21st International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
27-30 September 2022, Aveiro, Portugal

ACCELERATION OF SIMULATIONS BY APPLICATION OF A KERNEL METHOD IN A HIGH-RESOLUTION LAGRANGIAN PARTICLE DISPERSION MODEL

Daniela Barbero[1,2], Bruno Ribstein[3], Maxime Nibart[3] and Gianni Luigi Tinarelli[1]

[1]ARIANET S.R.L., 20159 Milan, Italy
[2]Department of DICA, Politecnico di Milano, 20133 Milan, Italy
[3] ARIA Technologies, 92100 Boulogne-Billancourt, France

Abstract: The role of microscale atmospheric dispersion modelling is becoming increasingly important in air quality assessment, especially in residential areas, for regulatory purposes and to project pollution control strategies. Consequently, the use of these models in long-term studies and forecasting systems is growing. However, this modelling can be challenging because of the amount of time and CPU required for the simulations, especially if the computational domain has a significant extension.
This article describes the application of a concentration calculation methodology to reduce the computational time needed by microscale Lagrangian Particle Dispersion Models (LPDMs). These models normally estimate the concentration with the box counting method: a 3D mesh is set, and the density is computed counting particles in each box. An alternative method based on the use of the statistical technique of kernel density estimation is proposed to determine the concentration from the particles' position. The kernel method allows for a reduction of computational particles emitted during the simulation, guaranteeing a similar accuracy to that of box counting method. It enables therefore to optimize the overall simulation time and the required CPUs in order to improve time and cost, enhancing the efficiency of models and widening their application fields.
The use of the kernel method to perform high-resolution simulations took the first steps with an application inside the LPDM of PMSS (Parallel-Micro-SWIFT-SPRAY) system to evaluate road traffic gases emissions in urban areas, enabling an 80% simulation time reduction. In this article, additional features of this method are developed within the Micro-SPRAY model and validated against two new test cases. The test cases consider microscale simulations of industrial sources emitting gases and particles, evaluated both inside a domain divided in tiles and inside a nested domains configuration. The existing kernel method is enhanced in order to estimate the pollutant concentrations of point sources and to compute also the corresponding deposition at building-resolving scale and in nested domains with different horizontal resolution.

Key words: *kernel method, microscale simulation, lagrangian particle dispersion model, PSPRAY*

INTRODUCTION

Obtaining a correct and punctual assessment of air quality is increasingly requested to provide a better prevention of the population exposure and to define effective abatement strategies in advance. In this context, microscale atmospheric models can play a fundamental role, because they could assess, and moreover predict, air pollution levels at very high resolution. However, models that work at microscale generally have significant computational costs, in term of both CPU and time demanded for the simulations; especially for forecasting systems, which have the constrain of time, these costs can become limiting. Consequently, it is still crucial to reduce the amount of time needed for a simulation keeping the same computational resources.

Inside the microscale Lagrangian Particle Dispersion Model (LPDM) PSPRAY of the model suite PMSS (Parallel-Micro-SWIFT-SPRAY, Oldrini et al., 2011, 2017) an alternative method to compute concentrations to improve the computational time of a long-term simulation over a large urban area (Barbero et al., 2021) was already added. In addition to the box-counting method, originally applied in PSPRAY to calculate concentrations, a new algorithm called "kernel method" has been implemented, which allows to reduce the number of computational particles used for a simulation and therefore to optimize the time cost. In fact, PSPRAY is the parallel version of the LPDM SPRAY and it is based on a 3D form of the Langevin equation for the random velocity (Thomson, 1987); it simulates the dispersion of a pollutant following the trajectories of virtual particles, producing 3D concentrations fields and 2D dry

and wet depositions fields. Furthermore, PSPRAY can exploit MPI instructions to run in parallel mode, using several processors to split particles and/or to split the domain in several tiles or to use a nested configuration. The use of the kernel method has appeared promising for urban microscale simulations; therefore, some new numerical experiments have been performed considering industrial sources to test the method with point sources emitting gas and particles.

MATERIALS AND METHODS

The most common method used in LPDMs to calculate concentrations is the box-counting method, that consists in imposing a 3D mesh and computing the concentration field as the total mass of the particles falling inside a cell of the mesh divided by the cell volume. This method intrinsically requires emitting a large number of computational particles during the simulation, to avoid relatively large statistical errors.
The number of particles used in a simulation strongly influences the overall computational time, therefore an alternative method that allows to produce more regular fields with fewer particles could be useful, as in the case of the kernel method (Lorimer, 1986). This latter is based on kernel density estimation, which is a set of techniques for the nonparametric estimation of functions, and each particle is considered as the centre of a probabilistic distribution of its pollutant mass. Therefore, the concentration at a given time *t* and in a certain location in the domain defined by *x*, *y*, *z* could be computed as:

$$C(x,y,z;t) = \sum_{p=1}^{Ntot} \frac{m_p}{h_x h_y h_z} K\left(\frac{x_p - x}{h_x}\right) K\left(\frac{y_p - y}{h_y}\right) K\left(\frac{z_p - z}{h_z}\right) \quad (1)$$

where N_{tot} is the total number of particles in the domain, x_p, y_p, z_p and m_p are the position and the mass of the p^{th} particle, *K* is a generic kernel function, which must be greater than 0 and its integral over the domain must be equal to 1, and h_i are the so-called bandwidths. *K* defines the shape of kernel, while the bandwidths adjust its width in the three directions of space, defining the so-called volume of influence of a particle. To obtain a field equivalent to that of the box-counting, the concentration is computed in the centres of the 3D mesh cells. The kernel method has been integrated in LPDMs at local scale since the late 1980s (Lorimer, 1986), but the inclusion in PSPRAY (Barbero et al., 2021) is the first application at microscale. The next paragraph summarizes the main specifications chosen for the kernel method, whose implementation has been adapted to the microscale.

Kernel method specifications

Kernel function and bandwidths definition

As shown in Equation (1), the kernel method in PSPRAY uses the 3D product kernel (de Haan, 1999), which is the product of independent one-dimensional kernels, defined for PSPRAY with the bi-weight kernel function in all the three space directions:

$$K(x) = \begin{cases} \frac{15}{16} * \left[1 - \left(\frac{x_p - x}{h_x}\right)^2\right]^2 & for \left|\frac{x_p - x}{h_x}\right| \leq 1 \\ 0 & for \left|\frac{x_p - x}{h_x}\right| > 1 \end{cases} \quad (2)$$

However, the shape of the kernel function does not have a crucial influence on the results, but the role of bandwidths is of great importance (de Haan, 1999). The three bandwidths are computed as constants, directly proportional to PSPRAY 3D mesh sizes (Uliasz, 1994):

$$\begin{aligned} h_x &= 3.5 * \Delta x \\ h_y &= 3.5 * \Delta y \\ h_z &= 1.2 * h_0 \qquad (h_0 = \text{height of ground level concentration}) \end{aligned} \quad (3)$$

The concentration is obtained by PSPRAY by counting N_{snap} times the mass contained in each cell. With the box-counting method, the smallest modeled concentration equals m_p/N_{snap}, i.e. when a numerical particle is counted only once inside a cell. With the kernel method, the smallest modeled concentration is smaller, depending on the chosen bandwidths. In PSPRAY, with the same number of emitted particles, the smallest modeled concentration is about $K(x_p + 3.5 * \Delta x)K(y_p + 3.5 * \Delta y)K(z_p + 1.2 * \Delta z) \approx 4\%$ of the previous value. This is one of the reasons that explains why the number of emitted particles can be reduced.

Impermeable boundaries

A special treatment of the particles that are close to the ground or to obstacles is applied to prevent a mass lost due to the cut of the kernel function.

For all particles whose distance from the ground is less than its vertical bandwidth, a reflection term is introduced, which modifies the kernel function in the z direction as follows:

$$K(z) = K\left(\frac{z_p - \mathrm{z}}{h_z}\right) + K\left(\frac{z_p + z}{h_z}\right) \quad (4)$$

This is based on the idea of considering, for each *x,y,z* point, a contribution of a virtual particle that is actually the original release reflected over the boundary, in addition to the normal contribution of the particle. Therefore, what actually goes over the boundary due to the kernel function is taken back by means of a reflected particle, whose quantity is the same as the quantity previously lost below the ground.

A different handling is necessary for particles that are close to an obstacle because it must be considered that the mass related to a particle must not contribute to the concentration both within a building and also in all the points behind the obstacle with respect to the particle itself. Therefore, the volume of influence of each particle is modified at each timestep based on the obstacles surrounding it, and its mass is distributed inside this new volume.

Deposition

In addition to the computation of concentration fields, it has been integrated also the calculation of dry and wet deposition 2D fields. The deposition fields are computed using a 2D product kernel, defined using the same kernel function of Equation (2) and the same bandwidths in horizontal directions of Equation (3).

Tiled and nested configuration

The existing kernel method has been applied also in nested configuration, as well as in tiled configuration. In these two configurations, each particle contributes only inside the domain in which it is located and the borders of adjacent tiles or nested domains are treated as impermeable boundaries. This is an approximation, but it allows to avoid an increase of time of MPI communications.

Test cases specifications

In order to perform a comparison of the methods, two different test cases have been considered, both with point sources, hereinafter referred to as “TEST CASE 1” and “TEST CASE 2”.

The TEST CASE 1 is simpler, considering one point source located at 1.5 meters from the ground and emitting a gaseous pollutant. The simulation is run with a tiled configuration, using 6 rectangular tiles, with a horizontal resolution of 1 meter. It allowed to firstly verify the same kernel method already applied with linear sources, but considering point sources.

The TEST CASE 2 was used to test the newer features of the kernel method. This test case considers a point source that emits both gases and particles, computing also the deposition. The simulation is run with a nested configuration, with two domains: the inner one with a horizontal resolution of 1 meter, and the external one with a horizontal resolution of 2 meters.

Both the test cases consider the presence of buildings in the domain and consider an averaging time for concentrations and depositions of 1 hour.

RESULTS AND DISCUSSIONS

For both the test cases, the hourly concentration and deposition fields were computed firstly using the box-counting method and then applying the kernel method using 1/5 of particles; for both methods the same number of CPU were used.

Figure 1 and **Figure 2** show the comparison between the hourly concentration fields of the gaseous pollutant at the ground level computed with the box-counting method and the kernel method respectively in TEST CASE 1 and TEST CASE 2. The obtained concentration fields of the particulate pollutants and the corresponding deposition fields in TEST CASE 2 closely resemble the ones shown. From the figures it

can be seen that the kernel method is able to provide qualitatively similar fields, emitting 1/5 of the particles that the box-counting method needs to provide a reliable statistical result. The fields produced with the kernel method appear smoothed enough and of the same shape as that of the box-counting method.

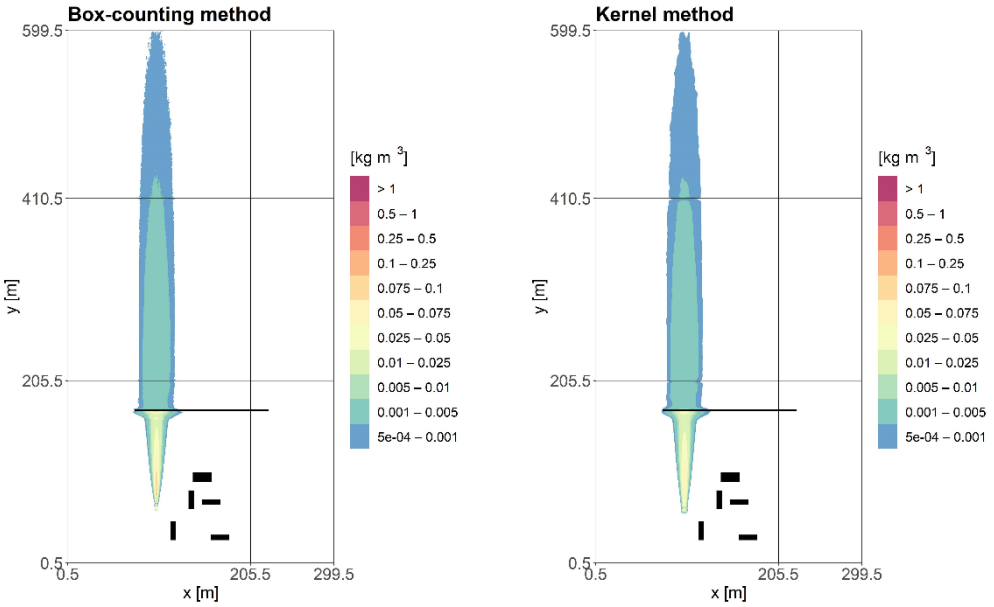

Figure 1. Hourly ground concentration fields of a gaseous pollutant computed with box-counting method (left) and kernel method (right), TEST CASE 1. The internal lines indicate the tiles of the domain, while the black rectangular shapes are the obstacles.

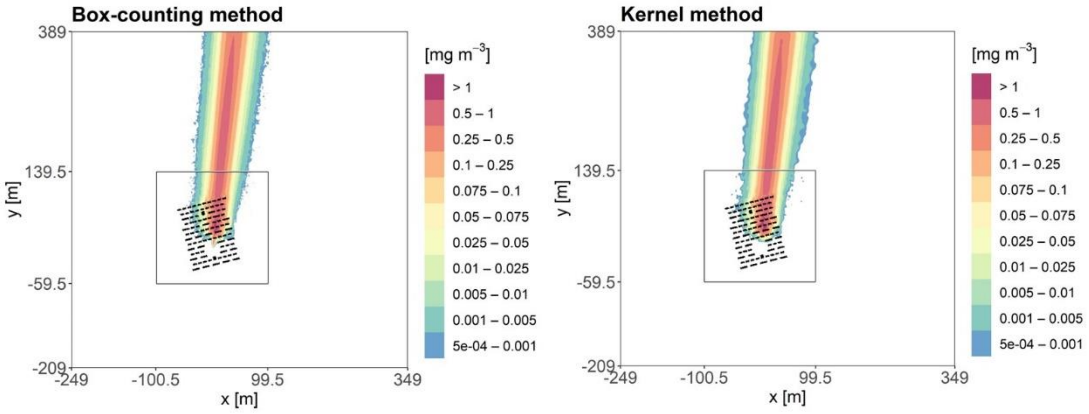

Figure 2. Hourly ground concentration fields of a gaseous pollutant computed with box-counting method (left) and kernel method (right), TEST CASE 2. The internal lines indicate the nested domain, while the black rectangular shapes are the obstacles.

The concentration fields are compared using as statistical indexes the fraction of predicted values within a factor of two of observations (FAC2), the Pearson correlation coefficient (r) and the Index of Agreement (IA) by Willmott (1981). The indexes are calculated by considering the concentration (and deposition) values at each point in the domain grid. However, all points where both resulting fields are 0 or lower than a threshold value set at 1/100 of the maximum value in the domain have been excluded, in order to avoid a non-significant bias in the statistics. Finally, it is also verified the overall reduction of time for the simulation obtained using the kernel method compared to the box-counting method, computed as:

$$Time\ reduction\ [\%] = \frac{time_{box-counting} - time_{kernel}}{time_{box-counting}} * 100 \qquad (5)$$

The **Table 1** shows the statistical indexes and the reduction of time obtained in both the test cases. For TEST CASE 2 three separate analysis were carried out, for the concentration fields of gaseous pollutant and particulate pollutant and for the deposition field of the particulate pollutant.

Table 1. Statistical indexes of hourly concentration and deposition fields for comparison of box-counting method and kernel method results and obtained time reduction with kernel method

	TEST CASE 1	**TEST CASE 2**		
		Gas Concentration	**Particulate Concentration**	**Particulate Deposition**
Particles reduction [%]	80%	80%		
Time reduction [%]	68%	80%		
FAC2	0.95	0.98	0.98	0.95
Correlation	0.94	0.94	0.95	0.90
Index of agreement	0.96	0.97	0.97	0.95

The statistical analysis confirms the qualitative analysis of the figures, the concentration and deposition fields computed with the kernel method are statistically similar to those computed with the box-counting method, for both gaseous and particulate pollutants. All the statistical indexes are greater than 0.9 and close to their ideal value of 1. Regarding the reduction of time, for TEST CASE 2 the use of the kernel method allows for a reduction equal to the reduction of the particles emitted during the simulation: with a particles reduction of 80% we obtain an almost identical percent of computational time reduction. Considering the same reduction of emitted particles, the time saved is lower for TEST CASE 1. This is a result of the tiled configuration that requests more MPI communications than the nested configuration of TEST CASE 2; in that case, even if PSPRAY manages the nested domains as the tiled ones, we have only two sub-domains instead of six. The MPI communication load is independent from the calculation of concentration and constitutes a fixed part of the overall computational time. Even if this is still significant compared to the total amount of time, the advantage of using the kernel is less visible.

CONCLUSIONS

This work presents the use of kernel method as alternative to the common box-counting method to compute concentration inside the microscale Lagrangian Particle Dispersion Model PSPRAY. The method is tested to compute hourly ground concentration and deposition fields of gaseous and particulate pollutants emitted by point sources. The obtained results, although still preliminary, are promising and in line with those already obtained for linear sources, both in term of reduction of the overall computational time requested by the simulation and the statistical comparison to the box-counting method results.

REFERENCES

Barbero, D., Tinarelli, G., Silibello, C., Nanni, A., Gariazzo, C., Stafoggia, M., Viegi, G., 2021: A microscale hybrid modelling system to assess the air quality over a large portion of a large European city. *Atmospheric Environment*, **264**, 118656.

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

Lorimer, G. S., 1986: The kernel method for air quality modelling – mathematical foundation. *Atmospheric Environment*, **20**, 1447-1452.

Oldrini, O., Olry, C., Moussafir, J., Armand, P., Duchenne, C., 2011: Development of PMSS, the parallel version of Micro – SWIFT – SPRAY, paper presented at the 14th International Conference on *Harmonisation Within Atmospheric Dispersion Modelling for Regulatory Purposes*, Kos, Greece.

Oldrini O., Armand P., Duchenne C., Olry, C., Moussafir J., Tinarelli G., 2017: Description and preliminary validation of the PMSS fast response parallel atmospheric flow and dispersion solver in complex built-up areas. *Environmental Fluid Mechanics*, **17**, 997-1014.

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

Uliasz, M., 1994: Lagrangian particle dispersion modeling in mesoscale applications. *Environmental Modeling*, **11**.

Willmott CJ., 1981: On the validation of models. *Physical Geography*, **2**, 184 194.