ACCEPTED MANUSCRIPT
Published at: doi:10.1016/j.atmosenv.2011.10.053

Comparison of CFD and operational dispersion models in an urban-like environment

G. Antonioni [a,*], S. Burkhart [d], J. Burman [e], A. Dejoan [c], A. Fusco [b], R. Gaasbeek [i], T. Gjesdal [f], A. Jäppinen [g], K. Riikonen [h], P. Morra [a], O. Parmhed [e], J.L. Santiago [c]

[a] University of Bologna, Dept. of Chemical, Mining and Environmental Engineering, via Terracini 28, 40131, Bologna, Italy
[b] MoD, General Directorate of Land Armaments, Italian Government, Via Marsala 104, 00185, Roma, Italy
[c] CIEMAT, Research Center for Energy, Environment and Technology, Avda Complutense 22, 28040, Madrid, Spain
[d] DGA Maîtrise NRBC, 5 rue Lavoisier, 91710, Vert-le-Petit, France
[e] FOI, Swedish Defence Research Agency, Cementvägen 20, 901 82 Umeå, Sweden
[f] Norwegian Defence Research Establishment (FFI), P.O.Box 25, NO-2027, Kjeller, Norway
[g] Defence Forces Technical Research Centre, Explosives and CBRN Protection Technology, P.O. Box 5, FI-34111, Lakiala, Finland
[h] Finnish Meteorological Institute, Air Quality Research, P.O. Box 503, FI-00101 Helsinki, Finland
[i] TNO, Netherlands Organisation for Applied Scientific Research, P.O.Box 45, NL-2280 AA, Rijswijk, The Netherlands

A B S T R A C T

Chemical plants, refineries, transportation of hazardous materials are some of the most attractive facilities for external attacks aimed at the release of toxic substances. Dispersion of these substances into the atmosphere forms a concentration distribution of airborne pollutants with severe consequences for exposed individuals. For emergency preparedness and management, the availability of assessed/validated dispersion models, which can be able to predict concentration distribution and thus dangerous zones for exposed individuals, is of primary importance.

Air quality models, integral models and analytical models predict the transport and the turbulent dispersion of gases or aerosols after their release without taking into account in detail the presence of obstacles. Obstacles can modify the velocity field and in turn the concentration field. The Computational Fluid Dynamics (CFD) models on the other hand are able to describe such phenomena, but they need to be correctly set up, tested and validated in order to obtain reliable results.

Within the project Europa-ERG1 TA 113.034 "NBC Modelling and Simulation" several different approaches in CFD modelling of turbulent dispersion in closed, semi-confined and urban-like environment were adopted and compared with experimental data and with operational models. In this paper the results of a comparison between models describing the dispersion of a neutral gas in an idealized urban-like environment are presented and discussed. Experimental data available in the literature have been used as a benchmark for assessing statistical performance for each model. Selected experimental trials include some water channel tests, that were performed by Coanda at 1:205 scale, and one full-scale case that was tested in the fall of 2001 at the Dugway Proving Grounds in Utah, using an array of shipping containers. The paper also suggests the adoption of improved statistical parameters in order to better address differences between models, and to have a more straightforward method for comparing models suitable for emergency preparedness aims.

1. Introduction

During the last decades the number of terrorist attacks and of the related security issues remarkably grew, therefore there are increased concerns about releases of Toxic Industrial Chemicals (TICs), due to terrorist activities or accidents, in congested industrial sites or downtown urban areas. This in turn increased the request of tools for understanding and predicting concentration distribution of hazardous substances within densely populated areas such as built-up urban areas (Milliez and Carissimo, 2007; Britter and Hanna, 2003). In these areas, traditional models for predicting pollutant concentration, such as Gaussian plume modelling or integral box models for heavy gas dispersion, can fail

* Corresponding author. Tel.: +39 (0) 51 2090230; fax: +39 (0) 51 2090247.
E-mail address: giacomo.antonioni3@unibo.it (G. Antonioni).

due to critical assumptions not met, e.g. complex morphology creating non-uniform flow field (Gailis and Hill, 2006).

Thanks to recent advances in computing power, Computational Fluid Dynamics (CFD) models can now resolve the flow field around individual buildings and predict wind pathways through complex terrain as an urban centre. Such models are increasingly being used to simulate the transport of pollutants within urban areas, where the population is at risk (Milliez and Carissimo, 2007). Besides these advantages however, CFD models can be extremely computationally expensive and need to be accurately set up and tested for each specific field of application because they need the correct definition of many adjustable parameters. For these reasons and also if such models have to be used in emergency contexts, the European Defense Agency (EDA) approved the project Europa-ERG1 TA 113.034 "NBC Modelling and Simulation" (EUROPA-ERG1 TA 113.034, 2008)[1] where operational models and CFD models performance have been evaluated and compared by testing them on experimental data in order to evaluate whether traditional, fast-responding models can also be reliable for the purpose of predicting concentrations in an urban environment. However, the project was mainly focused on a benchmark among the approaches used by different countries' specialists (using both CFD or operational models), thus also some differences due to modellers' choices can be present.

For dispersion models, validation comes traditionally in the form of tracer release and capture studies. Whilst the optimum tracer study for validation of an urban model occurs in a real urban centre within the planetary boundary layer (PBL), high costs and difficulties for a correct experimental set up have limited the number of such full scale investigations to a handful (Grimmond, 2006; Batchvarova and Gryning, 2006).

In the same way that homogeneous terrain studies have invaluably aided understanding of more complex situations (Fernando et al., 2001), so the study of flows within a stylised area aids to understand flow in more complex geometries, such as a real urban area.

The Mock Urban Setting Test (MUST) investigations provide a simplified or stylised urban area, thus allowing the investigation of the physics involved in urban dispersion phenomena, without the overwhelming complexities observed in a real urban area. Results from measurements in an urban obstacle array at different scales in different experimental configurations have been presented in the literature (Yee et al., 2006). The full scale case was tested in the fall of 2001 at the Dugway Proving Grounds in Utah with a 10 × 12 array of shipping containers (Biltoft, 2001). Water channel tests were performed by Coanda at 1:205 scale (Hilderman and Chong, 2004). Both data sets have been used for the comparison of different models, from Gaussian plume models to more sophisticated Large Eddy Simulations (LES), in order to identify discrepancies between models, influential factors, and to improve or discard respective sub-models for the given purpose.

2. Description and selection of experimental data

The MUST trials were designed to test the effects of an array of roughness elements (buildings) on the flow and dispersion of pollutants within an idealised urban morphology under a range of

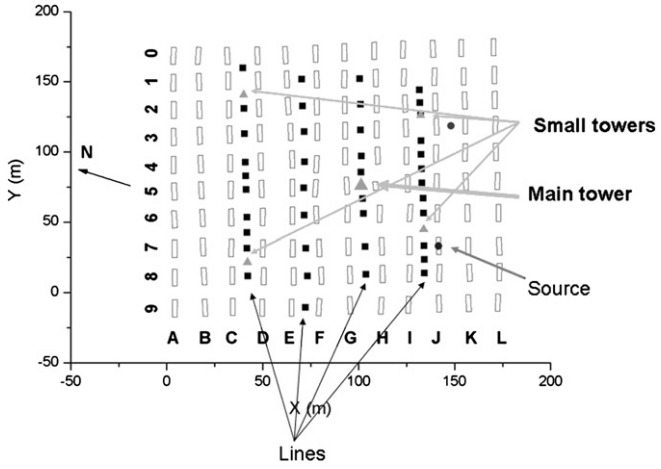

Fig. 1. MUST full scale experiment.

atmospheric stabilities for the full scale case and under a fully developed inlet flow for water channel (small scale) case. MUST set up and details on all trials conducted are described in Biltoft (2001) and Hilderman and Chong (2004). A brief description of the main parameters of the study is given below.

For the full scale trials (see Fig. 1), 120 shipping containers (each 12.2 m × 2.42 m × 2.54 m) were placed in a regular formation of 10 lines of 12 containers forming an approximately 200 m × 200 m square array. Meteorological data was sampled at a number of locations, including four 6 m towers (small towers) which were distributed within the array, one in each quadrant, each holding two 3-D sonic anemometers, one at 4 m and the other at 6 m. One trial with an average wind direction of 196° and an average wind speed of 4.54 m s^{-1} at 4 m height has been selected. Propylene was used as the tracer and 40 photoionisation detectors (digiPIDs) were

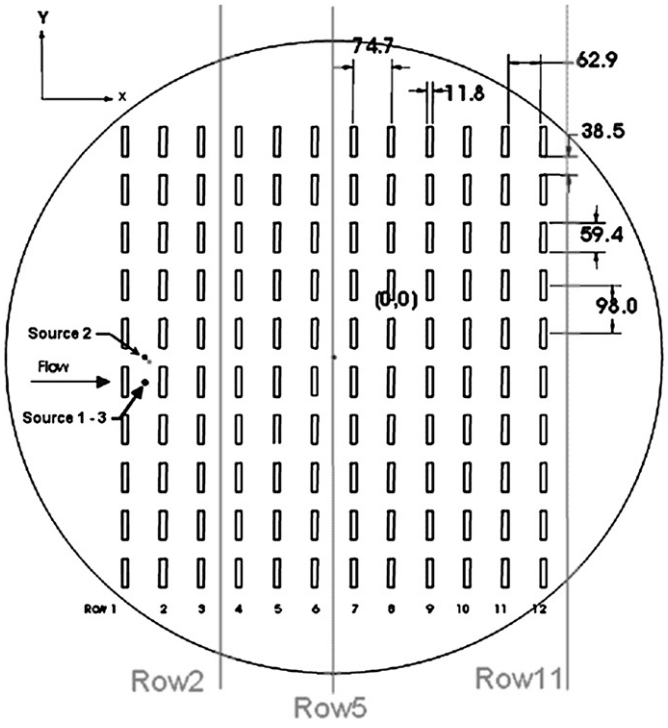

Fig. 2. Water channel experiment configuration taken from Hilderman and Chong (2004).

[1] In the project a number of reference NBC-terrorist scenarios have been established for the evaluation and comparison of the model chains used by participating members of different countries. The main goals within this project were to obtain a better understanding of the uncertainties in NBC-simulations through the identification of differences and weaknesses in the model chains and through the comparison of the results of calculations from different models, based on the same scenarios, by searching for the key reasons for the observed differences.

arranged in four sampling 'lines' or 'arcs' approximately 25, 60, 95 and 125 m from release locations (release height 2.7 m); detectors were placed at a height of 1.6 m above ground level. Detectors were placed on small towers too and on one higher tower (main tower in Fig. 1) for the measure of at least one experimental vertical profile. The sampling calibration range of these detectors was 0.04–1000 ppm (v) (Biltoft, 2001; Hilderman and Chong, 2004).

In water channel tests (Fig. 2) flow direction is perpendicular to the obstacle rows. Three different scenario configurations designated as TIC001, TIC002 and TIC003 were selected and adopted in the project (EUROPA-ERG1 TA 113.034, 2008). In TIC001 water channel scenario the source is behind the 5th obstacle from the bottom; in TIC002 scenario the source is in the canyon between the 5th and the 6th obstacle columns; in TIC003 the source location is the same as in TIC001, but without obstacles. Water channel experiment provides concentration measurements by horizontal linescan at 11 downstream positions and several heights, and by vertical linescan at 11 downstream positions on the centreline.

In this paper, only the results of the simulation of the field scale case, and of the TIC001 water channel case will be presented and discussed since these cases and in particular the second one, are the most critical ones, due to the position of the source, which is located immediately upwind from an obstacle. Indeed, in this configuration, effects of obstacle downwash became important especially in the near field.

3. Description of models

Several models were investigated for the purpose of ranking their capability of reproducing data sets. Some of them were used only for full scale dispersion modelling or water channel modelling, while other models were used both at full scale and at laboratory scale. Models can be classified into three main classes: RANS models, LES models and operational models.

RANS (Reynolds Averaged Navier–Stokes) models are a general class of Computational Fluid Dynamics (CFD) models widely used to reproduce flow conditions (Hinze, 1975); they are all based on the time-averaged form of the Navier–Stokes equation and when coupled with a mass transport model, they are capable of reproducing concentration distribution within a domain of interest. They mainly differ from each other for the turbulence model used (e.g. standard $k-\varepsilon$ or other methods) but also for transport parameters needed by the mass transport equation, such as turbulent Schmidt number, used in order to better fit experimental data. Moreover the boundary conditions used for reproducing actual incoming flow into the computational domain can result in differences between model outputs.

Large Eddy Simulation (LES) is another class of CFD models where rigorous Navier–Stokes equations are solved only for large scales of the flow domain and modelled for small scales, which require the highest computational resources (Deardorff, 1970).

While LES models solve numerically flow equations at least at large scale, RANS models do not solve them at any scale but provide a time averaged solution everywhere in the computational domain.

Finally operational models are among the most popular dispersion models for practical applications mainly due to their low computational cost (Hanna and Britter, 2002). There are a wide variety of them, but the ones used in this work are mainly general Gaussian models (Pasquill and Smith, 1983) and one Lagrangian stochastic particle model (Näslund et al., 1994; Schönfeldt, 1997 and Sehlstedt, 2000). They provide a physically consistent solution (mass is conserved) to turbulent dispersion equations but do not solve any flow equation, all the uncertainties inflow field prediction are left to some few adjustable parameters, or, in other words, the flow field is fixed and modelled with similarity profiles.

Table 1
Summary of models used.

Model Name	Model Class	Scenarios
RANS1	RANS	TIC001, TIC002, TIC003
RANS2	RANS	TIC001, TIC002, TIC003
RANS3	RANS	TIC001, TIC002, TIC003
RANS4	RANS	TIC001, TIC002, TIC003
RANS5	RANS	TIC002
LES1	LES	TIC001
LES2	LES	TIC001, TIC002
RL1	Operational/Gaussian	Field scale
RL2	Operational/Gaussian	Field scale
MOD1	Operational/Gaussian	Field scale
MOD2	Operational/Gaussian	Field scale
WD1	Operational/Stochastic	Field scale
WD2	Operational/Stochastic	Field scale
LE1	LES	Field scale
LE2	LES	Field scale
S1	RANS	Field scale

In the followings, for the sake of simplicity, models will be named as shown in Table 1, which also reports the class of the models and the scenarios where they were applied. Water channel scenarios have been modelled by means of five different RANS models (i.e., RANS1 to RANS5) and two LES models (i.e., LES1 and LES2).

The computational domain does not differ very much among scenarios, as can be seen from Table 2, however, even if different grids have been built for models to run, for each model the grid independency of results have been checked by each modeller (EUROPA-ERG1 TA 113.034, 2008). Main differences in the application of models and in their implementation are summarized in Table 2 for TIC001 scenario, small differences can be found for other water channel scenarios. Assumptions for the field scale case are in Tables 3a and 3b, respectively for operational models and CFD models. A description of each model is given below.

In RANS1 model turbulence has been modelled by means of standard $k-\varepsilon$ model (Launder and Spalding, 1972) in the implementation of Fluent 6.1 software. 1.8 M cells were non-uniformly distributed over the computational domain which ranges from 0.208 m (about 17 times the height of the obstacles $h = 12.4$ mm) upstream from the first obstacle row to 0.208 m downstream from the last row. In cross-stream direction the computational domain is simply the water channel width (1.5 m) and finally it is 0.3 m high (water depth in the channel). Boundary conditions were set to "velocity inlet" for the inflow surface, with a fully developed turbulent profile derived from simulations on an empty domain with the same roughness length, and to "pressure outlet" for the outflow condition. Standard wall functions were applied to buildings, ground and side walls (smooth side walls were considered)

Table 2
Models applied in TIC001 simulation.

Model name	Software implementation	Grid	Turbulence model	Turbulent Schmidt number for passive scalar transport
RANS1	FLUENT	1.8 M cells	Standard $k-\varepsilon$	0.7
RANS2	FLUENT	3.6 M cells	Realisable $k-\varepsilon$	0.7
RANS3	FLUENT	1.3 M cells	Standard $k-\varepsilon$	0.9
RANS4	PHOENICS 2006	0.75 M cells	Chen-Kim	1
LES1	User	1.5 M cells	Explicit SGS modelling - Smagorinsky type	0.6
LES2	User	5.4 M cells	OEEVM	n.s.

Table 3a
Operational models applied in TIC field scale simulation.

Model name	Model type	Atmospheric stability	Wind speed	Surface roughness
RL1	Gaussian	Pasquill class D	5 m s^{-1} @ 10 m	0.5 m
RL2	Gaussian	Pasquill class D	5 m s^{-1} @ 10 m	0.7 m
MOD1	Gaussian (SUKEVA)	Pasquill class E	4.54 m s^{-1} @ 4 m	0.3 m
MOD2	Integral model (ESCAPE 3.5)	Pasquill class E	4.54 m s^{-1} @ 4 m	0.254 m
WD1	Stochastic particle model (Langevin)	Neutral	5 m s^{-1}	n.s.
WD2	Same as WD1 with tilted wind direction			

and finally the symmetry conditions were applied to the top surface of the domain.

In RANS2, realisable $k-\varepsilon$ was employed from the Fluent 6.3 software implementation (Shih et al., 1995), with respectively two different mesh sizes, 1.9 M cells and 3.6 M cells (structured hexaedrical mesh excepted close to sources, where tetrahedral meshes were employed). Here results for the 3.6 M cells are reported. Domain size was chosen as one building length upstream, two downstream and one half on each side. "Wall" conditions were taken on top, bottom and lateral domain faces, "Velocity Inlet" on entrance face, and "Pressure Outlet" on exit face. For the input profile, the flat terrain profiles from 'must03' from Coanda (Hilderman and Chong, 2004) for mean and fluctuating velocity data have been used to determine k and ε inlet profiles.

In RANS3 for TIC001 scenario, the turbulent closure used was standard $k-\varepsilon$ (Launder and Spalding, 1972). In addition a transport equation for passive scalar is considered with a turbulent Schmidt number of 0.9 (Santiago et al., 2007). Standard wall functions were used for building and ground. At top and lateral boundaries of domain, symmetry boundary conditions were applied. The inflow conditions were taken from a fully turbulent flow in a periodic channel. The upper limit of domain is located at 11 h and the inflow and outflow boundaries at 10 h and 15 h respectively. The number of grid cells is approximately 1.3 M with a resolution of $15 \times 5 \times 10$ grid points close to the containers. A test about the grid-independence of results was performed.

In RANS4 for TIC001 scenario, the turbulent closure used was $k-\varepsilon$ with the Chen-Kim modification (Chen and Kim, 1987). The dispersed tracer is calculated with a scalar-transport equation with the Schmidt-number taken to be unity. Wall functions were used for obstacles and ground with roughness as indicated in the experiment. Laterally and at the upper boundary slip conditions is assumed. Inlet velocity profile was as measured in the experiment see Hilderman and Chong, 2004). Inlet value for k is approximated from measurements, the value for ε is calculated from boundary layer profile relations and the turbulent kinetic energy assumed. The upper boundary is at 24 h. A stretched grid is used above obstacle height with geometrical expansion of 1.12 ratio. Inlet is at 95 h upstream of the first obstacle row, using stretched grid from the inlet and vice versa for the outlet. Around the obstacles the resolution is $14 \times 3 \times 5$. The number of grid cells is approximately 0.75 M. The grid is staggered for the velocity compared with pressure and scalars.

Table 3b
CFD models applied in TIC field scale simulation.

Model name	Software implementation	Grid quality	Turbulence model	Turbulent Schmidt number for passive scalar transport
S1	FLUENT	1.6 M cells	Standard $k-\varepsilon$	0.9
LE1	User	1.5 M cells	Explicit SGS modelling - Smagorinsky type	0.6
LE2	Same as LE1 with tilted inlet flow direction			

In LES1 for TIC001 scenario, a Smagorinsky sub-grid scale model was used (Smagorinsky, 1963; Xie and Castro, 2009) and the concentration is modelled by the filtered passive scalar equation with a turbulent sub-grid scale Schmidt number of 0.6. Wall, top and lateral boundaries were similar to those used in RANS3. A mean velocity profile to which is added a random noise is used as inflow conditions. The number of grid cells is approximately 1.5 M with a resolution of $16 \times 5 \times 12$ grid points close to the containers. A test about the grid-independence of results was performed.

In LES2 for TIC001 scenario, an OEEVM[2] (Kim and Menon, 1999) is used describing the sub-grid turbulence. All variables are cell centred with finite-volume discretisation and PISO pressure-velocity coupling with ICCG (1E-5 for pressure) and BICCG (1E-7 for other variables). A Hexahedral grid is used ($1.31 \times 0.39 \times 0.3$ m^3) with 5.4 M cells. The timestep is 0.00025 s. The inlet profile is taken as $3 \times (z/0.15) \times 0.16$. Free-slip walls and symmetry conditions on sides. Fixed pressure is prescribed on the outlet boundary. k is fixed to 0.1 at the inlet.

In RANS for field scale TIC (S1 model), the model and boundary conditions are similar to those used in RANS3 for TIC001. Standard wall functions were used for building and ground. At top of domain, symmetry boundary conditions were applied. In this case where the inflow is not orthogonal to the obstacles, two lateral boundaries are considered as inflow and two as outflow. The mean inflow velocity profiles were taken from a power law profile that matches the measurements given at a meteorological mast upstream of the array. Inlet turbulent kinetic energy is interpolated from experimental data and dissipation profile deduced from equilibrium hypothesis. The upper limit of domain is located at 8 h and the inflow and outflow boundaries are located at more than 20 h from containers. The number of grid cells is approximately 1.6 M with a resolution of $11 \times 4 \times 10$ grid points close to the containers. A test about the grid-independence of results was performed.

In LES for field scale TIC (LE1 and LE2), the models are the same as LES1 for TIC001. The domain, grid and boundary conditions are similar to those used for RANS in field scale TIC, except the inlet turbulent fluctuations which are represented by random noise. The difference between LE1 and LE2 is the inlet wind direction (10°).

In the simulation of TIC001 main differences about turbulence model adopted are observed among RANS and LES models, with RANS models being more affected by this choice than LES models. Differences in turbulent Schmidt number can also be considered significant since concentration predictions are very sensitive to this parameter as shown by Blocken et al. (2008).

4. Statistical performance measures

Models have been evaluated and compared following the guidelines for measuring atmospheric dispersion model performance suggested by Hanna et al. (1993) and summarized by Chang and Hanna (2004). The evaluation should take into account a wide range of observed and predicted concentration including null values (both observed and predicted). Performance measures are required to be applicable at low and high concentration levels, thus

[2] One Equation Eddy Viscosity Model.

Table 4
Model performance measured by statistical parameters.

	Factor 2 Underestimation	Perfect Model	Factor 2 Overestimation	Acceptable model performance (Chang and Hanna, 2004)
MRB	−2/3	0	+2/3	|MRB|<0.3
MRSE	+4/9	0	+4/9	<0.1
FA2	–	100%	–	>50%
MG	0.5	1	2	0.7/1.3
VG	1.6	1	1.6	<1.6

they must weight all pairs of observations and predictions equally, independently on the absolute concentration value. They must be consistent and capable of distinguishing between model performances and of indicating if a model over- or under-predicts the measured values. They must also indicate the level of scatter or random deviation from this average under- or over-prediction. Another requirement, if spatially-paired data have to be used, is that they must be capable of accepting zero predicted or measured concentrations.

The following equations define the statistical performance measures, which include the fraction of predictions within a factor of two of the observations (*FA2*), the mean relative bias (*MRB*), the geometric mean bias (*MG*), the mean relative square error (*MRSE*) and the geometric variance (*VG*) (Hanna and Chang, 2001). In these definitions, ψ_p and ψ_o indicate the predicted and observed concentrations, respectively, and the angle brackets <…> denote an average over *N* observed/predicted pairs.

$$FAn = \frac{N_{1/n<\psi_p/\psi_o<n}}{N}, \quad n = 2 \tag{1}$$

$$MRB = \left\langle \frac{\psi_p - \psi_o}{1/2 \cdot \left(\psi_p + \psi_o\right)} \right\rangle \tag{2}$$

$$MRSE = \left\langle \frac{\left(\psi_p - \psi_o\right)^2}{1/4 \cdot \left(\psi_p + \psi_o\right)^2} \right\rangle \tag{3}$$

$$MG = \exp\left\langle \ln\left(\frac{\psi_p}{\psi_o}\right) \right\rangle \tag{4}$$

$$VG = \exp\left\langle \left[\ln\left(\frac{\psi_p}{\psi_o}\right)\right]^2 \right\rangle \tag{5}$$

A perfect model would have *MG*, *VG* and *FA2* = 1, *MRB* and *MRSE* = 0. It has to be noted that because of the influence of random turbulent processes, an over/under estimation by a factor of 2 has been assumed as a sufficient model performance (Table 4), however, more restrictive acceptance criteria (Chang and Hanna, 2004) have been also reported in the table.

MRB and *MG* measures are symmetric in terms of under and over predictions, as over-predictions and under-predictions compensate each other. Thus it is possible to still have a *MRB* = 0 or a *MG* = 1 for a simulation whose predictions are completely out of observations. A possible solution to this problem is to consider separately the two error components of *MRB* and *MG*, or the over-predicting and under-predicting components. Thus, MRB_n and MG_n (false negatives) only consider under-predictions, while MRB_p and MG_p (false positives) only consider over-predictions.

Table 5
FA2 for TIC001 scenario.

	RANS1	RANS2	RANS3	RANS4	LES1	LES2
FA2	61%	57%	49%	25%	52%	47%
FA2 – horizontal profiles	58%	51%	45%	24%	49%	46%
FA2 – vertical profiles	81%	94%	71%	33%	63%	54%

$$\begin{aligned} MRB_n &= \left\langle \frac{|\psi_p - \psi_o|}{1/2 \cdot \left(\psi_p + \psi_o\right)} \right\rangle, \quad \psi_p < \psi_o \\ MRB_p &= \left\langle \frac{\psi_p - \psi_o}{1/2 \cdot \left(\psi_p + \psi_o\right)} \right\rangle, \quad \psi_p > \psi_o \end{aligned} \tag{6}$$

$$\begin{aligned} MG_n &= \exp\left\langle \ln\left(\frac{\psi_o}{\psi_p}\right) \right\rangle, \psi_p < \psi_o \\ MG_p &= \exp\left\langle \ln\left(\frac{\psi_p}{\psi_o}\right) \right\rangle, \psi_p > \psi_o \end{aligned} \tag{7}$$

Clearly overall measures are related to their positive and negative component by the following expressions: $MRB = MRB_p - MRB_n$, $MG = MG_p/MG_n$.

If the relative bias is nearly constant over observed/predicted pairs, the relation $MRSE_{\min} = MRB^2$ defines the minimum possible value of *MRSE* for a given *MRB*. A model with a *MRSE* close to its minimum value provides a nearly constant deviation from the measured data in the whole range of concentrations.

If we consider separately the two error components of *MRB* (the over-predicting MRB_p and under-predicting MRB_n components) and *MRSE*, the measure of a model performance can be represented in a three-dimensional space (MRB_p, MRB_n, $\sqrt{MRSE}$), and the distance from this point to the point representing the perfect model (0,0,0) can be considered an overall measure of the model performance:

$$D_{MRB-MRSE} = \sqrt{MRB_p^2 + MRB_n^2 + MRSE} \tag{8}$$

Similar considerations can be made for the measures based on a log scale (*MG* and *VG*), leading to the definition of another overall statistical parameter (equation (9)) for model performance evaluation. However this parameter, along with its components, is greatly influenced by low and extremely low values, and so it is less important since peak concentrations (high values) are of primary

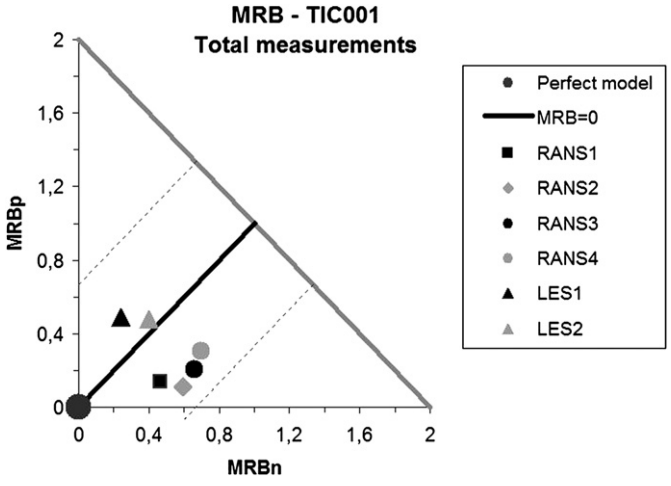

Fig. 3. MRB representation for TIC001 total measurements comparison.

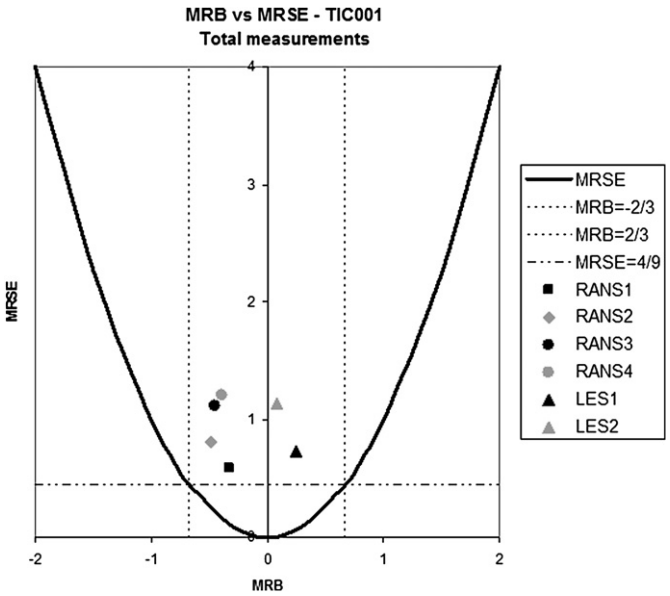

Fig. 4. MRB vs MRSE representation for TIC001 total measurements comparison.

importance when evaluating the effects of toxic substances on exposed humans.

$$D_{MG-VG} = \sqrt{(\ln MG_p)^2+(\ln MG_n)^2+\ln VG} \quad (9)$$

Parameters defined by equations (8) and (9) have been introduced in this work in order to obtain a more concise measure of model performance, which could take into account other parameters and the variability of predictions with respect to observations. Their application will be shown in the following section.

5. Comparison of models

5.1. Data filtering

In the small scale laboratory tests, linescans provided thousands of observed concentrations. For each test, observations (and thus also predictions) have been normalized with respect to the source concentration used for that test (source concentration was variable among tests), thus values can also be interpreted as the concentrations produced by a unitary concentration source.

The most common and simplest approach in order to avoid errors at very low concentrations is to invoke a threshold for all predicted and observed concentrations. Considerations about asymmetric data sets observed in cross-stream concentration profiles given by water channel linescans and about the high sensitivity of some statistical measures to low values, yielded to the selection of symmetric y cross-stream concentration values (for horizontal measurements) and to discard observed and predicted values falling below a threshold value of 1E-8 (minimum measured value in TIC001: 1.3E-8).

For the selection of data sets for the field scale scenario, all data were included for the parameter Factor of 2 evaluation, while only non-zero observations or predictions were selected for estimating *MRB* and *MRSE* parameters and finally a selection of both non-zero observations and predictions lead to very few data points for *MG* and *VG* average values.

In order to allow for valid statistics, the number of data points in each subset (horizontal and vertical) can be considered sufficient for the scenarios in water channel (the minimum number of filtered data measurements is 299 for a single linescan) but rather insufficient for the field scale scenario in the estimation of the *MRB* and *MRSE* parameters (25 horizontal and 22 vertical filtered data points). Moreover most of field scale measurements are available only at locations quite far from the plume centreline, thus, rather than statistical parameters, other comparison methods have been used as shown in the next paragraph.

5.2. Results and discussion

The analysis performed for TIC001 scenario was mainly focused on differences between results of RANS and LES models and evidenced for all models a significant difference between predictions of vertical and horizontal profiles, with the latter having worst performances than predictions for vertical profiles as confirmed by the *FA2* calculated separately for horizontal and vertical data sets (Table 5). However the number of observations from horizontal linescans is much higher than observations from vertical linescans and moreover horizontal profiles, especially at ground level, should be considered more important for a further evaluation of effects on exposed individuals.

MRB values, shown in Fig. 3 confirmed that overall good performances are achieved almost by all models, as a matter of fact, according to *MRB* parameter, 6 of 6 models predict within factor 2 range. In particular all RANS models under-predict and LES models over-predict but LES are more balanced (false positives compensate false negatives) and satisfy more restrictive criteria (last column of Table 4). Nevertheless the analysis of *MRSE* vs. *MRB* (Fig. 4) showed that the relative bias is far from being nearly constant and so, if low values became important, no model could be considered acceptable in the whole range of concentrations.

Statistical measures can also be used for model comparison and ranking, which is reported in Table 6, on the basis of both overall indexes of model performance previously proposed (equations (8) and (9). Model ranking is very similar both on the basis of $D_{MRB-MRSE}$ and of D_{MG-VG}, showing that single measures can lead to different classification, but they can compensate each other if considered together.

Statistical parameters calculated for single horizontal linescans at different distances from the source and different scaled heights from the ground (see also Fig. 2 for the locations of individual linescans within the domain of interest) are reported in Table 7, where bold values represent best performance for each linescan on the basis of different parameters. RANS1, LES1 and RANS2 models showed the overall best performances, but with some differences from near field (row 2) to far field calculations (row 11). As could be expected, in general LES models provide a better prediction of near field concentration distribution since their turbulence models are more accurate in taking into account effects of bluff obstacles on the flow field (large separation flows) and consequently in the mass transport of the contaminant. On the other hand in the far field, where flow field is more stable, RANS models have a better behaviour than LES models.

Table 6
Models ranked by distances from the perfect model for TIC001 scenario.

$D_{MRB-MRSE}$ – total measurements		D_{MG-VG} – total measurements	
RANS1	0.91	RANS1	1.10
LES1	1.01	LES1	1.38
RANS2	1.09	RANS2	1.44
LES2	1.24	RANS4	1.65
RANS3	1.26	LES2	1.94
RANS4	1.33	RANS3	1.96

Table 7
Statistical parameters at different distances from the source for TIC001 scenario.

		ROW2					ROW5					ROW11				
		RANS1	RANS2	RANS3	RANS4	LES1	LES2	RANS1	RANS2	RANS3	RANS4	LES1	LES2	RANS1	RANS2	RANS3	RANS4	LES2
$H = 9.3$ mm	MRB	−0.91	−0.87	−1.07	−0.56	0.45	**−0.05**	−0.40	−0.49	−0.55	**−0.38**	0.67	0.75	−0.74	−0.92	−0.84	−0.75	**0.50**
	MRSE	1.26	1.19	1.99	1.32	**0.53**	1.00	**0.47**	0.72	1.08	0.98	0.68	0.76	**0.73**	1.25	1.29	1.49	0.88
	FA2	41%	48%	38%	21%	**62%**	37%	**70%**	67%	51%	29%	37%	56%	51%	40%	46%	20%	**59%**
	MG	0.30	0.31	0.13	0.52	1.71	**1.09**	0.63	0.53	0.43	**0.67**	2.12	2.38	**0.43**	0.29	0.29	0.40	1.96
	VG	11.13	11.83	>100	7.77	**2.30**	7.13	**1.90**	3.61	13.17	4.07	2.52	3.17	2.74	13.19	25.23	12.44	**5.04**
$H = 18.6$ mm	MRB	−0.53	−0.62	−0.79	**−0.27**	0.56	0.32	−0.57	−0.72	−0.69	−0.91	**0.38**	0.51	−0.51	−0.78	−0.70	**−0.46**	0.69
	MRSE	0.74	0.85	1.65	1.23	**0.59**	1.05	0.60	0.96	1.17	1.19	0.45	**0.43**	**0.44**	1.03	1.15	1.41	1.07
	FA2	0.56	0.51	0.35	0.22	**0.60**	0.37	0.60	0.49	0.48	0.22	0.63	**0.68**	**0.66**	0.42	0.48	0.22	0.53
	MG	0.52	0.45	0.23	**0.76**	1.87	1.75	**0.52**	0.39	0.35	0.34	1.51	1.75	**0.57**	0.37	0.36	0.61	2.66
	VG	3.18	4.42	>100	6.32	**2.19**	9.94	2.44	6.80	18.95	5.24	**1.70**	1.71	**1.73**	6.35	12.60	11.10	10.64
$H = 74.4$ mm	MRB	1.64	1.57	1.45	1.70	−1.39	**1.29**	0.56	0.32	**0.26**	0.89	−1.20	−1.81	−0.40	−0.60	−0.70	**−0.17**	−0.91
	MRSE	2.70	2.46	2.09	2.91	2.18	**1.68**	0.45	**0.24**	0.44	0.90	1.88	3.40	**0.30**	0.55	0.87	0.96	1.60
	FA2	0.00	0.00	0.00	0.00	**0.03**	0.00	0.41	**0.91**	0.59	0.34	0.25	0.00	**0.73**	0.59	0.48	0.30	0.35
	MG	10.26	8.31	6.23	12.68	**0.10**	4.71	1.79	1.38	**1.27**	2.72	0.14	0.02	0.65	0.52	0.41	**0.90**	0.26
	VG	>100	89.27	28.67	>100	>100	**11.80**	1.64	**1.31**	1.78	3.40	>100	>100	**1.41**	1.93	4.82	4.88	35.36

Table 8
FA2 parameter – field scale scenario.

	S1	LE1	LE2	RL1	RL2	MOD1	MOD2	WD1	WD2
FA2	50%	49%	40%	46%	49%	15%	36%	42%	45%
FA2-lines	45%	57%	45%	40%	42%	15%	40%	42%	45%
FA2-towers	56%	37%	34%	53%	56%	16%	31%	n.a.	n.a.

For the field scale case, due to lack of observed concentrations only *FA2* parameter (Table 8) could be calculated with a significant number of data points. Results show that performances are pretty poor due to the strong wind meandering, always present in real situations, and because several data points are too far from the plume centreline. However a vertical profile quite near to the centreline was available (main tower in Fig. 1). Fig. 5 represents the comparison of the main tower profiles predicted by different models with the concentration measurements observed in the experimental trial. Significant differences between the models are due to different estimations of vertical dispersion parameters between operational models and of the vertical wind profile between CFD models. The only RANS model applied to this case seems to reproduce experimental data better than all the other models including LES. This is also confirmed by Fig. 6, where calculated concentration contours at 1 ppm and 1.6 m height are compared with an estimation of the 1 ppm contour obtained from a polyharmonic spline interpolation of observed concentrations. Fig. 6 shows that the real plume centreline (black solid line) does not follow the average wind direction, because obstacle array is not aligned with the wind. This causes operational models to have difficulties in predicting concentrations since they do not calculate the flow field and cannot predict the tilt angle of the centreline caused by obstacle array (about 15° for this case), but an average wind direction must be provided. On the other hand CFD models are able to calculate the flow field, but LES models are very sensitive to inflow average wind direction and slightly exceed in the prediction of plume centreline deviation due to a small over-estimation of the wind speed caused by canyons between obstacles.

It is worth noting that simplified models could still have pretty good performances if the wind meandering effect is correctly estimated (e.g by means of an appropriate averaging time) and if an appropriate tilt angle for the wind direction was applied in order to take into account the deviation of the plume centreline due to the geometrical features of the obstacle array (length, width and height of a single obstacle, crosswind spacing between obstacles, etc.).

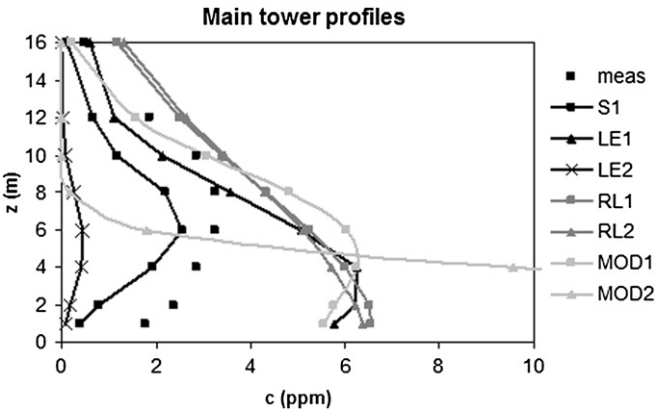

Fig. 5. Main tower profiles (no data in main tower for WD1 and WD2 models): Comparison of models results with concentration measurements ("meas" in the legend).

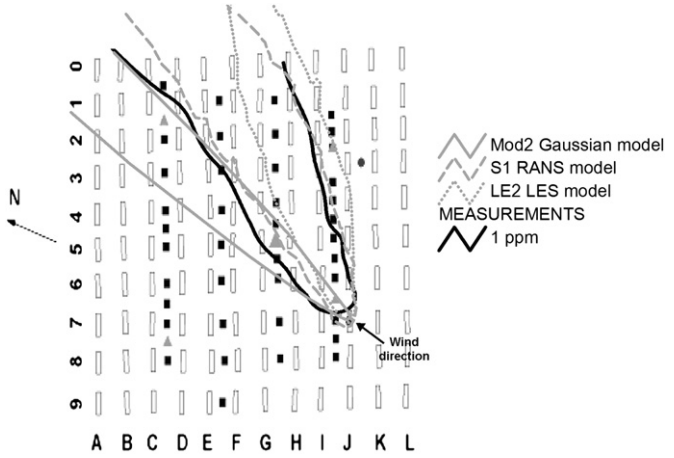

Fig. 6. Calculated concentration contours at 1.6 m compared with measured concentration contour at 1 ppm for field scale scenario.

6. Conclusions

The comparison performed with the aid of several statistical measures showed that the performances of models could be ranked differently on the basis of single parameters and so the choice of the correct parameter could be made keeping in mind the purpose of the application of the model, in this case the concentration prediction for the assessment of damage on exposed individuals. Thus two new parameters, given by the combination of the other ones, has been proposed in order to take into account more parameters at the same time. Ranking of models on the basis of these two parameters provided different results as well, but they could still be used for screening purposes.

Nevertheless it seemed quite clear that uncertainties and variability in atmospheric phenomena have not allowed to achieve acceptability criteria for all the parameters at the same time, and this was especially true in the field scale case (and it would become even more important in real situations with irregular obstacle location and with random wind direction). In the field scale situation statistical measures (especially if based on few available observations) could mislead the interpretation of models' results and other comparison criteria have been applied. When uncertainties become more important and predictions more difficult, even if computational expensive models could provide better predictions under controlled conditions, the results of this study for the field scale case showed that maybe small corrections to fast-responding operational models could bring a great improvement in the efficiency of concentration prediction especially if also operational parameters, such as exploitation time (and costs), are taken into account in the evaluation of model performances.

Acknowledgments

The authors would like to convey thanks to the Ministries of Defence of Finland, France, Italy, Norway, Spain, Sweden and The Netherlands for providing the financial means in the frame of the EDA NBC Modelling and Simulation project (EUROPA-ERG1 TA 113.034). The Members of the Italian team who participated in the project wish to express their deep gratitude to Lt. Col. Antonio Fusco, from the Italian Ministry of Defence, for the support, assistance and guidance he offered all along the project. He passed away prematurely recently, and the Italian team dedicates its contribution to this paper to him.

References

Batchvarova, E., Gryning, S.E., 2006. Progress in urban dispersion studies. Theoretical and Applied Climatology 84, 57–67.
Biltoft C.A., 2001. Customer Report for Mock Urban Setting Test, DTC Project No. 8-CO-160-000-052, DPG Document No. WDTC-FR-01-121, West Desert Test Center, U.S. Army Dugway Proving Ground, Dugway, Utah.
Blocken, B., Stathopoulos, T., Saathoff, P., Wang, X., 2008. Numerical evaluation of pollutant dispersion in the built environment: comparisons between models and experiments. Journal of Wind Engineering and Industrial Aerodynamics 96, 1817–1831.
Britter, R.E., Hanna, S.R., 2003. Flow and dispersion in urban areas. Annual Review of Fluid Mechanics 35, 469–496.
Chang, J.C., Hanna, S.R., 2004. Air quality model performance evaluation. Meteorology and Atmospheric Physics 87, 167–196.
Chen, Y.S., Kim, S.W., 1987. Computation of Turbulent Flows Using an Extended $k-\varepsilon$ Turbulence Closure Model. NASA. CR-179204.
Deardorff, J., 1970. A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. Journal of Fluid Mechanics 41 (2), 453–480.
EUROPA-ERG1 TA 113.034, Unpublished results, EDA-B0155-GEM3-ERG EUROPA-ERG1 TA 113.034 "NBC Modelling and Simulation" – Final Report, 2006–2008.
Fernando, H.J.S., Lee, S.M., Anderson, J., Princevac, M., Pardyjak, E., Grossman-Clarke, S., 2001. Urban fluid mechanics: air circulation and contaminant dispersion in cities. Environmental Fluid Mechanics 1, 107–164.
Gailis, R., Hill, A., 2006. A wind-tunnel simulation of plume dispersion within a large array of obstacles. Boundary-Layer Meteorology 119, 289–338.
Grimmond, C.S.B., 2006. Progress in measuring and observing the urban atmosphere. Theoretical and Applied Climatology 84, 3–22.
Hanna, S.R., Chang, J.C., Strimaitis, D.G., 1993. Hazardous gas model evaluation with field observations. Atmospheric Environment 27, 2265–2285.
Hanna, S.R., Chang, J.C., 2001. Use of the Kit Fox field data to analyze dense gas dispersion modelling issues. Atmospheric Environment 35, 2231–2242.
Hanna, S.R., Britter, R.E., 2002. Wind Flow and Vapor Cloud Dispersion at Industrial and Urban Sites, first ed. Wiley-American Institute of Chemical Engineers.
Hilderman, T., Chong, R., 2004. A Laboratory Study of Momentum and Passive Scalar Transport and Diffusion within and above a Model Urban Canopy – Final Report. Coanda R&D Corp., Canada.
Hinze, J.O., 1975. Turbulence. McGraw-Hill Publishing Co., New York.
Kim, W.-W., Menon, S., 1999. A new Incompressible Solver for large-eddy simulations. International Journal for Numerical Fluid Mechanics 31, 983–1017.
Launder, B.E., Spalding, D.B., 1972. Lectures in Mathematical Models of Turbulence. Academic Press, London, England.
Milliez, M., Carissimo, B., 2007. Numerical simulations of pollutant dispersion in an idealized urban area, for different meteorological conditions. Boundary-Layer Meteorology 122, 321–342.
Näslund, E., Rodean, H.C., Nasstrom, J.S., 1994. A comparison between two stochastic diffusion models in a complex three-dimensional flow. Boundary-Layer Meteorology 67, 369–384.
Pasquill, F., Smith, F.B., 1983. Atmospheric Diffusion. E. Horwood.
Santiago, J.L., Martilli, A., Martin, F., 2007. CFD simulation of airflow over a regular array of cubes. Part I: three-dimensional simulation of the flow and validation with wind-tunnel measurements. Boundary-Layer Meteorology 122, 609–634.
Schönfeldt, F.A., 1997. Langevin Equation Dispersion Model for the Stably Stratified Boundary Layer. Technical Report. FOI Swedish Defence Research Agency, Division of NBC Defence.
Sehlstedt, S.A., 2000. Langevin Equation Dispersion Model for the Unstably Stratified Boundary Layer. Technical Report. FOI Swedish Defence Research Agency, Division of NBC Defence.
Shih, T.-H., Liou, W.W., Shabbir, A., Yang, Z., Zhu, J., 1995. A New eddy-Viscosity model for high Reynolds number turbulent flows – Model Development and validation. Computers Fluids 24 (3), 227–238.
Smagorinsky, J., 1963. General circulation experiments with the Primitive equations. I. The Basic experiment. Monthly Weather Review 91, 99–164.
Xie, Z., Castro, I.P., 2009. Large-eddy simulation for flow and dispersion in urban streets. Atmospheric Environment 43, 2174–2185.
Yee, E., Gailis, R., Hill, A., Hilderman, T., Kiel, D., 2006. Comparison of wind-tunnel and water-channel simulations of plume dispersion through a large array of obstacles with a scaled field experiment. Boundary-Layer Meteorology 121, 389–432.