Available online at www.sciencedirect.com
ScienceDirect
Atmospheric Environment 40 (2006) 8020–8033
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

Validation of a Lagrangian dispersion model implementing different kernel methods for density reconstruction

Lina Vitali[a], Fabio Monforti[a,*], Roberto Bellasio[b], Roberto Bianconi[b], Valentina Sachero[a], Sonia Mosca[b], Gabriele Zanini[a]

[a]*ENEA PROT-INN section-via Martiri di Monte Sole 4, I-40129 Bologna, Italy*
[b]*ENVIROWARE srl-C.D. Colleoni Andromeda 1, I-20041 Agrate Brianza (MI), Italy*

Received 6 February 2006; received in revised form 7 June 2006; accepted 7 June 2006

Abstract

In this paper the inert version of a Lagrangian particle model named photochemical Lagrangian particle model (PLPM) is described and validated. PLPM implements four density reconstruction algorithms based on the kernel density estimator. All these methods are fully grid-free but they differ each other in considering local or global features of the particles distribution, in treating the Cartesian directions separately or together and in being based on receptors or particles positions in space. Each kernel has been shown to have both advantages and disadvantages, but the overall good performances of the model when compared with the well known Copenhagen and Kincaid data sets are very encouraging in view of its extension to fully chemically active simulations, currently under development.

Keywords: Photochemical pollution; Lagrangian chemical transport model; Grid-free model; Complex chemical mechanism; Kernel density estimator

1. Introduction

Particle Lagrangian dispersion modelling in atmosphere has been extensively employed in last decades. In these models the pollutants emitted by a source are modelled as discrete parcels each carrying a finite mass amount. These parcels are successively transported in the atmosphere by both the average air flows and the turbulence, usually modelled by means of a Monte Carlo approach. For this class of models, the density reconstruction problem is a crucial one. The problem consists in finding an algorithm allowing reconstruct a continuous concentration three-dimensional field on the basis of the particle masses and positions. The simpler approach is the *box counting* method, also known as "particles-in-the-box" method: a grid of boxes is superimposed to the modelling domain and the concentration field is computed simply as the total mass of the particles falling in a box divided by the box volume. In this approach, concentrations computed in boxes containing a small number of particles are affected by relatively large statistical errors that can be reduced only by increasing the number of the particles emitted throughout the simulation. Furthermore, in this approach the advantage of a grid-free dispersion treatment typical of the Lagrangian approach is partially lost.

*Corresponding author. Tel.: +39 0516098650; fax: +39 0516098675.
E-mail address: monforti@bologna.enea.it (F. Monforti).

1352-2310/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2006.06.056

To achieve a full independence from Eulerian grids, a technique for computing the concentration fields different from the usual box counting method can be employed: the *kernel density estimator* (Silverman, 1986). The kernel method can be understood in intuitive terms if one states that particles identify the centre of mass of a "cloud". The density profile of the "cloud" is given by the *kernel* function while the volume of mass spreading depends on the *bandwidths* associated to each particle or receptor.

In this paper a dispersion Lagrangian model implementing four different density reconstruction approaches based on the kernel method is introduced and partially described and validated. The model, called photochemical Lagrangian particle model (PLPM) has been developed with the final goal of achieving a fully Lagrangian treatment of complex photochemistry, but in this paper only the inert version of the model is discussed as in PLPM dynamic and chemical approaches are clearly separated: turbulence acts only on the particle movement in the atmosphere whereas the efficiency of the chemical reactions between particle are derived only from particles positions and masses, independently from turbulence features of the atmosphere.

Here a detailed description of the dynamical and density reconstruction modules of PLPM is given and the model is fully validated in the case of inert pollutants dispersion on the basis of the two data sets of the classical Copenhagen and Kincaid experiments (Gryning, 1981; Gryning and Lyck, 1984, 2002; Bowne and Londergan, 1983). In both cases the Model Validation Kit methodology was applied (Cuvelier, 1994; Olesen, 1995a, b, 1998).

Generally speaking, the inert case validation is a necessary step to be fulfilled by a model dealing with atmospheric chemistry before approaching more complex situations in order to assess the description of the physic of the dispersion implemented in the algorithm. Furthermore, in the case of PLPM, the inert case validation has the further goal of comparing model performances achieved by implementing the different kernel bandwidths computation approaches.

2. PLPM model description

PLPM implements the Lagrangian approach to the atmospheric dispersion. Particles are generated to represent each a fraction of the emitted mass, and then moved in the space accordingly to wind and turbulence features. Transport and diffusion are independent from any grid: parcels motion is described using the whole information on the meteorological fields and time and space interpolation is applied to calculate meteorological variables in the actual particle positions.

2.1. Domain and meteorological pre-processing

PLPM is interfaced with the diagnostic meteorological model CALMET (Scire et al., 2000), acting as a meteorological pre-processor. Domain and meteorological parameters are directly read from CALMET, with the domain of PLPM being slightly reduced at the lateral and top borders for interpolation. Moreover, as the bottom layer in CALMET is 20 m high from the terrain (and then meteorological variables are computed 10 m high) some additional levels have been added for which meteorological data are obtained according to similarity theory.

2.2. Sources

Several sources of different types (point, area, volume and linear) can be located anywhere inside the simulation domain. Area and volume sources can have arbitrary extensions with various shapes since they are not referred to a grid mesh. Each source marks the particles it emits so that, at any time during the simulation it is possible to trace back a given particle to its source.

2.3. Initial and boundary conditions

Initial conditions are treated as instantaneous volumetric sources activated at the beginning of the simulation that generate particles randomly and homogeneously inside the volumes used to characterise initial concentration values.

Boundary conditions are described through an appropriate number of volume sources overlapping the domain boundaries with a transverse extension depending on the component of wind velocity normal to the boundary, entering or exiting the domain. Particles emitted by the initial and boundary pseudo-sources are marked too.

2.4. Turbulent particles motion

In the random walk Lagrangian models family, turbulence is supposed to act on particles moving through a stochastic velocity component added to the average velocity from meteorological flows. The random walk induced by turbulence is supposed also to be Markovian: given the position of a particle at time t, its position at the time $t+\Delta t$ is given by

$$x_i(t+\Delta t) = x_i(t) + \Delta t(u_i + u_i'), \tag{1}$$

where $i = 1,2,3$ indicates, respectively, the x, y and z-direction, u_i is the mean wind component along the ith direction and u'_i represents the turbulent velocity fluctuation along the same ith direction. The time evolution of the velocity fluctuation is described in the most general terms by the non-linear Langevin equation introduced by Thomson (1987):

$$\mathrm{d}u_i' = a_i(x,u',t)\,\mathrm{d}t + b_{ij}(x,u',t)\,\mathrm{d}\xi_j(t), \tag{2}$$

where a_i and b_{ij} are functions of space, velocity and time, and $\mathrm{d}\xi_j(t)$ is a random increment of a Wiener process with independent components. The fluctuating turbulent term at time t is correlated to the one at time $t+\Delta t$ and a Lagrangian time scale T_L can be defined as the value at which the autocorrelation coefficient is equal to $1/e$. Both Lagrangian time and a_i and b_{ij} coefficients in Eq. (2) are linked to the structure of turbulence through functional relations with the meteorological variables. In PLPM particles move independently in each direction and their trajectories are constructed using Eq. (1). In the horizontal directions homogeneous Gaussian turbulence is supposed so that Eq. (2) reduces to

$$\mathrm{d}u' = -\frac{u'}{T_\mathrm{L}}\mathrm{d}t + \left(\frac{2\sigma^2}{T_\mathrm{L}}\right)^{1/2}\mathrm{d}\xi. \tag{3}$$

On the contrary, different turbulence models are applied in the vertical direction, depending on stability. Under convective conditions three models can be alternatively used to describe the turbulent vertical motion: the homogeneous skewed model of Hurley and Physick (1993), the quasi-homogeneous model described in Bianconi et al. (1999), and the non homogeneous skewed model of Luhar and Britter (1989). Evolution equations for particles velocity fluctuations are solved by a forward in time scheme.

2.5. Concentration fields

In the kernel approach the concentration c in every point (x,y,z) at time t is computed as:[1]

$$c(x,y,z,t) = \sum_{i=1}^{n} \frac{m_i}{\lambda_x\lambda_y\lambda_z}\, d\left(\left|\frac{x_i-x}{\lambda_x}\right|\right) \times d\left(\left|\frac{y_i-y}{\lambda_y}\right|\right) d\left(\left|\frac{z_i-z}{\lambda_z}\right|\right), \tag{4}$$

where n is the total number of particles in the domain, (x_i,y_i,z_i) and m_i are the position and the mass of the ith particle, respectively, and $d(u)$ is a *kernel* function satisfying the two following requirements:

$$d(u) \geqslant 0, \quad \forall u \in \Re, \tag{5}$$

$$\int_{\Re} d(u)\,\mathrm{d}u = 1. \tag{6}$$

The λ_j parameters are the *bandwidths* and control the volume in which the mass of the particle is spread in the domain.

The kernel method is a well known non parametric density estimation tool, widely employed in general statistics problems (see e.g. Silverman, 1986). It has been applied several times in Lagrangian dispersion modelling, starting from Lorimer (1986) and Lorimer and Ross (1986) that employed a kernel estimator to efficiently solve the advection-diffusion equation. Yamada et al. (1987, 1989), Yamada and Bunker (1988) have used a Gaussian kernel estimator where bandwidths are determined as the time integration of the velocity variances encountered by the particle during its life. More recently, De Haan (1999) has compared the efficiency of a number of kernel functions other than Gaussian in the frame of puff dispersion when bandwidths are set following the "optimal bandwidth" approach described by Eq. (8).

Finally, Davakis et al. (2003) have compared again different kernel functions efficiency when dealing with dispersion experiments over highly complex terrains (TRANSALP90 and ETEX experiments). In this frame, authors have also compared two different approaches to bandwidths computation: i.e., the time integration approach employed by Yamada et al. (1987) against setting

[1] In Eq. (4) a *product kernel* density estimator is described. The more general *three-dimensional kernels* are not considered in this paper. An extensive review of their application in pollutant density reconstruction can be found in De Haan (1999).

bandwidths equal to the variance of the particle distribution around each observation point.

2.5.1. Kernel methods in PLPM

Following the results of the cited literature and the outcomes of a number of tests (Peverieri, 2000), the so-called Epanechnikov kernel estimator (Epanechnikov, 1969) has been implemented in PLPM:

$$d(u) = \begin{cases} \frac{3}{4}(1-u^2), & |u| \leqslant 1, \\ 0, & |u| > 1. \end{cases} \quad (7)$$

Experience in statistical applications of kernel estimators (Jones et al., 1996) suggests that correctly choosing bandwidths in applying (4) is crucial for the estimator performances: too small bandwidths can lead to a concentration field more irregular than real one whereas bandwidths overestimation can result in a large bias between the reconstructed and the real field. Thus, an optimal method for bandwidth calculation is needed, aimed to minimize both variance and bias. In PLPM four different methods can be employed for bandwidth calculation:

(1) *PG method (particles-based and globally-defined)*. Bandwidths are *associated to the particles*. The same set of bandwidth, λ_x, λ_y and λ_z, are associated to all particles. Bandwidths are calculated from global features of particle distribution as (De Haan, 1999):

$$\lambda_j = \alpha A(K) n^{-(1/7)} \min\left(\sigma_j, \frac{R_j}{1.34}\right) \quad \text{with } j = [x, y, z], \quad (8)$$

where σ_j and $R_j$are the standard deviation and the interquartile range of the particles distribution in each space direction, n is the total number of particles, α is a tuning parameter (set equal to 0.85) and $A(K)$ is a parameter depending on the particle distribution and the kernel function.

(2) *PL3 method (particles-based, locally-defined, three-dimensionally ordered)*. Bandwidths are *associated to the particles* and a different set of bandwidths λ_{ij} (with $i = 1,\ldots,n$ and $j = [x, y, z]$) is associated to each particle. A *neighbourhood* for each particle is defined (see Section 2.5.2). In each direction the bandwidth is set as the maximum projection of the distances of the nearest neighbours in that direction.

(3) *PL1 method (particles-based, locally-defined, one-dimensionally ordered)*. Bandwidths are *associated to the particles* and a different set of bandwidths λ_{ij} (with $i = 1,\ldots,n$ and $j = [x, y, z]$) is associated to each particle. For each particle, *three different neighbourhoods are defined for each space direction,* based on the projection of particles distances on the three axis. For each direction the bandwidth is set as the projected distance of the first particle excluded from each of the neighbourhoods.

(4) *RL3 method (receptors-based, locally-defined, three-dimensionally ordered)*. Bandwidths are *associated to the receptors*, i.e., the points where the concentration has to be estimated and a different set of bandwidths is associated to each receptor. For each receptor *all particles are ordered* by their distance from the receptor. A neighbourhood of the receptor is defined (see Section 2.5.2) and the bandwidth associated to the receptor, in each direction, is the maximum projection of the distance of the neighbours in that direction.

2.5.2. Neighbourhood definition

In the case of locally defined kernels (i.e., all kernels tested here except PG) a method to define the extension of particles or receptors neighbourhood is needed. In the frame of the general reconstruction problem, it is usual to link the number k of particles contained in the neighbourhood to the total particles number N through a simple empirical relation. Three possibilities have been tested: $k = N/4$, $N/8$ and $k \propto N^{1/2}$. On the basis of a wide range of specifically designed tests (Monforti, 2001; Peverieri, 2000), both in one and in three dimensions, the second formulation was preferred. Then, in PLPM a neighbourhood is defined as the ensemble of closer particles containing 1/8 of the total mass, which coincides with 1/8 of the total particles if particles carry a constant mass amount. In PLPM kernel bandwidths are computed on the basis of the distribution of particles in space and their masses, without any quantity related to the physics of the dispersion being involved. In such an approach, pollutants dispersion is supposed to be fully described by the stochastic part of the particles motion, the problem of underlying concentration reconstruction being treated purely as a mathematical problem.

It is worth noticing as the four kernels chosen for the implementation in PLPM span a number of dichotomies in bandwidths calculation (i.e., global vs. local, three-dimensional vs. one-dimensional and

particles-based vs. receptors-based) and that most of these approaches were never applied in the frame of Lagrangian dispersion modelling.

These new approaches were introduced with a view to the development of the fully Lagrangian approach to photochemistry implemented in PLPM. Indeed, in the case of a chemically complete simulation, a density reconstruction algorithm should assure good performances not only in estimating ground concentrations, as it is usually required in dispersion models validation, but it should also give satisfactory and numerically stable results in the whole three dimensional dispersion field, where chemical reactions take place. A set of tests of kernel performances in full chemical simulations are currently under development.

2.5.3. Comparison with box counting

All the authors cited in Section 2.5 agree that the kernel method gives better results than the usual box counting method for a fixed number of particles. Nevertheless, some tests were performed even in the case of PLPM to verify the extent real precision advantage if compared with box counting. As an example, Fig. 1 (top) shows the concentration obtained with the four kernels from a plume miming the dispersion situation in the Copenhagen experiment (see Section 4) in 20 September 1978 on a receptor arc about 2 km distant from the source and the concentrations obtained in the same conditions with the box counting algorithm with four different box dimensions (bottom—see figure captions for details). Comparison with measurements clearly illustrates the typical problem encountered in applying the box counting method: for large box volumes (e.g BC3 and BC4) the concentration field is too smooth, while for small volumes (e.g. BC1 and BC2) the concentration field is perturbed by a strong numeric noise. On the contrary, the results obtained with the kernel method are shown to be more realistic. Tests performed on other arcs and simulation days have confirmed these findings, even from the quantitative point of view.

3. Model validation

The inert version of PLPM has been validated in a real case by means of two well-known dispersion experiments performed in Copenhagen, Denmark in 1978 and 1979 and in Kincaid, Illinois, USA in 1980 and 1981. Both experiments concerned the release of SF_6, a gaseous inert tracer with very low background concentrations, from a stack located in an almost flat zone. A number of receptors was placed at the height of around 2 m from the ground, along some arcs centred on the stack at increasing distance. Simulation domains were set up according to receptors spatial distribution for both Copenhagen and Kincaid experiments. For both Copenhagen and Kincaid simulations, the diagnostic model CALMET (Scire et al., 2000) was used as meteorological pre-processor. A preliminary sensitivity analysis enabled, for both data sets, the better choice of running options and parameters, summarized in Table 1.

3.1. Sources and time schedule

In both cases, the only tracer source in the domain was the stack. Background contribution and inflows were assumed as negligible and boundary and initial conditions were not needed. For the Kincaid experiment, hourly emissions have been divided in 120 puffs released every 30 s, each containing 30 particles, for a total emission rate of 3600 particles h^{-1}. On the contrary, for Copenhagen releases, puffs including 90 particles, released every 20 s, for a total emission of 16 200 particles h^{-1}, where employed. The higher hourly release adopted in Copenhagen simulation is motivated by the shorter time span covered by the experiment and the smaller domain extension. With such particles emission rates, the typical number of particles present in the simulation domain (i.e., emitted but not yet transported outside the domain) ranges from 6000 to 10 000 in the case of Copenhagen and from 3000 to 8000 in the case of Kincaid. Generally speaking, these numbers are smaller by at least one order of magnitude than particle numbers usually employed in typical Lagrangian simulations (Tinarelli, 2005).

3.2. Plume rise

Whereas Copenhagen experiment deals with ambient temperature tracer releases, in the Kincaid experiment tracer is released at a temperature around 100 °C higher than the ambient one and plume is expected to rise because of the buoyancy force.

In the present version of PLPM a plume rise treatment rooted in the Eulerian scheme proposed by Ooms (1972) is implemented. The scheme is

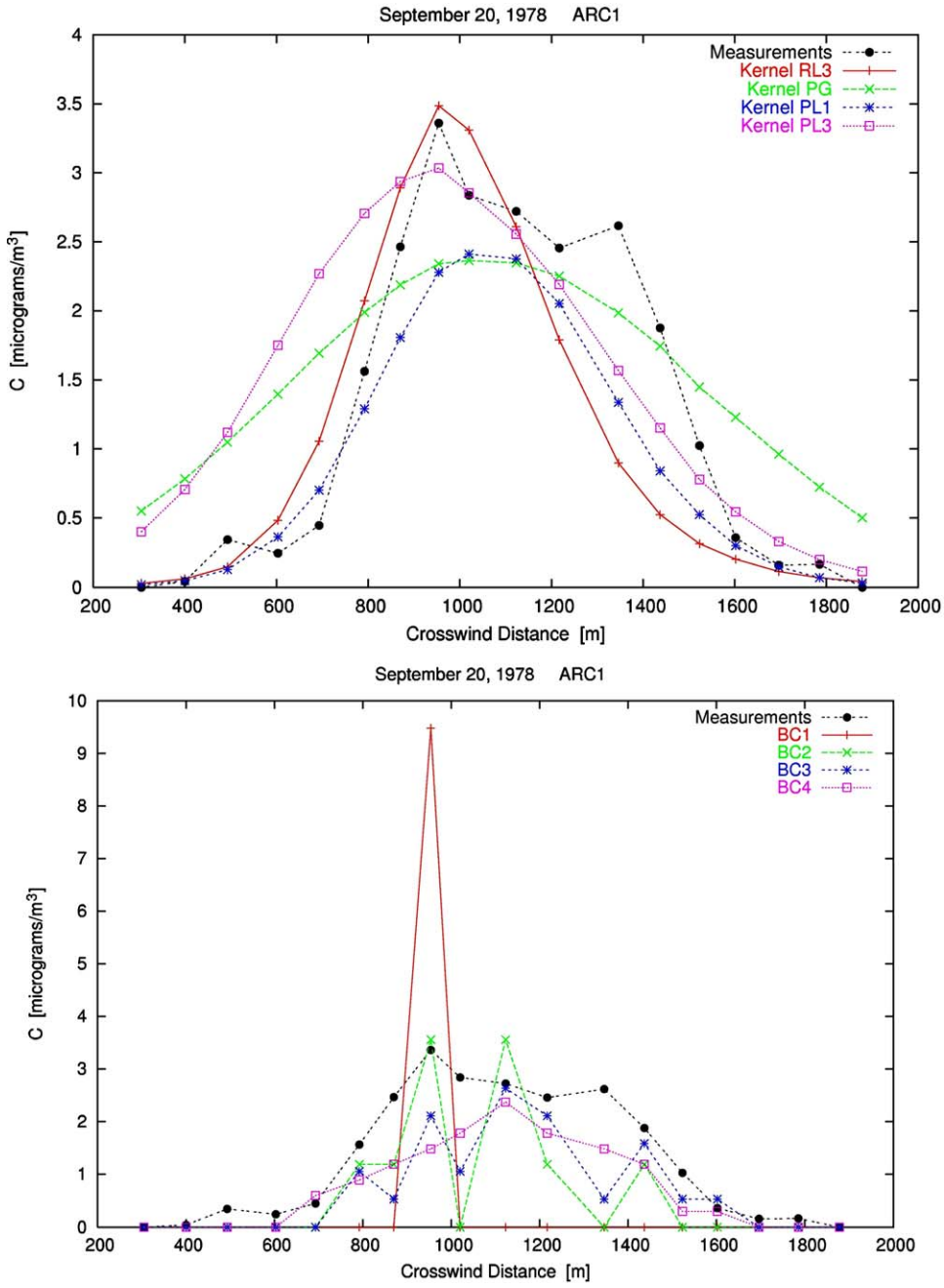

Fig. 1. Concentrations pattern on an arc distant about 2 km from Copenhagen stack in 20 September 1978. Top: different kernel methods compared with measurements. Bottom: box counting method with different box dimensions compared with experiments. (BC1:dx = dy = 50 m; BC2: dx = dy = 100 m BC3: dx = dy = 150 m; BC4: dx = dy = 200 m; dz = 5 m in all cases).

explained in full detail elsewhere (Sachero et al., 2004; Sachero, 2003) and a short summary of their main features follows.

The plume rise scheme implemented in PLPM consists in a set of scalar ordinary differential equations (ODEs) describing the plume rise phenomenon ruled by the conservation of mass, momentum and heat (enthalpy) of a plume slice. The ODEs system has been discretized in order to obtain the final implemented equation, resolved with a simple forward-in-time numerical technique. Following the examples of Anfossi et al. (1993) and

Table 1
Main features of CALMET runs for Copenhagen and Kincaid simulations

	Copenhagen	Kincaid
Model resolution	400 m	2000 m
Horizontal domain geometry	Squared	Squared
Horizontal size	12 000 m	60 000 m
Vertical size	3160 m	2800 m
Surface meteorological measuring points	• Ground meteorological measurements at Gladsaxe TV power (tracer release point)	• National Weather Service surface meteorological measurements at Springfield station.
Upper Air Meteorological measuring points	• Meteorological measurements at several heights at Gladsaxe TV power (tracer release point) • Routine upper air meteorological measurements at Upper Air Station at Koebenhavn (no. 06181)	• Meteorological measurements at several heights at a 'Central Site', located around 650 m east of tracer release point • Routine upper air meteorological measurements at Upper Air Station at Peoria

Webster et al. (2002), this Eulerian plume rise scheme has been nested within a Lagrangian framework. Including plume rise within a Lagrangian model poses the problem that the rise of each particle is influenced by the buoyancy of the plume *as a whole*, in contrast with the principle of particle independence on which Lagrangian modelling is based. This problem has been overcome by the introduction of the concept of the *plume-particle*: each particle is considered as a plume and an integral system based on the governing conservation equations following each particle separately is solved, using local mean flow properties. Each particle rises according to local conditions and it is modelled as a small plume driven by local ambient conditions, the deterministic particle velocity being replaced by the plume-particle velocity, achieved by solving the conservation equation system.

Thus, when plume rise is acting, the equation describing particle motion becomes

$$\boldsymbol{x}(t+\Delta t) = \boldsymbol{x}(t) + \Delta t(\boldsymbol{u}_{\mathrm{p}} + u'), \tag{9}$$

where the deterministic term $\boldsymbol{u}_{\mathrm{p}}$ represents the velocity of each particle, including also plume rise effect. Eq. (9) is valid as long as a buoyancy force acts on the particle. When the plume-particle reaches the equilibrium with the surrounding atmosphere, the particle starts to be transported passively by the wind. For each model iteration the plume rise routine is triggered only if the vertical component of $\boldsymbol{u}_{\mathrm{p}}$ is significantly larger than the vertical component of the wind field $\boldsymbol{u}_{\mathrm{a}}$ or if the particle is not more aged than 30 min (Webster et al., 2002). Those conditions allow the evaluation of the phase of plume dispersion, controlling the end of the rising phase and the beginning of the passive advection of the particle.

3.3. Concentration fields

Ground concentrations of SF_6 are calculated from particle positions in the domain by means of all the kernel methods. As suggested by Olesen (1995a, b), comparison between measurements and model output has been performed in the case of the Kincaid experiment by means of the method of identifying the ArcMax concentration in both measured and simulated data, whereas in the case of Copenhagen experiment also cross-wind integrated concentration on each arc has been computed thanks to the availability of corresponding measurements.

4. Validation results for copenhagen experiment

4.1. Performance indexes

Tables 2 and 3 show the usual performance indexes for model validation, respectively for normalized (to emission rate) arc maximum concentrations ($C_{\max}/Q$) and normalized cross-wind integrated concentrations (C_y/Q), respectively.

The analysis of Tables 2 and 3 shows as:

- All the density reconstruction methods, both for $C_{\max}/Q$ and C_y/Q, lead to a positive bias, i.e. a general underestimation of the concentration data;

Table 2
PLPM performance indexes on the Copenhagen data set for normalized maximum concentrations (C_{max}/Q) obtained with different kernels

($N = 23$)	Mean	Sigma	Bias	NMSE	R	FAC2	FB	FS
Measures	632.66	450.25	0.00	0.00	1.000	1.000	0.000	0.000
RL3	321.92	244.61	310.74	1.03	0.681	0.435	0.651	0.592
PG	356.52	203.71	276.13	0.84	0.710	0.652	0.558	0.754
PL1	292.70	182.94	339.95	1.26	0.714	0.478	0.735	0.844
PL3	311.97	200.14	320.68	1.14	0.664	0.391	0.679	0.769

Table 3
PLPM performance indexes on the Copenhagen data set for normalized cross-wind integrated concentrations (C_y/Q) obtained with different kernels

($N = 23$)	Mean	Sigma	Bias	NMSE	R	FAC2	FB	FS
Measures	448.70	239.29	0.00	0.00	1.000	1.000	0.000	0.000
RL3	245.74	106.36	202.96	0.73	0.569	0.652	0.585	0.769
PG	410.65	157.89	38.04	0.21	0.603	0.870	0.089	0.410
PL1	259.78	89.46	188.91	0.67	0.533	0.739	0.533	0.912
PL3	340.70	150.26	108.00	0.33	0.569	0.870	0.274	0.457

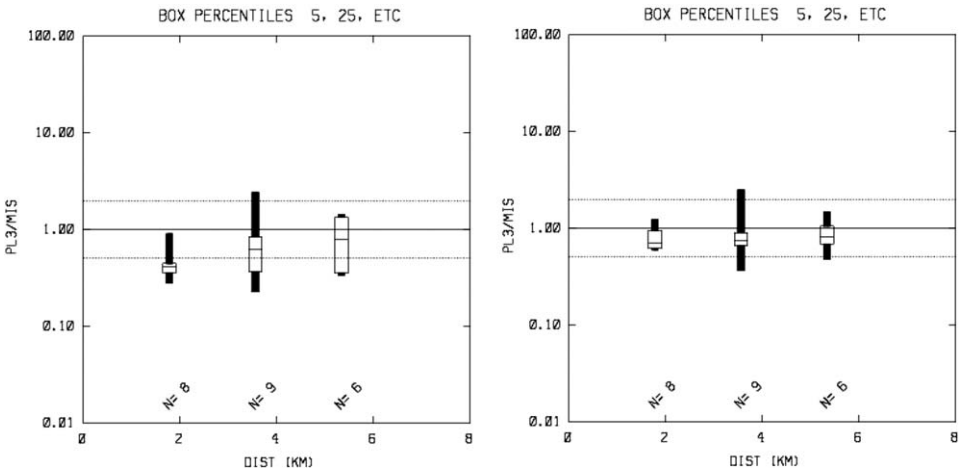

Fig. 2. Residual analysis for Copenhagen experiment arc maximum concentration (left) and integrated concentration (right) for varying downwind distance from the source (PL3 density reconstruction employed).

- PLPM, for all the density reconstruction methods, gives better performances in estimating integrated concentration rather than maximum concentration.

4.2. Residual analysis

A residual analysis has been performed on C_{max}/Q and C_y/Q to evaluate model performances in function of the downwind distance from the source. In Fig. 2 (left) percentiles (5th, 25th, 50th, 75th and 95th) of the C_{max}/Q residual distributions, for different values of the downwind distance when PL3 kernel is employed are shown (results for other density reconstruction methods are qualitatively similar). Plotted results show clearly the improvement of the estimation of arc maximum concentration at increasing distances from the source.

On the contrary, as far as cross-wind integrated concentration C_y/Q is concerned (Fig. 2, right), no

clear correlation with the downwind distance has been found and model performances are good also on the arc closest to the source. According to residual analysis, the correct evaluation of the maximum concentration near the source seems to be a critical issue.

4.3. Concentration patterns

Some more considerations on the suitability of the kernel methods employed can be driven from the visual analysis and the comparison of the reconstructed concentration fields.

Fig. 1 (top) shows the modelled and measured concentration values on an arc distant about 2 km from Copenhagen stack in 20 September 1978 while Figs. 3 and 4 show the concentration fields obtained by means of different kernels on the ground and on the x–z section plan including the stack, respectively.

For example, one can notice as kernels PG and PL3 tend to disperse the larger fraction of mass in the distribution tails and for this reason their performance improve considerably moving from maximum concentrations (Table 2) to integrated concentration (Table 3), when the whole dispersed mass is taken into account, regardless to its location within the arc. Furthermore, some irregularities in the concentration fields obtained with PL1 are also evident, together with better smoothness of other kernels.

It is worth also noticing that all the density reconstruction methods provide a suitable estimation of the maximum position, although this model skill is not evaluated by Model Validation Kit protocol.

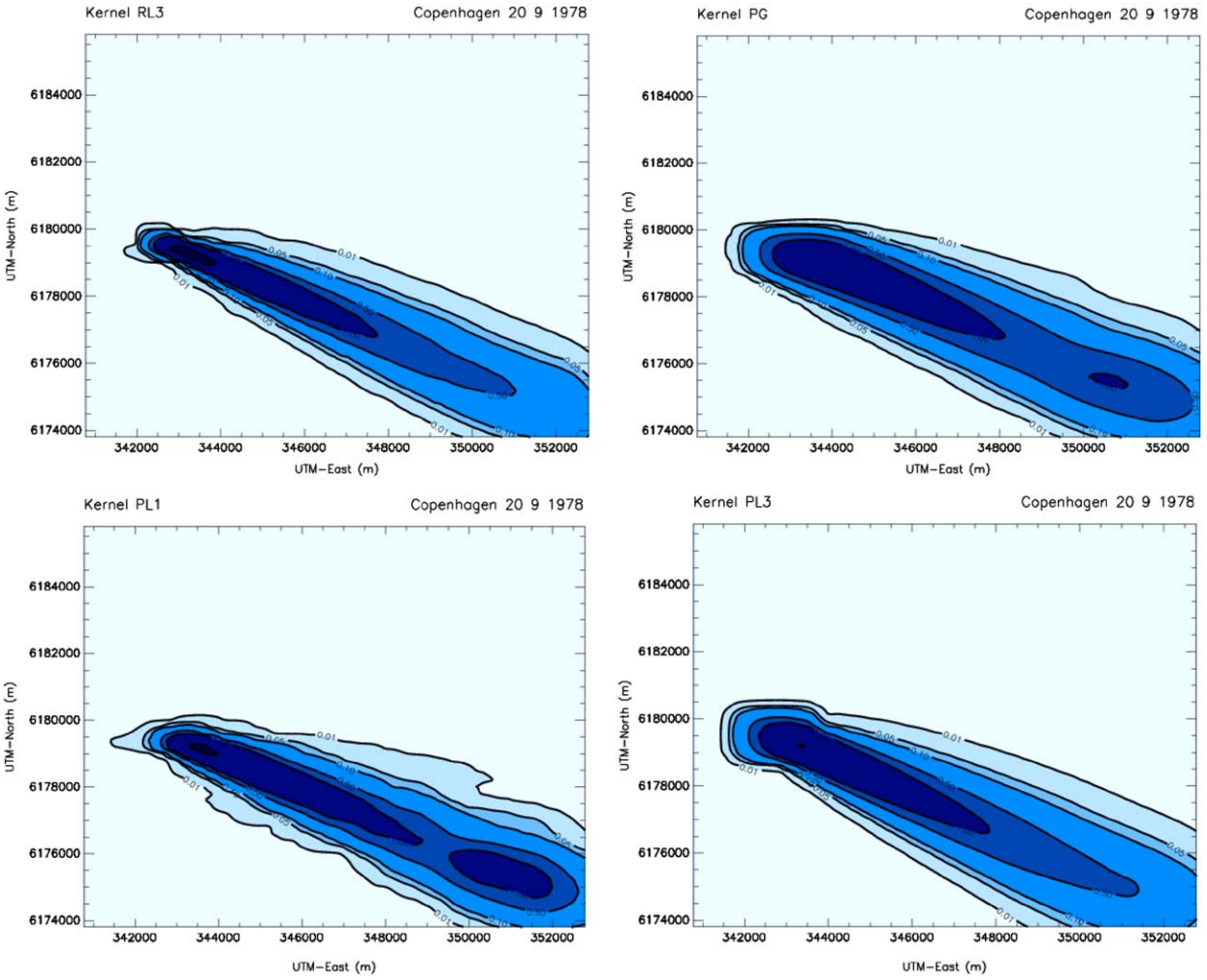

Fig. 3. Concentration fields computed on the ground by means of different kernels methods (Copenhagen, 20 September 1978).

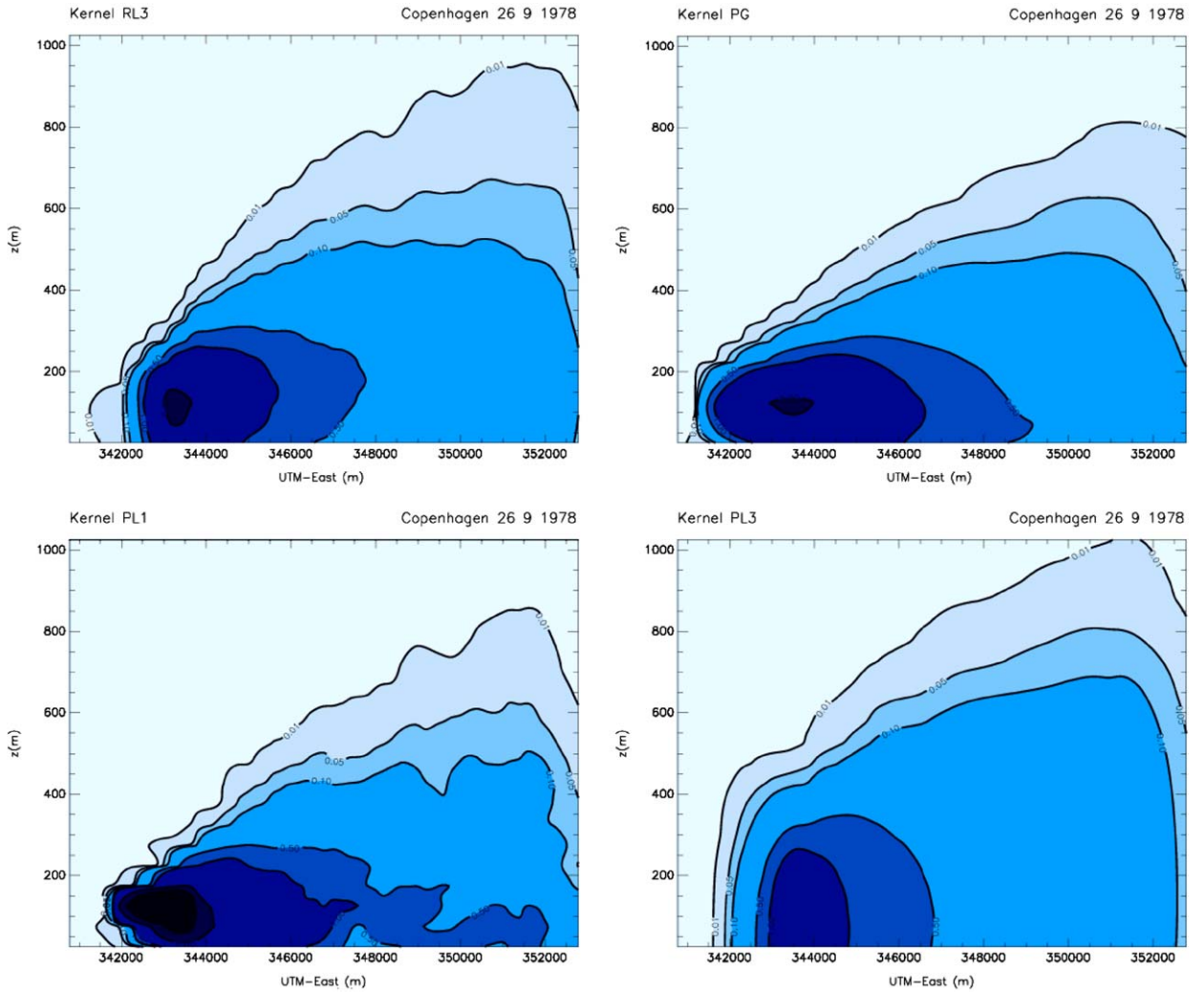

Fig. 4. Concentration fields computed in a x–z section plan including the stack by means of different kernels methods (Copenhagen, 26 September 1978).

5. Validation results for kincaid experiment

5.1. Performance indexes

Table 4 shows the usual performance indexes achieved by PLPM simulations performed using the highest quality data subset ($Q = 3$) of the Kincaid available data set and applying the four different kernel density reconstruction methods. Results obtained do not support evidence for a kernel method performing outstandingly better than other ones, except for PL1 that seems to obtain slightly worst results.

It is worth noticing that all the methods lead to negative bias, i.e., a general overestimation of the concentration data is pointed out. On the contrary, an overall underestimation was achieved for the Copenhagen data set. One possible explanation of this different behaviour can reside in the data features: first of all, Kincaid experiment is by large more complete than the Copenhagen one (338 data points, against 23, spanning a larger set of meteorological situations) and PLPM might have a true tendency to (slightly) overestimate ground concentrations. According to this assumption, the Copenhagen experiment could be an "unlucky" small data set on which underestimation prevails by hazard.

Anyway, it is also worth noticing as the two experiments have different space scales, both from

the point of view of stack height (115 m in Copenhagen against 187 m in Kincaid, where the plume rise effect results in an effective height between 250 and 300 m) and receptors location. Furthermore, even the temporal scale of the two experiments is different: the Kincaid tracer releases and concentration measurements lasting normally 5–8 h for each day, while in Copenhagen they usually covered only 1–2 h every day. For this reason, the comparison of the model performances on the two data sets could be also interpreted as the proof of the tendency of PLPM to (slightly) underestimate concentrations during transients against a tendency to (slightly) overestimate concentrations in more stationary situations.

Table 4
PLPM performance indexes on the Kincaid Q3 data subset obtained with different kernels

($N = 338$)	Mean	Sigma	Bias	NMSE	R	FAC2	FB	FS
Measures	54.34	40.25	0.00	0.00	1.000	1.000	0.000	0.000
RL3	56.76	50.29	−2.43	0.91	0.331	0.577	−0.044	−0.222
PG	58.50	50.24	−4.17	0.64	0.522	0.604	−0.074	−0.221
PL1	78.08	88.08	−23.75	1.71	0.337	0.512	−0.359	−0.745
PL3	59.54	49.57	−5.20	0.80	0.337	0.577	−0.091	−0.208

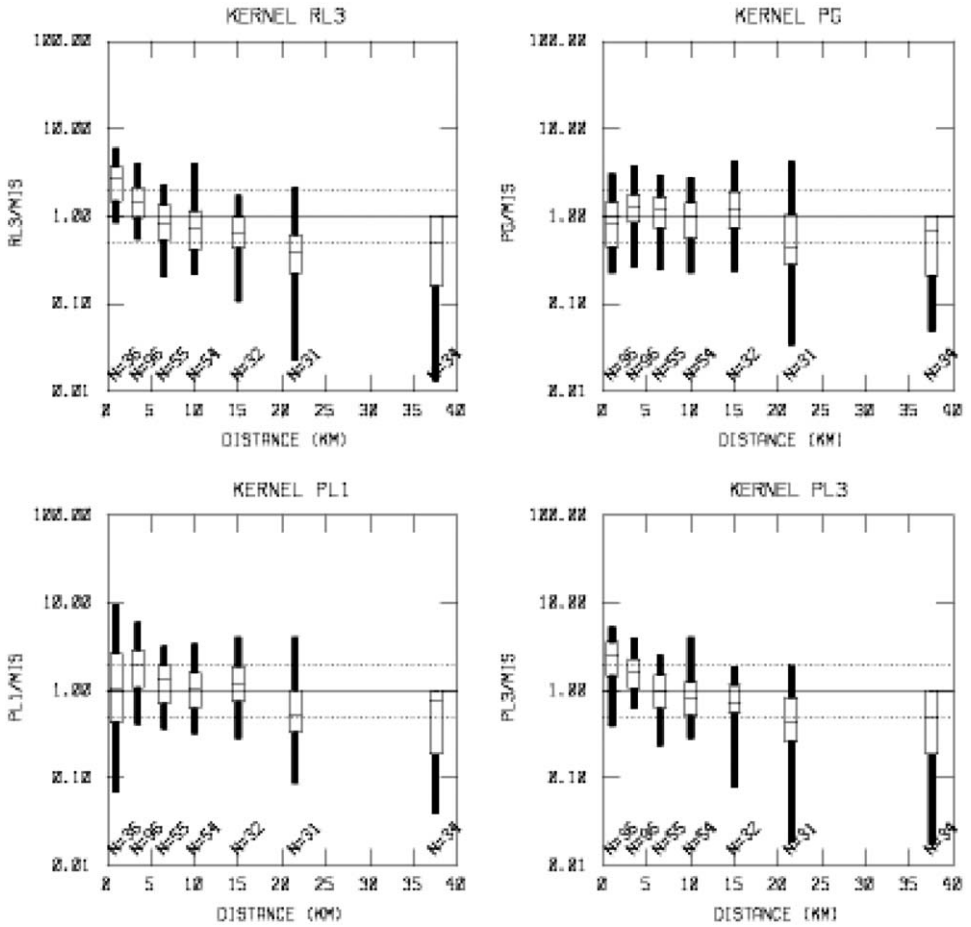

Fig. 5. C_{max}/Q residual distributions in function of the downwind distance from the source (Kincaid experiment—all density reconstruction methods).

5.2. Residual analysis

In Fig. 5, C_{max}/Q residual distributions, for different values of the downwind distance are shown. Results for different density reconstruction methods are qualitatively similar except for RL3 and PL3 that seem to have worse behaviour at smaller distances. In these two applications model performances improve with increasing distance along to 7 km (a similar behaviour was observed on Copenhagen data set). Then model performances, for all kernel methods, get worse at greater distances (more than 20 km).

Finally, a residual analysis has been performed in order to assess PLPM performances in the case of different atmospheric stability classes. In Fig. 6 the output achieved with kernel PL3 are collected (very similar results are obtained by the application of the other kernels): any dependence of quality of results from the stability class is proved.

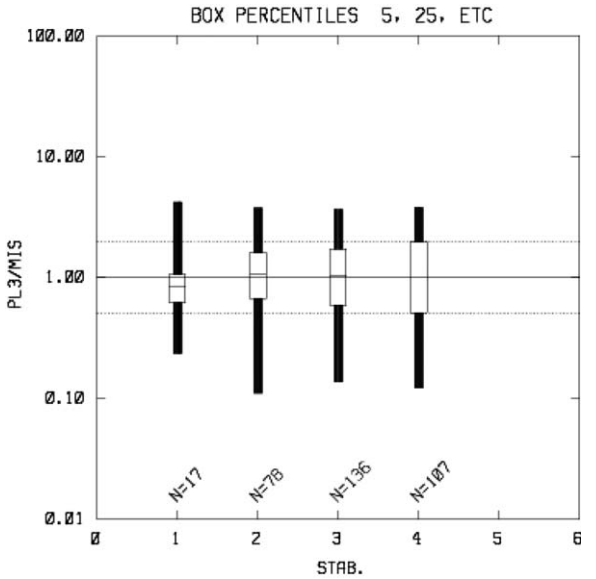

Fig. 6. C_{max}/Q residual distributions in function of stability classes (Kincaid experiment—PL3 density reconstruction employed).

6. Comparison with other models performances

Copenhagen and Kincaid experiments are reference data bases employed in dispersion models evaluation; Tables 5–7 show performance indexes obtained using the Copenhagen and the Kincaid Q3 data sets by a number of well known dispersion models.

A comparison of Tables 2–4 with Tables 5–7, highlights as performance indexes achieved with all kernel methods are comparable with the ones obtained by well-known dispersion models; in some cases, in particular with Kincaid data set, PLPM performs even better. Focusing on the more statistically significant case of Kincaid experiment, it is worth noticing as PLPM, whatever the kernel employed, is the only model showing a slight tendency to overestimate concentrations. From the regulatory and healthcare point of view such a model feature could be considered an asset, suggesting a more conservative and precautionary approach.

7. Discussion and conclusions

The inert version of the Lagrangian particle model PLPM has been presented, tested and validated against two well known data sets dealing with inert pollutants dispersion on flat terrain. The particles dynamics in PLPM adopts the classical Thomson approach based on the non-linear Langevin equation with different turbulence parameterization schemes available. Density reconstruction in PLPM is based on the kernel method, overcoming

Table 5
Performance indexes on Copenhagen data set for normalized maximum concentrations (C_{max}/Q) obtained by a number of well-known dispersion models

($N = 23$)	Mean	Sigma	Bias	NMSE	R	FAC2	FB	FS
Measures	632.66	450.25	0.00	0.00	1.00	1.00	0.00	0.00
HPDM	358.23	268.09	274.42	0.61	0.874	0.652	0.554	0.507
IFDM	551.87	345.27	80.79	0.19	0.843	0.870	0.136	0.264
INPUFF	560.55	352.65	72.10	0.50	0.490	0.739	0.121	0.243
OML	283.61	251.05	349.05	1.12	0.823	0.217	0.762	0.568
UK-ADMS	177.12	138.48	455.53	2.84	0.891	0.043	1.125	1.059

The statistics are reproduced from Olesen (1995a).

Table 6
Performance indexes on Copenhagen data set for normalized cross-wind integrated concentrations (C_y/Q) obtained by a number of well known dispersion models

($N = 23$)	Mean	Sigma	Bias	NMSE	R	FAC2	FB	FS
Measures	448.70	239.29	0.00	0.00	1.00	1.00	0.00	0.00
HPDM	382.32	161.62	66.37	0.16	0.778	1.00	0.160	0.387
IFDM	443.26	193.38	5.43	0.16	0.681	0.957	0.012	0.212
INPUFF	339.59	180.43	109.10	0.46	0.361	0.696	0.277	0.280
OML	249.17	131.70	199.52	0.52	0.893	0.565	0.572	0.580
UK-ADMS	207.05	110.68	241.64	0.86	0.912	0.348	0.737	0.735

The statistics are reproduced from Olesen (1995a).

Table 7
Performance indexes on the Kincaid Q3 data subset obtained by a number of well-known dispersion models

($N = 338$)	Mean	Sigma	Bias	NMSE	R	FAC2	FB	FS
Measures	54.34	40.25	0.00	0.00	1.00	1.00	0.00	0.00
ADMS 3	51.7	34.7	2.7	0.6	0.45	0.67	0.05	0.15
AERMOD	21.8	21.8	32.6	2.1	0.40	0.29	0.86	0.59
ISCST3	30.0	60.0	24.3	2.8	0.26	0.28	0.58	−0.39
NAME	38.7	47.2	15.6	1.45	0.272	0.562	0.335	−0.159

The statistics are reproduced from NAME validation summary (Webster et al., 2002).

the problem of box dimensions setting encountered when dealing with box counting density estimator. As of bandwidths setting, four methods are available in PLPM, one based on global features of particles distribution and three considering local distribution features by means of the definition of a particle (or receptor) neighbourhood.

Validation results are satisfactory and the model achieves performance indexes comparable or better than the ones obtained with other dispersion models, both relying or not on the Lagrangian approach.

The different kernels have different strong and weak points and that it is not possible to address a kernel method as the best or the worst one.

Anyway, all the methods are worth to be implemented even in the case of the chemically active model version, currently under development.

The overall good performance of PLPM, regardless of the kernel employed, confirms the suitability of the physical description of particles dispersion and plume rise implemented. Such a reliability of PLPM in reproducing dispersion dynamics in very diverse atmospheric stability conditions, both for buoyant and non buoyant plumes, is proved in this paper and the encouraging results described are a solid basis to continue in developing the future model improvements.

Acknowledgements

Authors would like to thank S.E. Gryning for his very keen and active support when dealing with the Copenhagen data set. Contribution of ENVIROWARE to this work was granted under contract ENEA/51713/Bologna. The contribution of Valentina Sachero was made possible thanks to a PhD fellowship of Università degli Studi di Milano, Department of Mathematics and Università degli Studi di MIlano Bicocca, DISCo.

References

Anfossi, D., Ferrero, E., Brusasca, G., Marzorati, A., Tinarelli, G., 1993. A simple way of computing buoyant plume rise in Lagrangian stochastic dispersion models. Atmospheric Environment 27, 1443–1451.

Bianconi, R., Mosca, S., and Graziani, G., 1999. PDM: a Lagrangian particle model for atmospheric dispersion. European Communities report 17721 EN, Ispra, Italy.

Bowne, N.E., Londergan, R.J., 1983. Overview, Results, and Conclusions for the EPRI Plume Model Validation and development Project: Plains Site, EPRI Report EA-3074.

Cuvelier C. (Ed.), 1994. Proceedings of the Workshop "Intercomparison of Advanced Practical Short-Range Atmospheric Dispersion Models". August 30–September 3, 1993, (Manno-Switzerland), CSCS (Centro Svizzero di Calcolo Scientifico). Joint Research Centre, European Commission, EUR 15603 EN.

de Haan, P., 1999. On the use of density kernels for concentration estimations within particle and puff dispersion models. Atmospheric Environment 33, 2007–2021.

Davakis, E., Nychas, S.G., Andronopoulos, S., Bartzis, J.G., 2003. Validation study of the dispersion Lagrangian particle model DIPCOT over complex topographies using different concentration calculation methods. International Journal of Environment and Pollution 20, 33–46.

Epanechnikov, V.K., 1969. Non-parametric estimation of a multivariate probability density. Theory of Probability and its applications 14, 153–158.

Gryning, S.E., 1981. Elevated source SF_6 tracer dispersion experiments in the Copenaghen area. Report Risø-R-446, Risø National Laboratory.

Gryning, S.E., Lyck, E., 1984. Atmospheric dispersion from elevated sources in an urban area: Comparison between tracer experiments and model calculations. Journal of Applied Meteorology 23, 651–660.

Gryning, S.E., Lyck, E., 2002. The Copenhagen Tracer Experiments: Reporting of Measurements. Report Risø-R-1054, Risø National Laboratory.

Hurley, P., Physick, W., 1993. A skewed homogeneous Lagrangian particle model for convective conditions. Atmospheric Environment 27A, 619–624.

Jones, M.C., Marron, J.S., Sheather, S.J., 1996. A brief survey of bandwidth selection for density estimation. Journal of American Statistical Association 91, 401–407.

Lorimer, G.S., 1986. A kernel method for air quality modelling-I: Mathematical foundation. Atmospheric Environment 20, 1447–1452.

Lorimer, G.S., Ross, D.G., 1986. The kernel method for air quality modelling—II. Comparison with analytic solutions. Atmospheric Environment 20, 1773–1780.

Luhar, A.K., Britter, R.E., 1989. A random walk model for dispersion in inhomogeneous turbulence in a convective boundary layer. Atmospheric Environment 23, 2311–2330.

Monforti, F., 2001. Numerical treatment of advection and diffusion schemes in atmospheric modeling. PhD thesis, University of Milan.

Olesen, H.R., 1995a. The model validation exercise al Mol: overview of results. Workshop on Operational Short-Range Atmospheric Dispersion Models for Environmental Impact Assessment in Europe, Mol, Belgium, Nov. 1994. International Journal of Environment and Pollution 5, 761–784.

Olesen, H.R., 1995b. Toward the establishment of a common framework for model evaluation. In: Proceedings of the 21st International Meeting on Air Pollution Modelling and its Applications, Baltimore, USA, 1995.

Olesen, H.R., 1998. Model Validation Kit-status and outlook. In: Proceedings of the Fifth Workshop on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purpose, Rhodes, Greece.

Ooms, G., 1972. A new method for the calculation of the plume path of gases emitted by a stack. Atmospheric Environment 6, 899–909.

Peverieri, S., 2000. Kernel density estimators for dispersion models. Graduation thesis. University of Bologna, Italy (in Italian).

Sachero, V., 2003. Software architecture learning in air pollution modelling. Ph.D. Thesis in Mathematics, Statistics, computational sciences and Informatics, Milan State University.

Sachero, V., Vitali, L., Monforti, F., Zanini, G., 2004. PR-PLPM (Plume Rise Photochemical Lagrangian Particle Model): Formulation and Validation of the new plume rise scheme. In: Proceedings of the Ninth International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Garmisch-Partenkirchen, Germany, 2004.

Scire, J.S., Robe, F.R., Fernau, M.E., Yamartino, R.J., 2000. A User's Guide for the CALMET Meteorological Model (Version 5), Earth Tech, Inc.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall Ltd., London.

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

Tinarelli, G., 2005. Private communication.

Webster, H.N., Thomson, D.J., Redington, A.L., Ryall, D.B., 2002. Validation of a Lagrangian model plume rise scheme against the Kincaid data set. In: Proceedings of Eighth International Conference on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes, Sofia, Bulgary, 2002.

Yamada, T., Bunker, S., 1988. Development of a nested grid, second moment turbulence closure model and application to the 1982 ASCOT brush creek data simulation. Journal of Applied Meteorology 27, 562–578.

Yamada, T., Bunker, S., Niccum, E., 1987. Simulations of the ASCOT brush creek data by a nested-grid, second-moment turbulence closure model and a kernel concentration estimator. In: Proceedings of AMS Fourth Conference on Mountain Meteorology, 25–28 August 1987, Seattle, WA.

Yamada, T., Jim Kao, C.-H., Bunker, S., 1989. Airflow and air quality simulations over the western mountainous region with a four-dimensional data assimilation technique. Atmospheric Environment 23, 539–554.