21	0.76	−0.04	0.38	0.34	−9	0	−0.61	−0.13
DK5	0.39	−0.36	0.72	−0.43	0.06	0.67	6	0	−0.94	−0.90
H2	0.11	0.05	0.91	−0.11	0.50	0.16	3	0	0.43	0.55
PL3	0.22	0.11	0.58	0.28	0.41	0.36	−3	0	0.50	0.68
Av.Arch 2	0.27	−0.11	0.74	−0.08	0.34	0.39	−1.29	−2.14	−0.11	0.11
Average	0.23	−0.10	0.74	−0.09	0.31	0.38	−0.55	−3.82	−0.22	0.02

Notations: $\bar{O}$, $\bar{C}$—mean observed and computed concentrations; R_t—time correlation coefficient; RelD—relative deviation; FMT—figure of merit in time [0, 1]; RMSET—root-mean-square error in time, dT_{arr}, dT_{dep}—time difference of modelled and observed plumes arrival and departure; subscripts "cum" and "max" refer to cumulative and maximum concentrations.

Table 3
Examples of time-related verification statistics for all sites with $N>5$ valid observations

Model run	$\bar{C}-\bar{O}$ (ng m^{-3})	R_t	RelD	FMT	RMSET (ng m^{-3})	dT_{arr} (h)	dT_{dep} (h)	$(C_{cum}-O_{cum})/O_{cum}$	$(C_{max}-O_{max})/O_{max}$
EC_comb_fc6_dt5_1 50m	−0.01	0.57	−0.11	0.37	0.33	4.4	−4.1	0.11	0.79
H2_comb_fc6_dt5_1 50m	−0.03	0.60	−0.11	0.39	0.28	4.4	−4.4	−0.12	−0.32
H5_HirAbl_fc0_dt5_1 50m	−0.04	0.63	−0.15	0.40	0.28	4.6	−3.4	−0.24	−0.17
H5_HirAbl_fc 1 dt5 150m	−0.05	0.54	0.01	0.33	0.26	1.2	−3.5	−0.27	−0.43
H5_HirAbl_fc3 dt5 150m	**−0.05**	**0.55**	**0.00**	**0.33**	**0.26**	**1.3**	**−3.3**	**−0.30**	**−0.51**
H5_HirAbl_fc6_dt5_1 50m	−0.05	0.59	−0.05	0.37	0.25	2.7	−3.6	−0.31	−0.51
H5_Ri_fc0_dt5_1 50m	0.06	0.59	−0.14	0.38	0.47	6.2	−2.3	0.65	0.90
H5_const fcO dt5 150m	−0.04	0.48	−0.10	0.31	0.32	3.2	−4.1	−0.25	−0.43
H5_comb_fc0_dt5_1 50m	−0.04	0.62	−0.16	0.39	0.29	5.6	−3.0	−0.20	0.25
H5_HirAbl_fc0_dt5_adv_2d_1 50m	−0.07	0.63	−0.20	0.39	0.26	4.8	−3.8	−0.50	−0.27
H5_HirAbl_fc0_dt15_150m	0.05	0.63	−0.06	0.39	0.40	4.5	−2.6	0.68	0.41
EC_comb_fc6_dt5_5 50m	0.04	0.57	−0.06	0.35	0.41	3.2	−4.9	0.61	1.00
H2_comb fc6 dt5 550m	−0.01	0.60	−0.08	0.39	0.28	4.1	−4.4	0.02	0.23
H5_HirAbl_fc0_dt5_5 50m	−0.01	0.63	−0.10	0.40	0.31	4.3	−3.3	0.12	0.32
H5_HirAbl fc3_dt5 550m	**−0.02**	**0.54**	**0.07**	**0.34**	**0.27**	**1.2**	**−2.9**	**−0.05**	**0.05**
H5_HirAbl fc6 dt5 550m	−0.01	0.56	0.00	0.36	0.28	2.2	−3.7	0.00	0.20

Note: The observed mean over all stations for all runs is $O = 0.12$ ng m^{-3}. Notations are the same as in Table 2. Bold lines correspond to an operational SILAM setup.

interpretation of the various statistical measures can be found in Mosca et al. (1998) and Sofiev (1999), and additional numerical results of SILAM are presented at http://silam.fmi.fi. Table 2 shows the performance of the best setup (that was selected after this study for the operational runs) for the two ETEX arcs of stations. Table 3 shows the time statistics averaged over all ETEX sites and also includes some sensitivity runs discussed in Section 4. Fig. 3 presents the evolution of the Figure of Merit in Space (FMS) for the three SILAM operational setups (corresponding to the three meteorological drivers) in comparison with the other models that are taking part in the on-going ENSEMBLE–ETEX evaluation (Galmarini et al., 2004a).

Based on the quantitative results presented in ETEX (1998), Graziani et al. (1998), Brandt et al. (2000) and Galmarini et al. (2004a), the performance of SILAM against the ETEX data can be considered to be amongst the best models participating in the exercise. As Fig. 3 shows, the model performance depends on the input data and setup. However, even for the operational SILAM setups with 3- and 6-h forecast fields (which were slightly worse than the runs forced by the meteorological analysis), the mean correlation coefficient for the arcs ($R_t = -0.74$) is one of the highest among the models. The same is true for FMS (Fig. 3): the model is in the top-5 list for most of the time, and deteriorates only close to the end of the 60-h period; this deterioration is caused by the above-mentioned

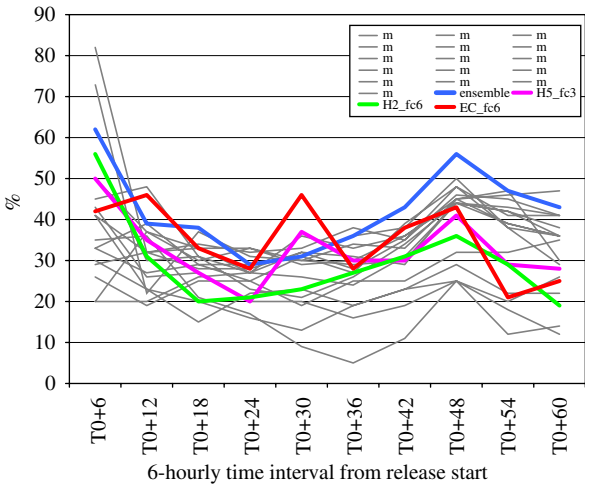

Fig. 3. Time evolution of the FMS of SILAM ("H2_fc6", "H5_fc3", "EC_fc6") compared with a number of other models ("m") that participated in the ENSEMBLE project and with the median-based ensemble fields ("ensemble").

too narrow plume. The number of sites with FMT > 0.2, $R_t > 0.7$, and a simulated mean within a factor of 2 of the observed one (the high-quality thresholds given by Graziani et al., 1998) are 55, 37, and 40, respectively.

4. SILAM sensitivity to input data and model setup

Overall, a few hundreds of runs were performed for most of the reasonable combinations of the methods of Table 1. The scores for some of them are summarized in Table 3. The model demonstrated a stable performance in most of the setups—there were no drastic losses of accuracy, except for a few a priori poor ones (e.g., with constant ABL). Thus, variations in the FMT and R_t for the best 10 setups are within 0.1.

The key question for this study was the *quality of the advection scheme*, its price–performance ratio, and compatibility with the random-walk mechanism. The scheme is resource-consuming due to its iterative algorithm and the 4D interpolation of all parameters to the locations of the Lagrangian particles. However, these costs appeared justified, as follows from the stable and high correlation coefficient. It does not deteriorate in case of longer time steps, i.e., to 15 (_dt15_ case in Table 4) and even 30 min (not shown). The higher mean value for the longer time steps shows only that the simple well-mixed random-walk approach is sensitive to the model time step.

We found such a combination of accurate advection and simple random-walk mechanisms to be reasonable because, for emergency simulations, the most important task is to describe the location of the plume, while the absolute value of concentration is of secondary importance. Indeed, in a real emergency case, the concentration will in any case be high, and its absolute value will be determined by

Table 4
Mean values and standard deviations for the main statistics for H5c_hirABL_fc3hr_dt5_150

Statistics	Mean for 93 non-zero sites	Standard deviation of the statistics
R_t	0.59	0.014
$\bar{C} - \bar{O}$	$-0.07\,\mathrm{ng\,m^{-3}}$	$0.008\,\mathrm{ng\,m^{-3}}$
RelD	−0.06	0.070
dT_{arr}; dT_{dep}	1 h; −3 h	7 h; 7 h

Values are computed only for sites with non-zero values in both observed and simulated data sets (93 sites in total).

the poorly known parameters of the release, such as the emission rate, height and composition. In this light, the accuracy and robustness of the advection scheme is clearly the most important feature of the model.

Another test of the advection and random-walk combination was a reduction in the advection to two dimensions, thus ignoring the vertical wind. Again, the cloud shape seems to have not much deteriorated (the correlation stays the same), due to the weak wind shear along the vertical in the case studied. However, the mean modelled concentration appeared to be the smallest of all the runs (by almost a factor of 3 from the 3D advection run). This confirmed that the influence of the vertical wind should still be taken into account.

Extended *vertical averaging*—up to 550 m—in almost all cases led to an increase in the mean values, being insignificant otherwise. From a general point of view, this parameter can be recommended to be as large as possible while not exceeding the ABL depth.

An analysis of the performance of the *different meteorological drivers* leads to somewhat surprising conclusions. The runs driven by the old HIRLAM 2 and ECMWF T213 data showed nearly the same scores as the results based on the new HIRLAM 5 model. The ECMWF runs tend to slightly lose the time correlation, but show nearly the best FMS (Fig. 3) and higher absolute concentrations. Such differences can be attributed to the coarse resolution of the ECMWF fields and the errors originating from their interpolation from 75 to 50 km by the ECMWF MARS system during data downloading. This also caused an extra horizontal dilution of the cloud (compare the ~50 always-zero sites for ECMWF runs with the 60–70 for the HIRLAM runs), which improved coverage of the stations located near the plume edge.

Variation of the *time resolution of the forecast fields* showed that in practically all cases, the runs driven by meteorological fields with a better temporal resolution (1 or 3 h) were somewhat worse than those with a 6-h meteorological time step. The reason is the considerable spin-up process of HIRLAM after data assimilation. A reasonable trade-off between the necessity to follow rapid meteorological developments, time constraints, and the spin-up problem is probably to use the 3-h forecast.

The *ABL height assessments* significantly affected the results. The differences between input data and algorithms themselves are illustrated in Fig. 4. Panel (a) represents the heights computed from the ECMWF and HIRLAM 2 input data, while panel (b) is based on HIRLAM 5. Data in Table 3 show that the "simple" approach of a constant ABL height causes a significant reduction in the correlation coefficient and FMT compared to the parcel, combination or HIRLAM ABL approaches. The Richardson number method also performed poorly, which is not surprising, because the very existence of a critical Richardson number is doubtful in many cases (Seibert et al., 2000).

The constant ABL (shown in panel (a)) is too high, especially for nighttime conditions, thus

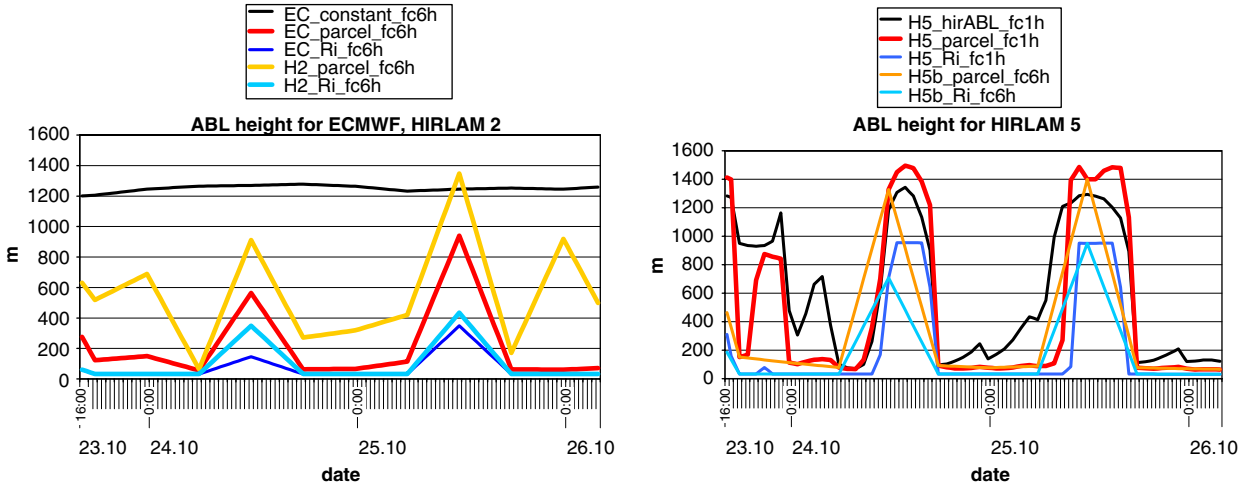

Fig. 4. Examples of ABL height estimates at the source location (2°00′30″W, 48°03′30″N) made by different methods and using different meteorological data sets.

damaging the time correlation and FMT. Admittedly, this result is in some disagreement with ETEX (1998), where several models reported better results with a cruder ABL assessment. The explanation might lie in the high accuracy of the SILAM advection routine, which is sensitive enough to use the advantages of an accurate ABL height and detailed wind profiles along the vertical.

An extra complexity was introduced by the deep ABL during the night of 23–24 October as claimed by HIRLAM 5. However, according to the sounding data (http://rem.jrc.cec.eu.int/etex/), the nighttime ABL was not deep, so that the estimates based on the ECMWF or HIRLAM 2 data are closer to the observations. This could be the point at which HIRLAM 5 lost its "natural" advantage of high resolution.

It is also seen from Fig. 4 that the 6-h meteorological time step is inadequate for following the dynamic weather developments. Indeed, whether the deep ABL was real or not, the 6-h data sets would anyway miss it. In the current case, this lack of resolution helped to avoid a wrong development, thus improving the scores of the 6-h runs, but this is "being right for the wrong reasons".

4.1. Statistical significance and generalization of the results

Since none of the statistics used in ETEX is robust from a mathematical standpoint (see e.g. Huber, 1981), we explicitly assessed the statistical significance of the results (expressed via relative standard deviations of corresponding measures). The outcome is shown in Table 4 for the operational SILAM setup. The standard deviations of the key measures—mean values, absolute deviation and correlation coefficient—are much smaller than the values themselves (the significance of the integrated measures, such as FMT, FMS and RMSE, can be deduced from that of the basic ones) and thus the above results are statistically significant.

Clearly, any set of selected dispersion tracer studies cannot provide a complete assessment of the model performance. However, (i) the results of other evaluation exercises for the SILAM model (listed in Section 1) are in qualitative agreement with the current one and (ii) the scores of most of the SILAM modules tested during the sensitivity study agreed well with theoretical expectations. Several important conclusions of this evaluation will therefore be likely to persist in other cases, and corresponding approaches in other models will produce a similar outcome. Firstly, both the high accuracy of the advection scheme and the sensitivity of the random-walk method to the time step are not dependent on the model and application. Secondly, the combination of critical Richardson number and parcel methods for the ABL height evaluation can be recommended if the ABL height is not available from the NWP.

The vertical distribution of pollution mass could not be evaluated using the near-surface ETEX observations (an experiment with 2D advection was encouraging but insufficient, as well as the three aircraft flights made during the 3-day campaign). Other parts of the model that were not verified in this study are the dry and wet deposition modules. However, they are based on state-of-the-art methodologies, so that an extra validation, although desirable, is not crucial for conclusions regarding the overall model applicability.

5. Conclusions

The SILAM modelling system (version 3.5) was evaluated against the ETEX-1 database. The system contains a wide variety of ways of handling meteorological data as well as representations of the terms in the dispersion equation. This made it possible to evaluate the relative advantages and limitations of these methodologies. The following setup was found to be optimal for the operational runs: 3D iterative advection with 4D interpolation of all parameters to actual particle locations, a well-mixed ABL random-walk method, a model time step of 5–15 min for a 25 km spatial resolution, a combination of dry parcel and critical Richardson number methods for the ABL height assessment (unless this parameter is directly available from the meteorological input), a meteorological input with a 3-h temporal resolution and the best available spatial one (currently, 18 km).

The results of this setup were in compliance with the main ETEX quality criteria. Comparing the presented results with other quantitative model evaluation studies against ETEX-1 data, the performance of SILAM can be considered to be among the best.

However, the model tends to overestimate the concentration peaks inside the pollution plume, with simultaneous underestimation of the horizontal plume size, partially missing the low-concentration areas. The effects are believed to originate from

the simplified approach to the random-walk computations and are similar to phenomena reported by many ETEX participants. However, these inaccuracies are within the acceptable limits in view of the model's emergency applications.

The *sensitivity runs* with different sources of input data showed that high spatial and temporal resolutions do not automatically lead to better results. Thus, the SILAM runs with the new HIRLAM 5.2.1 do not show scores better than those obtained using the HIRLAM 2.6 or the ECMWF model, despite the twice-better spatial and 6-times better temporal resolutions of the former model. On the contrary, the 1-h resolution data suffer from the spin-up problem, while the other advantages were offset by a mistaken deep nighttime ABL at the release site. This problem with HIRLAM was also reported in other evaluation studies.

It was found that use of the ABL height estimate from the NWP is the best option, if data are available. Otherwise, the combination of the dry parcel and the critical Richardson number approaches is a good compromise. The Richardson method alone and a constant ABL both perform poorly.

The two strong advantages of the above configuration are a precise (although resource-consuming) advection routine and a combination of two methods for the ABL height assessments. These methods showed a good and stable performance in the ETEX-1 experiment and can therefore be recommended for other regional-scale short-term dispersion models.

Acknowledgements

The authors gratefully acknowledge the contribution by Mika Salonoja to the formulation of the SILAM standards and in coding the first version of the system. Part of the SILAM evaluation and comparison with other models was carried out within the EU-funded ENSEMBLE and FUMAPEX, NMR-funded HIRLAM-Baltic and Academy of Finland funded POLLEN projects.

References

Brandt, J., Christensen, J.H., Frohn, L.M., Zlatev, Z., 2000. Numerical modelling of transport, dispersion, and deposition—validation against ETEX-1, ETEX-2 and Chernobyl. Environment Modelling and Software 15 (6–7), 521–531.

Directions, 1994. Directions for actions in a nuclear weapons explosion. Publication of the Ministry of the Interior Department for Rescue Services, Series A:48.(in Finnish: Toiminta ydinräjähdystilanteessa Sisäasianministeriö, pelastusosaston julkaisu, Sarja A:48).

Eerola, K., 1990. Experimentation with a three-dimensional trajectory model. FMI Meteorological Publication 15, 33pp.

ETEX, 1998. Atmospheric Environment 32 (24), 4089–4375 (special issue dedicated to the ETEX experiment).

Galmarini, S., Bianconi, R., Addis, R., Andronopoulos, S., Astrup, P., Bartzis, J.C., Bellasio, R., Buckley, R., Champion, H., Chino, M., D'Amours, R., Davakis, E., Eleveld, H., Glaab, H., Manning, A., Mikkelsen, T., Pechinger, U., Polreich, E., Prodanova, M., Slaper, H., Syrakov, D., Terada, H., Van der Auwera, L., 2004a. Ensemble dispersion forecasting—part II: application and evaluation. Atmospheric Environment 38 (28), 4619–4632.

Galmarini, S., Bianconi, R., Klug, W., Mikkelsen, T., Addis, R., Andronopoulos, S., Astrup, P., Baklanov, A., Bartniki, J., Bartzis, J.C., Bellasio, R., Bompay, F., Buckley, R., Bouzom, M., Champion, H., D'Amours, R., Davakis, E., Eleveld, H., Geertsema, G.T., Glaab, H., Kollax, M., Ilvonen, M., Manning, A., Pechinger, U., Persson, C., Polreich, E., Potemski, S., Prodanova, M., Saltbones, J., Slaper, H., Sofiev, M.A., Syrakov, D., Sørensen, J.H., Van der Auwera, L., Valkama, I., Zelazny, R., 2004b. Ensemble dispersion forecasting—part I: concept, approach and indicators. Atmospheric Environment 38 (28), 4607–4617 http://ensemble.e.

Graziani, G., Klug, W., Mosca, S., 1998. Real-time long-range dispersion model evaluation of the ETEX first release. Office for official publications of the European Communities, L-2985, Luxembourg, 215pp.

Graziani, G., Galmarini, S., Mikkelsen, T., 2000. RTMOD: Real-Time MODel evaluation. Joint Risø-Report-1174 (EN)/JRC-Ispra Report TN.I.00.11. ISBN:87-550-2682-6 87-550-2687-7, ISSN:0106-2840, 47pp., http://www.risoe.dk/rispubl/VEA/ris-r-1174.htm

Hicks, B.B., Baldochi, D.D., Meyers, T.P., Hosker Jr., R.P., Matt, D.R., 1987. A preliminary multiple resistance routine for deriving dry deposition velocities from measured quantities. Water, Air and Soil Pollution 36, 311–330.

Horn, H.-G., Bonka, H., Maqua, M., 1987. Measured particle bound activity size-distribution, deposition velocity, and activity concentration in rainwater after the Chernobyl accident. Journal of Aerosol Science 18, 681–684.

Huber, P.J., 1981. Robust Statistics. Wiley, New York, Chichester, Brisbane, Toronto, 303pp.

Jylhä, K., 1991. Empirical scavenging coefficients of radioactive substances released from Chernobyl. Atmospheric Environment 25A (2), 263–270.

Kocher, D.C., 1980. Dose–rate conversion factors for external exposure to photon and electron radiation from radionuclides occurring in routine releases from nuclear fuel cycle facilities. Health Physics 38, 543–621.

Kocher, D.C., 1983. Dose–rate conversion factors for external exposure to photons and electrons. Health Physics 45, 665–686.

Lindfors, V., Joffre, S.M., Damski, J., 1993. Meteorological variability of the wet and dry deposition of sulphur and nitrogen compounds over the Baltic Sea. Water, Air and Soil Pollution 66, 1–28.

Marchuk, G.I., 1982. Mathematical modelling in the environmental problems (in Russian). Nauka Publisher, Moscow 320 pp.

Maryon, R.H., Buckland, A.T., 1994. Tropospheric dispersion: the first ten days after a puff release. Quarterly Journal of the Royal Meterorological Society 121 (528), 1799–1833.

Maryon, R.H., Smith, F.B., Conway, B.J., Goddard, D.M., 1991. The UK nuclear accident model. Progress in Nuclear Energy 26 (2), 85–104.

Mosca, S., Graziani, G., Klug, W., Bellasio, R., Biaconi, R., 1998. A statistical methodology for the evaluation of long-range dispersion models: an application to the ETEX exercise. Atmospheric Environment 32 (24), 4307–4324.

Nieuwstadt, F.T.M., Van Dop, H., (Eds.), 1982. Atmospheric Turbulence and Air Pollution Modelling. D. Reidel Publishing Company, Dordercht, The Netherlands, 360pp.

Olesen, H.R., 1995. Data sets and protocol for model validation. Workshop on operational short-range atmospheric dispersion models for environmental impact assessments in Europe. International Journal of Environment and Pollution 5, 693–701.

Phipps, A.W., Kendall, G.M., Stather, J.W., Fell, T.P. (1991). Committed equivalent organ doses and committed effective doses from intakes of radionuclides. National Radiological Protection Board, NRPB-R245/NRPB-M288, Chilton.

Pirazzini, R., Vihma, T., Launiainen, J., Tisler, P., 2002. Validation of HIRLAM boundary-layer structures over the Baltic Sea. Boreal Environmental Research 7, 211–218.

Pobanz, B.M., Vogt, P.P., Aluzzi, F.J., Pace, J.C., 1999. Comparison of gridded versus observation data to initialize ARAC dispersion models for the Algeciras, Spain Steel Mill Cs-137 Release. In: Proceedings of the American Nuclear Society Seventh Topical Meeting on Emergency Preparedness and Response, 14–17 September, Santa Fe, NM, UCRL-JC-131200.

Pöllanen, R., Valkama, I., Toivonen, H., 1997. Transport of radioactive particles from the Chernobyl accident. Atmospheric Environment 31 (21), 3575–3590.

Robertson, L., Langner, J., Engardt, M., 1999. An Eulerian limited-area atmospheric transport model. Journal of Applied Meteorology 38, 190–210.

Rodean, H.C., 1996. Stochastic Lagrangian models of turbulent diffusion. American Meteorological Society, Boston, MA.

RSIC, 1981. Calculation of dose-rate conversion factors for exposure to photons and electrons. Radiation Shielding Information Centre, DOSFACTER II, computer code collection, CCC-400. Oak Ridge National Laboratory, Oak Ridge, TN.

RSIC, 1988. Dose-rate conversion factors for external exposure to photons and electrons. Radiation Shielding Information Centre, DOSFACTER DOE, computer code collection, CCC-536. Oak Ridge National Laboratory, Oak Ridge, TN.

RSIC, 1989. Calculation of dose-rate conversion factors for exposure to photons and electrons. Radiation Shielding Information Centre, DOSDAT-DOE, data library collection, DLC-144. Oak Ridge National Laboratory, Oak Ridge, TN.

Saltbones, J., Foss, A., Bartnicki, J., 1996. A real time dispersion model for severe nuclear accidents, tested in the European tracer experiment. Systems Analysis Modelling Simulation 25, 263–279.

Savijärvi, H., Kauhanen, J., 2001. High resolution numerical simulations of temporal and vertical variability in the stable wintertime boreal boundary layer: a case study. Theoretical and Applied Climatology 70, 97–103.

Seibert, P., Beyrich, F., Gryning, S.E., Joffre, S., Rasmussen, A., Tercier, P., 2000. Review and inter-comparison of operational methods for the determination of the mixing height. Atmospheric Environment 34, 1001–1027.

Smith, F.B., Clark, M.J., 1989. The transport and deposition of radioactive debris from the Chernobyl nuclear power plant accident with special emphasis on consequences to the United Kingdom. Meterorological Office Scientific Paper, N42, HMSO, London.

Sofiev, M., 1999. Validation of model results on different scales. Approaches to scaling of trace gas fluxes in ecosystems. In: Bouwman, A.F. (Ed.), Developments in Atmospheric Science, vol. 24. Elsevier, New York, pp. 235–255.

Sofiev, M., Atlaskin, E., 2004. An example of application of data assimilation technique and adjoint dispersion modelling to an inverse dispersion problem based on the ETEX experiment. In Air Polution Modelling and its Applications, vol. XVII, in press, also in pre-prints of 27th International Technical Meeting on Air Pollution Modelling and its Applications, 23–30 October 2004, Banff, Canada, pp. 405–412.

Sørensen, J.H., 1998. Sensitivity of the DERMA long-range Gaussian dispersion model to meteorological input and diffusion parameters. Atmospheric Environment 32 (24), 4195–4206.

Stohl, A., Thomson, D.J., 1999. A density correction for Lagrangian particle dispersion models. Boundary Layer Meteorology 90, 155–167.

Stohl, A., Hittenberger, M., Wotawa, G., 1998. Validation of the Lagrangian particle dispersion model FLEXPART against large scale tracer experiments. Atmospheric Environment 24, 4245–4264.

Thomson, D.J., 1987. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. Journal of Fluid Mechanics 180, 529–556.

Troen, I., Mahrt, I., 1986. A simple model of the atmospheric boundary layer: sensitivity to surface evaporation. Boundary-Layer Metorology 37, 129–148.

Unden, P. (Ed.), 2002. HIRLAM-5 Scientific Documentation. HIRLAM Project, SMHI, S-601 76, 2002, Norrkoping, Sweden.

Zannetti, P., 1992. Particle modelling and its application for simulating air pollution phenomena. In: Melli, P., Zannetti, P. (Eds.), Environmental Modelling. Computational Mechanisms Publications, Southampton, UK.