Meteorol Atmos Phys (2018) 130:107–124
https://doi.org/10.1007/s00703-017-0506-0

ORIGINAL PAPER

# Prediction of dosage-based parameters from the puff dispersion of airborne materials in urban environments using the CFD-RANS methodology

G. C. Efthimiou[1] · S. Andronopoulos[1] · J. G. Bartzis[2]

Received: 9 May 2016 / Accepted: 27 January 2017 / Published online: 1 February 2017


**Abstract** One of the key issues of recent research on the dispersion inside complex urban environments is the ability to predict dosage-based parameters from the puff release of an airborne material from a point source in the atmospheric boundary layer inside the built-up area. The present work addresses the question of whether the computational fluid dynamics (CFD)–Reynolds-averaged Navier–Stokes (RANS) methodology can be used to predict ensemble-average dosage-based parameters that are related with the puff dispersion. RANS simulations with the ADREA-HF code were, therefore, performed, where a single puff was released in each case. The present method is validated against the data sets from two wind-tunnel experiments. In each experiment, more than 200 puffs were released from which ensemble-averaged dosage-based parameters were calculated and compared to the model's predictions. The performance of the model was evaluated using scatter plots and three validation metrics: fractional bias, normalized mean square error, and factor of two. The model presented a better performance for the temporal parameters (i.e., ensemble-average times of puff arrival, peak, leaving, duration, ascent, and descent) than for the ensemble-average dosage and peak concentration. The majority of the obtained values of validation metrics were inside established acceptance limits. Based on the obtained model performance indices, the CFD-RANS methodology as implemented in the code ADREA-HF is able to predict the ensemble-average temporal quantities related to transient emissions of airborne material in urban areas within the range of the model performance acceptance criteria established in the literature. The CFD-RANS methodology as implemented in the code ADREA-HF is also able to predict the ensemble-average dosage, but the dosage results should be treated with some caution; as in one case, the observed ensemble-average dosage was under-estimated slightly more than the acceptance criteria. Ensemble-average peak concentration was systematically underpredicted by the model to a degree higher than the allowable by the acceptance criteria, in 1 of the 2 wind-tunnel experiments. The model performance depended on the positions of the examined sensors in relation to the emission source and the buildings configuration. The work presented in this paper was carried out (partly) within the scope of COST Action ES1006 "Evaluation, improvement, and guidance for the use of local-scale emergency prediction and response tools for airborne hazards in built environments".

Responsible Editor: M. T. Prtenjak.

✉ G. C. Efthimiou
gefthimiou@ipta.demokritos.gr

[1] Environmental Research Laboratory, INRASTES, NCSR Demokritos, Patriarchou Grigoriou and Neapoleos Str., 15310 Agia Paraskevi, Greece

[2] Department of Mechanical Engineering, University of Western Macedonia, Sialvera and Bakola Str., 50100 Kozani, Greece

## 1 Introduction

One of the key issues of recent research on the dispersion inside complex urban environments is the ability to predict dosage-based parameters from the puff release of an airborne material from a point source in the atmospheric boundary layer. Due to the stochastic nature of turbulence, the puff pathway is unknown and depends on the instantaneous atmospheric conditions at the time of the release.

This can be observed by examining the concentration time series of a sensor downwind of the source. For example, in Fig. 3 of Harms et al. (2011), seven puffs were released consecutively from a point source (indicated by the black bars) during a wind-tunnel experiment and a sensor downwind of the source measured the concentration (red line). One can notice that some puffs hit the sensor, while the two puffs released at times 15 s and 35 s bypassed the sensor altogether, following the different pathways.

The problem is more complicated when the puff release occurs inside an urban environment where the buildings affect the dispersion.

A consistent set of dosage-based parameters for puff dispersion characterization was introduced recently by Berbekar et al. (2015), Fig. 1.

The dosage-based parameters are:

- the dosage (ppmVs): the time-integrated concentration of tracer gas over the detection period,
- peak concentration (ppmV): the highest concentration measured at the highest temporal resolution setting of the instrument occurring at the measurement location during the detection period (this is not a dosage-based parameter but is used for the characterization of a puff),
- arrival time (s): the time between the beginning of the puff release and when 5% of the total dosage of the puff has reached the measurement location,
- peak time (s): the time between the beginning of the puff release and when the peak concentration occurs at the measurement location (this is not a dosage-based parameter but is used for the characterization of a puff),
- leaving time (s): the time after the beginning of the puff release when 95% of the total observed dosage of the puff is recorded at the measurement location,

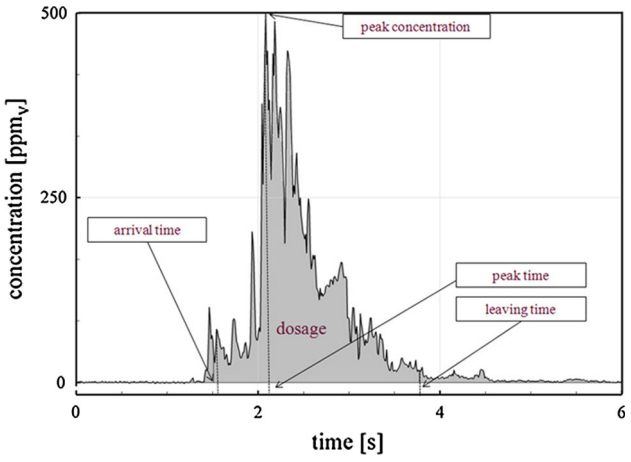



**Fig. 1** Puff parameters for statistical study. The arrival time is defined as the time after release when the dosage exceeds the threshold of 5%. The leaving time is defined as the time when 95% of the total dosage is reached (Baumann-Stanzer et al. 2015)

- ascent time (s): the time interval between the arrival time and the peak time,
- descent time (s): the time period between the peak time and the leaving time,
- duration (s): the time interval between the arrival time and the leaving time.

In the literature, the great majority of studies that attempt to model puff dispersion apply Gaussian and Lagrangian approaches (e.g., Hanna et al. 1982; Yim et al. 2010; Cui et al. 2011; Thomson and Jones 2011; Wan et al. 2014). The puff dispersion models of this type are numerous. Typical examples are the CALPUFF, PUFF-PLUME, RIMPUFF, SCIPUFF, TSCREEN, Puff, UDM, etc. Attempts to model puff dispersion through computational fluid dynamics (CFD) techniques are reported in only two studies (Ching et al. 2009; Harms et al. 2011). In Ching et al. (2009), a CFD–statistical hybrid modelling method was developed. The source emission was represented by emissions of many infinitesimal puffs. Statistical analysis was performed to obtain first the statistical properties of the puff trajectories, and subsequently, the most probable distribution of the puff trajectories that represent the macroscopic dispersion behaviours. In two case studies of ambient dispersion, the numerical modelling results obtained agree reasonably well with both experimental measurements and the conventional $k$–$\varepsilon$ modelling results published in the literature. Concerning the prediction of dosage-based parameters from puff dispersion, some mention is only made by Harms et al. (2011). In particular, the frequency distributions for the 'peak time' parameter were compared between experimental data from 200 wind-tunnel releases and predictions from a simulation of 60 puff releases using the large eddy simulation-based numerical model Fast3D. The comparison concerned the Central Business District of Oklahoma City test case.

In the present paper, authors attempt a more thorough examination of the efficiency of the CFD Reynolds-averaged Navier–Stokes (RANS) methodology for the prediction of dosage-based parameters from a puff release. In Sect. 2, the proposed methodology is presented, while in Sects. 3 and 4, the wind-tunnel data sets and the CFD-RANS numerical simulations are described, respectively. The results are presented and discussed in Sect. 5. It should be noted that the present work is part of the COST Action ES1006 "Evaluation, improvement, and guidance for the use of local-scale emergency prediction and response tools for airborne hazards in built environments".[1]

[1] http://www.elizas.eu.

## 2 Methodology

The computational simulations have been performed in two steps:

1. Simulation of the wind field;
2. Simulation of the puff dispersion using the results of the first step.

The simulation of the wind field consisted of the numerical solution of the RANS equations for mass, the three components of momentum, and the transport equations required for the turbulence model employed. No energy equation has been solved, since the simulated experiments were performed for neutral stability conditions. Constant in time and variable in height inlet flow boundary conditions were assumed (in accordance with the experimental data sets); therefore, steady-state wind and turbulence parameters fields have been calculated in this step. More information on the wind field simulations of the wind-tunnel experiments is given in the following sections of this paper and in the previous work by Hertwig et al. (2012). This study focuses on the results of the puffs dispersion simulations.

For the puff dispersion simulations, the partial differential equation governing the evolution of the mean concentration has been solved (Efthimiou et al. 2015) using the wind and turbulence fields calculated in the previous step. A short-duration release of the tracer has been considered, with the released mass and duration being equal to the experimental ones. This RANS methodology simulates the dispersion of a single puff which is representative of the ensemble average of a series of identical puffs released and dispersed in the specific urban environment under identical wind field conditions. A typical calculated concentration time history representing an ensemble-average puff at a randomly selected sensor is presented in Fig. 2.

To validate this RANS methodology by comparison with experimental data, the experiment should provide ensemble-averaged dosage-based parameters. The experimental sample of puffs should ideally be infinite large. The experiments simulated in this study involved a sufficiently large number (i.e., in the order of 200) of repetitions of puff releases under identical flow conditions. The calculation of the ensemble-averaged parameters from the experimental data can be performed alternatively: (1) by calculating the ensemble average of the observed concentrations from all individual puffs over each sensor and thus producing one average concentration time-history per sensor. Then, the ensemble-averaged dosage-based parameters are calculated from the average concentration time-history, or (2) by calculating the dosage-based parameters of each individual puff concentration time history at each sensor and then finding the ensemble average for each parameter at each

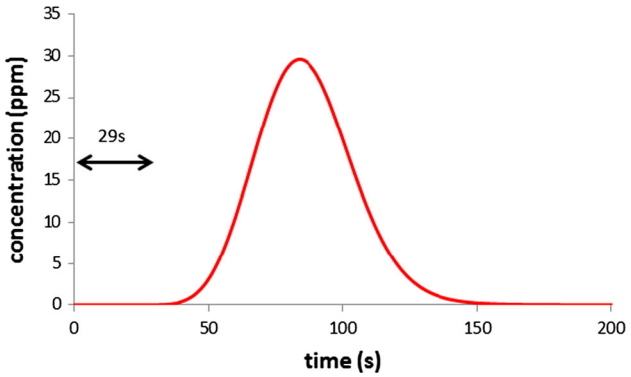



**Fig. 2** Typical calculated concentration time-history representing an ensemble-average puff at a randomly selected sensor. The release duration was 29 s. The release of tracer is indicated by the *black arrow*, while the *red line* is the concentration predicted at a sensor downstream of the source

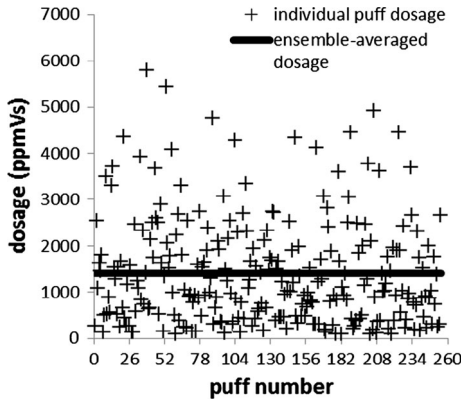



**Fig. 3** Observed dosage calculated from each individual puff (with *cross*) and the ensemble-averaged dosage (with *solid line*), at a particular sensor

sensor. In Fig. 3, the dosage calculated from the observed concentration of each individual puff at a particular sensor is indicated with a cross, together with the ensemble-averaged dosage (solid line) at this sensor.

In the present paper, the experimental data were averaged in the second way. The repeatability of the puff dispersion that is required for a reliable calculation of ensemble averages can only be ensured in a wind tunnel where the mean flow field can be kept stationary. There are field experiments where puff releases have been performed (e.g., the MUST experiment, Biltoft 2001), but their disadvantage is that it is practically impossible to repeat the same experiment as many times as in a wind tunnel.

## 3 Test cases and data descriptions

To validate the proposed methodology, experimental data sets were utilized which pertain to the "Michelstadt" and "CUTE" (Complex Urban Terrain Experiment) test cases

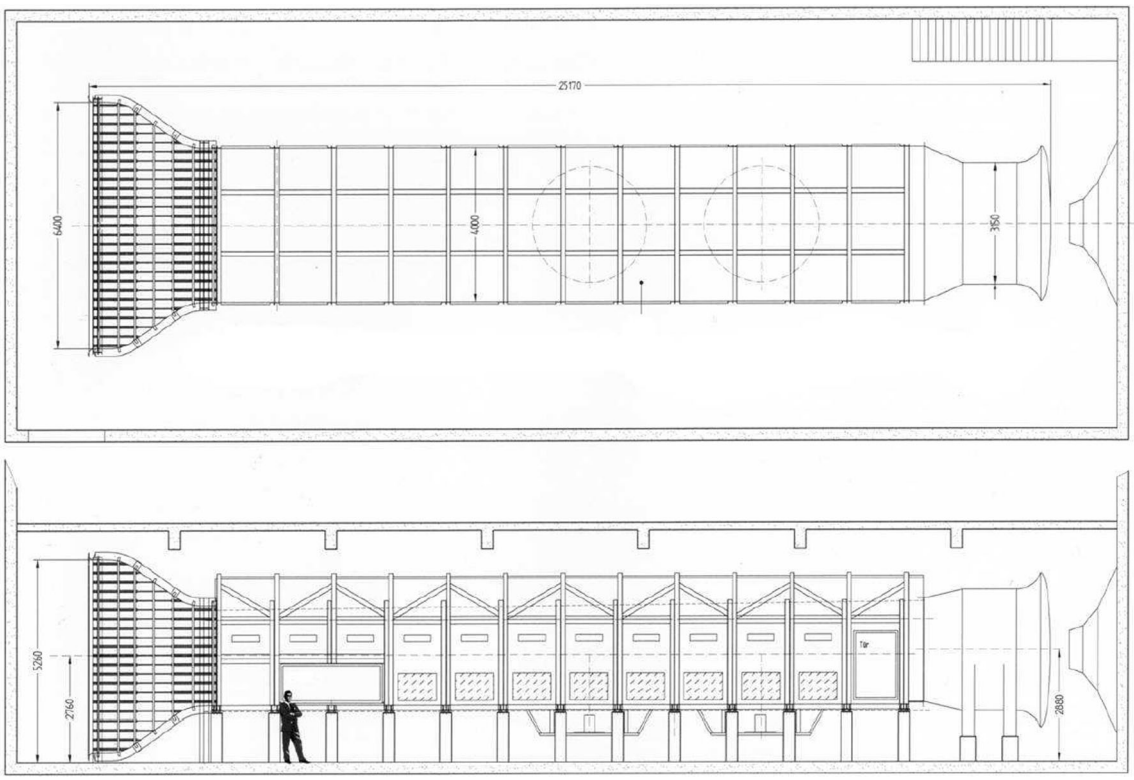



**Fig. 4** Drawing of the large boundary layer wind-tunnel facility 'WOTAN' of the Universität Hamburg (pictures courtesy of Professor Bernd Leitl)

studied in the framework of the COST Action ES1006. The test cases and the relevant data will be described below. For both test cases, authors used only data obtained from wind-tunnel experiments, although the CUTE experiment involved also field measurements.[2]

### 3.1 Wind-tunnel configuration and experimental procedure

Wind-tunnel experiments can be used to study atmospheric dispersion, in particular with respect to turbulent motion and flow dynamics within the lower atmospheric boundary layer. The advantage of a wind-tunnel measurement over field measurements lies in its statistical representativeness. The boundary conditions in a wind tunnel can be controlled and systematically varied precisely. In contrast to full-scale experiments, mean boundary conditions in the laboratory can be kept stationary. Thus, statistically representative data for model validation can be generated, by carrying out a sufficient amount of repetitive measurements for short-term releases.

Within the scope of COST Action ES1006, the large boundary layer wind-tunnel facility "WOTAN" at the Environmental Wind-Tunnel Laboratory of Universität Hamburg was used for the measurements (Fig. 4). A neutrally stratified model boundary layer flow is generated by a carefully optimized combination of turbulence generators (so-called 'spires') at the inlet of the test section, and a compatible floor roughness.

Flow measurements were carried out with 2D fibre-optic laser-Doppler-anemometry (LDA) as exemplarily shown in Fig. 5. Data sampling rates were in the range of at least 100–500 Hz (model scale). The size of the measurement

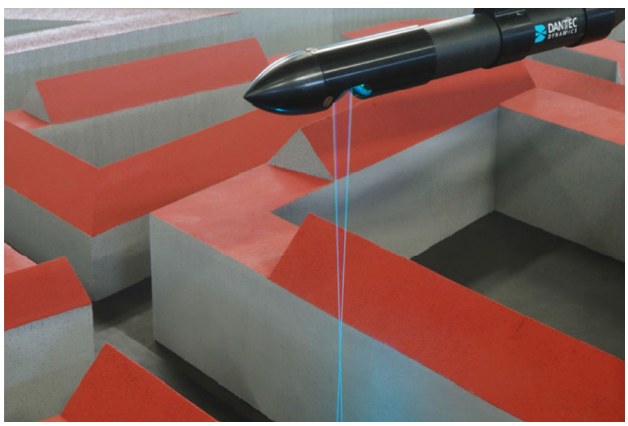



**Fig. 5** 2D fibre-optic laser-Doppler-anemometer-system measuring the *U*- and *V*-components of the wind vector

[2] Data set contacts are available in http://elizas.eu/index.php/data-set-contacts.html.

control volume of the flow measurements was 1.65 mm × 0.07 mm × 0.07 mm at wind-tunnel scale. To assure an adequate representativeness of statistical quantities derived from the measured time series, the convergence of results was tested and a minimum measurement time of 270 s (model scale) was defined at a free-stream velocity $U_\infty$ of about 9 $ms^{-1}$. Reference wind speed was permanently monitored close to the tunnel inlet through Prandtl tube measurements.

Concentration measurements were carried out to study the dispersion of passive airborne hazardous materials in the modelled areas. Ethane tracer gas was emitted from ground-level point sources. Solenoid valves representing the sources were used to release tracer in short-term mode. A fast flame ionization detector (FID) measured concentration time series with 140 Hz resolution (model scale) at numerous measurement locations. The background concentration was constantly monitored with an additional FID.

Measurement devices were checked and calibrated frequently to maintain a high accuracy of measured data. The overall uncertainty of the measured data was determined based on a sufficient number of independent repetitive measurements.

The results in the model validation data sets are provided as full-scale values. The conversion of the different variables from model scale to full scale was based on dimensionless numbers.

For the short-term (puff) releases, distributions of puff parameters and their statistical descriptions (such as median, mode, percentiles, skewness, etc.) are provided. To ensure sufficient statistical representativeness of the puff dispersion results, data from at least 200 individual releases were collected for each measurement location. For each release, typical puff dispersion parameters were calculated and converted to full-scale conditions. The validation data set is completed by a comprehensive documentation facilitating the safe use of the provided data.

### 3.2 The Michelstadt wind-tunnel experiment

The "Michelstadt" wind-tunnel experiment (Fischer et al. 2010) was designed as the first application-specific test case for the validation of local-scale emergency response models. The building structure named "Michelstadt" represents an idealized Central-European urban environment. Figure 6

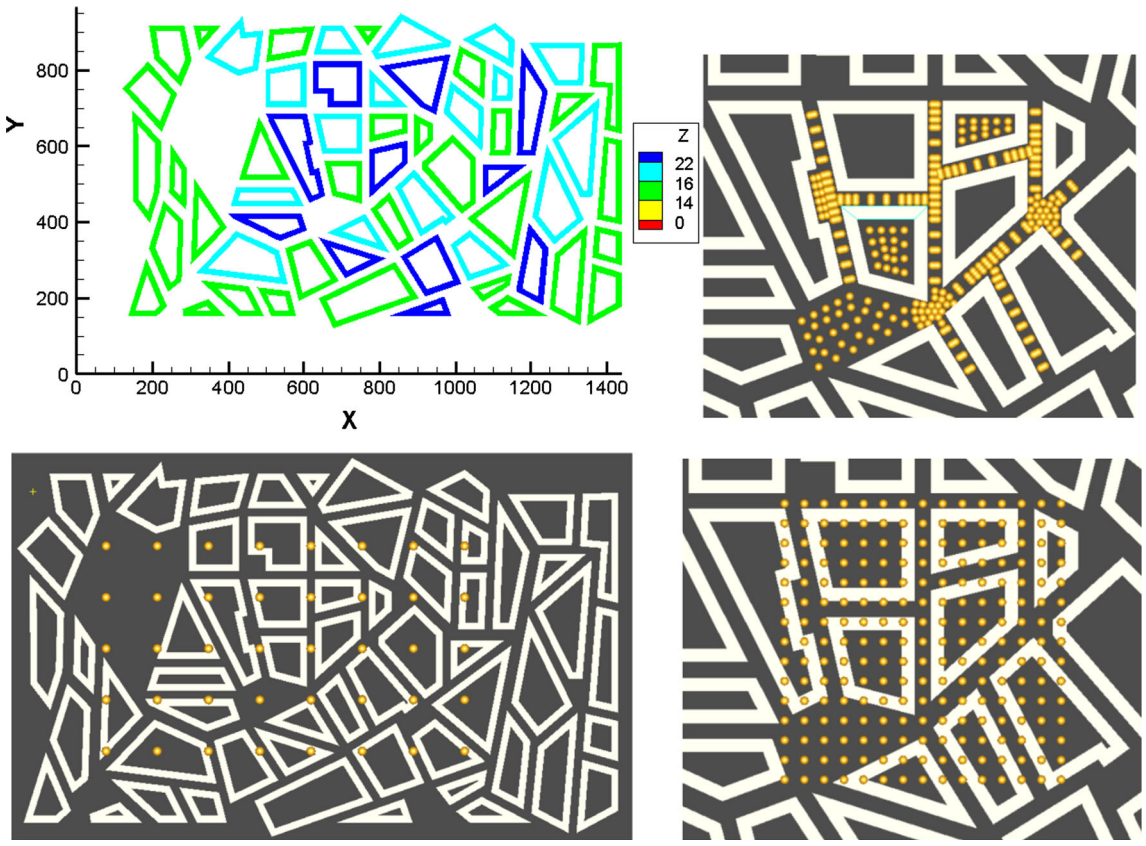



**Fig. 6** Model of "Michelstadt". Measurement locations are indicated by *yellow circles*. *Top left* the idealized city domain dimensions and building heights (meters, in full scale); *top right* densely spaced measurement locations within the urban canopy layer at elevations of 2, 9, and 18 m; *bottom left* vertical profile locations; *bottom right* densely spaced measurement locations above roof top at heights of 27 and 30 m

indicates the urban layout developed and used for model evaluation. The geometric scale of the wind-tunnel model is 1:250. Flow and concentration measurements were carried out in selected relevant locations with a higher density of data close to the ground. Measurements were collected for seven release scenarios corresponding to different point source locations and two different wind directions. Short-term (puff) releases were carried out. Flow and concentration data were made available in a first "open" test case for the modelling exercise. In a second "blind" test, only minimum information on inflow data and the emission description were provided to the modellers.

Information about the urban wind field is available from 40 vertical velocity profiles measured all across the urban model (Fig. 6, triangles). Within the core of the model, densely spaced measurements at five different heights above ground were conducted. The three lowermost measurement heights are located within the urban canopy layer at elevations of 2, 9, and 18 m (full scale at various ($x$, $y$) locations), which are indicated by black dots in Fig. 6. The other two measurement heights are located above the rooftop at elevations of 27 and 30 m. Their horizontal locations are indicated with squares in Fig. 6 (225 positions per height). Each horizontal measurement plane contains 383 measurements points.

The measurement results were transferred to non-dimensional values (on the basis of the reference wind speed measured at a height of $z_{\mathrm{ref}} = 100$ m above ground). The non-dimensional mean values were scaled to full-scale conditions with a reference wind speed of 6 $\mathrm{ms}^{-1}$ at a reference height of 100 m above ground.

For dispersion experiments, seven point sources were used consecutively in short-term-release mode (Fig. 7). Two opposite wind directions were simulated.

The design of the experiments took into account that hazardous releases may occur at very different locations in an urban area: in open squares, narrow or wide streets, streets aligned perpendicular or parallel to the prevailing large-scale flow, or even in courtyards. In general, the distribution of the plume from a release in a rather open part of the urban building structure (e.g., source S2 in Fig. 8) should be well reproduced by the model, while releases from a street perpendicular to the mean approaching flow (e.g. source S5 in Fig. 9) or a courtyard (source S08 in Fig. 10) are affected by more complex flow structures in the vicinity of the source, and larger differences are expected between measurements and corresponding model simulations.

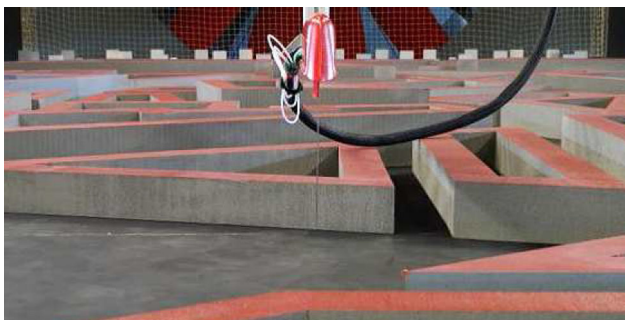


**Fig. 7** Fast FID system used for concentration measurements inside

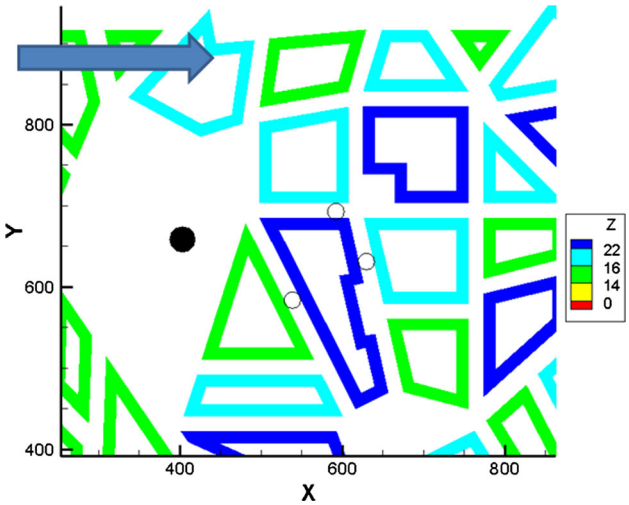



**Fig. 8** Positions of the source location S2 (*filled black circle*) situated at an open yard, and of the receptor points (*open circles*), where concentration was measured, for the non-blind test run with puff release. The wind direction is from left to right, as indicated by the *arrow*. Dimensions given in full scale, in meters

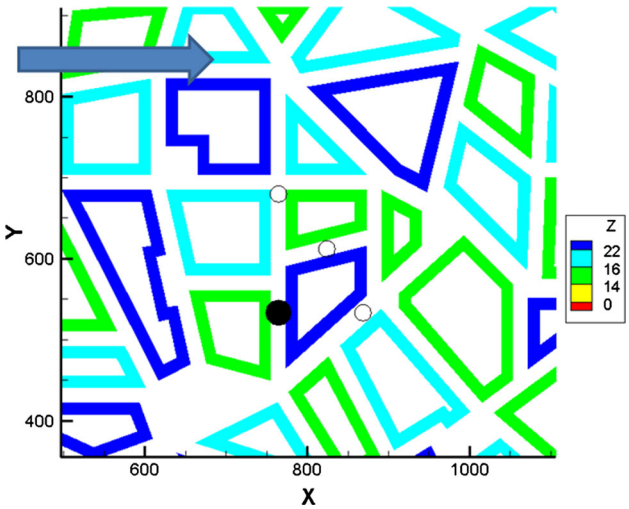



**Fig. 9** Positions of the source location S5 (*filled black circle*) situated in a street perpendicular to the wind direction in this experiment, and of the receptor points (*open circles*), where concentration was measured, for the non-blind test run with puff release. The wind direction is from *left* to *right*, as indicated by the *arrow*. Dimensions given in full scale, in meters

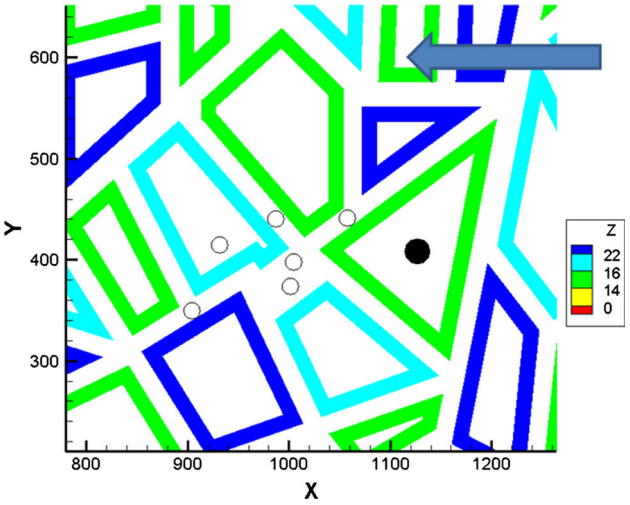



**Fig. 10** Positions of the source location S08 (*filled black circle*) situated in a courtyard, and of the receptor points (*open circles*), where concentration was measured, for the blind test run with puff release. The wind direction is from *right* to *left*, as indicated by the *blue arrow*. Dimensions given in full scale, in meters

On the other hand, the affected area may also comprise various building configurations. Therefore, receptor points (locations of concentration measurements) are also situated in different surroundings as depicted in Figs. 8, 9, and 10.

### 3.3 The complex urban terrain experiment (CUTE)

The CUTE data set includes results from both field and wind-tunnel measurements. The data set is dedicated to test emergency response tools/atmospheric dispersion models predicting dispersion processes in urban areas.

The tests were done in the downtown area of a typical central-European city. The area of interest is densely built-up with building heights between 25 and 35 m. There is no significant ground elevation and no major urban greenery in the modelled area. Figure 11 shows the heterogeneous structure of the selected city.

The second part of the CUTE data set (CUTE Cases 2–4) consists of measurements in the wind tunnel, where a scaled model of the European city centre was constructed (Fig. 12). The geometric scale of the wind-tunnel model is 1:350.

The flow-related data in the CUTE wind-tunnel data set included only the wind speed and direction of the approaching flow at the reference height to simulate a real accidental release scenario when only limited information about the boundary conditions would be available. The wind direction was chosen to be 235° (south-westerly winds).

The CUTE wind-tunnel data set consists of concentration data only. Concentration time series of tracer gas from puff releases were measured at pedestrian level.


**Fig. 11** Selected test site in a densely built-up urban environment of a typical central-European city (image source: Google Earth). The *star* indicates the source location

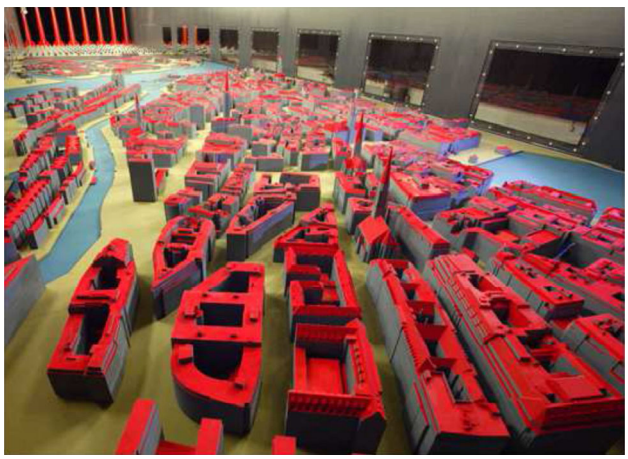


**Fig. 12** Wind-tunnel model of a European city centre (CUTE)

Concentration measurements are available for three cases. Each case represents a different source location. However, only one case ("Case 3") was used during the model evaluation presented in this paper. In the selected case, the source is located between houses near the river on the opposite side of the harbour (see also Figs. 11 and 14).

## 4 The CFD-RANS numerical simulations

The present results were produced in the frame of the COST Action ES1006. According to the strategy of this Action, during an emergency situation, a little information is available to the modeller and the available time for computations is limited. Thus, the selection of various parameters for the numerical simulations, e.g., the discretization of the grid, the turbulence model, the numerical schemes, the inlet boundary conditions, etc., was left to the modellers' judgement. Furthermore, different strategies

were followed concerning the dissemination of the experimental results to the modellers. For the Michelstadt experiment, the modeller had access a priori to the results (i.e., concentration measurements) of three cases/emission sources (S2, S4 and S5). Such cases were referred to as "non-blind" tests. The source S2 was located in an open space. The source S4 was located in an along-the-wind street canyon, while the S5 in a perpendicular-to-the-wind street canyon. The concentration measurements of the other four cases/emission sources (S05, S06, S07, and S08) were not provided a priori to the modellers but only after the end of the Action. Such cases were referred to as "blind" tests. The source S05 had the same location as the source S5. The sources S06 and S07 were located in street intersections. The source S08 was located in a courtyard. The CUTE experiment was treated as a blind test not only concerning the results but also concerning the inlet boundary conditions. Only a single velocity at a reference height was provided to the modellers, while for Michelstadt, a vertical velocity profile was provided for the inlet boundary conditions.

### 4.1 The ADREA-HF code

The computational simulations presented in chapter 2 were performed with the in-house code ADREA-HF (Efthimiou et al. 2016b; Bartzis et al. 1991; Venetsanos et al. 2010). The ADREA-HF is a three-dimensional CFD code. Using the finite volume methodology in Cartesian structured grids, it calculates the unsteady flow and dispersion of pollutants into arbitrarily complex geometries. It was developed primarily for atmospheric applications at local scale and generally for pollutant dispersion problems, but it is sufficiently generic to be used for a wide range of CFD problems.

In general, the working medium can be considered a multicomponent mixture, where each component may be simultaneously in a gaseous or non-gaseous state, i.e., liquid or solid. The mixture in a cell is considered to be in thermodynamic equilibrium. Equations of state of both ideal and real gases (e.g., Peng–Robinson) are available. The phase distribution is estimated using Raoult's law. The conservation equations for the total mass, the mass of each component, the momentum, and the energy can be solved in the general case. The code can also solve a transport equation for the concentration variance. The discretized momentum equations are solved on a staggered grid. There are turbulence models of zero, one and two equations available (LVEL, mixing length, anisotropic of one equation, $k$–$\varepsilon$ and RNG $k$–$\varepsilon$), as well as LES. The pollutant(s) emission rate(s) can have any variation with time, the pollutant(s) can be gas, liquid, or two-phase with physical properties (e.g., density, temperature, specific heat) defined by the user, and the code can handle expanded or under-expanded jets.

The grid can be independent of the geometry of solid surfaces, which are allowed to cross it in any way. This is accomplished using the technique of medium porosity and surface permeability (Sha 1980; Moult et al. 1979). According to this, the computational cells are classified into fully active with porosity 1, to inactive with porosity 0, and partially active (when they are intersected by solid boundaries) having a porosity between 0 and 1. The porosity is defined as the ratio of the fluid volume to the total volume of the cell. Calculations are not performed in inactive cells and they are not stored in memory. The conservation equations are solved in the fully active cells, while in the case of partially active cells, the integration of conservation equations is performed in their free (fluid) volume. The solid bodies in each partially active cell are taken into consideration approximately using the geometric data for each solid surface that penetrates the cell. Such data are the orientation of the surface, its location, the distance from the centre of the cell, and the surface permeability in the three axes (Bartzis et al. 1991). Surface permeability is the ratio of the free surface to the total surface (front) of the cell.

The available numerical schemes for the convection terms of conservation equations include a large number of classic and more recent schemes of first-, second-, and third-order accuracy, like the upwind, central, linear upwind, cubic upwind, and Fromm and QUICK schemes. There is the scheme family of type $k$ ($k$-schemes) and schemes which are based on the concept of restriction functions (limiters) as VanLeer, Van Albada, OSPRE, MinMod, Super-B, MUSCL, SMART, Umist, and CHARM. The implementation of schemes has been made for non-equidistant grids. For the time discretization, there are schemes of first and second order: the implicit Euler method, the Crank–Nicolson method, and the implicit second order method. The pressure is obtained from the equation of continuity using the algorithm ADREA/SIMPLER (Bartzis et al. 1991).

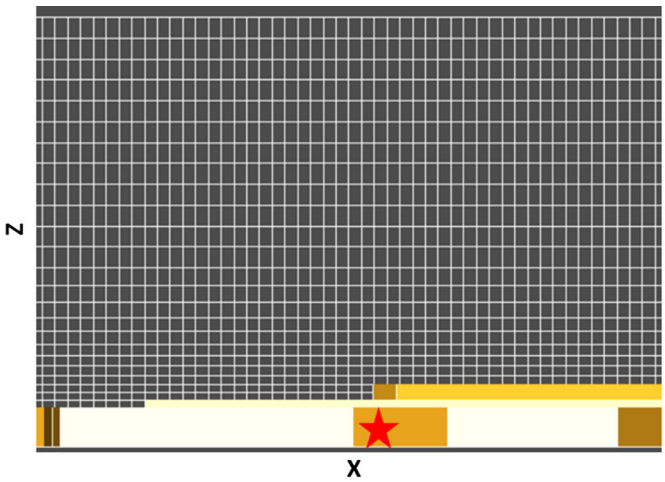
The linear systems resulting from the discretization of the conservation equations can be solved using a wide variety of solvers, from simple iterative methods like Gauss–Seidel and Line Gauss–Seidel to complex Krylov methods like bi-conjugate gradient stabilized method and GMRES (with various preconditioners like ILU (0), ILU (1), MILU (0), MILU (1), and Kovalets et al. 2008). The code has recently been parallelized for both shared memory architectures and distributed memory systems, and is available in 32-bit and 64-bit versions for both Windows and Linux. There is the capability of implementing an automatic process for the variation of the integration time step, while the user can control the solution stability by

setting the CFL (Courant) number and/or the maximum time step.

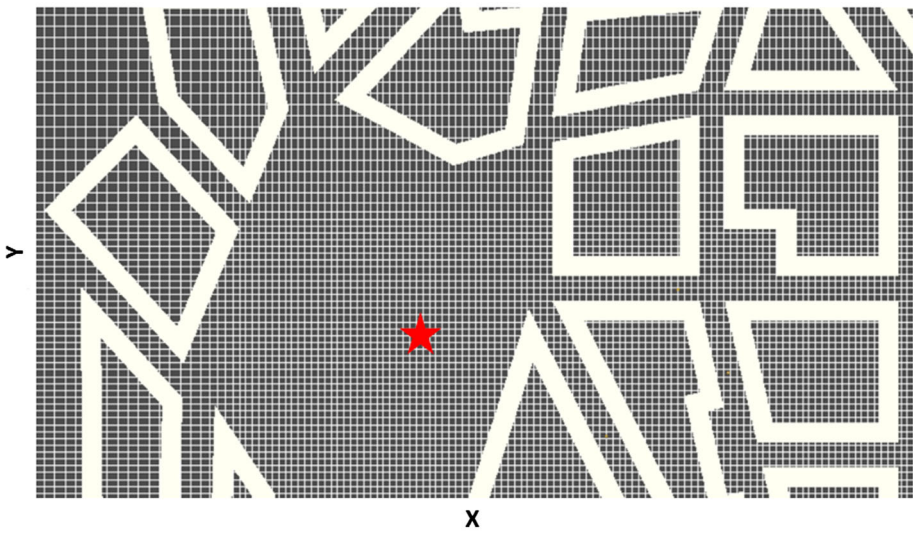
### 4.2 Simulations of the Michelstadt experiment

The computational domain had a total spatial extent of $68H_{max} \times 44.6H_{max} \times 6.85H_{max}$ ($H_{max}$ is the maximum height of buildings equal to 24 m) in the streamwise, spanwise, and vertical directions, respectively. An upstream fetch of $5H_{max}$ (distance between the inlet—$x$ plane and the windward face of the first obstacle) was used in the simulations. A coarse grid of 259 × 154 × 29 control volumes was used for the discretization of the domain. Figure 13 shows the grid used in the simulations. Due to the limited available time, no sensitivity of the solution to the discretization of the computational domain was performed in this study. However, in a recent publication concerning ADREA-HF (Efthimiou et al. 2016a), grid independency of the solution was tested, and it was found that the spatial resolution used in the current study delivers results fitting well to validation data. Considering that in emergency response applications, where time is critical, and no grid sensitivity tests can be performed; the grid resolution used here is realistic for practical users.

At each solid surface (e.g., ground, obstacle walls, and roof) in the computational domain, zero-flux boundary conditions were used for the concentration field. Furthermore, at these solid boundaries, standard wall functions were used for the mean velocities and turbulence quantities. These log-law-based 'wall laws' are used with high-Reynolds number turbulence models to relate the flow quantities at the first grid point away from the solid surface




**(a)** *x-z* view




**(b)** *x-y* view

**Fig. 13** Portion of the grid system for the Michelstadt simulations. The star indicates the source location of the non-blind test case S2

to the friction velocity at the associated surface. A state of local equilibrium of turbulence, in which the production and dissipation of turbulence energy are balanced, is assumed near a solid surface.

At the inlet boundary plane ($-x$) of the three-dimensional domain, the inflow conditions were computed by separate solution of the one-dimensional (1D) momentum equation in the vertical direction. For the 1D simulation, the horizontal velocity at the top of the domain was adjusted to obtain the vertical profiles of velocity, turbulent kinetic energy, and dissipation (for the $k$–$\varepsilon$ model) in agreement with the experimental measurements. Since the selected incident wind direction was 0° relative to the domain $x$-axis, the momentum equation for the $u$ velocity component had to be solved.

For the velocity component $u$, a constant experimental value was given at the upper free surface of the computational domain, while a zero value was set for velocity component $w$ and zero gradient for the velocity component $v$. A constant value was given also for the TKE and its dissipation rate was calculated from the 1D simulations. For the concentration, a zero value was set at inflow boundaries. At outflow boundaries and at the lateral planes ($-y$ and $+y$), zero gradient was imposed on all variables. At the lateral planes, the velocity $v$ was set to zero.

It is important to note that the tracer source was enclosed in only one grid cell. The source was modelled by a surface placed at the same location and height inside the computational domain as in the experiment. The area of the surface and the physical properties of the passive tracer were the same as in the experiment. The release duration of the single puff was set equal to 29 s and the flow rate equal to 0.34 kg $s^{-1}$ (as for the experiment, full-scale values). The problem was treated as a transient case with a total simulation period of 1600 s for the single puff release.

The simulations were performed in full scale. The calculated three-dimensional (3D) fields of mean mass fractions were transformed to ppmV to be comparable with the experimental measurements. Subsequently, they were interpolated to obtain the concentration values at the exact locations of the experimental sensors. The dose-related parameters were then calculated in the same way as the corresponding experimental values.

Furthermore, the following choices were made for the 3D simulations:

- For turbulence modelling, the standard $k$–$\varepsilon$ (modified to include buoyancy terms) and the $k$–$l$ (Bartzis 1989) models were used.
- The maximum time step for concentration computations was set to 0.666 s to obtain the same temporal resolution of the calculated concentration time series as the experimental of 1.5 Hz.
- For the discretization of the convective terms in the momentum and mass fraction equations, the upwind scheme was used. This selection was made for reasons of computational time economy. The acceptance of the relatively larger numerical error of the upwind scheme in comparison to higher order schemes is compensated by its fast and robust convergence, which is important in emergency response applications.

### 4.3 Simulations of the CUTE experiment

The domain with its building arrangement for the CUTE experiment is shown in Fig. 14a. With respect to the ($x$, $y$, $z$) coordinate system shown in that figure ($x$, $y$: horizontal, $z$: vertical), the computational domain extends from $x = -540$ m to $x = 2105$ m, from $y = -540$ m to $y = 1755$ m, and from $z = 0$ m to $z = 6\,H_{\max}$, $H_{\max}$ being the maximum building height equal to 108 m. The origin of the coordinate system is shown in Fig. 14a, and is located, so that the smallest $x$-coordinate of any obstacle (building) is at $x = 0$, and likewise for the $y$-coordinate. The pollutant source is located at $x_s \equiv (x_s, y_s, z_s) = (229\text{ m}, 295\text{ m}, 0)$. The incident wind direction is at 35° relative to the $x$-axis.

In the spirit of the COST exercise, to imitate conditions of urgency, no sensitivity of the solution to the discretization of the computational domain was performed. A coarse grid of 157 × 129 × 25 control volumes (in the $x$-, $y$-, and $z$-directions, respectively) was used for the simulations. Partial views for the discretization of the computational domain, in the ($x$, $z$) (vertical) and ($x$, $y$) (horizontal) planes, are displayed in Fig. 14b and c, respectively. The grid is equidistant between the buildings and increases logarithmically outside the urban area.

At the inflow boundary planes ($-x$ and $-y$) of the three-dimensional domain, the inflow conditions were computed by separate solution of the one-dimensional (1D) momentum equation in the vertical direction. For the 1D simulation, the horizontal velocity at the top of the domain was adjusted to obtain the given wind velocity at the reference height. It should be noted that, since the selected incident wind direction was 35° relative to the domain $x$-axis, the momentum equation for both the $u$ and $v$ velocity components had to be solved.

At the outflow boundary planes ($+x$ and $+y$), the gradients of all flow variables were assumed to vanish; viz., $\partial\bar{u}/\partial n = \partial\bar{v}/\partial n = \partial\bar{w}/\partial n = \partial\bar{k}/\partial n = \partial\bar{\varepsilon}/\partial n = 0$where $n$ is $x$ and $y$. At the upper boundary, zero gradient boundary conditions were used for all the flow variables except from the $w$ velocity which value was calculated from cell mass balance. Finally, at all solid boundaries (ground, roof, and walls of buildings), standard wall functions are used for mean velocities and turbulence quantities.

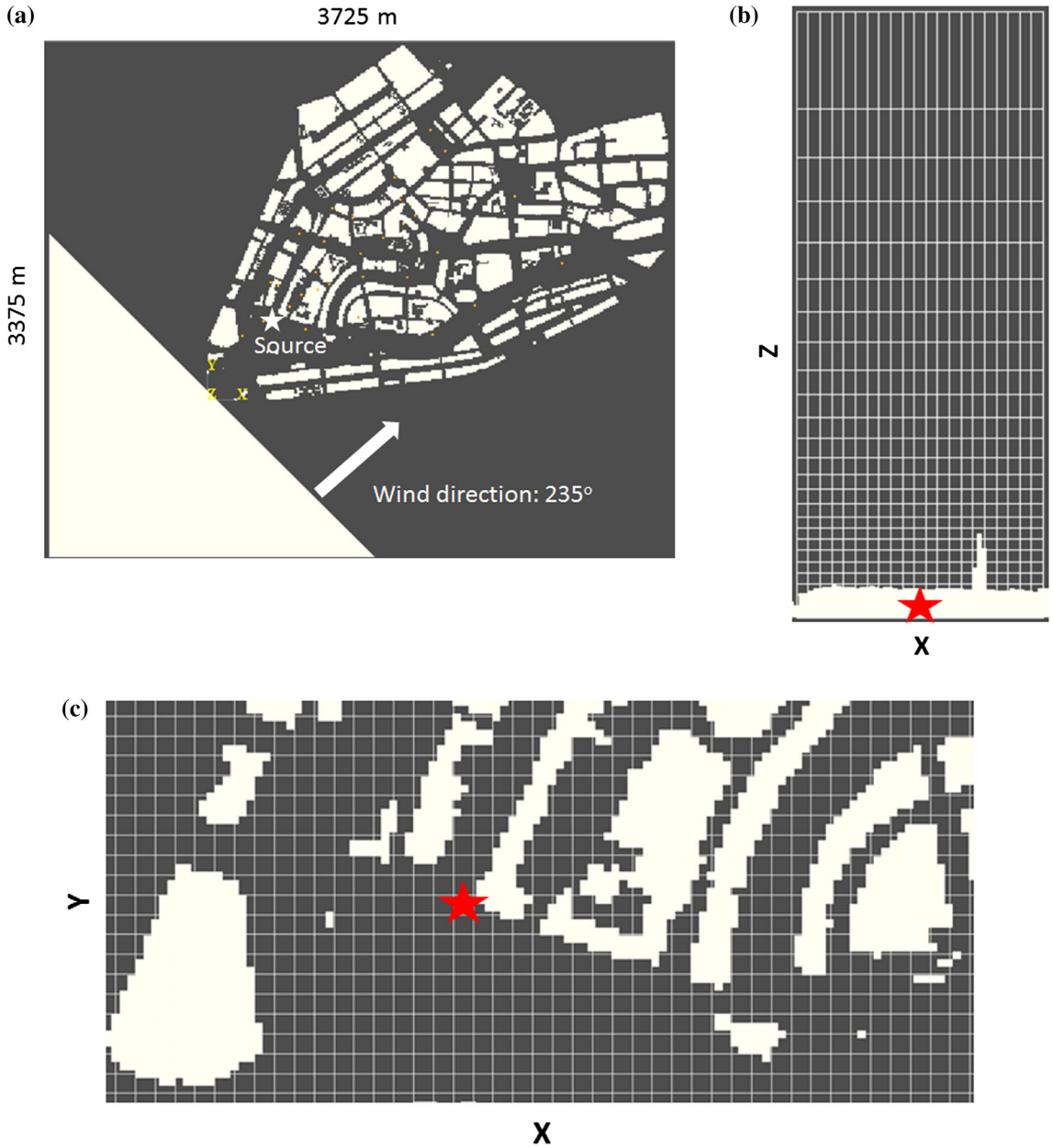



**Fig. 14** **a** Geometry of the buildings in the wind tunnel. **b** Partial *x*–*z* view (*vertical plane*) of the computational grid. **c** Partial *x*–*y* view (*horizontal plane*) of the computational grid. The location of the ground-level source is marked by a star. It is noted that the size of the *star* is exaggerated, so that the source is visible on the scale of the figure. The actual source is contained in only one grid cell of the computational domain

For the mean concentration at all solid boundaries (e.g., ground, walls, and roofs of buildings) and at the computational (flow) domain boundaries (i.e., inlet, outlet, and top boundaries), zero Neumann conditions were used. The source was modelled by a surface placed at the same location and height inside the computational domain as in the experiment. The area of the surface and the physical properties of the passive tracer were the same as in the experiment. The duration of the single puff was set equal to 31 s and the flow rate equal to 1.61 kg $s^{-1}$ (as in the experiment, full-scale values). The problem was treated as a transient case with a total simulation period of 1000 s for the single puff release.

The simulations were performed in full scale. The calculated three-dimensional (3D) fields of mass fractions were transformed to ppmV to be comparable with the experimental measurements. Subsequently, they were interpolated to obtain the values at the exact locations of the experimental sensors. The dose-related parameters were then calculated in the same way as the corresponding

experimental values. Furthermore, the temporal resolution of the numerically produced concentration time series was set equal to the experiment (1 Hz in full scale).

Furthermore, the following choices were made for the 3D simulations:

- For turbulence modelling, the standard $k$–$\varepsilon$ (modified to include buoyancy terms) model was used.
- The time step was automatically adapted according to the desired Courant–Friedrichs–Lewy (CFL) number whose maximum value was set to 2.
- For the discretization of the convective terms in the momentum and mass fraction equations, the upwind scheme was used. The rationale for this selection has been presented in Sect. 4.2.

## 5 Numerical results

In this paper, the comparison of the model versus the experiment is performed with the aid of scatter plots and validation metrics. The scatter plots depict the calculated versus the measured values at the same locations. In the scatter plots that follow, auxiliary lines are also drawn: the 1:1 line (compact diagonal line), and the factor-of-two lines (dotted lines), to visualize the spread of the points. Specifically for the dosage, the scatter plots are plotted in logarithmic scale to make the points more easily visible, since the values vary by many orders of magnitude. Concerning the validation metrics, the fractional bias (FB), the normalized mean square error (NMSE), and the fractions within a FACtor of 2 (FAC2) were used. These measures were suggested by Hanna and Chang (2012) and are widely used for atmospheric dispersion model evaluation (e.g., Schatzmann et al. 2010):

$$\mathrm{FB} = 2\frac{\bar{C}_\mathrm{O} - \bar{C}_\mathrm{P}}{\bar{C}_\mathrm{O} + \bar{C}_\mathrm{P}} \tag{1}$$

$$\mathrm{NMSE} = \frac{\overline{(C_\mathrm{O} - C_\mathrm{P})^2}}{\left(\bar{C}_\mathrm{O}\bar{C}_\mathrm{P}\right)} \tag{2}$$

$$\mathrm{FAC2} = \text{ fraction of data that satisfy } 0.5 \leq \frac{C_\mathrm{P}}{C_\mathrm{O}} \leq 2.0, \tag{3}$$

where $C_\mathrm{P}$ denotes the model predictions, $C_\mathrm{O}$ denotes experimental measurements/observations, and overbar ($\bar{C}$) denotes the average over the data set. The values of the FB indicate the tendency of the model to over- or under-estimate the observed values. The values of the NMSE represent the scatter between the individual pairs of observed and predicted values. A perfect model would have FAC2 = 1.0 and FB, NMSE = 0.0. FAC2 can be given also in percentage.

The following reference acceptance criteria for atmospheric dispersion modelling of accidental releases in built environments are chosenHanna and Chang (2012):

- $|\mathrm{FB}| < 0.67$,
- $\mathrm{NMSE} < 6$,
- $\mathrm{FAC2} > 0.30$ (or 30%), i.e., the fraction of $C_\mathrm{P}$ within a factor of two of $C_\mathrm{O}$ must exceed 0.30.

For Michelstadt, the non-blind test includes results of the S2, S4, and S5 cases, while the blind test includes results of the S05, S06, S07, and S08 cases. For CUTE, the examined case was blind. It is reminded that during a non-blind test, the concentration measurements were known to the modellers, while during a blind test, they were not, until the end of the COST Action. The basic reason for this distinction is that the non-blind tests would give the modellers the opportunity to test and find their model settings that would give the best agreement with the experimental data. This is the usual practice for any model evaluation study. The blind tests on the other hand were meant to evaluate the models performance in simulating a real-world emergency situation where limited information would be available to the modellers and certainly concentration measurements would not initially exist. In such a case, the model settings are left to each modeller's expert judgement. Therefore, higher discrepancies between model predictions and experimental measurements are expected in a blind test.

### 5.1 Results for the Michelstadt exercise

The continuous release variant of the Michelstadt experiment was studied computationally in the past (e.g., Efthimiou et al. 2015). However, for the present puff release variant, significantly more variables must be considered for comparison than the mean concentration used for the continuous release case. Here, the values are compared not only in space, but also in time, which is very demanding for the model.

Scatter plots comparing calculated and measured dosage for the non-blind and blind tests are presented in Figs. 15 and 16, respectively.

Figure 15 aims at comparing the performance of the two selected turbulence models, $k$–$\varepsilon$ and $k$–$l$. Concerning the non-blind S2 case, Fig. 15a ($k$–$\varepsilon$ model) shows that two of the three sensors are close to the ideal line (1 to 1). The third sensor presents a large underprediction. The same situation is observed for the $k$–$l$ model (Fig. 15b) for the same sensors, but in this case, the third sensor presents a smaller underprediction of the measurement. Concerning the S4 case, the $k$–$\varepsilon$ model predicted the values at both sensors rather well and within a factor of two of observations (FAC2), while the $k$–$l$ model underpredicted those

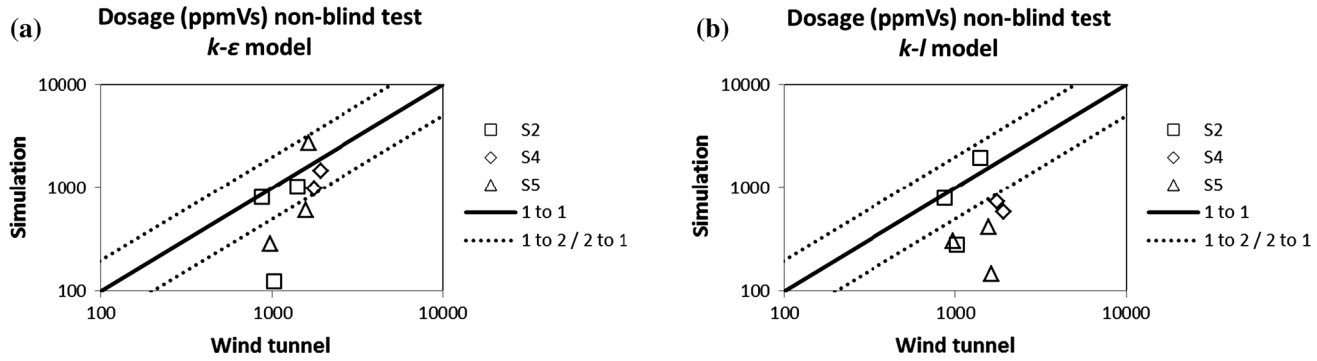



**Fig. 15** Comparison of calculated and measured dosage for the non-blind tests and all available receptor points, **a** *k–ε* model, **b** *k–l* model

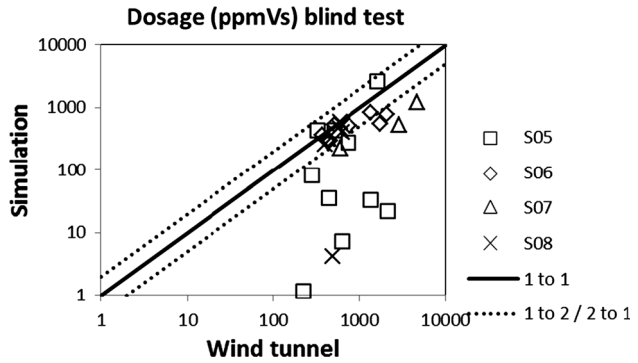



**Fig. 16** Comparison of calculated and measured dosage for the blind tests and all available receptor points

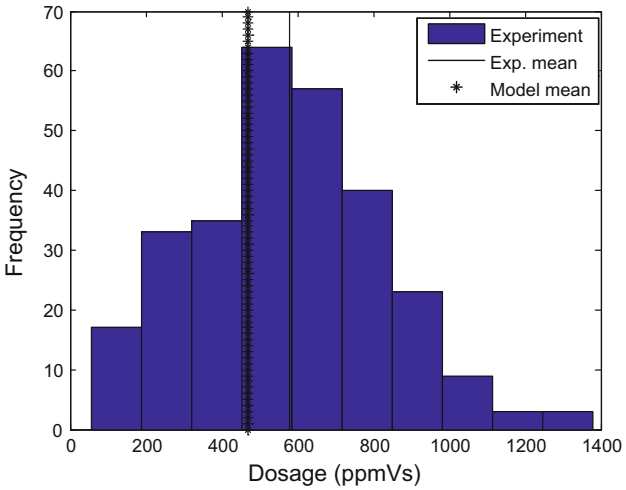



**Fig. 17** Histogram of experimental dosages for a sensor of the blind case S08. The experimental mean and model dosages are also presented

values. For the S5 case, the *k–ε* model overpredicted the dosage at one sensor—however, less than a factor-of-two—and underpredicted the dosage at other two sensors. On the other hand, the *k-l* model underpredicted the dosage at all sensors by more than a factor-of-two. Overall, it appears that the *k-ε* model presents a better performance (FAC2 = 62.5%) than the *k–l* model (FAC2 = 25%). In addition, it is obvious that for the more complex cases (i.e., S4 and S5) where the sources are located in streamwise and spanwise street canyons, the *k–l* model underpredicts the results by more than a factor-of-two. When the source is located at an open space (i.e., S2), both models present similar performance. Based on the above analysis, the *k–ε* model was selected and only results of this turbulence model are presented in the following analysis.

For the blind tests (Fig. 16), there is a general underprediction of the measurements especially for the S05 case for which the FAC2 is only 30%. The S06 case presents a better performance with an FAC2 equal to 75%. The lowest FAC2 value is obtained for case S07 (FAC2 = 0) and the highest is achieved for S08 with FAC2 = 80%.

The evaluated performance of the model depends on the position of the sensors in relation to the source location. At sensors located downwind of the source near the average puff pathway, the agreement of model results with experimental data is better than at sensors located farther from the average pathway. In addition, it should be noted that the measurement results show a high variability. An example is presented in Fig. 17 where the model-calculated dosage for a sensor of case S08 is overlaid on the distribution of the measured dosages for all the puffs released in S08. It is seen that, in this case, the model successfully calculated the average dosage (which here is close to the most probable value).

The next puff parameter to be considered is the puff duration, i.e., the time interval between the arrival time and the leaving time. Corresponding scatter plots are provided in Figs. 18 and 19.

According to Fig. 18, most of the results differ by less than a factor-of-two from the observations. The best results are achieved for the case S4, where the model predicts almost exactly the experimental puff duration at both sensors. The smaller experimental puff duration is 1.6 min and the higher is 3.7 min, while the smaller model puff duration is 1.1 min and the higher 9.2 min.

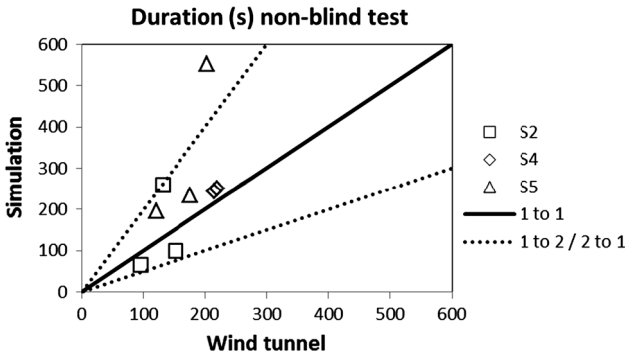



**Fig. 18** Comparison of calculated and observed puff duration values for the non-blind tests and all available receptor points

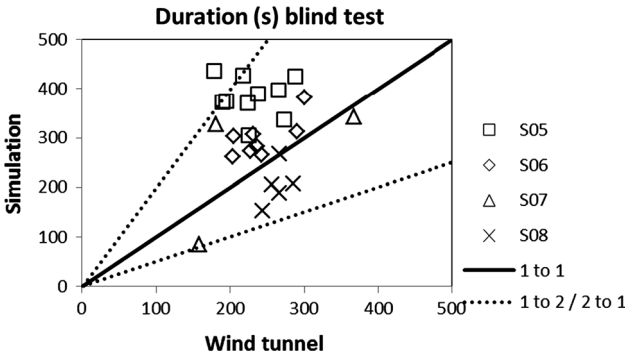



**Fig. 19** Comparison of calculated and measured puff duration for the blind tests and all available receptor points

In Fig. 19, it can be seen that the performance of the model for the blind cases is rather good as nearly all the calculated durations lie within a factor-of-two from the corresponding observations. For the cases S05 and S06, an overestimation is observed for the model, while an under-estimation is observed for the S08 case. The best agreement is achieved for the cases S06 and S08. The model performance appears similar to the non-blind test. The smaller experimental puff duration is 2.6 min, while for the model, it is 1.4 min. On the other hand, the higher experimental puff duration is 6.1 min, while for the model, it is 7.3 min.

In Figs. 20, 21, and 22, the FB, NMSE, and FAC2 metrics for dosage, peak concentration averaged over 15 s, and puff duration are plotted.

According to the FB, the model tends to under-estimate the dosage. For the blind test, the FB exceeds the acceptance limit 0.67, but the NMSE lies below the acceptance criterion value of 6. The FAC2 acceptance criterion is fulfilled for both tests (non-blind, blind). The quality of the results is better for the non-blind tests than for the blind ones.

The quality of the results for the 15-s-mean peak concentration is again better for the non-blind tests than for the blind ones. However, for both tests, the FB exceeds the acceptance limit of 0.67 indicating a high underprediction. The NMSE is lower than 6 only for the non-blind test. In addition, the FAC2 is acceptable only for the non-blind test. Overall, the blind test does not fulfil the acceptance criteria.

The puff duration is generally well predicted by the model, as shown in Fig. 22. The FAC2 is very high for both tests and the NMSE very low. The FB indicates a model overprediction of the measurements. It is surprising that, in this case, the model presents better performance for the blind test. The metrics for both tests are inside the acceptance limits, indicating that the model was able to predict successfully the duration of the puff dispersion.

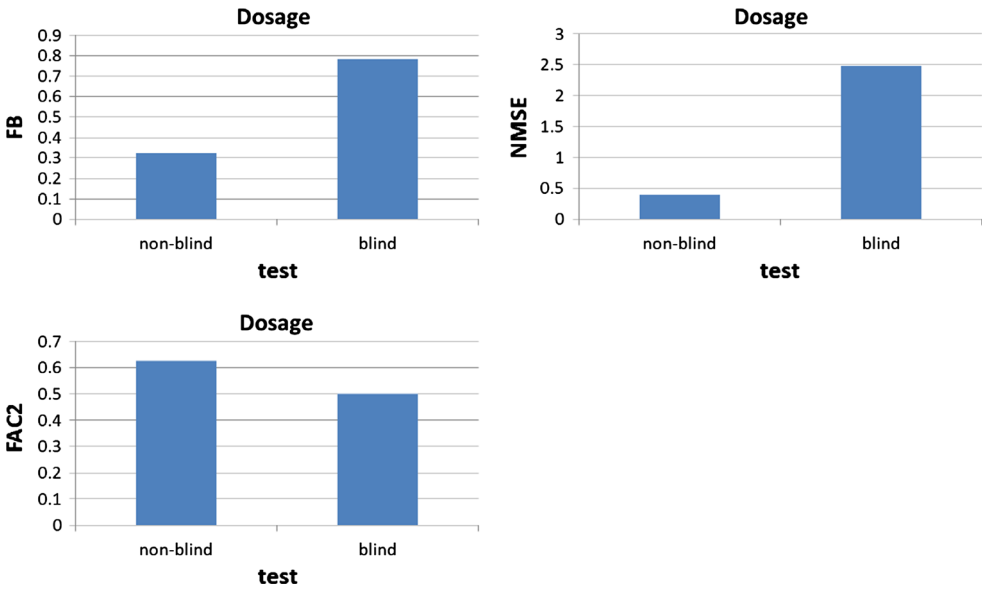



**Fig. 20** Comparison of the statistical metrics FB, NMSE, and FAC2 for dosages from all non-blind and all blind experiments with puff releases

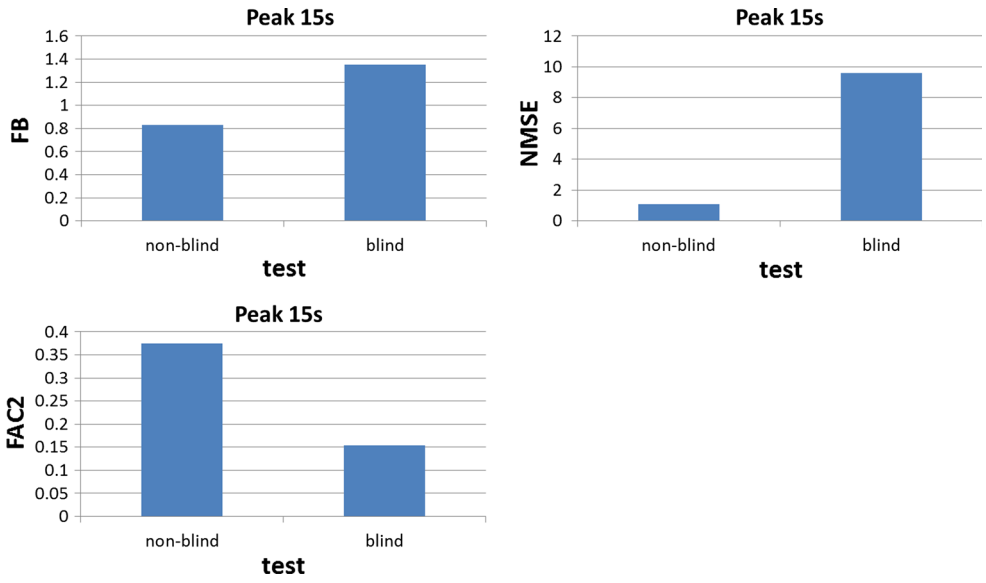



**Fig. 21** Comparison of the statistical metrics FB, NMSE, and FAC2 for 15 s peak concentrations from all non-blind and all blind experiments with puff releases

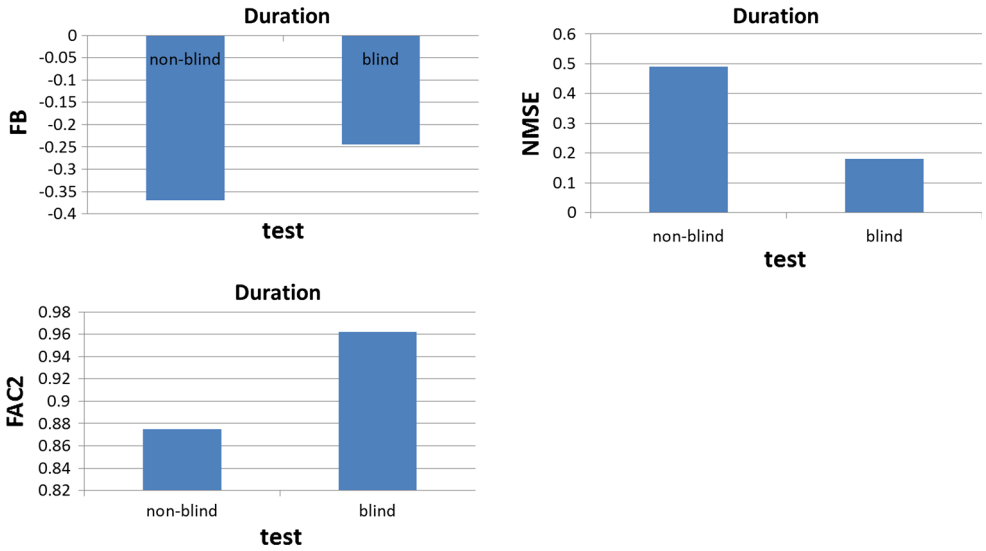



**Fig. 22** Comparison of the statistical metrics FB, NMSE, and FAC2 for puff duration from all non-blind and all blind experiments

The validation metrics for all the dosage-based parameters for both non-blind and blind tests are presented in Table 1. Concerning the non-blind tests, it is generally obvious that the model presents a better performance for the parameters related with time than for those related with concentration. Not all validation metrics fulfil the acceptance criteria. The model presents a high underprediction of the peak concentration (FB = 0.83) and a high overprediction of the ascent time (FB = -0.823). The latter means that the model predicts a longer time to reach the peak concentration than what is observed on the average. The least satisfactory results are obtained for the peak concentration, while the most satisfactory ones, for the arrival time. The scatter of results revealed by NMSE is smaller for the arrival and the peak time, while it is again largest for the peak concentration and the ascent time. For the dosage and the temporal parameters, the FAC2 is good and above 60%, while for the peak concentration, it is close the acceptance limit (37.5%).

The results of the blind tests indicate a rather good performance for the model especially for the temporal parameters. However, the dosage and peak concentration predictions are less satisfactory than in the non-blind tests and are both under-estimated. For the peak concentration, in particular, none of the metrics fulfil the acceptance criteria, while for the dosage, only the FAC2 is above the acceptance limit. Except for the ascent time, the metrics for all other temporal parameters are good and fulfil the acceptance criteria. In this case, the results are slightly better than for the non-blind test. According to NMSE, the temporal parameters present the same scatter for both the blind and non-blind tests. In general, the model

**Table 1** Validation metrics for the dosage-based parameters of Michelstadt for the non-blind and blind tests

| Dosage-based parameters | Non-blind | | | Blind | | |
|---|---|---|---|---|---|---|
| | FB | NMSE | FAC2 | FB | NMSE | FAC2 |
| Dosage (ppmVs) | 0.326 | 0.39 | 0.625 | **0.784** | 2.49 | 0.5 |
| Peak concentration 15 s averaged (ppmV) | **0.83** | 1.09 | 0.375 | **1.354** | **9.59** | **0.154** |
| Arrival time (s) | 0.083 | 0.06 | 1 | −0.102 | 0.1 | 0.923 |
| Peak time (s) | −0.197 | 0.22 | 1 | −0.264 | 0.2 | 0.923 |
| Leaving time (s) | −0.206 | 0.23 | 0.875 | −0.195 | 0.14 | 0.962 |
| Duration (s) | −0.369 | 0.49 | 0.875 | −0.244 | 0.18 | 0.962 |
| Ascent time (s) | **−0.823** | 2.14 | 0.625 | **−0.688** | 0.95 | 0.538 |
| Descent time (s) | −0.216 | 0.25 | 0.875 | −0.131 | 0.11 | 0.923 |

Bold numbers indicate the values that do not satisfy the reference acceptance criteria listed in the beginning of Sect. 5

**Table 2** Validation metrics for the dosage-based parameters of CUTE Case 3

| Dosage-based parameters | Validation metrics | | |
|---|---|---|---|
| | FB | NMSE | FAC2 |
| Dosage (ppmVs) | 0.222 | 0.23 | 0.688 |
| Peak concentration 15 s averaged (ppmV) | 0.611 | 0.95 | 0.5 |
| Arrival time (s) | 0.273 | 0.1 | 0.938 |
| Peak time (s) | 0.099 | 0.04 | 1 |
| Leaving time (s) | 0.06 | 0.02 | 1 |
| Duration (s) | −0.058 | 0.03 | 1 |
| Ascent time (s) | −0.375 | 0.25 | 0.938 |
| Descent time (s) | 0.023 | 0.03 | 1 |

All values satisfy the reference acceptance criteria listed in the beginning of Sect. 5

overpredicts the temporal parameters. This means that the model predicts the tracer to stay for a longer time at the sensors than it is observed in the experiment.

It is concluded that for both non-blind and blind tests, the CFD-RANS methodology is able to predict temporal puff dispersion parameters with a good accuracy. Less satisfactory results have been obtained for the dosage and especially the peak concentration. For the estimation of the latter, the authors have already presented further innovative ideas (Efthimiou et al. 2015).

### 5.2 Results for the CUTE exercise

The CUTE experiment consisted of four cases: one field trial and three wind-tunnel experiments. Case 1 was the field trial with a source located on a boat on the river; Case 2 was the corresponding wind-tunnel experiment. In wind-tunnel Case 3, the source was located inside the urban area near the river bank, and in Case 4, the source was located in the middle of the city centre. Here, exemplary results are presented and discussed, for one wind-tunnel experiment, i.e., Case 3. The case was run as blind test, having minimum input information available for the model simulations, as would be the case during a real incident.

The validation metrics for all the dosage-based parameters are presented in Table 2. The performance of the model is evidently good. A first general conclusion is that in this case, similarly to the Michelstadt case, temporal quantities are much better predicted than concentration quantities, as they have in most cases higher FAC2 and lower FB and NMSE values. For the dosage, the FB indicates an under-estimation of the measurements and the NMSE a rather low scatter of the values. Most of the calculated values lie inside a factor-of-two of observations (FAC2 = 68.75%) and many results follow the ideal line (for example, the experimental maximum dosage 3935 ppmVs was simulated very well by the model 3821 ppmVs). In this case, the acceptance criteria are well fulfilled by the metrics. For the peak concentration, the under-estimation is higher than the dosage, and the FB (0.611) is lower than, but close to, the acceptance criterion (0.67). A high scatter of data is indicated by the NMSE, while only half of the data are inside the factor of two of observations. A very good performance is achieved for the arrival time as the FB and NMSE are low, and the FAC2 approaches its

ideal value (93.8%). These results indicate that the model predicted well the travel time of the tracer from the source to the sensors. The results are even better for the peak time (i.e., the time of the peak concentration) and the leaving time. In these two cases, all the metrics present almost ideal values with a very slight underprediction of the measurements. The duration of the puff at the sensors presents also an ideal prediction by the model, and in this case, there is a slight overestimation, i.e., the calculated puff stays slightly longer at the sensors than in the experiment. The prediction of the ascent time (the duration from the arrival time to the peak time) is very good with a small overprediction. The descent time (and the ascent time) is directly affected by the performance of the parameters from which it is calculated. This means that the good prediction of the peak and leaving times also gives a good prediction for the descent time. The most satisfactory results are achieved for the leaving and ascent times and the less satisfactory for the peak concentration. In summary, it is concluded that the model in this case predicted much better the temporal parameters than the concentration parameters. Finally, it is noted that these results are better than those of the Michelstadt experiment, as the validation metrics are closer to their ideal values.

## 6 Conclusions

In this paper, the CFD-RANS methodology, and in particular the ADREA-HF code, was employed for predicting dosage-based parameters related to short-duration releases (puffs) of an airborne material emitted from a point source inside an urban/densely built-up environment. The prediction of dosage-based parameters from puff dispersion is a very challenging task for a model as the values are compared not only in space, but also in time. In addition, high variability is experimentally observed between different realizations of a puff dispersion experiment under identical conditions. The model validation was based on two wind-tunnel experimental data sets (Michelstadt and CUTE), which examined the dispersion of puffs in urban environments for different positions of the emission sources and were studied during the course of the COST Action ES1006. In each test case of the above experiments, a series of puffs was released to calculate ensemble-average values for the dosage-based quantities. Therefore, the experimental results were suitable for comparison with the predictions of a RANS model, as the one used in this study, which represent ensemble-average values. For a number of Michelstadt experiments, dosage-related measurements were provided to the modellers ("non-blind" tests) during the Action, while for others, the dosage-related measurements were kept unknown until the end of the Action ("blind" tests). For the CUTE experiment, only limited information concerning the inlet wind boundary conditions was provided to the modellers, aiming to evaluate the models' performance under conditions similar to a real emergency situation. The dosage-related measurements of CUTE were kept unknown to the modellers during the computational simulations.

The model performance was evaluated by comparing calculated and measured ensemble-average values of the following puff-related quantities at specific sensor locations inside the urban built-up environment: dosage, peak concentration, puff arrival time, ascent time, peak time, descent time, leaving time, and duration. In general, the ADREA-HF model presented a better performance for the prediction of temporal parameters (i.e., puff arrival time, peak time, leaving time, puff duration, etc.) than for the dosage and the peak concentration. A tendency of model overprediction was observed for the temporal parameters, indicating that the tracer is predicted to stay for a longer time at the sensors—or that the modelled puffs are more spread in space—than in the experiment. Nevertheless, the majority of the statistical indices that have been employed for the quantitative model validation had values inside the acceptance limits established in the literature. In particular, all the statistical indices for the CUTE case satisfied the acceptance criteria. In the Michelstadt case, the statistical indices for the temporal quantities satisfied the acceptance criteria, except for the FB of the ascent time. Furthermore, in the Michelstadt case, the statistical indices that did not satisfy the acceptance criteria were the FB of the dosage for the blind case, the FB, NMSE, and FAC2 of the peak concentration for the blind case, and the FB of the peak concentration for the non-blind case. Therefore, based on the obtained model performance indices, the CFD-RANS methodology as implemented in the code ADREA-HF is able to predict the ensemble-average temporal quantities related to transient emissions of airborne material in urban areas within the range of the model performance acceptance criteria established in the literature. The CFD-RANS methodology as implemented in the code ADREA-HF is also able to predict the ensemble-average dosage, but the dosage results should be treated with some caution, as in one case (Michelstadt blind), the observed ensemble-average dosage was under-estimated slightly more than the acceptance criteria (FB = 0.78 against the allowable FB = 0.67). Ensemble-average peak concentration proved to be the most challenging parameter for the model to predict as it was systematically underpredicted, although only in the Michelstadt cases to a degree higher than the allowable by the acceptance criteria (FB = 0.83 in the non-blind case, corresponding to 59% underprediction, FB = 1.354 in the blind case, corresponding to 81% underprediction; the acceptance criterion of FB $<$0.67

corresponds to a 50% underprediction). The reliable prediction of the peak concentration is a subject of further investigation by the authors. In addition to the above, it is noted that the evaluated performance of the model varied between simulated experiments and it depended on the examined sensors position in relation to the source location and to the buildings/streets configurations.

**Acknowledgements** The work presented in this paper was carried out (partly) within the scope of COST Action ES1006. The scientific exchange between Action Members and support provided by COST are gratefully acknowledged. The reference data set ("Michelstadt"/"CUTE") was compiled by members of COST Action ES1006. The provision of reference data is gratefully acknowledged.

## References

Bartzis JG (1989) Turbulent diffusion modelling for wind flow and dispersion analysis. Atmos Environ 23:1963–1969

Bartzis JG, Venetsanos A, Varvayani M, Catsaros N, Megaritou A (1991) ADREA-I: a three-dimensional transient transport code for complex terrain and other applications. Nucl Technol 94:135–148

Baumann-Stanzer K, Andronopoulos S, Armand P, Berbekar E, Efthimiou G, Fuka V, Gariazzo C, Gasparac G, Harms F, Hellsten A, Jurcacova K, Petrov A, Rakai A, Stenzel S, Tavares R, Tinarelli G, Trini Castelli S (2015) COST ES1006—model evaluation case studies: approach and results. COST Action ES1006

Berbekar E, Harms F, Leitl B (2015) Dosage-based parameters for characterization of puff dispersion results. J Hazard Mater 283:178–185

Biltoft CA (2001) Customer report for Mock Urban Setting Test. DPG Document No. WDTC-FR-01-121, West Desert Test Center, U.S. Army Dugway Proving Ground, Dugway

Ching WH, Leung MKH, Leung DYC (2009) An efficient approach to transient turbulent dispersion modeling by CFD–statistical analysis of a many-puff system. Fluid Dyn Res 41:17

Cui H, Yao R, Xu X, Xin C, Yang J (2011) A tracer experiment study to evaluate the CALPUFF real time application in a near-field complex terrain setting. Atmos Environ 45:7525–7532

Efthimiou GC, Berbekar E, Harms F, Bartzis JG, Leitl B (2015) Prediction of high concentrations and concentration distribution of a continuous point source release in a semi-idealized urban canopy using CFD-RANS modeling. Atmos Environ 100:48–56

Efthimiou GC, Andronopoulos S, Bartzis JG, Berbekar E, Harms F, Leitl B (2016a) CFD-RANS prediction of individual exposure from continuous release of hazardous airborne materials in complex urban environments. J Turbul. doi:10.1080/14685248.2016.1246736

Efthimiou GC, Andronopoulos S, Tolias I, Venetsanos A (2016b) Prediction of the upper tail of concentration distributions of a continuous point source release in urban environments. Environ Fluid Mech. doi:10.1007/s10652-016-9455-2

Fischer R, Bastigkeit I, Leitl B, Schatzmann M (2010) Generation of spatio-temporally high resolved datasets for the validation of LES-models simulating flow and dispersion phenomena within the lower atmospheric boundary layer. In: Proceedings of fifth international symposium on computational wind engineering (CWE2010), Chapel Hill

Hanna S, Chang J (2012) Acceptance criteria for urban dispersion model evaluation. Meteorol Atmos Phys 116:133–146

Hanna SR, Briggs GA, Hosker RP (1982) Handbook on atmospheric diffusion, national technical information service. U.S. Department of Commerce, Spring-field

Harms F, Leitl B, Schatzmann M, Patnaik G (2011) Validating LES-based flow and dispersion models. J Wind Eng Ind Aerodyn 99:289–295

Hertwig D, Efthimiou GC, Bartzis JG, Leitl B (2012) CFD-RANS model validation of turbulent flow in a semi-idealized urban canopy. J Wind Eng Ind Aerodyn 111:61–72

Kovalets IV, Andronopoulos S, Venetsanos AG, Bartzis JG (2008) Optimization of the numerical algorithms of the ADREA-I mesoscale prognostic meteorological model for real-time applications. Environ Model Softw 23:96–108

Moult A, Spalding DB, Markatos NCG (1979) Solution of flow problems in highly irregular domains by the finite difference method. Trans Inst Chem Eng 57:200–204

Schatzmann M, Olesen H, Franke J (2010) COST 732 Model evaluation case studies: approach and results. ISBN: 3-00-018312-4

Sha WT (1980) An overview on rod-bundle thermal-hydraulic analysis. Nucl Eng Des 62:1–24

Thomson DJ, Jones AR (2011) A new puff modelling technique for short range dispersion applications. Int J Environ Pollut 44:156–163

Venetsanos AG, Papanikolaou E, Bartzis JG (2010) The ADREA-HF CFD code for consequence assessment of hydrogen applications. Int J Hydrogen Energy 35:3908–3918

Wan J, Sui J, Yu H (2014) Research on evacuation in the subway station in China based on the combined social force model. Phys A 394:33–46

Yim SHL, Fung JCH, Lau AKH (2010) Use of high-resolution MM5/CALMET/CALPUFF system: SO2 apportionment to air quality in Hong Kong. Atmos Environ 44:4850–4858