Chapter 4

On the use of Density Kernels for Concentration Estimations Within Particle and Puff Dispersion Models*

Abstract—Stochastic particle models are the state-of-science method for modelling atmospheric dispersion. They simulate the released pollutant by a large number of particles. In most particle models the concentrations are estimated by counting the number of particles in a rectangular volume (box-counting). The effects of the choice of the width and of the position of these boxes on the estimated concentration is investigated. For the estimation of the concentration at a given point in space, it is shown that this numerical procedure can cause either oversmoothed predictions or too much scatter. As an alternative approach, the density kernel method to estimate concentrations is proposed, which optimizes bias and variance. It allows for a reduction of the number of particles simulated for the same accuracy. The efficiency of several density kernel shapes is compared, and methods for choosing their bandwidths are proposed. The relationship between the numerically motivated bandwidths and the description of the growth of a cluster of pollutant particles (puff dispersion) is discussed.

Key words: Kernel density estimation, box-counting, particle models, puff models, relative diffusion.

4.1 INTRODUCTION

At present, stochastic particle or random-walk models are the most advanced approach to simulate atmospheric dispersion, especially for convective atmospheric conditions (for a review, see Wilson and Sawford 1996). In such particle models, the emitted pollutant is simulated by a large number of particles, which are assumed to exactly follow the flow. This approach has several advantages over dispersion models like Gaussian plume or puff models.

---

* this chapter has been accepted for publication (editorial changes may still be made):
de Haan, P. (1999): On the use of Density Kernels for Concentration Estimations Within Particle and Puff Dispersion Models. Accepted for publication in *Atmospheric Environment*

However, simulating the emitted pollutant by many discrete particles brings up one difficulty: the correct estimation of the concentration at a certain location.

In atmospheric dispersion modelling with stochastic particle models, it is common practice to calculate such averages over a grid cell in space, i.e. by counting all particles in a box (see, f.e., Borgas and Sawford 1994; Luhar and Rao 1994; Hurley and Physick 1991; Luhar and Britter 1989; Rotach *et al.* 1996). The estimation of concentration is then obtained by multiplying the number of particles with their mass, and dividing this total mass by the size of the grid box. This way of counting the number of particles in a box is identical to calculating a three-dimensional histogram. Of course, if the volume average over such a box is what the modeler wants, such box-counting methods are most efficient.

Histogram estimations in general depend, however, on the choice of the width and the centre of the averaging interval, area or volume. To estimate point concentrations in the context of atmospheric dispersion modelling, there are no physical restrictions determining either the centre of a numerical averaging volume, or its size (besides the sampling volume of field instruments, which would lead to very small averaging volumes). This means that when choosing large averaging volumes, important details might get lost, and the estimation of the concentration density simulated by the particles will be oversmoothed. On the other hand, when choosing small averaging volumes, one runs at risk having random fluctuations in the number of particles per sampling volume.

The stochastic particle modeller will thus try to choose ”reasonable” sizes and positions of the averaging volumes. The differences in the resulting concentration predictions can be significant (see section 4.2) and of the same order of magnitude as effects which are of interest (i.e. are to be modelled). It should be stressed that this uncertainty originates from a numerical procedure only. Since atmospheric dispersion modelling already has to cope with considerable uncertainties, it is desirable to have a numerical method for concentration estimation within particle models which does not lead to much additional uncertainty.

Another approach to obtain concentrations with stochastic particle models is the backward-modelling approach of Flesch *et al.* (1995). In this approach box-counting is used, but instead of having an exact source location, the backward method has an exact receptor position (and thus sampling box position), and all backward trajectories coming ”near” to (i.e., originating from) the source are related to that source. This approach is suited for area sources together

with small receptor volumes. But even then, in principle the uncertainities related with either box-counting or kernel methods can only be circumvented when the source is located exactly at ground level, and all particles being reflected at the ground surface within the source area being related to the source. Flesch *et al.* (1995) discuss this special case into detail.

Clearly, for very high particle densities, the uncertainties of box-counting vanish; e.g. Borgas and Sawford (1994) theoretically assume the number of particles be ”high enough” for the choice of the dimensions of the particle counting box to be ”negligible”. But at present, computing power still drastically limits the number of particles which are normally emitted. With computing power currently available, only for one-dimensional models the effects might become negligible (see section 4.2).

For these reasons, another method is proposed in this paper. It relies on the concept of density distributions of different shape which are ”added” to the particle’s position, i.e. the mass represented by the particle is spread out in space. Such a density distribution around the centre of mass is called the density kernel. The kernel method has been widely applied since it was introduced in 1950 (for a good review, see Scott 1986, Ch. 6), whereas in atmospheric dispersion modelling the box-counting method is still used despite its deficiencies; Lorimer (1986) was the first, to the author’s knowledge, to use a kernel method for atmospheric dispersion modelling.

In contrary to the choice of box size and position, which makes the box-counts a bad choice for concentration estimation, the shape of the kernel and its width as a function of the particle distribution have to be specified. Of these two, the latter is chosen such that the bias and variance of the concentration estimation are minimised jointly. The number of particles normally simulated at present is more than sufficient to let this method become nearly independent of the shape of the kernel. This makes the kernel density method a suitable approach for particle modellers to estimate the concentrations simulated by their model.

The RAPTAD model of Yamada *et al.* (1987) and Yamada and Bunker (1988) combines a particle model with Gaussian shaped density kernels. They use the physical concept of absolute dispersion to estimate the appropriate size of these kernels. This approach is discussed in more detail in Section 4.6.

The box-counting method to estimate concentrations in particle models is reviewed in section 4.2. The concept of density kernel estimators is introduced in section 4.3, and the important

issue of the determination of the bandwidth of such kernels is treated in section 4.4. Different kernel shapes are compared in section 4.5. The analogy between these numerically determined density kernels and the physical description of the expansion of a cluster of particles is discussed in section 4.6.

4.2 CONCENTRATION ESTIMATION BY BOX-COUNTING

4.2.1 Box-counting in particle models

The smoothing of any given data actually is a non-parametric method to estimate the ”true” curve (being a density or a function), which is polluted by random noise. Most particle model simulations are one- or two-dimensional. The histogram (i.e., box-counting, see above) still is the commonly used density estimator for particle models. In the vertical direction, the size of the intervals (the averaging volume) is often constant with height. For their one-dimensional model, Luhar and Britter (1989) and Luhar *et al.* (1996), with 15 000 and 20 000 particles released, respectively, use vertical box heights Δz of $\Delta z/z_i = 0.05$, and a ”horizontal” extent $\Delta X = 0.1$, where X is the dimensionless travel distance, $X = w_* t/z_i$, and w_* denotes the convective velocity scale. Luhar and Rao (1994) release 12 000 particles per hour in their fully three-dimensional model, and use box dimensions $\Delta x \times \Delta y \times \Delta z$ of $200\ \mathrm{m} \times 50\ \mathrm{m} \times 10\ \mathrm{m}$ for elevated and $200\ \mathrm{m} \times 25\ \mathrm{m} \times 5\ \mathrm{m}$ for surface releases (and take the average over 1 h). Hurley and Physick (1991) have 5000 particles released in a two-dimensional model, and take $\Delta z = 25\ \mathrm{m}$ (with the inversion height $z_i = 1000\ \mathrm{m}$) and $\Delta x = 100\ \mathrm{m}$. Rotach *et al.* (1996) present a two-dimensional model and use 10 000 particles; the dimensions of the averaging boxes are $\Delta z = 20\ \mathrm{m}$ (with inversion height varying from 820 m to 1980 m) and $\Delta x = 300\ \mathrm{m}$.

The amount of additional variability of the model predictions caused by the box-counting method differs. For one-dimensional models and the use of approx. 100 000 particles, this variability seems to be small, whereas for two-dimensional particle models, this same number of particles, when applying Taylor's frozen turbulence hypothesis, has to be simulated for far longer time periods, which is about the limit of today's possibilities. For three-dimensional models, the order of magnitude of the number of particles roughly is the square of the number used in one- and two-dimensional models.

For example, Venkatram and Du (1997) report that with the simulation of 50 000 particles in a one-dimensional model, model results are not sensitive to the size of the numerical sampling

box. On the other hand, Du and Wilson (1994) (who do not mention the number of simulated particles) show concentration contour plots from a two-dimensional model where the size of the sampling volume can still be detected from the figures shown in their paper. The same holds for results from the one-dimensional particle model of Thomson and Montgomery (1994), who use 90 000 to 120 000 particles. This means that for two- and three-dimensional models, the expected increases in computing power in the near future will not be sufficient to rely on box-counting concentration estimations. Anyway, all of the box sizes used by particle modelers mentioned above are averages over volumes far larger than any receptor sampling volume.

Also, Yamada and Bunker (1988) report that determining the concentration at a given time and location by counting the particles simulated by their fully three-dimensional RAPTAD meso-scale particle model yields considerably varying results, depending on the size of the sampling volume and the number of particles used.

4.2.2 Near-source concentration estimation

Two areas can be identified where the box-counting method is most likely to be sensitive. Firstly, concentration predictions for the near-source field, where changes in the particle density are sharp and large proportions of the total number of particles might still be within the volume of few sampling boxes. Secondly, the estimation of surface concentrations, which are of interest for most air pollution simulations, but are likely to show gradients near the surface over distances similar to the vertical extent of box sampling volumes. For concentration predictions in the far field and for smooth particle probability density functions, the box-counting method will produce the same results as any other method, although the kernel method remains more efficient in these cases as well.

As an example to illustrate the variability in concentration predictions while using the box-counting method, simulations have been performed for a non-buoyant release at the height $z_s = z_i/2$, with different widths and positions of the boxes in which the particles are counted (Fig. 1). For the simulation, 1000 particles are emitted per second, the total duration of the simulation was 1000 seconds. The stochastic Lagrangian particle model used fulfils the well-mixed condition of Thomson (1987). All details of the model are outlined in Rotach *et al.* (1996).

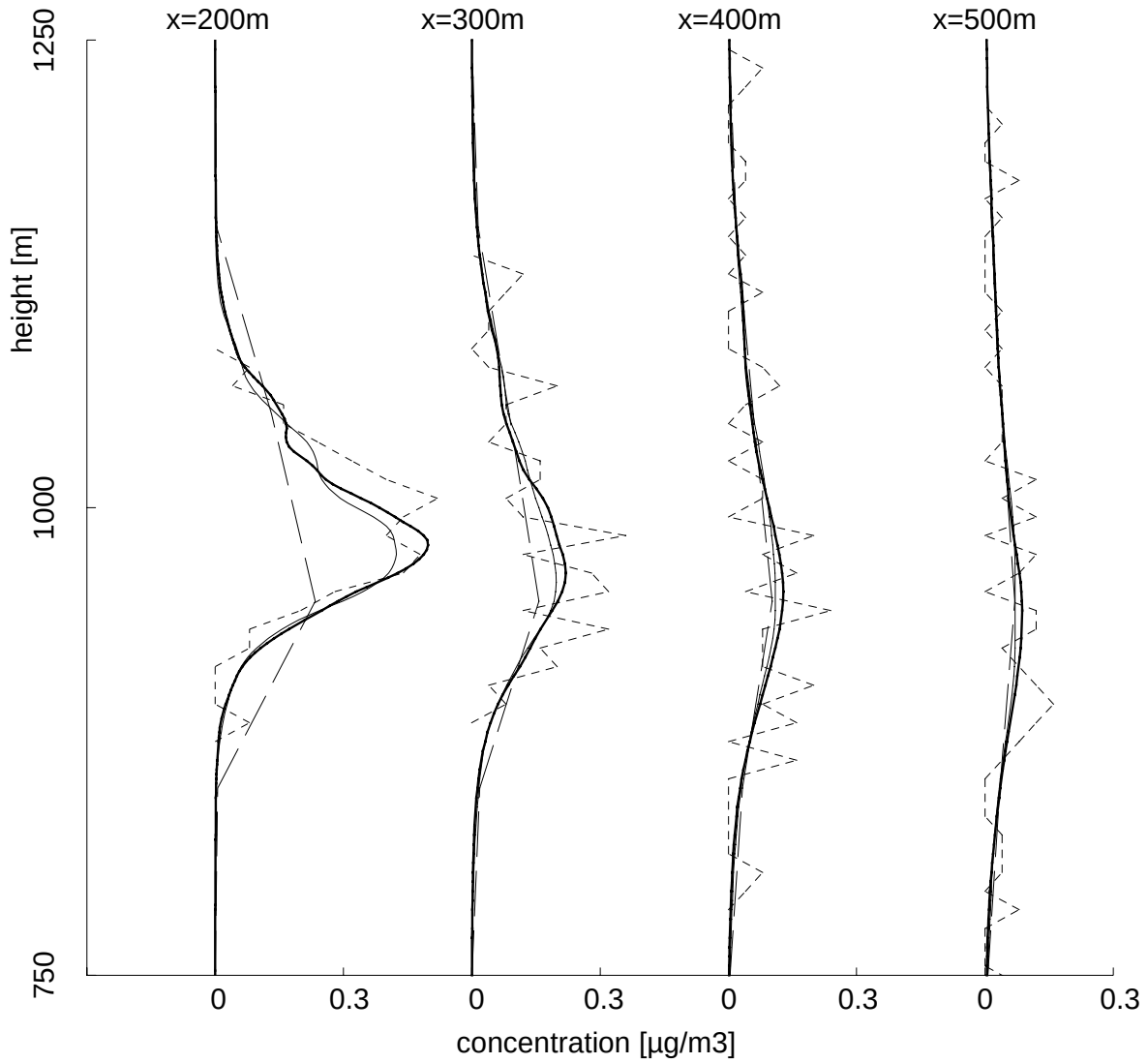

Figure 1 Vertical profiles of estimated concentration (in $\mu g \cdot m^{-3}$) for different sizes of the boxes in which the particles are counted. Profiles are taken at 200 m, 300 m, 400 m, and 500 m downwind of the source. The dashed line (– – –) shows the boxcounting prediction for vertical and horizontal box sizes of $0.05 \cdot z_i$. The dotted line (- - -) is for vertical and horizontal sizes of $0.005 \cdot z_i$. The thick line (▬) depicts the "true" concentration (see text). The thin solid line (—) shows the concentration predition using the kernel method as recommended in section 4.5.5. Release height is at $z/z_i = 0.5$; source strength $1\,g \cdot s^{-1}$; $z_i = 2000$ m, $u_* = 0.36\,m \cdot s^{-1}$, $L = -37$ m, $\overline{u}(z = 10\,m) = 2.1\,m \cdot s^{-1}$.

Concentration predictions for two sampling box sizes are shown in Fig. 1. The vertical size of $0.05 z_i$ (dashed line in Fig. 1) is often chosen by particle modellers (see above). The "true" concentration profile, based on 100 000 instead of 1000 particle trajectories, is depicted by a thick line in Fig. 1. It becomes clear that the near-source maximum concentration will be clearly underestimated for box sizes of $0.05 z_i$. The other density estimation shown in Fig. 1,

with box size $0.005z_i$ (short-dashed line), makes the concentration prediction fluctuating from one box volume to the other. However, the true value of concentration is approximated more adequately by the smaller box sampling volumes, despite the additional scatter. An additional smoothing procedure would improve the results (see, for example, Rotach *et al.* 1996). As an example, the results from a kernel density estimation following the recommendations of section 4.5.5 of this paper are also depicted in Fig. 1 (thin line).

4.2.3 Ground level concentration estimations

In Fig. 2 an example for the effects of box-counting on predicting ground-level concentrations is given. In this numerical experiment, the release took place at a height of 5 m, i.e. within the

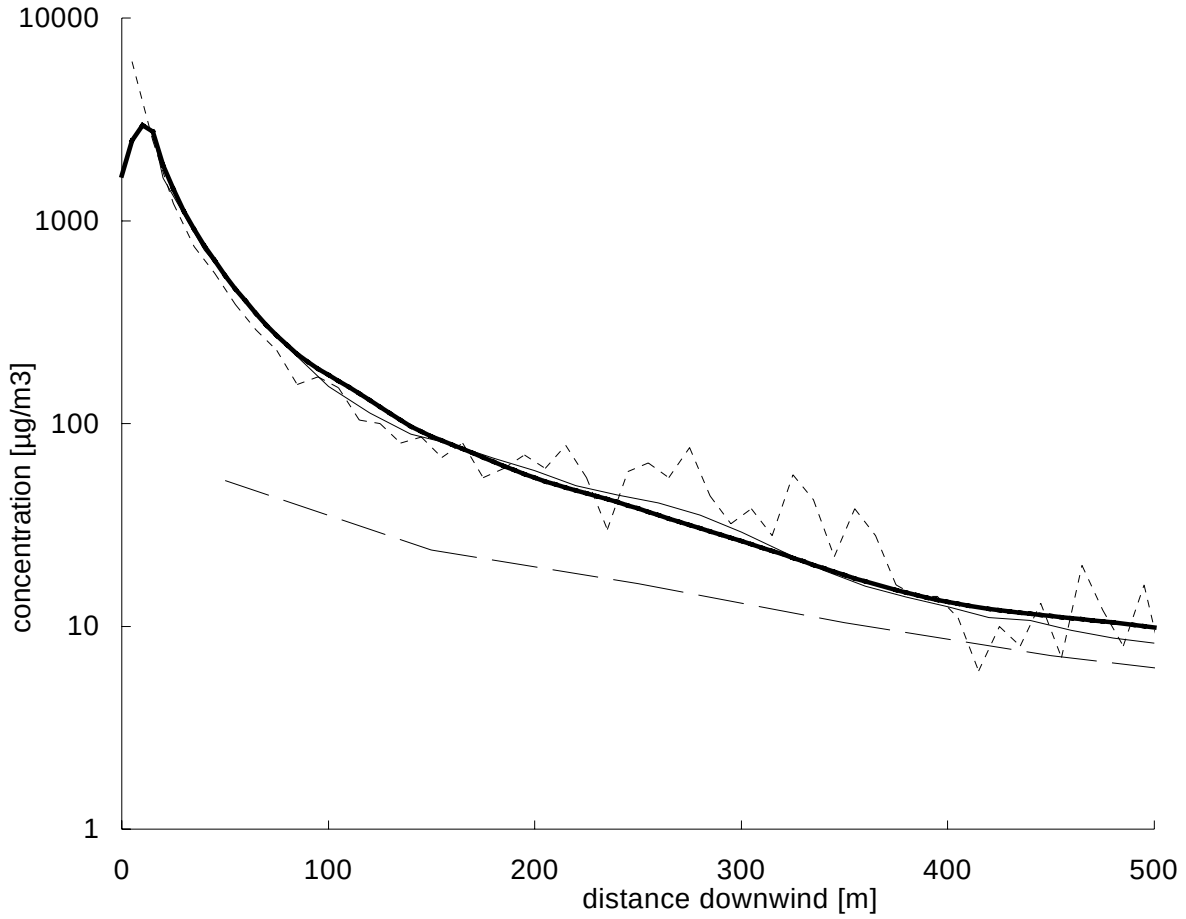

Figure 2 Same as Fig. 1, but for down-wind ground-level profiles of estimated concentration. Source is located at $z = 5$ m .

lowest sampling box. As for Fig. 1, results for vertical box sizes of $0.05z_i$ (dashed line) and $0.005z_i$ (dotted line) are depicted. Also shown are the ”true” concentration based on 100 000

particles (thick line) and the results from a triweight density kernel constructed after section 4.5.5 (thin line).

For the (widely used) $0.05z_i$ vertical box size, the ground-level concentrations, in this example, are oversmoothed, i.e. underestimated, by a factor of approximately 2 in the near-source field. The box size in the horizontal directions has been chosen equal to the vertical box size, leading to a first box center at a downwind-distance of approx. 5m ($0.005z_i$ box size) and 50m ($0.05z_i$ box size), respectively. For the smaller box with $0.005z_i$ vertical size, the estimates show a large random scatter (large variance) but no bias. The kernel estimation, on the other hand, provides a smooth estimate with almost no bias.

4.3 DENSITY KERNELS

4.3.1 Concentration estimation with density kernels

The kernel density estimator for the normalised concentration *c* of *n* given particles of equal mass at positions $\mathbf{x}_i$ is

$$c(\mathbf{x}) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{\mathbf{x}-\mathbf{x}_i}{h}\right), \tag{1}$$

where *h* is the width of the kernels (see section 4.4), and *K* is the kernel function, fulfilling $K(\mathbf{x}) \geq 0 \quad \forall \mathbf{x}$, and normalised so that

$$\int K(\mathbf{x})d\mathbf{x} = 1, \tag{2}$$

(thus making *c* a density distribution, i.e., $\int c(\mathbf{x})d\mathbf{x} = 1$). One of the most widely used kernels is the Gaussian kernel K_G,

$$K_G(\mathbf{r}) = \frac{1}{(2\pi)^{d/2}}\exp\left(-\frac{1}{2}\mathbf{r}^T\mathbf{r}\right), \tag{3}$$

where *d* denotes the dimension (for example, $d = 2$ for two-dimensional particle models).

The proper choice of the bandwidth *h* is of greater importance than the choice of the shape of the kernel (additional kernel shapes are introduced in section 4.3.2), since *h* plays the role of a smoothing parameter. There are several proposals for determining the value of *h* from the data. These methods aim at ”balancing” two counteracting measures, the bias (caused by the smoothing, which increases for large *h*) and the variance (increasing for small *h*).

In the following, the underlying distribution of particle positions $\mathbf{x}_i$ will be denoted as $f(\mathbf{x})$. Then, the measure on the over-all accuracy of the concentration estimate $c(\mathbf{x})$ adopted here, which is an estimator of f, is the mean integrated square error,

$$MISE(c) = \left[\int \left\{c(\mathbf{x}) - f(\mathbf{x})\right\}^2 d\mathbf{x}\right] \tag{4}$$

where the $[\ ]$-brackets denote the average over all possible data obeying $f(\mathbf{x})$. The right hand side of Eq. (4) can be formulated as

$$MISE(c) = \int \left\{\left[c(\mathbf{x})\right] - f(\mathbf{x})\right\}^2 d\mathbf{x} + \int \mathrm{var}\left\{c(\mathbf{x})\right\} d\mathbf{x} \tag{5}$$

i.e., as the sum of the integrated square bias and the integrated variance. Expressing $c(\mathbf{x})$ by Eq. (1) and taking the first term of the Taylor series expansion of $f(\mathbf{x})$, the integrated square bias can be approximated by (see Silvermann, 1986, for details)

$$\int \left\{\left[c(\mathbf{x})\right] - f(\mathbf{x})\right\}^2 d\mathbf{x} \approx \frac{1}{4} h^4 \alpha^2 \int \left\{\nabla^2 f(\mathbf{x})\right\}^2 d\mathbf{x} \tag{6}$$

and the integrated variance

$$\int \mathrm{var}\left\{c(\mathbf{x})\right\} d\mathbf{x} \approx \frac{\beta}{nh^d} \tag{7}$$

where $\beta = \int K^2(\mathbf{x}) d\mathbf{x}$ and $\alpha = \int x_1^2 K(\mathbf{x}) d\mathbf{x}$ $\left(\mathbf{x} = \left[x_1, x_2, x_3\right]^T\right)$. An "optimal" bandwidth h_{opt} for use in Eq. (1) can then be found by using the approximations (6) and (7) in Eq. (5), which becomes minimal for

$$h_{opt} = A(K) \cdot \sigma \cdot n^{-\frac{1}{d+4}} \tag{8}$$

for a particle distribution $f(\mathbf{x})$ with standard deviation σ (see section 4.4 for details). Thus the optimal bandwidth, h_{opt}, itself depends on the distribution of the data which are to be smoothed, *f*. The function $A(K)$ is defined as

$$A(K) = \left[d \frac{\beta}{\alpha^2} \left\{\int \left(\nabla^2 f\right)^2\right\}^{-1}\right]^{\frac{1}{d+4}} \tag{9}$$

In this paper, we assume that *f* be normal. Then,

$$\left\{\int \left(\nabla^2 f\right)^2\right\} = \frac{d(2+d)}{4\left(2\sqrt{\pi}\right)^d} \tag{10}$$

For Gaussian kernels K_G (Eq. 3),

$$\frac{\beta}{\alpha^2} = \left(\frac{1}{2\sqrt{\pi}}\right)^d . \tag{11}$$

So for Gaussian kernels, substituting Eqs. (10) and (11) into (9) leads to $A(K_G) = (4/[d+2])^{+1/(d+4)}$. The values of α and β for all other kernels discussed in the following sections are derived in Appendix A.

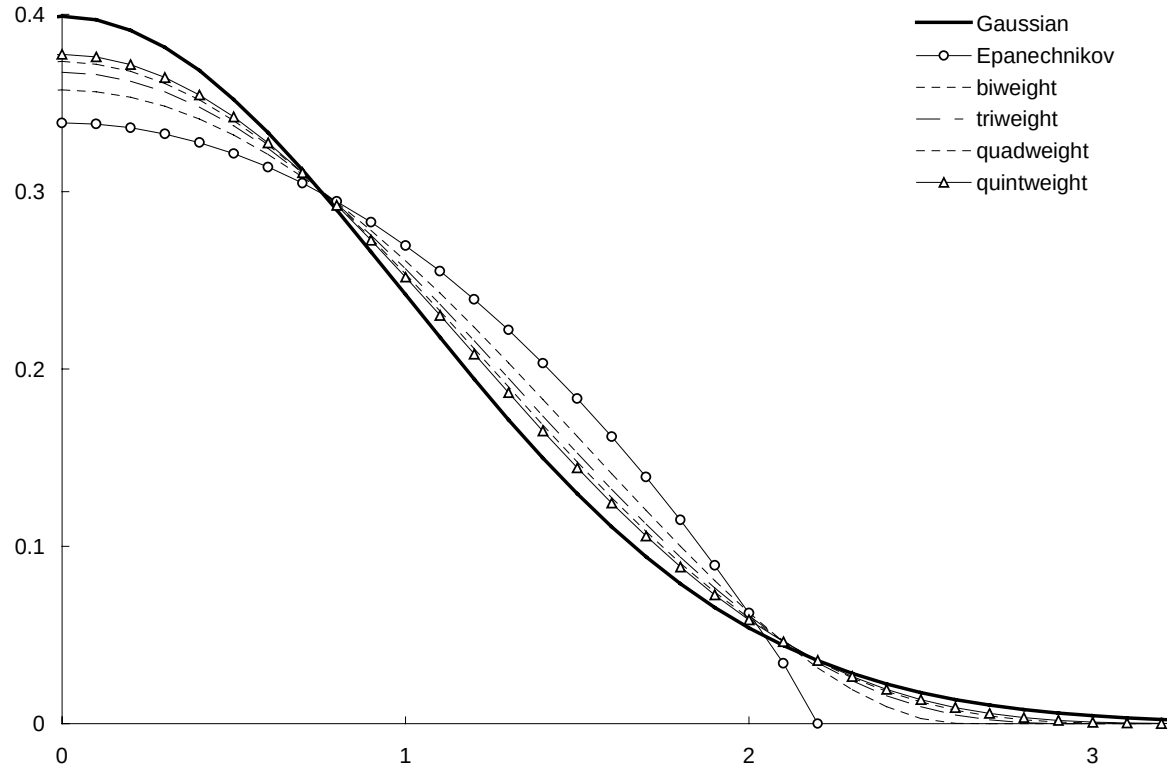

Figure 3 Comparison of one-dimensional Gaussian, Epanechnikov, bi-, tri-, quad- and quintweight kernels. The bandwidth of the Gaussian kernel is unity, and taken from Table II for the other kernels (see section 4.4.2). (Only the part for positive x is depicted; all kernels are symmetric around zero.)

4.3.2 Suitable kernels for particle models

In this paper, one family of kernels other than Gaussian is investigated, that of the form

$$K_a(\mathbf{x}) = \begin{cases} C_{d,a}\left(1-\mathbf{x}^T\mathbf{x}\right)^a & \left(\mathbf{x}^T\mathbf{x}<1\right) \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

In Eq. (12), $C_{d,a}$ are normalising factors ensuring that $\int K_a d\mathbf{x} = 1$, which are derived in Appendix A (see Table I). For $a=1$ up to $a=5$, these kernels are called Epanechnikov and bi-, tri-, quad- and quintweight kernels, respectively. Figure 3 depicts their one-dimensional

form, with different bandwidths h for every kernels. The relations between values of h for different kernels having approximately the same smoothing effect are given in section 4.4.2.

$C_{d,a}$	$a=1$	$a=2$	$a=3$	$a=4$	$a=5$
$d=1$	$\frac{3}{2c_1}$	$\frac{15}{8c_1}$	$\frac{105}{48c_1}$	$\frac{945}{384c_1}$	$\frac{693}{256c_1}$
$d=2$	$\frac{2}{c_2}$	$\frac{3}{c_2}$	$\frac{4}{c_2}$	$\frac{5}{c_2}$	$\frac{6}{c_2}$
$d=3$	$\frac{5}{2c_3}$	$\frac{35}{8c_3}$	$\frac{315}{48c_3}$	$\frac{1155}{128c_3}$	$\frac{3003}{256c_3}$

Table I Normalisation constant $C_{d,a}$ (Eq. alt 9) for Epanechnikov $(a=1)$, biweight $(a=2)$, triweight $(a=3)$, quadweight $(a=4)$ and quintweight $(a=5)$ kernels for up to three dimensions ($c_1=2$, $c_2=\pi$, $c_3=4\pi/3$).

Note that the support (the *d*-dimensional interval over which the kernel function is non-zero) is infinite for the Gaussian kernel (Eq. 3) and limited for the other five types (Eq. 12). The Epanechnikov kernel ($a=1$) is optimal in minimising the MISE (Epanechnikov 1969), but its derivative is discontinuous. As can be seen from Fig. 3, the differences between the Gaussian kernel and kernels of the $\left(1-\mathbf{x}^T\mathbf{x}\right)^a$-type decrease for increasing *a*, the only significant differences occurring at the tails of the distribution. Because of the small differences between the kernel shapes, other criteria such as the ease of computation are important as well as the question, which of the kernel functions resembles most the distribution of observed turbulent wind fluctuations. The latter are generally assumed to be normally distributed (with the exception of the vertical component for convective conditions), making the Gaussian kernel the most natural choice. However, the assumption of a normal distribution of wind fluctuations attributes a non-zero possibility (though very small) to unphysically large values. This would support the choice of a kernel with a limited support interval such as, say, the triweight kernel.

Another approach is the kernel method adopted by Lorimer (1986). He uses the product of a three-dimensional Gaussian with a three-dimensional Epanechnikov kernel. Because this kernel is still defined over an infinite domain, we only compare kernels after Eq. (12) with the Gaussian kernel.

It should be kept in mind that the concentration estimation using density kernels within particle models is a numerical task, and thus any form of kernel function can be used in principle. But the analogies which show up between these numerically motivated kernel density estimators and the concept of clusters of particles as adopted in puff models is remarkable, and is further discussed in section 4.6. Recommended kernel functions which perform well numerically and have a form which can be compared to physical observations are given in section 4.5.5.

4.3.3 Three-dimensional vs. product kernels

The approach formulated in section 4.3.1 is *d*-dimensional, and thus yields a single bandwidth for *d*-dimensional kernels. Since atmospheric dispersion in the logitudinal, cross-wind and vertical directions is, from the point of view of density estimation, very similar, this approach is justified.

In other scientific areas, however, often three-dimensional density kernels are being used which are the product of one-dimensional kernels. In such so-called product kernels, instead of Eq. (1),

$$c(\mathbf{x}) = \frac{1}{n h_x h_y h_z} \sum_{i=1}^{n} \left\{ \prod_{j=1}^{3} K\left(\frac{x_i - x_{ij}}{h_j} \right) \right\} \tag{13}$$

where the bandwidths h_i of the kernel distributions in the three directions are calculated in analogy to the one-dimensional case. Product kernels have the advantage of easier mathematics, and are easier to handle in the case of data points with large differences in the standard deviations in the three directions (dimensions).

For Gaussian kernels, Eqs. 1 and 13 are identical. This relationsship does not hold any longer for Eq.-(12)-type kernels. Fig. 4 depicts the differences in density distribution between a three-dimensional triweight (using Eq. 1) and the corresponding product kernel (Eq. 13). The product kernel formulation leads to higher densities in the edges and in the center of the carrier domain, while having smaller densities in the other regions. In the field of atmospheric dispersion, diffusion is approximately equal in all directions and no “edges” with higher probabilities do exist. This means that product kernels are less suited for atmospheric dispersion data than fully three-dimensional kernels. Therefore, only the latter are investigated in this paper. But for applications where the standard deviations of the data points

substantially differ according to direction, product kernels may avoid the necessity of scale transformations.

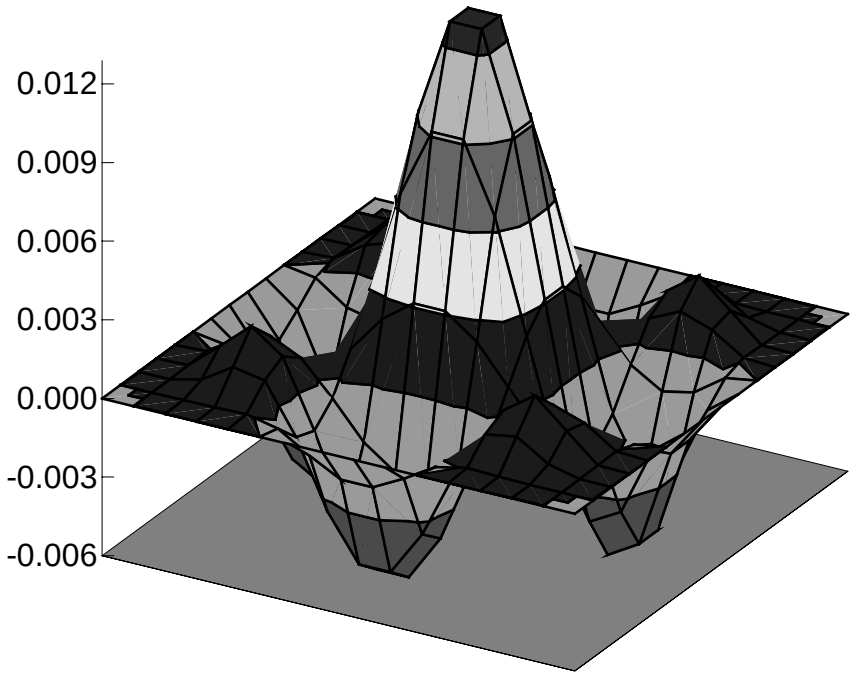

Figure 4 Density differences for a two-dimensional triweight vs. the corresponding product kernel.

4.4 BANDWIDTH ESTIMATION

4.4.1 Theory

The optimal bandwidth as given by Eq. (8) is only optimal under the assumption that the distribution of the data be normal. Therefore, more robust rules-of-thumb have been constructed. A recommended one is

$$h_{rob.} = \alpha \cdot A(K) \cdot n^{-\frac{1}{d+4}} \min\left(\sigma, \frac{R}{1.34}\right), \tag{14}$$

where R is the interquartile range, i.e. $R = R_{3/4} - R_{1/4}$, and α is taken as 0.85 in the present study. The upper quartile $R_{3/4}$ and lower quartile $R_{1/4}$ are defined as those values out of the data to be smoothed where 25% of the data has a larger or smaller value, respectively.

Lorimer (1986) used a slightly different approach to determine the kernel bandwidth. The bandwidth for the initial particle distribution was calibrated by visual comparision of the resulting density distribution. This provides an approximate value for the term $A(K)n^{-1/(d+4)}$ in Eq. (14), which remains constant. Then, Lorimer (1986) calculates the new standard deviation of the particle ensemble for every new time step, and thus calculates the bandwidth using Eq. (14).

At this point, the idea of having an estimate $\hat{\sigma}$ for the value of σ might come up. This would circumvent to some extent the fact that the optimal bandwidth depends on the density distribution of the particles to be smoothed, and would save computing time, as the standard deviation of the given particle distribution has not to be calculated. However, deriving expressions for $\hat{\sigma}$ means to theoretically solve the problem of atmospheric dispersion, and then no numerical dispersion model would be required as soon as correct, i.e. bias-free $\hat{\sigma}$ expressions (as a function of travel time or of downwind distance) are available.

4.4.2 Equivalent bandwidths for different kernels

It is possible to switch between kernels without reconsidering the optimal bandwidth. As soon as an optimal bandwidth has been chosen for one type of kernel, bandwidths for other kernels which show the same smoothing characteristics (called ”equivalent bandwidths”can be derived using Eq. (8)). The kernel dependence of the bandwidth is expressed by the function $A(K)$, and only the β/α^2 -part (Eq. 11) depends on the kernel itself. The ratio between two equivalent bandwidths for two kernels K_1 and K_2 of the same dimensionality is

$$\frac{h_1}{h_2} = \left[\frac{\beta_1/\alpha_1^2}{\beta_2/\alpha_2^2} \right]^{1/(d+4)} \tag{15}$$

For $d = 1$ and $d = 3$, Table II lists the corresponding conversion factors between equivalent kernel bandwidths.

	From/To	Gaussian	Uniform	Epanech.	Biw.	Triw.	Quadw.	Quintw.
one-	Gaussian	1	1.740	2.214	2.623	2.978	3.296	3.586
dimen-	Uniform	0.575	1	1.272	1.507	1.711	1.894	2.061
sional	Epanech.	0.452	0.786	1	1.185	1.345	1.489	1.620
	Biw.	0.381	0.663	0.844	1	1.136	1.257	1.367

	Triw.	0.336	0.584	0.743	0.881	1	1.107	1.204
	Quadw.	0.303	0.528	0.672	0.796	0.904	1	1.088
	Quintw.	0.279	0.485	0.617	0.731	0.830	0.919	1
three-	Gaussian	1	2.220	2.572	2.924	3.244	3.537	3.808
dimen-	Uniform	0.450	1	1.158	1.317	1.461	1.593	1.715
sional	Epanech.	0.389	0.863	1	1.137	1.261	1.375	1.481
	Biw.	0.342	0.759	0.880	1	1.109	1.210	1.302
	Triw.	0.308	0.684	0.793	0.901	1	1.090	1.174
	Quadw.	0.283	0.628	0.727	0.827	0.917	1	1.077
	Quintw.	0.263	0.583	0.675	0.768	0.852	0.929	1

Table II Bandwidth conversion factors for different kernel shapes. To obtain equivalent smoothing with the kernel shape K2 as with kernel K1, multiply the K1 bandwidths with the factor from row K1, column K2. Also shown are results for a uniform kernel.

4.4.3 Kernels with variable bandwidth

The use of individual bandwidths for every kernel rather than using the same h_{opt} for all kernels allows for adequate smoothing where particle density is high, whilst minimising the variance of the estimation in regions of low simulated particle densities. This is of special use for near-source as well as ground level concentration estimations, e.g. the most natural field of use of density kernels in atmospheric dispersion modelling.

For every kernel an individual bandwidth $h \cdot \lambda_i$ is used. To obtain the λ_i, a first-guess estimator $\hat{c}(\mathbf{x})$ is used (for example using h_{opt} from Eq. 8). The λ_i are then chosen proportional to $\hat{c}(\mathbf{x}_i)^{-\alpha}$, where $\mathbf{x}_i$ denotes the position of the *i*-th particle. Silvermann (1986) recommends a value $\alpha = 0.5$. However, such variable bandwidths are not investigated in this paper.

4.5 KERNEL SENSITIVITY

4.5.1 Kernel shape comparison

The effect of different shapes of density kernels is investigated and compared to the ”true” concentration prediction within a Lagrangian particle model. This ”true” concentration is obtained by using a very large number (500 000) of particle trajectories for the full three-dimensional model simulation. For 500 000 particles, the effect of kernel shape vanishes (no difference larger than 1.1%), and the arithmetric mean of the bi-, tri- and quadweight kernel results has been used as the ”true” estimate. These ”true” concentration predictions are then

compared to estimations by different kernel methods, for different numbers of particles released (500, 5000 and 50 000).

The meteorological conditions and the set-up of the Copenhagen tracer experiment are used for comparison. This experiment took place under atmospheric conditions of forced convection. Measurements of wind speed, friction velocity, wind fluctuation standard deviations, Obukhov length and mixing layer depth are available (see Gryning and Lyck, 1984, for more details on the experiment). The tracer was released from a 115 m high stack, and sampled at arcs of ground-level receptors typically 2 to 6 km down-wind from the source. The averaging time is one hour. For the present kernel shape comparison, the simulated concentrations have been evaluated for those locations where the receptors have been placed during the Copenhagen experiment. This was done to obtain a typical application for a Lagrangian particle dispersion model.

For the numerical simulation, the Lagrangian particle model of Rotach *et al.* (1996) is used, which has a probability density function (pdf) suitable for neutral to convective conditions. It has been extended to three dimensions (de Haan and Rotach 1998b) and fulfils the well-mixed condition of Thomson (1987). However, in the validation discussed in this section, the focus is not on finding the kernel shape obtaining the highest correspondence with the measured concentration values. Instead, the kernel concentration estimates are compared to the ”true” concentrations yielded for the simulation of 500 000 particles (see above). This way, we will be able to know which kernel shape has the lowest bias for a given number of released particles, and is best suited for telling the particle modeller what the model actually predicted. The near-source field is investigated in section 4.5.3, and the intermediate field (roughly between 1 and 6 km down-wind from the source) in section 4.5.4.

4.5.2 Experimental set-up

The simulations are performed by essentially rebuilding the experimental set-up. At the mixing height z_i as well as at the ground, perfect reflection without any entrainment or deposition of pollutants is assumed. To ensure that the inverse of the time steps for the simulation lies in the inertial subrange for the particle part of the model, the criteria after Rotach *et al.* (1996) are applied. The mean wind profile was determined based on an approach by Sorbjan (see Rotach *et al.* 1996 for details).

To ensure mass conservation, for the kernel concentration distributions belonging to each particle, image sources are assumed below the ground and above the inversion height. For Gaussian shaped kernels K_G (Eq. 3), six image sources were placed, i.e., at $z = -h_s$, $2z_i - h_s$, $-2z_i - h_s$, $4z_i + h_s$ etc., where h_s is the actual height of the particle. This is necessary to ensure that no mass gets lost (maximum loss is 0.10%). For the non-Gaussian kernel shapes with a limited carrier interval, only 4 image sources (2 below the ground, 2 above z_i) are needed at maximum (maximum loss is 0.06%).

For the present comparison, all data from the Copenhagen experiment (9 runs) have been simulated.

4.5.3 Near-source predictions

In Fig. 5 results are shown for one of the experimental runs. The ”true” estimate based on 500 000 particles acts as ”reference”. The effect of the number of particles upon the ground-level concentrations is larger than the choice of kernel shape. This result had to be expected, since a kernel, and its shape, is only a numerical method of translating particles positions into concentrations, whereas a tenfold increase in the number of simulated particles is likely to significantly increase model performance.

The kernel method will always produce ”smooth” concentration profiles, even for very low numbers of particles, because the kernel bandwidths adapt to the number of particles. For a similar effect using the box-counting method, it would be necessary to automatically determine the sampling box size from the number of particles. But the amount of unwillingly introduced additional smoothing (bias), i.e., artificial dispersion, will be higher for lower particle numbers. This will lead to an earlier and steeper increase of ground-level concentration near the source, and a faster decline of ground-level concentration after reaching the concentration maximum (Fig. 5).

The interesting question is which kernel shape approaches the ”true” estimate faster, that is, performs best for a given number of particles. To assess the near-source prediction performance, four ratios are calculated: The ratio of predicted to ”true” maximum ground-level concentration, $R_1 = \chi^{\max}_{pred.} \big/ \chi^{\max}_{'true'} - 1$: the corresponding ratio of down-wind distances,

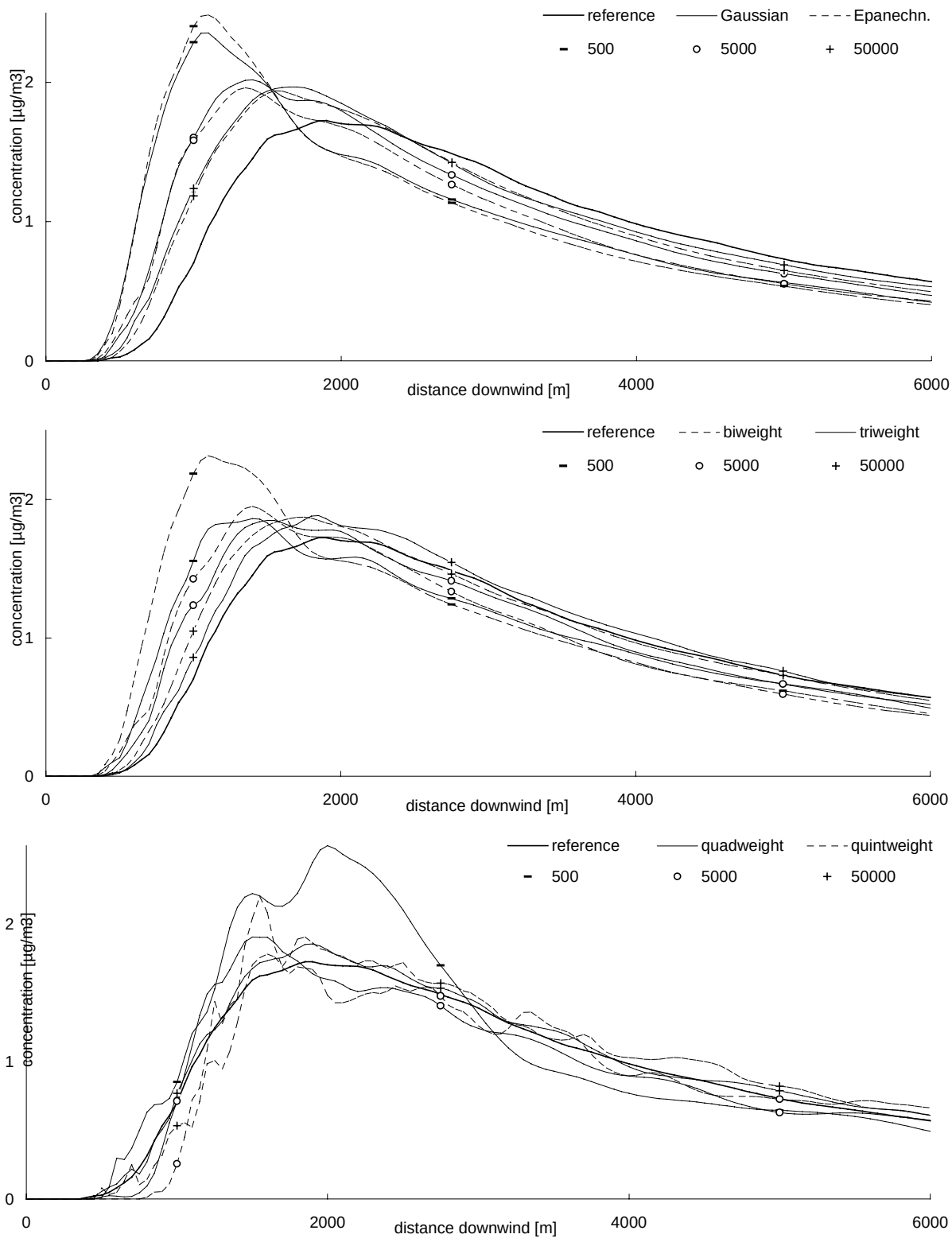

Fig. 5 Intercomparison of estimated ground-level concentration for different kernel types and numbers of particles. (a) Gaussian and Epanechnikov kernels; (b) bi- and triweight kernels; (c) quad- and quintweight kernels. Thick solid line (—): estimate for 500 000 particle trajectories using triweight kernels. Simulations with 500 particles are marked with minus signs (–), 5000 particles with circles (o), 50 000 particles with plus signs (+). Copenhagen exp. from Jun. 27, 1979 ($z_i = 1850$ m, $u_* = 0.64\ \mathrm{m \cdot s^{-1}}$, $L = -104$ m, $\bar{u}(z = 10\ \mathrm{m}) = 4.1\ \mathrm{m \cdot s^{-1}}$).

$R_2 = x^{\max}_{pred.} \big/ x^{\max}_{'true'} - 1$ (where the distances are defined by $\chi\left(x^{\max}_{pred.}\right) = \chi^{\max}_{pred.}$ and $\chi\left(x^{\max}_{'true'}\right) = \chi^{\max}_{'true'}$); the ratio of the down-wind plume-centerline integrated "over-/underpredicted" concentration to the total down-wind integrated "true" concentration, i.e.,

$$R_3 = \int_0^\infty \left|\chi_{pred.}(x) - \chi_{'true'}(x)\right| dx \Big/ \int_0^\infty \chi_{'true'}(x)\,dx \tag{16}$$

and a ratio which measures the "smoothness" of the prediction,

$$R_4 = \int_0^\infty \left|\frac{\partial}{\partial x}\chi_{pred.}(x)\right| dx \Big/ \int_0^\infty \left|\frac{\partial}{\partial x}\chi_{'true'}(x)\right| - 1 \tag{17}$$

Unity is subtracted from R_1, R_2 and R_4 so that all ratios are zero for a perfect prediction.

The cross-wind integrated plume-centerline groundlevel concentrations for experiment 7 of the Copenhagen data set were compared. The resulting ratios are listed in Table III. Two groups can be distinguished: the Gaussian, Epanechnikov and quintweight kernel on the one hand, and the bi-, tri- and quadweight kernels on the other hand. However, the first group is heterogeneous, since the performance of the Gaussian and Epanechnikov kernels is worsened because of "oversmoothing" (the maximum concentration is overpredicted, and too close to the source, which leads to higher values of R_1 and R_2, respectively). The performance of the quintweight kernel is most influenced by its "undersmoothing", i.e., the oscillating prediction (scatter) (see Fig. 5c).

Out of Table III, two tendencies can be distinguished. First, for an increasing number of particles, the quadweight kernel takes over the role of best predicting kernel from the triweight kernel. Second, the influence of the shape of the kernel vanishes for higher numbers of simulated particles, and is in fact negligible for the simulation of 50 000 particles. As a result, the triweight kernel can be recommended as the best kernel shape for near-source predictions; for high numbers of particles, the quadweight kernel performs slightly better.

4.5.4 Intermediate-field predictions

The intermediate-field prediction performance of the different kernel shapes is compared by calulating widely used statistical measures. From the predicted concentrations at the individual receptor locations, the arcwise maximum concentration and the cross-wind integrated concentration (on the arc) are determined. As statistical measures to describe the model performance, the fractional bias, $FB = \left(\overline{C}_{\text{"true"}} - \overline{C}_{\text{pred.}}\right) \Big/ \left(0.5\left(\overline{C}_{\text{"true"}} + \overline{C}_{\text{pred.}}\right)\right)$, and the

		Gaussian	Epanechn.	Biweight	Triweight	Quadw.	Quintw.
500 part.	R_1	0.3138	0.3875	0.2939	**0.0407**	0.4342	0.1836
	R_2	–0.4268	–0.4157	–0.3846	–0.2564	**0.0623**	0.1201
	R_3	0.2987	0.3441	0.2439	**0.0941**	0.1196	0.1379
	R_4	0.4182	0.5085	0.3792	**0.0762**	0.5844	1.1997
	sum	1.4575	1.6558	1.3017	**0.4674**	1.2006	1.6413
5000 part.	R_1	0.1270	0.0951	0.0889	**0.0331**	0.0641	0.2274
	R_2	–0.2654	–0.2829	–0.2564	–0.2102	**–0.2033**	–0.2267
	R_3	0.1076	0.1180	0.0801	0.0447	**0.0289**	0.0344
	R_4	0.1798	0.1525	0.1409	**0.0533**	0.0963	0.4377
	sum	0.6798	0.6486	0.5663	**0.3413**	0.3926	0.8762
50000 part.	R_1	0.0984	0.0438	0.0461	0.0528	**0.0350**	0.0634
	R_2	–0.1236	–0.2033	–0.0705	–0.0173	**0.0092**	–0.0273
	R_3	0.0488	0.0488	0.0241	0.0126	**0.0091**	0.0170
	R_4	0.1261	0.0703	0.0581	0.0581	**0.0300**	0.2055
	sum	0.3968	0.3661	0.1988	0.1408	**0.0833**	0.3132

Table III Ratios for the near-source performance of different concentration estimation kernel methods (using Gaussian, Epanechnikov, bi-, tri-, quad- and quintweight kernels) for the release of 50 000, 5000 or 500 particles, respectively. The reference estimation was obtained by releasing 500 000 particles. For definitions of $R_1 - R_4$ see text. The lowest absolute value in each row is shown in bold typeface. The sum denotes the sum of the absolute values of the ratios.

normalised mean square error, $NMSE = \overline{\left(C_{\text{"true"}} - C_{\text{pred.}}\right)^2} \Big/ \left(\overline{C}_{\text{"true"}} \overline{C}_{\text{pred.}}\right)$, are determined, where $C_{\text{"true"}}$ is the "true" (reference) concentration and $C_{\text{pred.}}$ the simulated one.

The behaviour found for the single run 7 as depicted in Fig. 5 can be found for all 9 runs. In Table IV, the NMSE and FB statistical measures for the performance on the whole data set are listed. Note that the statistical measures are given relative to the "true" estimate, not to field measurements.

Several conclusions can be drawn from the results in Table IV. First, even for 500 particles, the NMSE and FB values resulting are not that large. Other assumptions underlying the dispersion model could cause effects of a comparable magnitude. This means that for intermediate to far-field concentration predictions with particle models, a few thousand particles will do the job even in three-dimensional models (with kernel density estimation). Second, for all three different numbers of simulated particles, the quadweight kernel performs best, and the triweight kernel second best. These two kernel shapes also performed well in the near-source assesment. Third, for 50 000 simulated particles, the differences in performance

between different kernel shapes are minor, and certainly not significant anymore: regarding the low values of NMSE and FB, it is hard to determine which kernel actually performs best.

		ArcMax		CIC	
		NMSE	FB	NMSE	FB
500 000	reference	0.000	0.000	0.000	0.000
50 000	Gaussian	0.009	0.014	0.014	–0.008
	Epanechnikov	0.008	0.045	0.004	0.011
	biweight	0.005	0.029	0.002	0.008
	triweight	0.004	–0.021	0.002	–0.026
	quadweight	0.003	–0.039	0.001	–0.021
	quintweight	0.019	–0.083	0.006	–0.057
5000	Gaussian	0.020	0.068	0.007	0.001
	Epanechnikov	0.034	0.101	0.011	0.042
	biweight	0.022	0.069	0.007	0.029
	triweight	0.013	0.020	0.003	0.007
	quadweight	0.010	–0.003	0.003	0.007
	quintweight	0.060	–0.129	0.003	–0.034
500	Gaussian	0.136	0.253	0.033	0.107
	Epanechnikov	0.151	0.261	0.036	0.104
	biweight	0.101	0.206	0.026	0.084
	triweight	0.037	0.116	0.016	0.060
	quadweight	0.033	0.007	0.010	0.007
	quintweight	0.184	–0.249	0.019	–0.023

Table IV Statistical measures for the performance of different concentration estimation kernel methods (using Gaussian, Epanechnikov and bi-, tri-, quad- and quintweight kernels) for the release of 50 000, 5000 or 500 particles, respectively. The reference estimation is the arithmetic mean of the results obtained for 500 000 particles for bi-, tri- and quadweight kernels.

The results listed in Table IV show that the NMSE and FB values are reduced by a factor between 2 and 4 by simulating the tenfold number of particles. For 5000 particles, the NMSE and FB caused by the kernel procedure are already neglibible. This statement holds for concentration predictions at down-wind distances like those where the receptor arcs in the Copenhagen experiment were located (2 to 6 km down-wind from the source). For near-source predictions, before the maximum ground-level concentration occurs, this number probably is inaccurate, as can be seen in Fig. 5, but a number of 50 000 simulated particles should give a high accuracy. When conducting particle simulations for down-wind distances so near to the source that even a number of 50 000 particles (using the kernel estimate

method) could be considered insufficient, parametrizations within the particle model (e.g., for turbulence, the timestep and the pdf) are likely to have an even larger impact.

4.5.5 Recommended scheme

The triweight kernel approaches the near-source ”true” estimate faster than the other kernel shapes investigated (see previous sections). For intermediate to far-field estimations, the quadweight kernel is best. For the Copenhagen experiment, the simulation of 5000 particle trajectories has been shown to be more than sufficient when using a fully three-dimensional stochastic particle model. The same number has also been used by the author to simulate other tracer experiments which took place under stable and fully convective conditions. In fact, when the measurements available took place further down-wind than the location at which the ground-level concentration has its maximum (as is the case for the Copenhagen experiment), 500 particles together with quadweight kernels produce very good results.

It also is recommended to use the more robust formulation of Eq. (14) instead of Eq. (8). This produces clearly better predictions especially in situations where the assumption of normal distribution of the particle positions cannot be justified, i.e., for fully convective conditions, as well as for the vertical direction after the release has become well-mixed over the entire boundary layer (not shown). In future work, the use of kernels with variable bandwidths, as briefly outlined in section 4.4.3, will be further refined. This technique will eventually allow for more robust and even more efficient concentration predictions.

4.6 ANALOGY OF DENSITY KERNELS AND PUFFS

4.6.1 Interpretation of density kernels as clusters of particles

When interpreting density kernel distributions as models for clusters of pollutant particles (puffs), one has to be aware of the fact that the kernel approach originates from the need for a numerical non-parametric regression method. To fullfil this numerical task best, the bandwidths of density kernels should be chosen according to the procedure outlined in section 4.3. With the corresponding Eqs. (8) and (14), for the simulated particles obeying the same density distribution, kernels obtain smaller bandwidths as the particle number increases.

Density kernel bandwidths formulations derived from some physical description of cluster growth must not depend on the number of particles within the cluster, but only on the size of

the latter. The growth rate of such a cluster of particles depends on its size, i.e., the range of eddies small enough to be capable of dispersing the cluster.

Numerically obtained kernel bandwidths, on the other hand, depend on the number of particles released (Eq. 8). This fundamental difference could lead to the assumption that no relation can be established between cluster growth and kernel bandwidth, although such a relation would be desirable in order to be able to ”blend” puff and particle models. One way to indeed blend puff and particle models, which uses kernels with physically motivated sizes taken from dispersion theory, is the Puff-Particle approach by de Haan and Rotach (1998b).

4.6.2 Bandwidth estimates taken from dispersion theory

It should be distinguished between absolute and relative dispersion: Relative dispersion accounts for the dispersing effect of all turbulent eddies smaller than the separation of two particles, since these eddies will be able to enlarge their mean distance. Absolute dispersion describes the spread of a release with respect to a fixed point in space, and thus accounts for the dispersing effect of eddies of all sizes.

As the growth of a cluster of particles is described by relative dispersion, parametrisations of this growth, e.g. as a function of cluster size, could also be used as bandwidths of density kernels. Relative dispersion is applied in puff dispersion models which use frequently updated wind fields to account for all dispersion originating from eddies larger than the size of an individual cluster (thus not yet covered by the relative dispersion for such a cluster).

However, the size of a relatively dispersed cluster (i.e., puff) differs fundamentally from the numerical bandwidth for a density kernel. This can easily be verified by looking at two hypothetical cases: First, doubling the number of simulated particles will reduce the numerical bandwidth according to Eq. (8); the physical size of a cluster under the given turbulent conditions is not affected. Second, kernel bandwidths will only increase when the mean distance between particles increases, whereas relatively dispersed clusters grow with every timestep.

4.6.3 Use of absolute dispersion to describe kernel bandwidths

The RAPTAD model of Yamada *et al.* (1987, 1989) and Yamada and Bunker (1988) combines a particle model with Gaussian shaped density kernels. They use the physical concept of absolute dispersion to estimate the appropriate bandwidth of these kernels. This is

a physical constraint determining the size, in contrast to the numerically motivated bandwidths as outlined in section 4.4.

For a sufficiently high number of particles, the bandwidths of the numerical kernel will be smaller than the size of relatively dispersed puffs (see preceding section). This adds dispersion to the dispersion model as a whole, i.e. it widens the predicted concentration distribution. Therefore, this additional dispersion has to be removed from the particle model, by modifying (reducing) the turbulent energy spectrum representing the stochastic particle trajectories. This approach is realised in de Haan and Rotach (1998a), where a low-pass filter depending on the size of the relatively dispersed puff filters out dispersion from the particle trajectories.

Note that the ”dispersion” added to the model by the numerical kernel bandwidths should not be filtered out of the particle model. It is the smoothing necessary to interpolate between particle positions, and tends to zero as the number of particles approaches infinity (whereas the overdispersion caused by the physical sizes remains constant).

4.7 SUMMARY AND CONCLUSIONS

A method is proposed to estimate concentrations from stochastic Lagrangian particle models. Traditionally, most particle modellers use the box-counting method, where concentration is estimated by counting particles in imaginary sampling boxes. Using such box-averaged concentrations as an estimation for point concentrations, if boxes are chosen too small, the concentration distribution becomes noisy (having a large variance); if they are too large, the concentration is oversmoothed (having a large bias). It is shown for two hypothetical source-receptor configurations that the choice of box sizes can cause large differences in the predicted concentration, especially for near-field and ground level predictions. Only large particle numbers can minimize these effects.

The kernel density estimation method is proposed as an alternative which minimizes the sum of the variance and the bias of the predicted concentration distribution. This allows for the number of particles to be reduced by an order of magnitude as compared to predictions made with the box-counting method. The most important parameter, the bandwidth of the kernels, is determined from the standard deviations of the particle position distribution.

The effectiveness of six different kernel shapes in concentration estimation is investigated. The so-called quadweight kernel, used with a robust parametrisation for the calculation of

bandwidth, has optimal performance. It yields optimal results even if only a few thousand particle trajectories are simulated. The use of dispersion theory to describe kernel bandwidths, allowing an interpretation of the kernels as physically meaningful clusters of particles, is only possible if the corresponding amount of dispersion which has been added by the bandwidths of the kernels is removed from the stochastic particle trajectories (de Haan and Rotach 1998a).

Acknowledgements—The comments of Dr. R. Yamartino on a first draft of this paper have been very helpful. Thanks go to Dr. P. Manins for drawing my attention to the work of G. Lorimer. The present work was partly financed by the Swiss Agency of Education and Sciences (BBW) and the Swiss Agency of Environment, Forest and Landscape (BUWAL) through a project in the framework of COST 615 (citair), and by the Swiss National Science Foundation through Grant Nr. 21-46849.96.

APPENDIX—DETERMINATION OF $C_{d,a}$, α AND β

The normalizing factor $C_{d,a}$ for kernels of the form $\left(1-\mathbf{x}^T\mathbf{x}\right)^a$ (introduced in Eq. 12) is calculated by integrating in *d* dimensions the non-normalized kernel. Alternatively, the *d*-dimensional rotational volume can be calculated, $C_{d,a}{}^{-1} = c_d \int_0^\infty \left[K_a{}^{-1}(y)\right]^d dy$, with the substitution $y = \left(1-x^2\right)^a = K_a(x)$ and $K_a{}^{-1}(y) = \left(1-y^{1/a}\right)^{1/2}$:

$$C_{d,a} = \left\{\int \left(1-\mathbf{x}^T\mathbf{x}\right)^a d\mathbf{x}\right\}^{-1} = \left\{c_d \int_0^1 \left(1-y^{1/a}\right)^{d/2} dy\right\}^{-1} \tag{A1}$$

where c_d is the volume of the unit *d*-dimensional sphere (i.e. $c_1 = 2$, $c_2 = \pi$, $c_3 = 4\pi/3$). For $a = 1$, this gives $C_{d,1}{}^{-1} = 2c_d/(d+2)$, and the general solution is

$$C_{d,a} = \left\{\pi^{d/2} \frac{\Gamma(a+1)}{\Gamma(a+1+d/2)}\right\}^{-1} = \frac{\prod_{i=1}^{a}(d+2i)}{c_d \cdot 2^a \cdot a!} \tag{A2}$$

For the kernel functions used in this paper and for up to three dimensions, the corresponding normalising factors are listed in Table I.

The determination of the value of $\beta = \int K^2(\mathbf{x})d\mathbf{x}$ (introduced in Eq. 9) is straightforward using the above method and result, yielding

$$\beta = C_{d,a}^2 \int \left(1 - \mathbf{x}^T \mathbf{x}\right)^{2a} d\mathbf{x} = C_{d,a}^2 \left\{ \pi^{d/2} \frac{\Gamma(2a+1)}{\Gamma\left(2a+1+d/2\right)} \right\} = C_{d,a}^2 \frac{c_d 2^{2a} (2a)!}{\prod_{i=1}^{2a} (d+2i)} \tag{A3}$$

To obtain $\alpha = \int x_1^2 K(\mathbf{x}) d\mathbf{x}$ $\left(\mathbf{x} = \left[x_1, x_2, x_3\right]^T\right)$, we calculate

$$\begin{aligned} \alpha &= C_{d,a} \int_{-1}^{1} \left(\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \left(\int_{-\sqrt{1-x_1^2-x_2^2}}^{\sqrt{1-x_1^2-x_2^2}} x_1^2 \left(1 - \mathbf{x}^T \mathbf{x}\right)^a dx_3 \right) dx_2 \right) dx_1 \\ &= \frac{\Gamma(a+1)\pi^{(d-1)/2}}{\Gamma\left(a+(d+1)/2\right)} \int_{-1}^{1} t^2 \left(1 - t^2\right)^{(2a+d-1)/2} dt \end{aligned} \tag{A4}$$

With the solution

$$\alpha = C_{d,a} \frac{2^a a! c_d}{\prod_{i=1}^{a+1} (d+2i)} \tag{A5}$$

which has been proved for $d \le 3$ and $a \le 5$. Combining Eqs. (A4) and (A5) with Eq. (10), $A(K)$ from Eq. (9) can then be calculated.

REFERENCES

Batchelor, G. K. (1952): Diffusion in a Field of Homogeneous Turbulence—Part II: The Relative Motion of Particles. *Proc. Camb. Philos. Soc.*, **48**, 345–362

Borgas, M. S. and Sawford, B. L. (1994): A Family of Stochastic Models for Two-Particle Dispersion in Isotropic Homogeneous Stationary Turbulence. *J. Fluid. Mech.*, **279**, 69–99

Du, S., and Wilson, J. D. (1994): Probability Density Functions for Velocity in the Convective Boundary Layer, and Implied Trajectory Models. *Atmospheric Env.*, **28**, 1211–1217

Epanechnikov, V. K. (1969): Non-parametric Estimation of a Multivariate Probability Density. *Theory Probab. Appl.*, **14**, 153–158

Flesch, T. K, Wilson, J. D. and Yee, E. (1995): Backward-Time Lagrangian Stochastic Dispersion Models and Their Application to Estimate Gaseous Emissions. *J. Appl. Meteorology*, **34**, 1320–1332

Gryning, S.-E. and Lyck, E. (1984): Atmospheric Dispersion from elevated sources in an urban area: Comparison between tracer experiments and model calculations. *J. Clim. Appl. Met.*, **23**, 651–660

de Haan, P., and Rotach, M. W. (1998a): The Treatment of Relative Dispersion within a combined Puff-Particle Model (PPM). In: Air Pollution Modeling and Its Application XII, S.-E. Gryning and N. Chaumerliac (eds.), Plenum Press, New York, 389–396

de Haan, P., and Rotach, M. W. (1998b): A novel approach to atmospheric dispersion modelling: The puff-particle model (PPM). *Quart. J. Roy. Meteorol. Soc.*, **124**, 2771–2792

Hurley, P. and Physick, W. (1991): A Lagrangian Particle Model of Fumigation by Breakdown of the Nocturnal Inversion. *Atmospheric Env.*, **25A**, 1313–1325

Hurley, P. and Physick, W. (1993): A skewed homogeneous Lagrangian particle model for convective conditions. *Atmospheric Env.*, **27A**, 619–624

Lorimer, G. S. (1986): A kernel method for air quality modelling – I: Mathematical foundation. *Atmospheric Env.*, 20, 1447–1452

Luhar, A. K., and Britter, R. E. (1989): A Random Walk Model for Dispersion in Inhomogeneous Turbulence in a Convective Boundary Layer. *Atmospheric Env.*, **23**, 1911–1924

Luhar, A. K., and Rao, K. A. (1994): Lagrangian Stochastic Dispersion Model Simulations of Tracer Data in Nocturnal Flows over Complex Terrain. *Atmospheric Env.*, **28**, 3417–3431

Luhar, A. K., Hibberd, M. F., and Hurley, P. J. (1996): Comparison of Closure Schemes used to Specify the Velocity PDF in Lagrangian Stochastic Dispersion Models for Convective Conditions. *Atmospheric Env.*, **30**, 1407–1418

Mikkelsen, T., Larsen, S. E. and Pécseli, H. L. (1987): Diffusion of Gaussian Puffs. *Quart. J. Roy. Met. Soc.*, **113**, 81–105

Rotach, M. W., Gryning, S.-E. and Tassone, C. (1996): A Two-Dimensional Stochastic Lagrangian Dispersion Model for Daytime Conditions. *Quart. J. Roy. Met. Soc.*, **122**, 367–389

Scott, D. W. (1992): *Multivariate Density Estimation: Theory, Practice, and Visualisation*. Wiley, New York

Silverman, B. W. (1986): *Density Estimation*. Chapman and Hall, London

Thomson, D. J. (1987): Criteria for the Selection of Stochastic Models of Particle Trajectories in Turbulent Flows. *J. Fluid Mech.*, **180**, 529–556

Thomson, D. J., and Montgomery, M. R. (1994): Reflection Boundary Conditions for Random Walk Models of Dispersion in Non-Gaussian Turbulence. *Atmospheric Env.*, **28**, 1981–1987

Venkatram, A., and Du, S. (1997): An Analysis of the Asymptotic Behavior of Cross-Wind-Integrated Ground-Level Concentrations using Lagrangian Stochastic Simulation. *Atmospheric Env.*, **31**, 1467–1476

Wilson, J. D. and Sawford, B. L. (1996): Review of Lagrangian Stochastic Models for Trajectories in the Turbulent Atmosphere. *Boundary-Layer Meteorol.*, **78**, 191–210

Yamada, T., Bunker, S., and Niccum, E. (1987): Simulations of the ASCOT Brush Creek data by a Nested-Grid, Second-Moment Turbulence Closure Model and a Kernel Concentration Estimator. AMS 4th Conference on Mountain Meteorology, 25–28 Aug. 1987, Seattle, WA.

Yamada, T. and Bunker, S. (1988): Development of a Nested Grid, Second Moment Turbulence Closure Model and Application to the 1982 ASCOT Brush Creek Data Simulation. *J. Appl. Meteorol.*, **27**, 562–578

Yamada, T., Jim Kao, C.-H., and Bunker, S. (1989): Airflow and Air Quality Simulations over the Western Mountainous Region with a Four-Dimensional Data Assimilation Technique. *Atmospheric Env.*, **23**, 539–554