An urban large-eddy simulation-based dispersion model for marginal grid resolutions: CAIRDIO v1.0

Michael Weger[1], Oswald Knoth[1], and Bernd Heinold[1]

[1]Leibniz Institute for Tropospheric Research, Leipzig, Germany

Correspondence: Michael Weger (weger@tropos.de)

Abstract. The ability to achieve high spatial resolutions is an important feature of numerical models to accurately represent the large spatial variability of urban air pollution. On the one hand, the well-established mesoscale chemistry transport models have their obvious short comings due to the extensive use of physical parameterizations. On the other hand, obstacle resolving computational fluid dynamic models (CFD), although accurate, are still often too computationally intensive to be applied regularly for entire cities. The major reason for the inflated computational costs is the required horizontal resolution to meaningfully apply obstacle discretization, which is mostly based on boundary-fitted grids, like e.g. the marker-and-cell method. In this paper, we present the new City-scale AIR dispersion model with DIffuse Obstacles (CAIRDIO v1.0), in which the diffuse interface method, simplified for non-moving obstacles, is incorporated into the governing equations for incompressible large-eddy simulations. While the diffuse interface method is widely used in two-phase modeling, this method has not been used in urban boundary layer modeling so far. It allows to consistently represent buildings over a wide range of spatial resolutions, including grid spacings equal or larger than typical building sizes. This way, the gray-zone between obstacle-resolving microscale simulations and mesoscale simulations can be addressed. Orographic effects can be included by using terrain-following coordinates. The dynamic core is compared against a standard quality-assured wind-tunnel dataset for dispersion-model evaluation. It is shown that the model successfully reproduces dispersion patterns within a complex city morphology across a wide range of spatial resolutions tested. As a result of the diffuse obstacle approach, the accuracy test is also passed at a horizontal grid spacing of $40\,\mathrm{m}$. Although individual flow features within individual street canyons are not resolved at the coarse grid spacing, the building effect on the dispersion of the air pollution plume is preserved at a larger scale. Therefore, a very promising application of the CAIRDIO model lies in the realization of computationally feasible yet accurate air quality simulations for entire cities.

1 Introduction

The state-of-the-art in urban air quality modeling now almost routinely encompasses the scales at which processes governing the atmospheric dispersion within the urban planetary boundary layer (PBL) are explicitly represented (Benavides et al., 2019; Croitoru and Nastase, 2018; Kadaverugu et al., 2019). The reasons for this increase in physical detail are manifold. On the one hand, even though PBL mixing processes are often parameterized to a large extend, the parameterizations themselves must rely on a sound physical basis, for which detailed large-eddy simulations (LES) can be consulted (Noh and Raasch,

2003; Kanda et al., 2013). Direct benefits of more detailed numerical simulations include an increased ability to produce more representative air-quality forecasts for individual locations (Carlino et al., 2016) and the provision of high-resolution 4-dimensional data for research purposes, like, e.g., source attribution (Fernández et al., 2019) and exposure risk assessment (Chang, 2016). Exposure-relevant air pollution concentrations at pedestrian level are subjected to a considerable and also

30 complex spatio-temporal variability, as they are not only influenced by the relative location of pollution sources, but also very importantly by the urban morphology and associated meteorological conditions (Birmili et al., 2013; Paas et al., 2016; Harrison, 2018). For an accurate simulation, it is thus not only key to explicitly represent the urban canopy features, but also to consider the prevailing mesoscale meteorological conditions. Depending on the level of physical detail, a trade-off in the use of high-resolution numerical simulations is often their exclusive applicability in a limited area, which can be very restrictive.

35 Hence, an active topic of research is dedicated to improve the numerical efficiency of high-resolution modeling tools and their incorporation into a larger modeling framework (Baik et al., 2009; Jensen et al., 2017; Kurppa et al., 2020).

Commonly used multiscale approaches consist of nested domains and involve different types of models designed for a specific scale range. To address the global and regional scales, the use of chemistry transport models (CTM) in combination with numerical weather prediction (NWP) models is a well established practise. Examples of those coupled model systems include,

40 C-IFS (Flemming et al., 2015), WRF-Chem (Grell et al., 2005), CMAQ (Appel et al., 2017), ICON-ART (Rieger et al., 2015) and COSMO-MUSCAT (Wolke et al., 2004, 2012). In all of these models, subgrid-scale effects of temperature and moisture, as well as PBL dynamics are parameterized. The influence of the urban canopy on the PBL in NWP models can be considered through sophisticated canopy parameterizations (Martilli et al., 2002; Schubert et al., 2012). The improvements of these so-called urbanized NWPs are substantially reflected in the modeled pollutant concentrations (Kim et al., 2015; Wang et al.,

45 2019). Nevertheless, as the model domain does not allow for an explicit representation of buildings, the pollutants emitted at street level are diluted over the entire grid cell, which can considerably deviate from real conditions, where a large part of the physical volume may be impervious. As a result, pollutant concentrations modeled using NWP-based approaches are more representative for the urban background (Korhonen et al., 2019). Another practical limit to resolution and accuracy of NWP models arises from the use of parameterizations themselves. Using WRF, Haupt et al. (2019) observed that for horizontal grid

50 spacings below the typical PBL height, numerical results can become spurious. Based on their findings, they recommend not to apply NWPs on the sub-kilometer scale without careful replacement of the used parameterizations. In PBL meteorology, the microscale seamlessly follows the lower limit of the mesoscale (super-kilometer range) (Stull, 1988; Rakai and Gergely, 2013). However, adopting the modeling perspective, there is a clear segregation of micro and mesoscale. While the latter is constrained at the lower end of the resolution by the extensive use of parameterizations, attempting to model PBL processes

55 directly with microscale approaches requires sufficiently fine grid spacings. The landscape of microscale models is diverse (Fallah-Shorshani et al., 2017; Brown, 2014; Hanna et al., 2006) and it reflects the difficulty of finding a compromise between computational resources and accuracy. Computational fluid dynamics models (CFD) can be seen as the microscale pedants to NWPs. Among them, LES approaches are the most accurate but expensive ones. LES models resolve the turbulent spectra up to the filter cut-off size (often equivalent to the grid size) and rely on simplistic sub-scale parameterizations only (e.g.

60 Maronga et al., 2019). These two different approaches (extensive parameterization vs. explicit representation) are difficult

to merge at the bridging scale-range (few tens of meters to one kilometer), for which reason Haupt et al. (2019) coined the key word of "terra-incognita" to refer to the problem. An exemplary study, in which a LES model was applied for horizontal resolutions up to above $100\,\mathrm{m}$, is given by Efstathiou et al. (2016). However, their simulations did not include an urban canopy, whose discretization would obviously require a much finer grid. In fact, to fully resolve the energy-dominant eddies within street-canyons, a spatial resolution of 15-20 grid points over a typical obstacle dimension size is needed (Xie and Castro, 2006). This translates in a grid spacing of about $1\,\mathrm{m}$ for typically sized buildings. The use of such fine grids for entire city domains is still prohibitively expensive in LES modeling. E.g., Wolf et al. 2020 used a grid spacing of $10\,\mathrm{m}$ in their cutting-edge research study to simulate air pollutant dispersion in Bergen, Norway with the LES-model PALM. While this spatial resolution, according to the argument above, does not ensure a LES model to be eddy resolving in every part of the domain, the preference of a physically-based model over the more widely used parameterized models for this grid size (e.g. plume or street-canyon models) can nevertheless be a legitimate choice. Resolved physics can still be maintained outside the densely built urban canopy. Within the canopy, the dispersion pattern on a larger scale is mostly shaped by the channeling and blocking effects of buildings, which can be represented also with coarser grids.

In physical computing, it is well known that only a slight increase in the grid spacing results in a large computational saving (e.g. models using explicit time integration ideally perform 16 times faster on the same domain using a grid with doubled grid spacings in all 3 dimensions). These computational savings could in turn be spent in larger physical domains for more comprehensive yet accurate urban air-quality simulations. A key feature that sets the technical limits to the spatial resolution in urban boundary-layer modeling is the discretization method used for obstacles. A class of methods that allow for obstacle geometries not bound to the grid are the immersed boundary methods, summarized by Mittal and Iaccarino (2005). These are essentially Cartesian methods, but instead of relying on a commonly used marker-and-cell method (grid cells are assigned either to the building interiors or to the atmosphere), they represent rigid boundary conditions (e.g. Neumann boundary condition for pressure) on grid cell faces not coinciding with the obstacle boundary. Among these methods, the so-called direct forcing uses the ghost-cell interpolation technique, where image points from adjacent interior ghost points are mirrored across the rigid boundary and interpolated using surrounding fluid nodes. While this method greatly enhances the flexibility in the choice of the grid size, it still suffers from the empirical nature in the selection of the interpolation nodes and the interpolation method itself. On the other hand, an equivalent boundary forcing can be more rigorously deduced from a two-phase flow model (Drew, 1983) by neglecting the restoring source terms. Treating one of the phases as an undeformable solid, Kemm et al. (2020) derived a diffuse interface (DI) model for compressible fluids. One of the main advantages of their approach is the algorithmic simplicity, as the boundary-forcing term is analytically coupled to the DI, which is advected as a scalar.

In this work, we adopt the basic idea of DI and implement diffuse obstacle boundaries (DOB) in our new City-scale AIR dispersion model with DIffuse Obstacles (CAIRDIO v1.0), which will be used as a computationally feasible yet accurate downscaling tool for mesoscale air pollution fields over urban areas. DOB allow buildings to be represented as diffuse features, and thus the flexibility in the choice of grid resolution can be greatly increased. DOB is a simplified form of DI resulting from the static boundaries associated with buildings. DOB is incorporated in the discrete differential operators by considering the conservation laws in a finite volume framework. The equations are then solved with standard methods on a Cartesian grid. To

interpret our DOB approach from a physical point of view, it can be argued that the grid cells are interspersed with a porous and semi-permeable medium. Their detailed structure is only of concern insofar as it determines the mass and momentum budged at the grid level through two different types of interface fields. To the authors knowledge, this approach has not been used for air quality modeling so far, but very interestingly the concept of permeability finds application in geological science (Haga, 2011). In contrast to geometry-aligned discretizations, which preserve a high degree of accuracy near obstacle walls but require high resolutions, this approach suits more to the integral aspect of building shapes and configurations at marginal resolutions. However, by increasing the grid resolution, the approach seamlessly transitions to a traditional obstacle resolving Cartesian approach, as the interface eventually becomes sharply defined and imposes Neumann-boundary conditions on the pressure.

The paper is organized as follows: Section 2 provides a detailed description of the model CAIRDIO, including the spatial discretization method. Section 3 contains numerical tests to demonstrate the diffuse obstacle discretization, the dynamic core and the parallel scaling capabilities. In Section 4, we present a model-evaluation study by simulating a realistic wind-tunnel dispersion experiment. Finally, in Section 5, the benefits and limitations of the approach are summarized and concluded by an outlook for potential future applications.

2   Model description

2.1   Basic equations

The physical domain consists of an interspersed, stationary solid phase representing the buildings and a mobile fluid phase. The governing equations for the mobile phase are deduced from a simplified two-phase model by neglecting restoring forces. In a two-phase model, phase-fraction functions α_1, α_2 with $0 \leq \alpha_1 \leq 1$ and $\alpha_1 + \alpha_2 = 1$ are used to formulate the set of equations of both phases individually. The equation of motion of an incompressible phase is adopted from Drew (1983) (Eq. 41, 45 therein). By setting the interfacial force density and surface tension to zero, the simplified momentum equation of the mobile phase (indicated with α_1) is written as:

$$\partial_t \left(\alpha_1 \rho \boldsymbol{u} \right) = -\nabla \cdot \left(\alpha_1 \rho \boldsymbol{u} \otimes \boldsymbol{u} \right) - \alpha_1 \nabla \left(\alpha_1 p \right) + p_I \nabla \alpha_1 + \nabla \cdot \left(\alpha_1 \mathbf{T} \right) + \alpha_1 \rho \boldsymbol{b}. \tag{1}$$

Here, the 3D velocity vector is denoted by $\boldsymbol{u}$, ρ is the density of air and p_I the interface pressure, which reflects Newton's third law of motion near a fixed wall. As in Kemm et al. (2020), p_I is assumed to be in equilibrium with the surrounding fluid pressure p. The stress tensor $\mathbf{T}$ contains the contributions from subscale and surface fluxes in LES averaging. In this model, the implicit approach is used for spatial filtering (Schumann, 1975). Viscous stresses are neglected due to the high Reynolds numbers typically encountered in atmospheric flows. The sum of external body forces $\boldsymbol{b}$ contains the gravitational force and inertial forces resulting in a rotating frame of reference. The interface in our case is static, as buildings do not respond to the flow. This makes it possible to multiply the equation of motion with $\alpha_1^{-1} \rho^{-1}$, to obtain the tendency equation in single-flow

125 denotation. The reference density ρ_{ref} is kept constant in time for an incompressible fluid. The full set of model equations reads:

$$\partial_t \boldsymbol{u} = -\alpha_1^{-1} \nabla \left(\alpha_1 \boldsymbol{u} \otimes \boldsymbol{u}\right) - \frac{1}{\rho_{ref}} \alpha_1^{-1} \alpha_1 \nabla p - \alpha_1^{-1} \nabla \cdot \left(\alpha_1 \mathbf{T}\right) - \boldsymbol{f} \times \boldsymbol{u} + \boldsymbol{g} \frac{\overline{\Theta}_v - \Theta_v}{\overline{\Theta}_v}, \tag{2}$$

$$\alpha_1^{-1} \nabla \cdot \left(\alpha_1 \boldsymbol{u}\right) = 0, \tag{3}$$

$$\partial_t \Theta = -\alpha_1^{-1} \nabla \cdot \left(\alpha_1 \boldsymbol{u} \Theta\right) + \alpha_1^{-1} \nabla \cdot \left(\alpha_1 k_h \nabla \Theta\right) + \alpha_1^{-1} S_\Theta, \tag{4}$$

130 $$\partial_t Q_v = -\alpha_1^{-1} \nabla \cdot \left(\alpha_1 \boldsymbol{u} Q_v\right) + \alpha_1^{-1} \nabla \cdot \left(\alpha_1 k_h \nabla Q_v\right) + \alpha_1^{-1} S_{Q_v}, \tag{5}$$

$$\Theta_v = \Theta \left(1 + 0.61 Q_v\right). \tag{6}$$

In the momentum equation (Eq. 2), the body-force term is replaced by the Coriolis term, with the Coriolis parameter $\boldsymbol{f}$ for a mean geographic latitude, and the buoyancy term using the Boussinesq approximation. $\boldsymbol{g}$ is the gravity-acceleration vector in the local frame of reference and Θ_v the virtual potential temperature defined by Eq. 6. The overbar denotes the

135 horizontal mean state. Eq. 3 is the continuity equation derived in an analogous way as the momentum equation from the original formulation in Drew (1983). The transport equation for scalars, like the potential temperature Θ and specific humidity Q_v, contains source terms from parameterized surfaces fluxes. These have to be multiplied with α_1^{-1}. The scalar field k_h is the eddy-diffusion coefficient for heat. Note that the combinations $\alpha_1^{-1} \nabla \cdot \alpha_1$ and $\alpha_1^{-1} \alpha_1 \nabla$ can be identified as particular versions of the divergence and gradient operator, which incorporate diffuse boundaries. Using a staggered grid, the stencil of

140 α_1 (face-centered) differs from α_1^{-1} (volume-centered), so that the terms do not cancel each other out.

2.2 Computational grid

The computation uses a structured Arakawa-C grid, with the velocity components being defined at the cell faces and scalar fields at the cell centers. Vertical coordinate transformation allows for a curve-linear grid in the physical space to be adapted to a smoothly varying terrain function:

145 $$\tilde{x} = x, \quad \tilde{y} = y, \quad \tilde{z} = z - h(x, y). \tag{7}$$

Here, z is the mean-sea level height and $h(x,y)$ the terrain-height function. Elevated levels are simply given by adding a horizontally constant vertical increment to the first level.

The pressure gradient in terrain-following coordinates is modified to:

$$\nabla p = \tilde{\nabla} p - \left[\partial_x h(x,y) + \partial_y h(x,y)\right]\partial_z p. \tag{8}$$

The divergence operator is applied on the contravariant velocity components, which are parallel to the cell-face normal. In our simple case, only the vertical contravariant velocity component ω differs from the covariant non-transformed component. It is given by:

$$\omega = w - \partial_x h u - \partial_y h v. \tag{9}$$

Using Eq. 9, the advective tendency of a scalar q is written as:

$$\partial_t q = -u\partial_x q - v\partial_y q - \left[w - u\partial_x h - v\partial_y h\right]\partial_z q. \tag{10}$$

Similarly, the continuity equation in terrain-following coordinates follows from:

$$\partial_x u + \partial_y v + \partial_z(w - \partial_x h u - \partial_y h v) = 0. \tag{11}$$

Combining Eq. 11 and 8, the metric terms in the Laplace operator are obtained, which is later needed for the pressure equation:

$$\begin{aligned}
\triangle =& \tilde{\triangle} - \partial_x\left(\partial_x h \partial_z\right) - \partial_y\left(\partial_y h \partial_z\right) - \partial_z\left(\partial_x h \partial_x\right) \\
& - \partial_z\left(\partial_y h \partial_y\right) + \partial_z\left[(\partial_x h)^2\partial_z\right] + \partial_z\left[(\partial_y h)^2\partial_z\right].
\end{aligned} \tag{12}$$

As u, v and w are maintained as the prognostic model variables and $\boldsymbol{g}$ is invariant under the given coordinate transformation, no metric terms arise in the buoyancy term. However, the horizontal averaging of Θ_v is carried out on z-isosurfaces. Therefore, Θ_v is remapped to an auxiliary vertical grid and the calculated tendency is remapped back to the computational grid.

2.3 Diffuse obstacle boundaries

The spatial discretization uses a finite volume method, which allows a consistent treatment of the diffuse obstacle boundaries. To consider the conservation of a scalar q within two partitioned phases, the Gauss' theorem is formulated for a single grid-cell volume ΔV and its total surface area $\partial\Delta V$:

$$170 \quad \int_{\Delta V} \partial_t \left[q_m \chi_m + q_s (1 - \chi_m) \right] \mathrm{d}V' = - \int_{\partial(\Delta V)} \left[\boldsymbol{\eta_m} \cdot \boldsymbol{u_m} q_m + (1 - \boldsymbol{\eta_m}) \cdot \boldsymbol{u_s} q_s \right] \mathrm{d}A'. \tag{13}$$

Here, subscript m refers to the mobile phase and subscript s to the solid phase for the building interior. A' and V' are formal integration variables. χ_m is the volume-fraction function of the mobile phase and $\boldsymbol{\eta_m}$ the area-fraction function, over which the flux of the mobile phase is integrated. As already mentioned, the stationary solid phase is dropped, as it is $\partial_t \chi_m = 0$, $\partial_t q_s = 0$ and $\boldsymbol{u_s} = 0$. This simplifies the conservation form and finally leads to the particular differential form of the transport equation for the mobile phase:

$$\partial_t q_m = - \frac{1}{\chi_m \Delta V} \int_{\partial(\Delta V)} \boldsymbol{\eta_m} \cdot \boldsymbol{u_m} q_m \, \mathrm{d}A = - \frac{1}{\chi_m} \nabla \cdot \boldsymbol{\eta_m} (\boldsymbol{u_m} q_m) =: -\nabla_m \cdot (\boldsymbol{u_m} q_m). \tag{14}$$

The subscript m was only briefly introduced here and will be dropped again, as only the mobile phase is of interest.

Using the Cartesian grid structure, the discrete flux-divergence results from:

$$180 \quad \nabla \cdot \boldsymbol{F} = \frac{1}{\chi \Delta V} [(\eta_x \Delta A_x F_x)^L - (\eta_x \Delta A_x F_x)^R \tag{15}$$
$$+ (\eta_y \Delta A_y F_y)^L - (\eta_y \Delta A_y F_y)^R$$
$$+ (\eta_z \Delta A_z F_z)^L - (\eta_z \Delta A_z F_z)^R],$$

where ΔV and $\Delta A_{x,y,z}$ are the cuboid volumes and face areas, respectively. The superscripts R and L refer to the left and right cell face for each dimension, respectively. F_x, F_y, and F_z are fluxes, whose concrete form is determined by the partial differential equation. Note that the contravariant flux has to be used for F_z.

The pressure gradient components on the cell faces are discretized with second-order accuracy. The gradient components centred on the x- and z-orientated cell faces, respectively, are:

$$\partial_x p = \frac{2 \eta_x \Delta A_x}{(\chi \Delta V)^L + (\chi \Delta V)^R} \left(p^R - p^L \right) - \frac{1}{\Delta_x} \left(h^R - h^L \right) \mathbf{L}^{z \to x} \partial_z p, \tag{16}$$

$$190 \quad \partial_z p = \frac{2 \eta_z \Delta A_z}{(\chi \Delta V)^L + (\chi \Delta V)^R} \left(p^R - p^L \right), \tag{17}$$

where $\mathbf{L}^{z \to x}$ is the linear interpolation operator from a z-face to an x-face, and Δ_x is the cuboid size used for differencing the terrain function.

For a grid-conform surface, the area fractions $\eta_{x,y,z} = 0$, from which the required homogeneous zero Neumann-boundary condition follows. In this case, a grid cell is completely surrounded by grid-conform surfaces and the pressure value inside is decoupled from any neighboring values, reflecting its physical meaninglessness. For partially open semi-permeable cell-faces, the boundary condition is imposed on a fraction of the cell face area only.

The scaling fields $\eta_{x,y,z}$ and χ are derived from geometric building data. As an alternative to using terrain-following coordinates, it is possible to use diffuse boundaries for the terrain, in which case, the sub-surface is also represented by a geometric shape. Per definition, the volume-fraction field χ is expressed as the fraction of the obstacle-free volume in each grid cell. For numerical reasons, however, χ is limited to a small non-zero value. For well resolved buildings, the area-scaling fields $\eta_{x,y,z}$ are derived by calculating the intersections of the buildings with the grid-cell faces. For under-resolved buildings, however, this method fails to take the grid alignment into account. A resulting effect may be the entire missing of buildings if they do not intersect with a particular cell face, but nevertheless block the flow within a grid cell. Figure 1a) shows two possible scenarios, where grid cells are intersected by a building, which obviously blocks the flow in one dimension completely but does not intersect with the cell faces oriented in the direction of the blocking. In order to capture such occurrences more reliably, a modified method to calculate the area fractions is used. In this method, the grid-cell volume is partitioned in slices, with the slicing planes being displaced along the dimension considered (e.g. the $x-$dimension for $yz-$faces). The minimum value over the free-volume fractions of all slices is computed to define a cell-area-scaling factor, which is then assigned to the cell face in closer proximity to the obstacle. For a more robust capture of non-parallel building walls, the slicing can be repeated several times with a slight rotation of the plane normal. If, occasionally, a cell face is assigned to both values from the adjacent left and right grid cell, respectively, the minimum of both values is taken. For the resulting cell face left without assignment, the geometric intersection with buildings is calculated as in the resolved case. Figure 1b) demonstrates that this method preserves the blocking effect of a triangular-shaped building for spatial resolutions, which are too coarse to resolve individual building walls.

2.4 Numerics

2.4.1 Advection scheme

The advective tendency for a scalar q is obtained by adding the remnant velocity-divergence term from the approximate pressure solution to the flux-divergence term of the conservation law (Eq. 15):

$$(\partial_t q)_{adv} = -\nabla \cdot (\boldsymbol{u}q) + q\nabla \cdot \boldsymbol{u}. \tag{18}$$
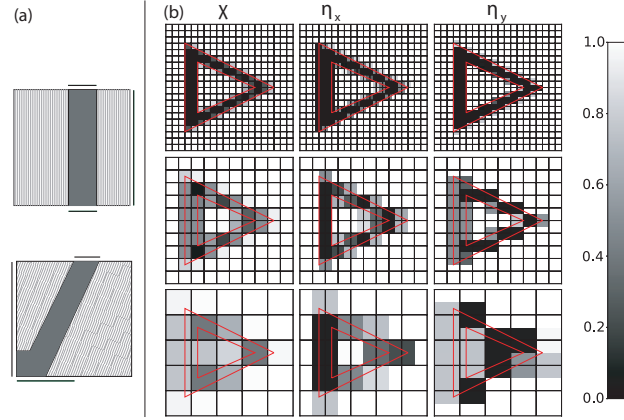
Figure 1. (a) Depiction of two different building sections (gray filled areas) inside a grid cell. The black lines mark the effectively blocked area of the respective side. The line pattern symbolizes the slicing of the grid cell volume, which is shown here only for the x-dimension. (b) Scaling factors for a complete building, now depicted as fields calculated for 3 different grid resolutions. χ is the volume-scaling field, and η_x, η_y are the horizontal components of the area-scaling field.

The flux components are linear and given by the product of the reconstructed values $\tilde{q}$ at cell faces and the exactly given momentum components. We use an upwind-biased stencil of 5th order, which results in two reconstructions $\tilde{q}^+$ and $\tilde{q}^-$ on each cell face. For an x-oriented face, the reconstructions are merged to the numerical flux by considering the wind direction:

$$(uq)_x = 0.5 \left[u_x(\tilde{q}^+ + \tilde{q}^-) - ||u_x||(\tilde{q}^+ - \tilde{q}^-) \right] \tag{19}$$

225   For non-uniform grid spacings, the coefficients of the reconstruction polynomials are pre-computed from the spatial derivative of the Lagrange polynomial which interpolates the primitive function at cell faces (see e.g. Shu, 1998). We use the pseudo-grid spacing calculated by $\Delta_x^{eff} = 2\Delta_x \chi/(\eta_x^L + \eta_x^R)$ instead of the computational grid spacing in order to automatically adjust the effective stencil width near obstacle boundaries. The behaviour of the pseudo-grid-based reconstruction of maximum 5th order is demonstrated in Fig. 2 for a positive flow direction (left to right). Within shown the circular obstacle, the pseudo-grid

230   spacing tends toward infinity, which effectively excludes such decoupled cells from interpolation. The resulting smaller stencils can be down-wind biased. An effective measure to prevent numerical instabilities is the application of a flux limiter (e.g. from Sweby 1984) near obstacle boundaries. To avoid non-physical results of positive scalars, a limiter has to be applied anyway. In the case of momentum advection, we use the absolute difference $|\eta_x(j-1) - \eta_x(j)|$ as a weighting function to merge the limited and unlimited reconstructions for obstacle-specific limiting. Note that this expression is upwind-biased (assuming a

235   positive wind direction). In the free boundary layer, the difference is zero and no limiting is applied.

The routine for scalar cell-centered advection is repeatedly used to advect a left-faced u^l and right-faced value u^r for each

momentum component (Hicken et al., 2005; Jähn et al., 2015), resulting in a total of 6 advection steps. The final momentum tendencies are obtained by interpolation of two centered tendencies on the face:

$$\partial_t^{adv} u = \frac{\left(\chi \Delta V \partial_t^{adv} u^l\right)^R + \left(\chi \Delta V \partial_t^{adv} u^r\right)^L}{(\chi \Delta V)^R + (\chi \Delta V)^L}. \tag{20}$$

240   Spatial accuracy of momentum advection is limited to the order of this interpolation procedure, which is of second order here.
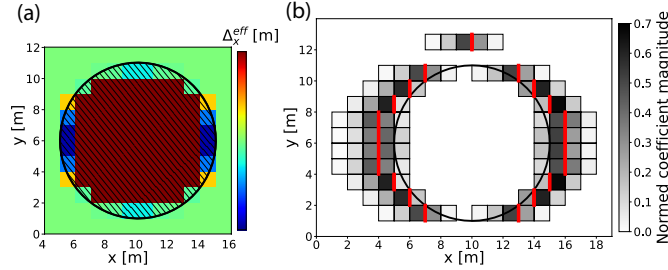
Figure 2. (a) Plot of the pseudo-grid spacing Δ_x^{eff} computed for a solid cylinder with diameter of $10\,\text{m}$ using a uniformly constant grid spacing of $1\,\text{m}$. Decoupled cells inside the cylinder are characterized by $\Delta_x^{eff} > 2$ in this plot. (b) Using same Δ_x^{eff} of (a), the reconstruction coefficients are computed for a 5-point upwind-biased stencil. The absolute values of the reconstruction coefficients are re-normed to one and depicted in shades of grey at individual reconstruction sites, which are marked by a vertical red bar. Stencil points are rendered invisible for values below 10^{-4}, as such have a negligible influence on the reconstructed value.

2.4.2   Model integration

In the incompressible-flow equations, the pressure gradient term is not directly coupled to a prognostic pressure equation. The projection method of Chorin (1968) is used to split the solution procedure in two steps. The first step integrates all the 245   momentum tendencies, except the stated pressure gradient term, explicitly in time to obtain a predicted velocity $\tilde{\boldsymbol{u}}$:

$$\tilde{\boldsymbol{u}} = \boldsymbol{u}_{t_0} + (\partial_t \boldsymbol{u})_{ex}|_{t_0} \Delta t. \tag{21}$$

The final velocity estimate after one integration step $\boldsymbol{u}_{t_1}$ has to fulfill the continuity equation:

$$\nabla \cdot \boldsymbol{u}_{t_1} = 0. \tag{22}$$

The required corrective tendency has to be associated with the neglected pressure gradient, which is formally integrated with an Euler-backward step to obtain the final velocity at $t_1 = t_0 + \Delta t$:

$$\boldsymbol{u}_{t_1} = \tilde{\boldsymbol{u}} - \frac{\Delta t}{\rho_{ref}} \nabla |p_{t_1}. \tag{23}$$

By applying the divergence operator on both sides and requiring that $\nabla \cdot \boldsymbol{u}_{t_1} = 0$, the well-known Poisson equation for pressure is obtained:

$$\frac{\rho_{ref}}{\Delta t} \nabla \cdot \tilde{\boldsymbol{u}} = \triangle p|_{t_1}. \tag{24}$$

After algebraic solution of this equation, the final state can now be composed of the fractional tendencies:

$$\boldsymbol{u}_{t_1} = \boldsymbol{u}_{t_0} + \Delta t \left[(\partial_t \boldsymbol{u})_{ex}|_{t_0} - \frac{1}{\rho_{ref}} \nabla p|_{t_1} \right]. \tag{25}$$

Equation 25 is only first-order accurate in time. For higher accuracy, it is instead used a 3rd-order strong-stability preserving Runge-Kutta scheme (SSP-RK3) for the advective and pressure-gradient tendencies, which also allows to use larger time steps. The pressure to correct the first two intermediate states is extrapolated from values given at previous time steps, and only for the final stage, the pressure solver is applied (Karam et al., 2019). We found that this combination supports stable integration up to a Courant number of $C = 0.7$.

2.4.3 Pressure solution

The discretization of the Laplace operator $\mathbf{P}$ in the Poisson equation (Eq. 24) is obtained by the product of the discrete divergence $\mathbf{D}$ and gradient in sparse matrix form. $\mathbf{D}$ is defined by the flux-balance in Eq. 15. Combining the operators to the pressure equation, the following sparse linear system is obtained:

$$\mathbf{P}p = \frac{1}{\Delta t} \mathbf{D}\tilde{u} =: b, \tag{26}$$

where $\tilde{u}$, p and b are the one-dimensional expanded arrays of the corresponding structured fields. Eq. 26 is solved with a geometric multigrid method in parallel using domain-decomposition in two dimensions. The multigrid algorithm consists of applying a smoothing method of choice which is accelerated by coarse-grid corrections. Therefore, a hierarchy of coarse grids is employed (Brandt and Livne, 2011). 3D coarsening of the grid is carried out by agglomeration of 8 grid cells to form a coarse grid cell of the next-level grid. For uneven grid sizes, a plane of grid cells with respective orientation is left uncoarsened. This particular multigrid used in combination with finite volume discretizations is often referred to as cell-centered multigrid in literature (Mohr and Wienands, 2004).

Particular challenges to the multigrid algorithm include grid stretching and, in our case, non-smoothly varying coefficients associated with the diffuse boundaries, both resulting in coefficient anisotropy. An odd grid size results in coefficient anisotropy

of the coarse grid operators. In such cases, plane smoothers are often much more robust than their point-wise pendants (Llorente and Melson, 2000). Nevertheless, most of the difficulties could be overcome by applying less elaborated methods in the current model.

Based on smoothing analysis, Larsson et al. (2005) give a condition for the optimal location of an uncoarsened plane in the case of an odd grid size. Galerkin coarse-grid approximation can result in a better approximation of the coarse-grid operator in this case, while discretization coarse-grid approximation can be more efficient for even-sized grid dimensions. Our algorithm employs a combination of both methods on different grid levels.

Yavneh (1996) found that for smoothing, successive over-relaxation (SOR) is generally superior to Gauss Seidel smoothing (even for isotropic coefficients), and he also derived approximately optimal over-relaxation factors for SOR with red-black ordering applied in multigrid for the solution of anisotropic elliptic equations.

In our implementation, also sparse-approximate inverse (SPAI) matrices (Tang and Wan, 2000; Bröker and Grote, 2001; Sedlacek, 2012) are available as suitable alternatives to SOR. These smoothers can inherently consider coefficient anisotropy through the algebraic method by which they are derived. Moreover, a variable number of non-zeros in the matrix allows a flexible control of the approximation quality and smoothing efficacy.
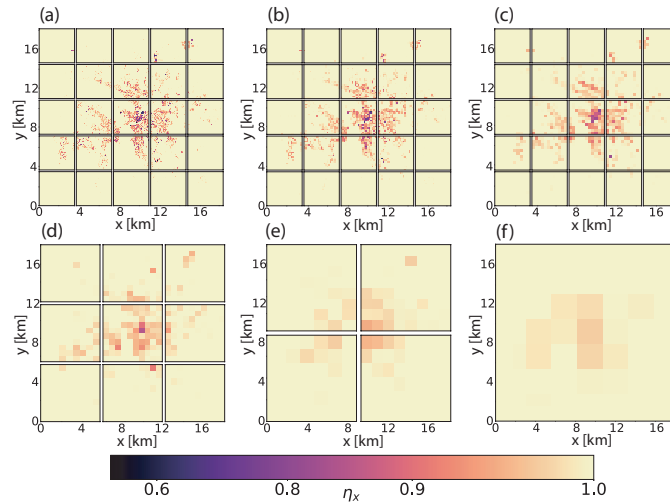
Figure 3. Example of a horizontal multigrid domain decomposition, involving 6 grid levels. Depicted are the area-scaling factors of yz-faces for the city of Leipzig at different resolutions of (a) 80 m, (b) 160 m, (c) 320 m, (d) 640 m, (e) 1280 m, and (f) 2560 m. The scaling factors are needed in the discretization of the Poisson equation. On the coarsest grid, a single processor is left for the computation.

2.4.4 Lateral boundary conditions

Before each pressure correction step, lateral boundary conditions of the intermediate velocity field $\tilde{\boldsymbol{u}}$ are updated. The update has to ensure global mass conservation:

$$\oint_{\partial V} \frac{\rho_{ref}}{\Delta t} \tilde{\boldsymbol{u}} \cdot \boldsymbol{n} \, dA' = 0. \tag{27}$$

This expression implies a homogeneous-zero Neumann boundary condition for pressure at all lateral boundaries, as the already updated boundary values should not be changed by the projection:

$$0 = \rho_{ref} \frac{\partial \tilde{\boldsymbol{u}}}{\partial t} \cdot \boldsymbol{n} = \nabla p \cdot \boldsymbol{n}. \tag{28}$$

Here, $\boldsymbol{n}$ is the unit normal vector on the boundary surface. A more general case arises with the use of terrain-following coordinates. By requiring not only $\partial_t \omega = 0$, but additionally $\partial_t w = 0$ at the bottom and top of the computation domain, the following boundary condition follows at these boundaries:

$$\partial_z p = 0 \tag{29}$$

$$\partial_x p = \begin{cases} 0 & \partial_x h \neq 0 \\ \text{not specified} & \text{else} \end{cases}$$

$$\partial_y p = \begin{cases} 0 & \partial_y h \neq 0 \\ \text{not specified} & \text{else} \end{cases} \tag{30}$$

In the discrete gradient operator, homogeneous Neumann boundary conditions are implemented by setting all coefficients associated with the node to zero where the condition applies.

The boundary condition for the velocity field has to be compatible with a dynamic mesoscale forcing and also satisfy Eq. 27. Firstly, inflow and outflow regions are dynamically determined in order to impose separate appropriate boundary conditions. Outflow regions are characterized by convective transport out of the domain. Therefore, a simple normalized convective transport speed $C_\perp = \frac{\Delta t}{\Delta x} \boldsymbol{u} \cdot \boldsymbol{n}$ is computed. Note that more elaborated formulations for $C_\perp$ exist. $C_\perp$ is further bounded to $[0,1]$ for numerical reasons. It is then $C_\perp > 0$ for outflow regions. For such regions, the radiation boundary condition by Miller and Thorpe (1981) is imposed:

$$\boldsymbol{u}_{l+1}^{t+1} = \boldsymbol{u}_{l+1}^t - C_\perp \left(\boldsymbol{u}_{l+1}^t - \boldsymbol{u}_l^t \right). \tag{31}$$

The order of the spatial indexing l is in normal direction to the boundary and not to be mixed up with the standard interior indexing. The index $l+1$ corresponds to the first ghost cell. At the remaining inflow boundaries, Dirichlet conditions are specified.

Instead of this flexible inflow-outflow boundary condition, a Rayleigh damping layer can be used as another prognostic tendency near the domain top:

$$\partial_t^{dmp} q = -\frac{R(z)}{\tau}(q - q_0).\tag{32}$$

Eq. 32 can be applied to gradually relax any prognostic variable toward a prescribed horizontal mean state at the top boundary. R is a ramp function with values between $[0, 1]$, τ the damping time scale and q_0 the prescribed boundary value.

Final mass conservation is enforced by making sure that the total inflow-mass flow is exactly balanced by the total outflow-mass flow. This is accomplished by computing an averaged correction velocity from the mass-flow difference. In the case, each spatial dimension is considered independently from each other, this results in 3 different correction velocities. The correction velocities are finally added to the boundary-perpendicular velocity component at the outflow regions.

Figure 4 shows that the outflow boundary condition with the proposed convective transport speed is well suited to our incompressible model even for highly unsteady flows, like the depicted vortex street in the wake of a cylinder. Individual vortices are not visibly reflected at the boundary, and also in the temporal mean, based on Fig. 4b), the influence of the boundary is not noticeable. The flexible boundary condition can principally be applied to any other scalar quantity, although, specifying Dirichlet conditions for advected scalars was found to be well suited.

2.5 Physical processes

2.5.1 Subgrid model

For numerical simplicity and efficiency, a static Smagorinsky subgrid model is used (Deardorff, 1970). Since the atmospheric model is operated in the limit of infinite Reynolds numbers, the principal purpose of the subgrid model is to dissipate enough energy at the shortest wavelengths to obtain a physically realistic energy cascade. In our case, the subgrid model also has to compensate for the under-resolved vertical mixing of tracers within the urban boundary layer.

In the Smagorinsky model, the rates of strain $s_{x,y}$ are approximated by the velocity gradients:

$$s_{x,y} = \frac{1}{2}\left(\partial_y u + \partial_x v\right).\tag{33}$$
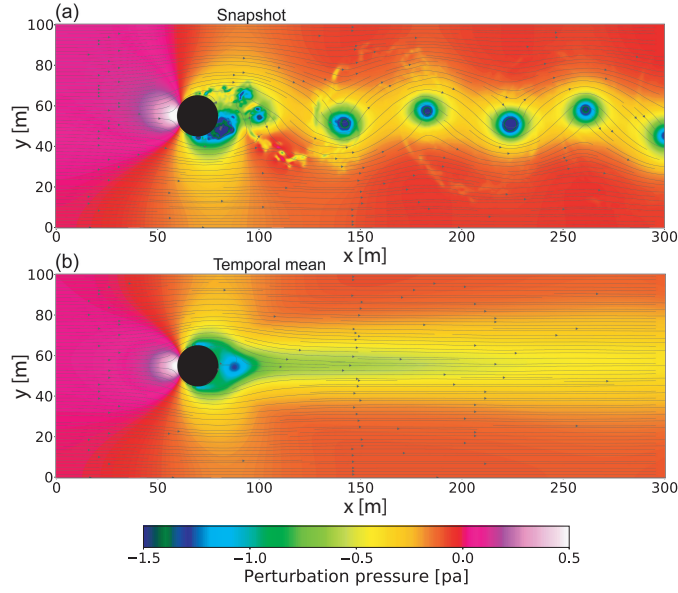
Figure 4. LES of 3D flow past a cylinder. The horizontal grid spacing is uniformly $0.5\,\mathrm{m}$. The cylinder has a diameter of $40\,\mathrm{m}$. The approaching flow is laminar with $u = 1\,\mathrm{m\,s^{-1}}$. Flexible Dirichlet-radiation conditions are imposed on all horizontal domain-boundaries, while periodic boundary conditions are used in z-direction. The contour plot depicts the pressure and the streamlines the horizontal velocity field. (a) shows a frame at the instance during a sharply defined vortex is about to cross the right boundary, (b) shows the temporal mean over a representative simulation period.

340   The subscripts x, y refer to the spatial components. The turbulent fluxes are derived in analogy to the viscous fluxes by assuming an eddy viscosity ϵ_k:

$$u_x' u_y' = 2\epsilon_{x,y} s_{x,y} \tag{34}$$

The often additionally mentioned anisotropic residual-stress tensor is ignored in the given incompressible case. In the most simple case, $\epsilon_{x,y}$ is also diagnosed from the rates of strain. $\epsilon_{x,y}$ is denoted in tensor form, as an anisotropic mixing length $l_{x,y}$
345   is used to reflect grid anisotropy:

$$\epsilon_{x,y} = l_{x,y}^2 |\mathbf{S}| f = (c_s \Delta_{x,y})^2 |\mathbf{S}| f_s \tag{35}$$

$|\mathbf{S}|$ is the Frobenius norm of the strain-rate tensor. c_s is the Smagorinsky constant, which is fixed in a static model. Tests with a boundary layer simulation revealed that the range $0.1 < c_s < 0.15$ gives good results in combination with the 5th-order

upwind scheme. Grid anisotropy enters via $\Delta_{x,y}$, which is further modified by half the mean distances to walls h_x and h_y,
350 respectively:

$$\Delta_{x,y} = \min\left(\sqrt{\Delta_x \Delta_y}, 1.8 h_x, 1.8 h_y\right). \tag{36}$$

Function f_s introduces the influence of the stratification on the eddy viscosity. It is assumed:

$$f_s = \begin{cases} 0 & Ri \geq 0.25 \\ \sqrt{1 - 16 Ri} & Ri < 0 \\ (1 - 4 Ri)^4 & else, \end{cases} \tag{37}$$

with the Richardson number Ri:

355 $$Ri = \frac{g \partial_z \Theta_v}{\Theta_v |S|^2}. \tag{38}$$

For scalar diffusion, the eddy viscosity is divided by the turbulent Prandtl number, which is assumed to be $Pr = 2/3$ here.

If not mentioned otherwise, the appearing spatial derivatives are discretized with second-order differences. To obtain the strain rate components and the eddy viscosity on different stencil points (cell faces areas and cell centers), linear interpolation 360 is used. The subgrid tendency is formed by the divergence of the diffusive fluxes. Shifted grids are introduced to account for the definition of the velocity components on the cell faces. For example, diffusion of the u-component requires a grid shifted by $\Delta_x/2$, for which the scaling fields are also obtained by linear interpolation.

The diffusive fluxes are given by:

$$F_{x,y} = 2 \epsilon_{x,y} s_{x,y}. \tag{39}$$

365 For diffusion of the u-component, the fluxes in the 3 spatial directions are $F_{x,x}$, $F_{x,y}$ and $F_{x,z}$. Those of the other components are obtained by permuting the subscripts. For a scalar quantity q, the required fluxes are $F_{x,x}$, $F_{y,y}$ and $F_{z,z}$, with the rates of strain being replaced by the gradient components.

2.5.2 Surface fluxes

Surface fluxes for momentum, heat and moisture are parameterized using Monin-Obhukov similariry theory (Louis, 1979).
370 An expression for the vertical transfer coefficient C_z can be obtained by transforming the logarithmic wind law:

$$C_z = \frac{k^2}{\log^2{(z/z_0)}} \tag{40}$$

$k = 0.4$ is the Von Kármán constant, z_0 the surface roughness length, and z the height difference from the modeled surface to the grid level where the parameterization is evaluated.

The momentum sinks from horizontal surfaces are:

$$\frac{\partial u'w'}{\partial z} = -\frac{A_z}{\Delta V} f_m C_z u \sqrt{u^2 + v^2} \tag{41}$$

and

$$\frac{\partial v'w'}{\partial z} = -\frac{A_z}{\Delta V} f_m C_z v \sqrt{u^2 + v^2}. \tag{42}$$

Analogously, the source terms for heat and moisture from horizontal surfaces are:

$$\frac{\partial \Theta'w'}{\partial z} = -\frac{A_z}{\Delta V} f_h C_z \left(\Theta - \Theta^s\right) \sqrt{u^2 + v^2} \tag{43}$$

and

$$\frac{\partial Q_v'w'}{\partial z} = -\frac{A_z}{\Delta V} f_h C_z \left(Q_v - Q_v^s\right) \sqrt{u^2 + v^2}. \tag{44}$$

Θ^s is the surface potential temperature, and Q_v^s the surface specific humidity. A_z is the total exposed horizontal surface within the grid cell, and ΔV the effective cell volume. f_m and f_h are stability functions, and z_0 is the surface roughness length. We adopt the expressions given in Doms et al. (2013) to calculate f_m and f_h for land surfaces. Sources and sinks from vertical building walls are treated similarly, but the stability function is set to unity in this case. For x-orientated surfaces, z is replaced by half of the average distance between surfaces in the equation for the transfer coefficient. Analogously, A_z is replaced by A_x for the total projected surface area with x-orientation. The surface fields Θ^s and Q_v^s are part of the external forcing and have to be provided either by the hosting mesoscale model or field-interpolated measurements.

2.5.3 Turbulence recycling scheme

Wu (2017) gives an overview of various turbulence generation methods to provide turbulent inflow conditions. Among the different methods, a turbulence recycling scheme is used in our model, as it is computationally efficient and can be applied to a wide range of different domains and flow types. In our implementation, turbulence can be extracted within a maximum of 4 vertical planes, each one properly displaced parallel to a particular inflow boundary. The resulting domain fetch for turbulence recycling has to extend several integral length scales in order to prevent a spurious periodic pattern of the recycled turbulent features.

At each model time step, a horizontal filter is applied on the velocity components within the recycling plane. In the case of coupling with a mesoscale model, the filter width w_r is set to just below the spatial resolution of the host model in order to

spare mesoscale variations. Vertical filtering is not feasible due to the strong vertical wind shear within the boundary layer. The filtered velocity component is subtracted from the original one to obtain the small-scale fluctuation component:

$$u(z,y)' = u(z,y) - <u(z,y)>_{w_r} .$$
(45)

Equation 45 assumes a boundary perpendicular to the flow in x-direction. The turbulent intensity is re-scaled to the target value, after which the fluctuation field can be added to the inflow boundary field u_{in} of the large scale flow u_{ls}.

$$u_{in}(z,y) = u_{ls}(z,y) + min\left[a_{max}, \frac{||u'_{tar}||_2(z)}{||u'||_2(z)}\right] u'(z,y)$$
(46)

a_{max} is used to limit the artificial amplification of turbulence shortly after model initialization.

The filtering operation, as well as the calculation of the turbulent intensities require communications in the parallel implementation. For optimal parallel scaling, filtering is performed on each subdomain containing a part of the recycling plane instead of the much simpler method of filtering all the data with a single processor. A box-shaped filter is used to minimize the amount of communication in the parallel filtering. A final communication may be necessary to transfer the recycled turbulence to the inflow boundary. Despite the communication-intense filtering, the computational costs of the recycling scheme were found to be 1% to 2% of total costs on average.

2.6 Programming language

The presented model CAIRDIO v1.0 is written in Python, a programming language which facilitates a straightforward implementation of numerical methods, code compactness, and code readability. Python packages like NumPy, SciPy, and mpi4py make the programming language also suitable for high-performance computing. Our model implementation can particularly benefit from NumPy, as all time critical numerical routines (e.g. all the routines for the computation of explicit tendencies) support vectorized computations. This is an inherent property of the diffuse interface approach, as all grid cells, except for the ghost cells at sub-domain boundaries, are computation cells and treated in the same manner. All tendencies are formulated in flux-divergence form, and as a result, obstacle boundaries have not to be specified particularly. For the multigrid pressure solver, SciPy provides efficient data structures, methods, and functions for sparse-matrix algebra. Parallel computation is realized through a 2D-domain decomposition, with each processing node running its own full-fledged simulation. Message passing interface (MPI) is used to successively exchange data between subdomains.

3 Numerical tests

Some numerical tests are conducted in order to examine grid-sensitivity of the diffuse obstacle interface, the dynamic core and parallel efficiency of the model. To test the diffuse obstacle interface, a similar advection test as reported in Calhoun and

LeVeque (2000) is performed. In order to show that the dynamic core can reproduce the expected evolution of an idealized setup, the rising-bubble experiment of Wicker and Skamarock (1998) is conducted once again. It also provides a benchmark to compare the anelastic approximation with a fully compressible model used in the original study. Finally, a third test is
430   conducted to demonstrate the strong scalability on a high-performance computing platform.

3.1   Advection through an obstacle field

In this 2D test, the computation domain contains randomly positioned circular obstacles of varying size. In a first step, an approximate potential flow solution is computed. Therefore, one step of the pressure projection method is applied on the initial wind field defined by $u = 1$ and $v = 0$. The resulting potential flow field is used to advect a test-tracer front, which is solely
435   characterized by the left inflow boundary condition:

$$c = \begin{cases} 1 & t \leq 40\,s \\ 0 & t > 40\,s \end{cases} \tag{47}$$

For the transversal-flow direction, periodic boundary conditions are used.

The reference simulation is carried out on a domain with 200×100 grid cells and a uniform grid spacing of $1\,\mathrm{m}$. The
440   obstacle radii range from $5\,\mathrm{m}$ to $10\,\mathrm{m}$. The simulation is repeated on coarser grids with dimension sizes of 100×50, 50×25, and 25×13, respectively. Figure 5 shows the simulation results at $t = 150\,\mathrm{s}$. The obstacles are well resolved on the grid with $1\,\mathrm{m}$ spacing, but become more and more diffuse with decreasing grid resolution towards $8\,\mathrm{m}$ for the coarsest grid. The initially planar test-tracer wave is delayed and deformed by the obstacles. The qualitative impression is that the shape of the wave is not very sensitive to the grid resolution. Even in the most diffuse case, the position and shape of the wave matches that of
445   the higher resolved simulations well. The wave dispersion can be quantified by considering the washout curves shown in Fig. 6, which are the spatially averaged concentrations at the outflow boundary versus time. For the case with the finest grid, the concentrations at the outflow boundary start to rise after $t = 145\,\mathrm{s}$ and peak at about $t = 185\,\mathrm{s}$. At $t = 250\,\mathrm{s}$, most of the wave is advected out of the domain. Remarkably, as already found by (Calhoun and LeVeque, 2000), the washout curve is not sensitive to the grid resolution up to $4\,\mathrm{m}$, which is at the transition when obstacles start to become diffuse. For the case with the coarsest
450   resolution of $8\,\mathrm{m}$, the peak is slightly broader, peak concentrations are lower, and the peak occurs earlier by about $5\,\mathrm{s}$. At this grid resolution, the numerical diffusion of the advection scheme becomes important, as the resolution capability is around 6 grid points, which is barely enough to resolve the wave.

3.2   Rising thermal

In the rising thermal simulation described in Wicker and Skamarock (1998), the Euler equations without diffusion are solved
455   on a 2D domain with a height of $10\,\mathrm{km}$ and a width of $20\,\mathrm{km}$. The grid spacing is uniformly $125\,\mathrm{m}$. The initially constant virtual potential temperature field of $\Theta_v = 300\,K$ is perturbed by a circular thermal:
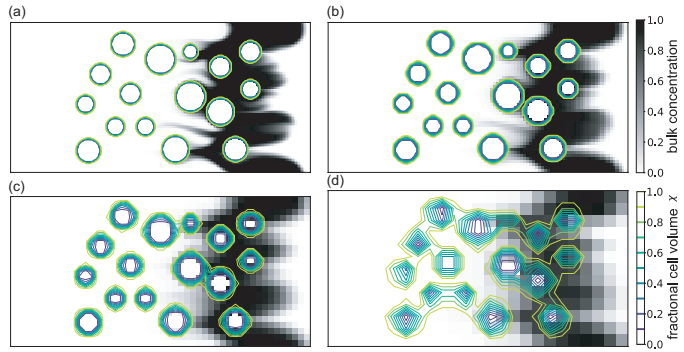
Figure 5. Results of the advection test with obstacles. Map plots of the concentration field of a test tracer after a simulation time of $150\,\mathrm{s}$. The obstacles are drawn by contours of the volume-scaling field χ. Shown are the results of the flow simulations for the grids (a) 200×100 cells, (b) 100×50 cells, (c) 50×25 cells, and 25×13 cells, respectively.
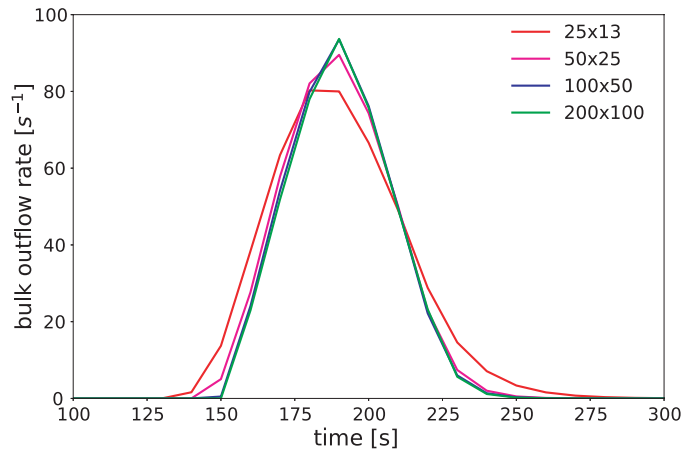
Figure 6. Washout curves of the test tracer for the different grid resolutions in Fig. 5. The values are computed by integrating over the modeled outflow at the downwind (right) domain border.

$$
\Delta\Theta_v = \begin{cases} 2\cos^2\left(\frac{\pi r}{2L}\right) & r \leq L \\ 0 & r > L, \end{cases}
$$
$$
r = \sqrt{x^2 + (z - 2\,\mathrm{km})^2},
$$
$$
L = 2\,\mathrm{km}
$$
(48)

The initial vertical velocity is set to $w = 0\,\mathrm{m\,s^{-1}}$ and the horizontal velocity to $u = 20\,\mathrm{m\,s^{-1}}$. Periodic lateral boundary conditions are used and a rigid boundary is placed at the domain top. Due to buoyant forces, the thermal starts rising while it is constantly advected to the right and eventually re-enters the domain at the left boundary. After a simulation time of $t = 1000\,\mathrm{s}$, the thermal is again situated in the center of the domain.

Figure 8a shows the evolution of the thermal based on contours of Θ_v at the time steps $t = 0\,\mathrm{s}$, $t = 350\,\mathrm{s}$, $t = 650\,\mathrm{s}$ and $t = 1000\,\mathrm{s}$. Two distinct and symmetric rotors develop. At the simulation time of $t = 1000\,\mathrm{s}$, the overall appearance of the thermal matches well that of the original simulation by Wicker and Skamarock (1998) as can bee seen from the comparison with the original plot in Fig. 7. In our simulation, however, the thermal is more compact as it is confined between $x = \pm 2300\,\mathrm{m}$ and the peak height is at about $8100\,\mathrm{m}$. In the original simulation, it is confined between about $x = \pm 2600\,\mathrm{m}$ and below $8500\,\mathrm{m}$. Whether this slight discrepancy stems from the Boussinesq approximation or the different numerical schemes used can not be finally clarified. By scale analysis, the magnitude of the buoyant acceleration is about one or two orders less than that of the inertial acceleration in the given example. So, the Boussinesq approximation should indeed apply well in this example. The used 5th-order upwind scheme is much less diffusive than the 3rd-order scheme used by Wicker and Skamarock (1998). On the other hand, the flux limiter introduces an adequate amount of diffusion near sharp gradients to prevent oscillations and to ensure a positive solution ($\theta_v \geq 300\,\mathrm{K}$). This can explain our smoother contour lines inside the rotor. A slight asymmetry from the lateral advection can be noticed, which is most evident in the contours of vertical wind speed in Figure 8b. This asymmetry can be slightly more reduced by decreasing the integration step size (not shown). The combination of the advection scheme with the 3rd-order SSP-RK3 time scheme gives stable results up to a Courant number of $C = 0.7$. Positivity of the solution is preserved up to $C = 0.5$.
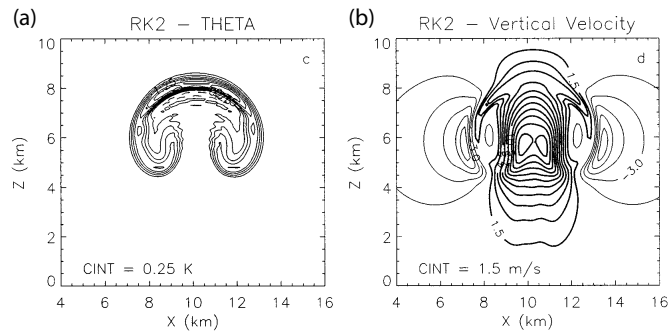
Figure 7. Original plot of Wicker and Skamarock (1998) (Figures 5c-d therein) ©American Meteorological Society. Used with permission. It shows the rising thermal problem computed using a third-order upwind scheme and a second-order Runge-Kutta scheme (Figures 5c-d therein): Shown are the potential temperature (a) and vertical velocity (b) after $1000\,\mathrm{s}$ of integration time.
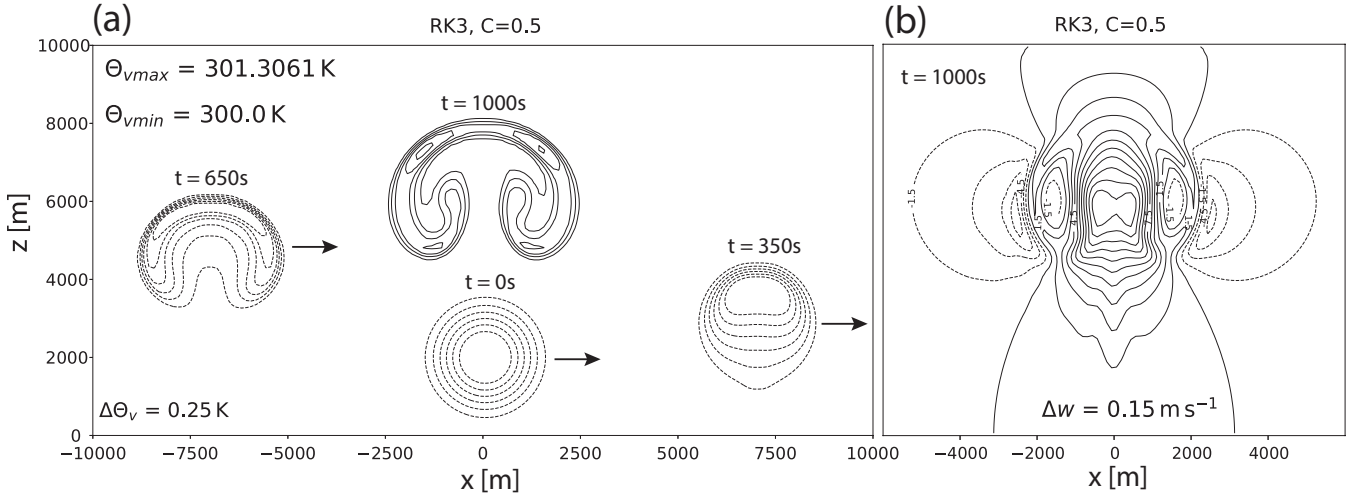
Figure 8. Rising thermal simulation with lateral advection: (a) Contours of virtual potential temperature at different time steps. The initial and intermediate states are drawn with dashed lines, the final state at $t = 1000\,s$ with solid lines. The contours with $\theta_v = 300\,\mathrm{K}$ are omitted. (b) Contours of vertical wind speed at $t = 1000\,s$. Dashed lines are used for negative values.

3.3 Strong scalability test

We tested parallel scalability of the model for a domain spanning 350×350 grid cells in the horizontal dimensions and 82 grid cells in the vertical, thus consisting of approximately 10 million cells in total. For realistic demands on the pressure solver, a grid of diffuse buildings was placed at the bottom of the domain and vertical grid stretching was applied. The high-performance computing platform we used for the scaling test is organized into nodes, each one equipped with two 12-core Intel Xeon E5-2680 v3 CPUs, making in total 24 cores per node available. The strong scalability test is carried out for a variable number of cores ranging from one to 400. 400 cores correspond to a minimum average horizontal sub-domain size of 17.5 grid cells. Fig. 9 demonstrates the scaling efficiency of the model, as well as of the individual tasks for pressure projection and advection computation of 3 momentum components and two scalars, which together are responsible for the bulk of the computation time. As a reference, the dashed lines show ideal scaling starting from a reference of one single core. Comparing the experimentally obtained curves with the idealized curves, firstly, a peculiar drop in scaling efficiency between 4 and 9 cores can be noticed. However, this drop cannot be explained by the parallelization design and it was also not observed using a different platform with less cores (not shown). The arithmetic-intense advection computation shows, apart from the already noted drop between 4 and 9 cores, excellent scaling, which becomes even super-linear above 100 cores due to cache effects. Not surprisingly, the pressure solver, which requires two communications for each smoothing iteration, does not scale so well. While cache effects can balance the increasing communication overhead up to 200 cores, above this number, scaling is no longer satisfactory. Due to the lower costs of the pressure solution in comparison with the advection computation, the model shows a very good

scaling up to 400 cores for this test case. Note, however, that the best decomposition to benefit from cache effects shifts toward a lower number of cores when the vertical dimension size is decreased. In theory, the implementation of the smoothing procedure supports overlapped communication and computation, as the matrix-vector product of the halo layer is computed independently from the inner matrix-vector product. In this test, we could not observe any true overlap, which would require further investigation. Addressing this feature in future can help to additionally improve scalability of the pressure solver.
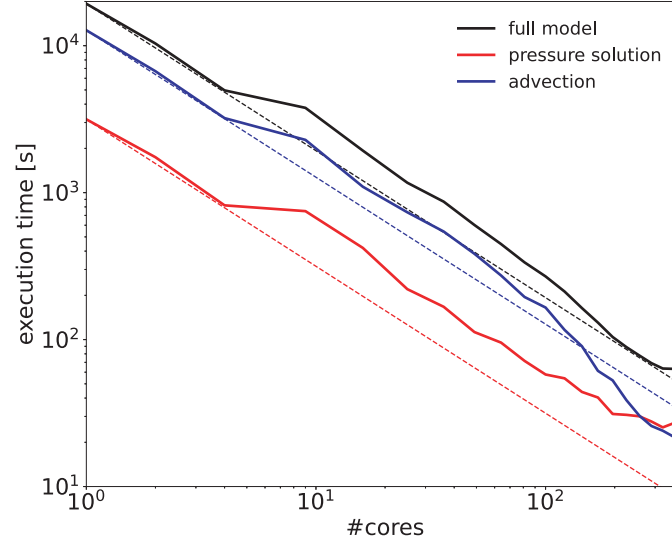
Figure 9. Strong scalability of CAIRDIO v1.0 tested for a domain with $350 \times 350 \times 82$ grid cells using 1 to 400 cores on a high-performance computing platform. Dashed lines show ideal scaling.

4 Model evaluation with wind-tunnel experiment and grid size sensitivity tests

The wind tunnel experiment "Michelstadt", which was carried out in the wind tunnel WOTAN (Lee et al., 2009) of the University of Hamburg, Germany, is used to evaluate the model accuracy and reliability of dispersion simulations within an urban canopy. Advantages of wind-tunnel data over field observations for model evaluation are the accurately controlled approaching-flow conditions and the high density of wind and concentration measurements in both space and time, which results in a high statistical significance of data (Schatzmann et al., 2017). The fictitious city district "Michelstadt" is based on a typical Central European downtown area with spacious polygonal courtyards surrounded by residential building units at a scale of 1:225 (see Fig. 10a). Building roofs are approximated by horizontal surfaces, and their heights range from $15\,\mathrm{m}$ to $25\,\mathrm{m}$ at the full scale. The approaching flow can be characterized by a neutrally stratified and horizontally homogeneous urban boundary layer with a parametric surface-roughness length of approximately $1.4\,\mathrm{m}$. Experiments were carried out using two different approach-flow directions of $0\,°$ and $180\,°$. This provides the opportunity to test the robustness of model results after model-parameter tuning

using the wind-tunnel results from the default $0\,°$ direction. The experimental dataset consists of time-resolved flow and disper-sion measurements, from which temporally averaged statistics were calculated. A dense array of sensors provided horizontal

515 wind measurements within planes at heights of $2\,\mathrm{m}$, $9\,\mathrm{m}$, $18\,\mathrm{m}$, $27\,\mathrm{m}$, and $30\,\mathrm{m}$ over a restricted area. For the dispersion mod-eling, neutrally-buoyant gas was continuously released at different locations on the floor (see Fig. 10 for the locations of the release points). Release points S2 and S4 were used for the approach-flow direction of $0\,°$, S6, S7 and S8 for the reverse direc-tion. S5 was used for both wind directions. Tracer-gas detectors to measure concentrations were all positioned at a full-scale height of $7.5\,\mathrm{m}$ and at different horizontal positions depending on the source. Further information on the experiment and the

520 datasets are provided in Baumann-Stanzer et al. (2015).

The numerical simulations are performed at the full scale using a series of grids with horizontal resolutions of $5\,\mathrm{m}$, $10\,\mathrm{m}$, $20\,\mathrm{m}$, $40\,\mathrm{m}$, and $80\,\mathrm{m}$. The $5\,\mathrm{m}$ horizontal resolution is used as the reference. The coarser domains are used to test the grid-sensitivity of the dispersion simulation. The vertical grid spacing near the surface ranges from $2\,\mathrm{m}$ for the finest grid to $7\,\mathrm{m}$ for the coars-est ones. It is increasingly stretched beyond $30\,\mathrm{m}$ above surface. The effective domain height is approximately $600\,\mathrm{m}$, which

525 corresponds to the scaled wind-tunnel height. Above this height, the wind components are dampened to the horizontally and temporally average state.

A first precursor simulation is run with periodic boundary conditions to obtain laminar and turbulent inflow profiles for the full vertical domain extend. These are used to drive the actual experimental simulations containing the buildings and test-tracer release points. In the precursor simulation, which uses a shorter domain with a total length of roughly $1\,\mathrm{km}$, buildings are not

530 present and the entire rigid bottom-domain boundary is covered with roughness elements. The geometries and arrangement of elements are adapted from the experiment. In order to model the effect of the lateral wind-tunnel confinement, rigid walls are placed at the flow-perpendicular boundaries. The modeled statistics of the established neutrally-stratified boundary layer flow are re-scaled to the reference wind speed of $6\,\mathrm{m\,s^{-1}}$ at 50 m height. The obtained horizontally averaged vertical profiles are vertically interpolated for the coarse-grid simulations. In the experimental simulations containing the model city, turbulent

535 approach-flow conditions with the correct target intensities are generated using the turbulence-recycling scheme. The recycling plane is placed well downstream of the model city and after a short pattern of roughness elements near the outflow boundary (Fig. 10a). This allows for a much shorter additional domain fetch necessary for turbulence generation. This particular posi-tioning of the recycling plane is also well supported by the similar parametric roughness length of the elements and the model city. Also, the extracted fluctuation intensities are re-scaled to the values of the target wind field, which further reduces the

540 impact of obstacles further upstream.

4.1 Inflow profile

Figure 11 shows the modeled and re-scaled horizontally averaged statistics of the boundary-layer flow generated by the precur-sor simulation. The experimentally obtained profiles are included for comparison. In the modeled mean horizontal wind speed, the roughness layer extends up to $20\,\mathrm{m}$ height, above which it is proceeded by the logarithmic Prandtl layer. The slope of the

545 modeled wind profile in the semi-logarithmic depiction matches the observed profile well. The modeled turbulent intensities are generally underestimated by up to $20\,\%$ when compared with the observations. This difference is most likely due to artificial
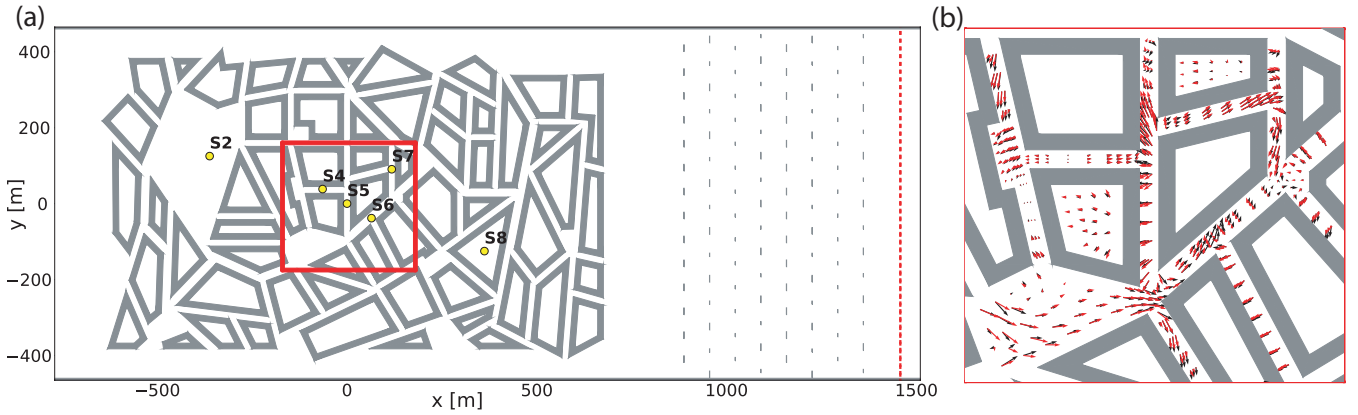
Figure 10. (a) Depiction of the model domain for the numerical simulation of the wind tunnel experiment "Michelstadt". Rigid boundaries, including the buildings, the side walls and the roughness elements are drawn using grey color. The yellow circles mark the tracer-gas release points. The red dotted line is at the position of the turbulence recycling plane. The area within the red bounding box contains the horizontal wind measurements arranged in horizontal planes. (b) Comparison of measured (black arrows) and modeled (red arrows) time-averaged horizontal wind vectors at $2\,\mathrm{m}$ height.

dissipation in the model. The peak values at about $400\,\mathrm{m}$ height are probably residual reflections at the domain top, which are not completely prevented by the dampening. The measurements below a height of $20\,\mathrm{m}$ are more difficult to compare, since they are located within the roughness layer and thus are strongly influenced by the relative location to the nearby roughness
550 elements.
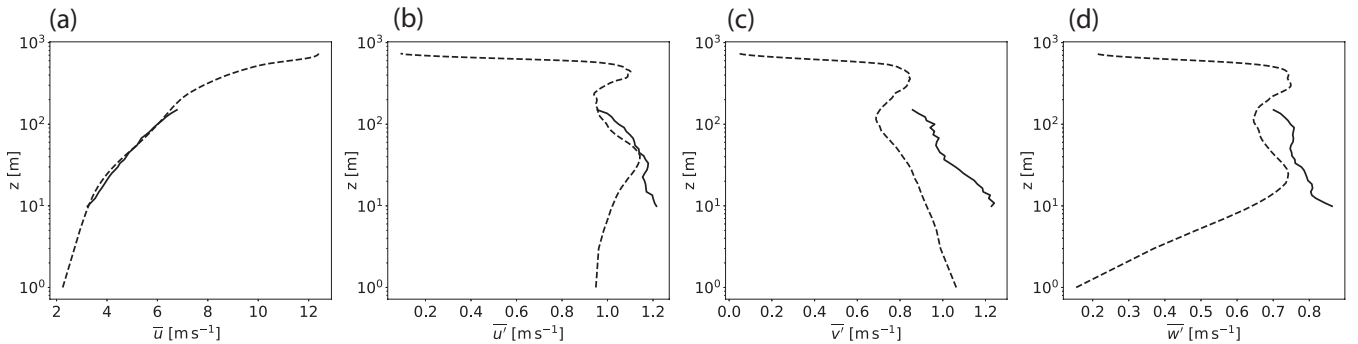
Figure 11. Modeled (dotted lines) and measured (full lines) mean and turbulent statistics of the approach flow in the "Michelstadt" wind tunnel experiment. Depicted are (a) the temporal mean velocity component $\overline{u}$ and (b-d) the grid-scale turbulent intensities $\overline{u'}$, $\overline{v'}$, and $\overline{w'}$, respectively.

4.2 Horizontal wind evaluation

The modeled horizontal wind is evaluated on the reference grid with $5\,\mathrm{m}$ horizontal resolution. Figure 10b) gives a first qualitative impression of the agreement. It shows the time-averaged horizontal wind vectors both for the model and the measurements within the plane of $2\,\mathrm{m}$ height. Overall, the model is capable of reproducing the measured flow pattern. However, the modeled wind direction does not always match the corresponding measured vector well, which is most notable near some intersections. The scatter plots for wind speed (Fig. 12) and wind direction (Fig. 13) give a more quantitative and conclusive picture. The accuracy of modeled wind speed is high taking the normalized mean standard error (NMSE) as a proxy. Its value is consistently below 0.1. The fractional bias (FB) shows only a very slight underestimation of the wind speed (0.02 - 0.07). This is not very surprising, as the model resolution of $5\,\mathrm{m}$ is not high enough to completely resolve the re-circulation zones within the urban canopy (Xie and Castro, 2006). Thus, the modeled wind speed within circulation zones tends to be lower as a result of more extensive spatial averaging.
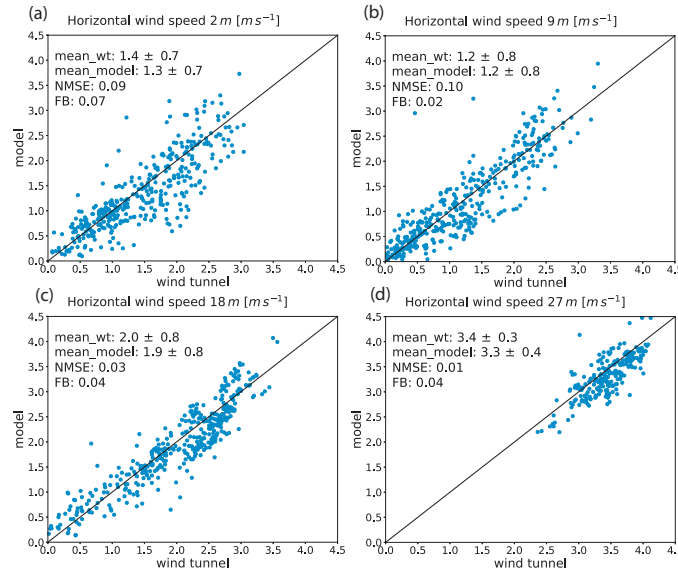
Figure 12. Scatter plot of modeled vs. measured horizontal wind speed within planes at different heights. For a quantitative comparison, NMSE and FB are calculated.

4.3 Dispersion evaluation

The resulting time-averaged modeled concentrations are interpolated to the detector sites and paired with corresponding measurements. As proxy for the quality of model results in comparison to the measurements, NMSE, FB and additionally the fraction of within factor 2 (FAC2) are calculated. Based on the guidelines presented in Hanna and Chang (2012), acceptance criteria for a valid simulation are $\mathrm{NMSE} < 6$, $|\mathrm{FB}| < 0.67$ and $\mathrm{FAC2} > 0.3$.
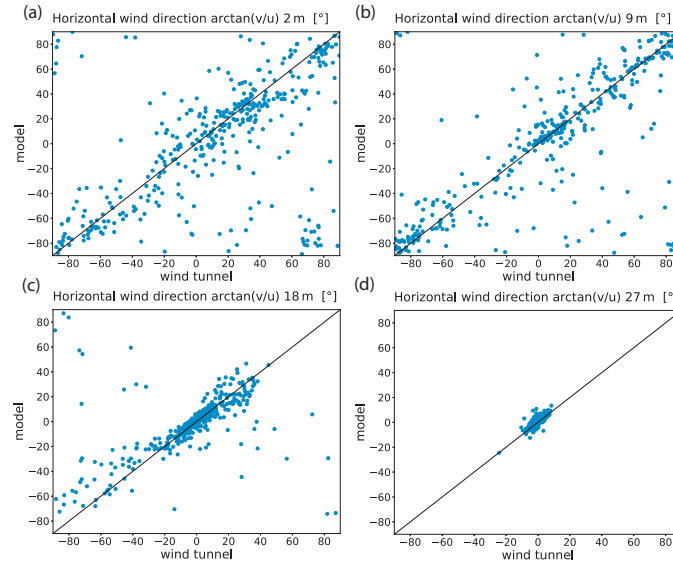
Figure 13. Scatter plot of modeled vs. measured horizontal wind direction within planes at different heights.

The simulations with the approach flow direction of $0\,^\circ$are used to optimize the model configuration with respect to these test criteria, while the reverse flow direction is used to validate the robustness of model results using the same parameter configuration. One important model tuning parameter is the Smagorinky constant, which was set to $c_s = 0.15$ for all simulations. However, for the simulations with $40\,\mathrm{m}$ and $80\,\mathrm{m}$ resolution, the vertical mixing length had to be increased to $20\,\mathrm{m}$ within the urban canopy to compensate for the poorly defined eddies important for vertical mixing.

Among the tested sources, S2 is the only one emitting into an open area, whereas all other sources are placed within street canyons or courtyards. Thus, S2 is probably the least difficult to simulate. Figure 14 gives an impression of the simulated time-averaged plumes resulting from S2 at the height of detector sites. Figure 14a shows the reference simulation with the highest grid resolution. The overall qualitative agreement with measurements seems very good, except for a street canyon in roughly -45 $^\circ$direction and in close proximity to the source. Therein, the modeled concentrations are too low. Expectedly, increasing the grid spacing results in a deterioration of the qualitative agreement with measurements and with the reference simulation. For example, increasingly less tracer gas is advected inside the flow-parallel street canyon, where most of the detectors reside. Instead, the plumes are increasingly smeared over a wider area. This is especially evident for the $20\,\mathrm{m}$ grid spacing, whereas for the even coarser resolutions of $40\,\mathrm{m}$, the agreement with measurements improves again. Notably, this behaviour is only observed for this particular source. The grid resolution of $80\,\mathrm{m}$ clearly shows the least accurate results. While the dispersion pattern still resembles those of the better resolved simulations and shows the imprint of buildings, concentrations are too high in the down-wind swath. At this resolution, a more elaborated mixing parameterization could still give improved results over the simple Smagorinsky model.
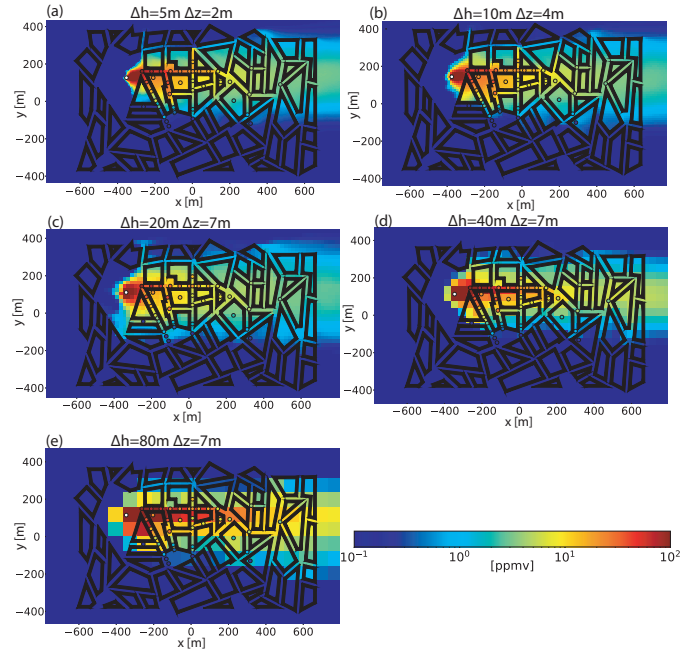
Figure 14. Map plots of modeled concentration fields in ppmv at $7.5\,\mathrm{m}$ height for source S2 and different grid resolutions. The black circles mark the location of measurements for this particular source, and the fill color indicates the measured concentration according to the color bar of the map plots.

For the quantitative evaluation of the presented simulations, modeled and measured values are compared in scatter plots in Fig. 15, and the aforementioned statistical acceptance parameters are derived. For the reference case, most of the model data is tightly distributed near the bisecting line. In fact, the model accuracy is very good ($\mathrm{NMSE} = 0.10$), with only a slight positive bias toward too low values ($\mathrm{FB} = 0.12$) and only few outliers present ($\mathrm{FAC2} = 0.84$). The decrease in model accuracy with increasing grid spacing is evident in the scatter plots, as modeled values tend to be too low for high concentrations measured and vice versa. This results in a steady increase in $\mathrm{NMSE} = 0.25$ for the $10\,\mathrm{m}$ grid spacing and $\mathrm{NMSE} = 1.35$ for the $20\,\mathrm{m}$ grid spacing. Since FB is more sensitive to deviations at the upper end of the logarithmic scale, the smearing also results in an increasingly positive bias ($\mathrm{FB} = 0.17$ and $\mathrm{FB} = 0.65$). The trend is, however, reversed at the even coarser resolution of $40\,\mathrm{m}$, resulting in an improvement of model results for this particular source. Finally, the $80\,\mathrm{m}$ case shows a large negative bias ($\mathrm{FB} = -0.57$) and a value of $\mathrm{FAC2} = 0.32$ at the verge of acceptance. Table A1 summarizes the statistical parameters for all other simulated cases. From the sensitivity results in Tab. A1, it can be concluded that the quality of model results generally declines with decreased grid resolution (mostly evident in NMSE value), but not to an extend to compromise model reliability at up to $40\,\mathrm{m}$ resolution, where buildings are only represented diffusely. Only one source located within a courtyard was problematic to model with $40\,\mathrm{m}$ grid spacing, as the resulting plume was not well contained within. We attributed this to the difficulty in representing diagonally orientated building walls as impermeable as they are at such coarse resolutions.
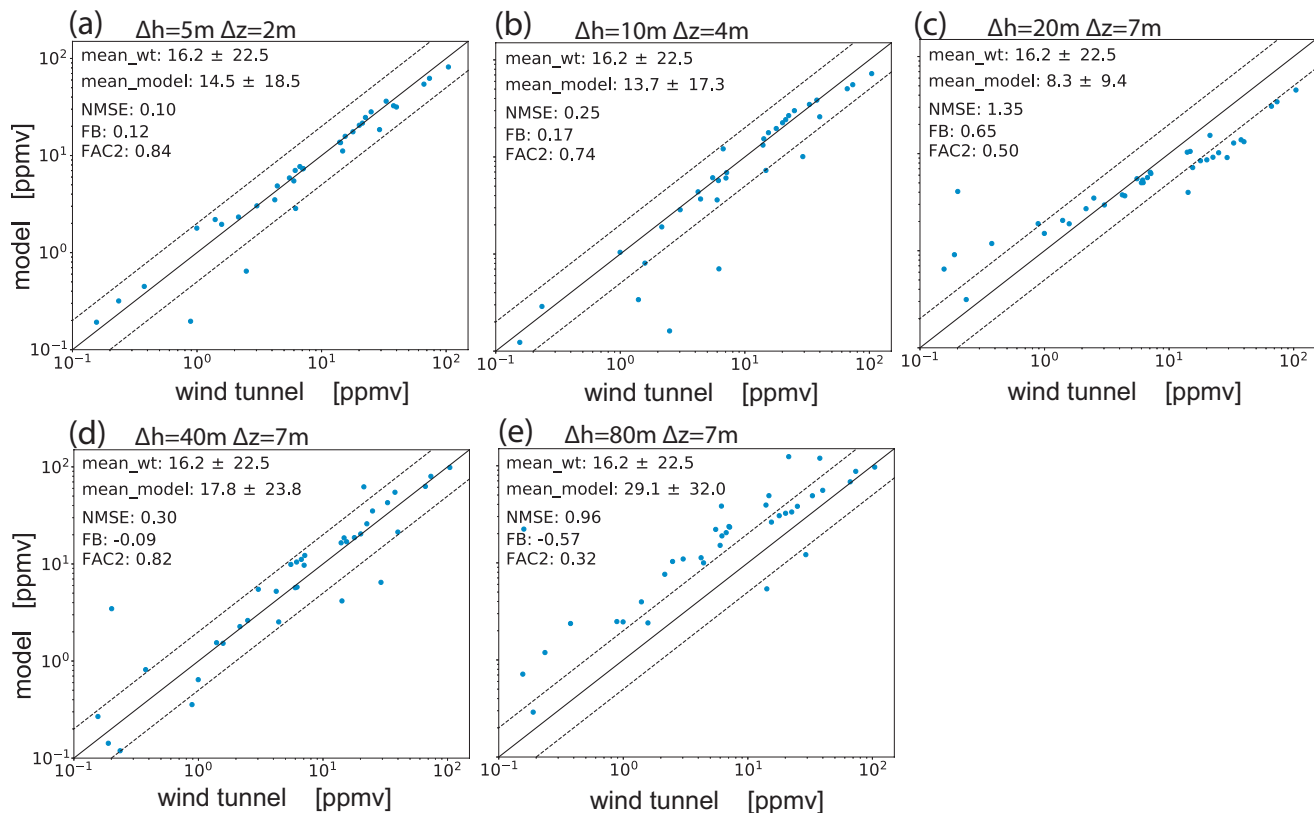
Figure 15. Scatter plots of measured vs. modeled concentrations for source S2 and different grid resolutions. The dotted lines confine the region within a factor of two of measurements.

Using the $80\,\mathrm{m}$ grid, the model still performs acceptable for some of the sources. Figure 16 shows scatter plots of modeled and observed data collected from all simulated cases with a given resolution. The according statistical results are again summarized in Table A2. It is shown that the reliability of model results is not very sensitive to the grid spacing down to $40\,\mathrm{m}$ resolution. For example, FAC2 decreases from a value of 0.8 at $5\,\mathrm{m}$ to 0.61 at $40\,\mathrm{m}$ grid spacing. Conversely, NMSE increases from 0.54 to 2.75, which is still well within the acceptable range. The average fractional bias is below $\mathrm{FB} = 0.2$ for all simulations, except for the $80\,\mathrm{m}$ case, where it is much larger ($\mathrm{FB} = -0.35$). In this regard, the increased vertical mixing length showed to be an effective tuning option in combination with the $40\,\mathrm{m}$ resolution to keep the bias comparatively low ($\mathrm{FB} = -0.17$). It has to be kept in mind that this test uses isolated point sources. When applied to a more realistic scenario with traffic emissions, modeled for example by emission lines, it can be hypothesized that scattering of the data would be of less a concern, because the pollution is more widely distributed horizontally. Therefore, ultimately, the most important reliability measure in our view is the FB value, as it is a proxy whether the model will under- or overestimate air pollution. In this regard, the $80\,\mathrm{m}$ resolution
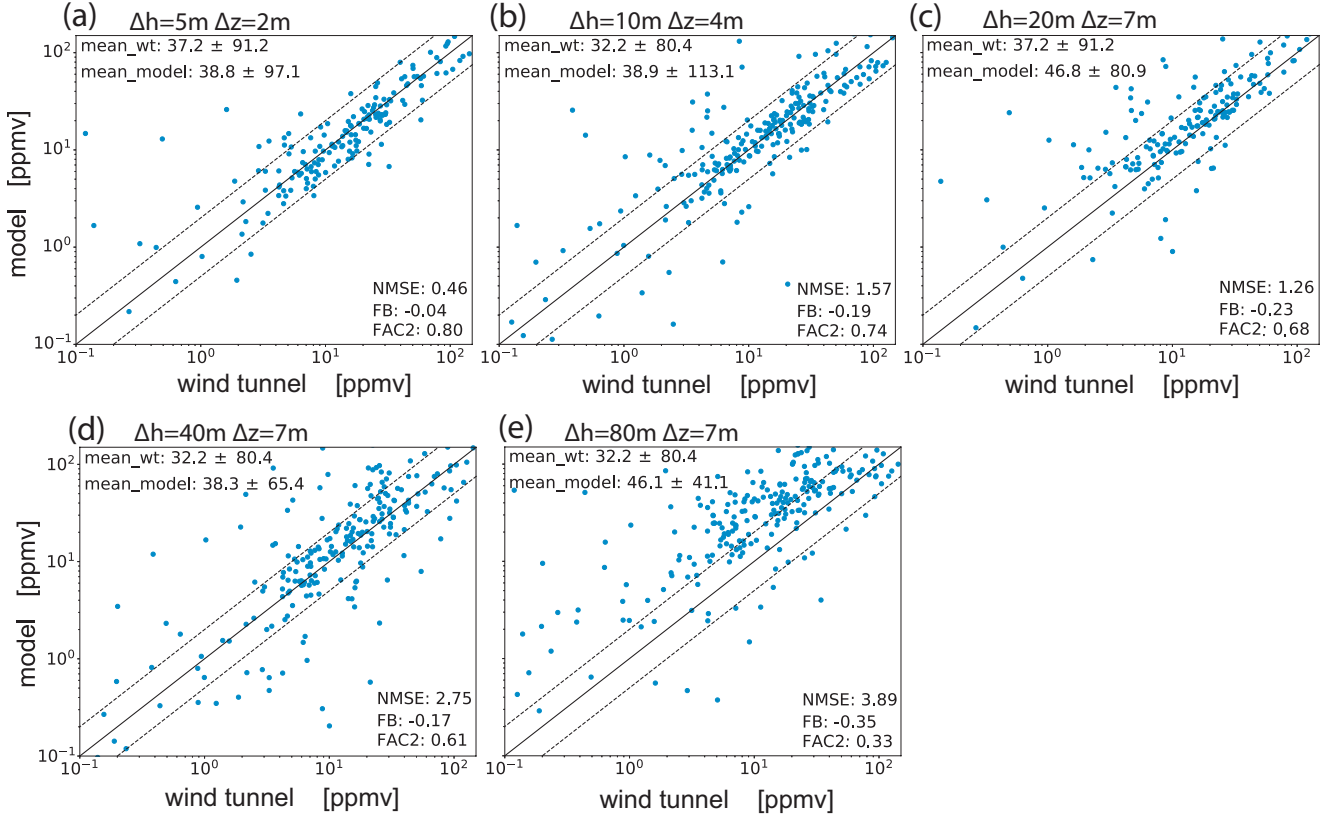
Figure 16. Scatter plots of measured vs. modeled concentrations combined for all sources and different grid resolutions. The dotted lines confine the region within a factor of two of measurements.

is the least accurate, and it is currently not aspired without the use of a more sophisticated mixing parameterization. Finally, when comparing all model results with only those from the $180°$-approach-flow tests using the same parameter configuration, it was found that these latter cases could be modeled not significantly less accurately. This is largely attributed to the simplicity of the model, as it requires little parameter tuning.

5   Summary

In this paper, the new large-eddy simulation-based model CAIRDIO for urban dispersion simulations was presented. The model uses diffuse obstacle boundaries in the framework of a finite volume discretization to represent building walls at a wide range of spatial resolutions. Diffuse obstacle boundaries allow for a consistent implementation of buildings in the model code, as they are essentially described by a scalar field for the volume-scaling factors and a vector field for the area-scaling

620 factors. Using these fields to discretize the differential operators, boundary conditions are incorporated automatically and the governing equations are solved for the entire computation grid without the need to distinguish different types of grid cells. This allows for a straightforward and vectorized implementation of spatial operations. The inherent option for under-resolved diffuse buildings enables the model to be applied at marginal grid resolutions inaccessible for conventional Cartesian-grid models. The computational savings can be invested in larger domains to model whole city areas and its surroundings. The large-

625 scale influence of orographic terrain can be adequately represented by curve-linear grids. To benefit from modern hardware architecture, the model is parallelized using a 2D-domain decomposition method, which is sufficient for the expected large grid-aspect ratios of typical boundary-layer applications. The numerical schemes presented are approved and efficient choices. Linear upwind schemes of selectable order of accuracy with optional limiting are used for advection, a static Smagorinsky model for subgrid turbulence modeling, and multi-stage higher-order time methods for model integration. The coupling with

630 the mesoscale meteorology can be obtained through different forcing methods by using data from a mesoscale host model. A simple numerical test, consisting of a test-tracer being advected by a potential-flow through an obstacle field, demonstrated the robustness of the obstacle discretization up to grid spacings where the resolution capability of the numerical schemes start to interfere. The results of the rising thermal experiment to test the dynamic core are plausible and similar to those presented in the original studies. Finally, we evaluated the model with data from the "Michelstadt" wind-tunnel experiment. In this study,

635 the model reproduced reliably the complex wind fields and embedded tracer dispersion. For the latter application, this was also true using spatial resolutions beyond $20\,\mathrm{m}$, at which buildings can only be represented as increasingly diffuse features. The sensitivity study researching grid-spacing of the dispersion test showed promising results for a future study with more realistic emission distributions and real-sized cities. In near future, also the coupling with mesoscale meteorology will be addressed. From previous and accompanying air-quality studies, simulations with the regional CTM COSMO-MUSCAT are available for

640 different German cities, including Berlin and Leipzig, for which also comprehensive measurement data are available for model evaluation. In this framework, a promising application could be a more comprehensive and holistic model evaluation with field data, as mobile measurements are available for the city of Leipzig in addition to operational air monitoring. Potential model improvements worth to be addressed in the future are the parameterization of air pollution sinks and the implementation of a simple urban atmospheric chemistry. Alternatively, it would be interesting to include diffuse obstacle boundaries in the aerosol-

645 chemistry code MUSCAT, and to investigate whether there is a benefit for the application on urban air pollution modeling.

Code and data availability. The source code of CAIRDIO model version 1.0, utilities for pre- and postprocessing, as well as evaluation data are accessible in release under the license GPL v3 and later at https://doi.org/10.5281/zenodo.4159497 (Weger et al., 2020).

Author contributions. Michael Weger contributed in model development, implementation, evaluation, and paper writing. Oswald Knoth and Bernd Heinold assisted in model development, paper writing and proof reading.

650   *Competing interests.*   All authors declare that they have no competing interests.

Acknowledgements.   Wind tunnel data of the "Michelstadt" case were kindly provided by Bernd Leitl and Frank Harms from the Meteorological Institute, University of Hamburg, Germany. The authors gratefully acknowledge the Centre for Information Service and High Performance Computing (Zentrum für Informationsdienste und Hochleistungsrechnen, ZIH) TU Dresden for computing time and their service.

Appendix A

source (Δh)	mean [ppmv]	wt mean [ppmv]	model NMSE	FB	FAC2	180° approach flow					
0° approach flow						S5* (5 m)	42.4 ±131.9	49.4 ±154.7	0.49	-0.15	0.76
						S5* (10 m)	42.4 ±131.9	69.8 ±227.2	3.19	-0.49	0.71
						S5* (20 m)	42.4 ±131.9	45.9 ±118.8	0.18	-0.08	0.68
						S5* (40 m)	42.4 ±131.9	43.8 ±87.2	2.41	-0.03	0.65
S2 (5 m)	16.2 ±22.5	14.5 ±18.5	0.10	0.12	0.84	S5* (80 m)	42.4 ±131.9	48.3 ±55.9	**7.10**	-0.13	0.50
S2 (10 m)	16.2 ±22.5	13.7 ±17.3	0.25	0.17	0.74	S6* (5 m)	54.1 ±126.8	57.5 ±127.7	0.21	-0.06	0.78
S2 (20 m)	16.2 ±22.5	8.3 ±9.4	1.35	0.65	0.50	S6* (10 m)	54.1 ±126.8	58.5 ±155.2	0.47	-0.08	0.89
S2 (40 m)	16.2 ±22.5	17.8 ±23.8	0.30	-0.09	0.82	S6* (20 m)	54.1 ±126.8	59.0 ±136.6	1.32	-0.09	0.78
S2 (80 m)	16.2 ±22.5	29.1 ±32.0	0.96	-0.57	0.32	S6* (40 m)	54.1 ±126.8	53.3 ±94.1	2.91	0.01	0.76
S4 (5 m)	4.0 ±3.6	4.5 ±4.0	0.18	-0.14	0.71	S6* (80 m)	54.1 ±126.8	41.5 ±33.6	**6.19**	0.26	0.54
S4 (10 m)	4.0 ±3.6	4.4 ±3.6	0.07	-0.11	1.00	S7* (5 m)	37.8 ±63.9	37.0 ±70.8	0.39	0.02	0.86
S4 (20 m)	4.0 ±3.6	5.2 ±4.7	0.50	-0.26	0.43	S7* (10 m)	37.8 ±63.9	44.1 ±72.3	0.58	-0.16	0.79
S4 (40 m)	4.0 ±3.6	6.3 ±5.8	0.60	-0.46	0.57	S7* (20 m)	37.8 ±63.9	53.8 ±89.9	1.76	-0.35	0.63
S4 (80 m)	4.0 ±3.6	18.9 ±12.1	1.46	**-1.30**	**0.0**	S7* (40 m)	37.8 ±63.9	43.3 ±60.0	2.20	-0.14	0.55
S5 (5 m)	25.7 ±40.8	27.1 ±72.8	2.58	-0.05	0.48	S7* (80 m)	37.8 ±63.9	49.9 ±33.8	1.66	-0.28	0.33
S5 (10 m)	25.7 ±40.8	30.1 ±43.0	1.02	-0.16	0.52	S8* (5 m)	11.4 ±7.8	9.6 ±7.1	0.05	0.18	0.88
S5 (20 m)	25.7 ±40.8	38.5 ±57.9	2.11	-0.40	0.67	S8* (10 m)	11.4 ±7.8	13.5 ±14.1	1.06	-0.16	0.55
S5 (40 m)	25.7 ±40.8	41.1 ±63.8	1.52	-0.48	0.62	S8* (20 m)	11.4 ±7.8	10.6 ±11.2	1.19	0.08	0.42
S5 (80 m)	25.7 ±40.8	55.6 ±56.2	2.30	**-0.73**	**0.24**	S8* (40 m)	11.4 ±7.8	31.3 ±40.8	4.06	**-0.93**	0.30
						S8* (80 m)	11.4 ±7.8	58.4 ±40.3	1.74	**-1.35**	**0.03**

Table A1. Statistical results of all dispersion simulations performed in the "Michelstadt" wind tunnel simulation study. The cases superscripted with a star were carried out using the reverse approach-flow direction of 180°. Missed acceptance criteria are highlighted in bold font.

source (Δh)	mean [ppmv]	wt mean [ppmv]	model NMSE	FB	FAC2
All (5 m)	32.2 ±80.4	33.0 ±89.5	0.54	-0.02	0.80
All (10 m)	32.2 ±80.4	38.9 ±113.1	1.57	-0.19	0.74
All (20 m)	32.2 ±80.4	38.5 ±89.5	1.72	-0.18	0.61
All (40 m)	32.2 ±80.4	38.3 ±65.4	2.75	-0.17	0.61
All (80 m)	32.2 ±80.4	46.1 ±41.1	3.89	-0.35	0.33
All* (5 m)	37.2 ±91.2	38.5 ±100.2	0.40	-0.03	0.83
All* (10 m)	37.2 ±91.2	46.2 ±129.5	1.48	-0.22	0.75
All* (20 m)	37.2 ±91.2	45.8 ±100.7	1.50	-0.21	0.63
All* (40 m)	37.2 ±91.2	43.2 ±71.2	2.72	-0.15	0.56
All* (80 m)	37.2 ±91.2	49.4 ±40.1	3.97	-0.28	0.35

Table A2. Statistical results derived from the combined data of all simulated sources at the given spatial resolution. The cases superscripted with a star are the combined results of the 180 °approach-flow cases only.

References

Appel, K. W., Napelenok, S. L., Foley, K. M., Pye, H. O. T., Hogrefe, C., Luecken, D. J., Bash, J. O., Roselle, S. J., Pleim, J. E., Foroutan, H., Hutzell, W. T., Pouliot, G. A., Sarwar, G., Fahey, K. M., Gantt, B., Gilliam, R. C., Heath, N. K., Kang, D., Mathur, R., Schwede, D. B., Spero, T. L., Wong, D. C., and Young, J. O.: Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1, Geoscientific Model Development, 10, 1703–1732, doi:10.5194/gmd-10-1703-2017, 2017.

Baik, J.-J., Park, S.-B., and Kim, J.-J.: Urban flow and dispersion simulation using a CFD model coupled to a mesoscale model, Journal of Applied Meteorology and Climatology, 48, 1667–1681, doi:10.1175/2009JAMC2066.1, 2009.

Baumann-Stanzer, K., Andronopoulos, S., Armand, P., Berbekar, E., Efthimiou, G., Fuka, V., Gariazzo, C., Gašparac, G., Harms, F., Hellsten, A., Jurcacova, K., Petrov, A., Rákai, A., Stenzel, S., Tavares, R., Tinarelli, G., and Trini Castelli, S.: COST ES1006 Model evaluation case studies: Approach and results, retrieved from: http://www.elizas.eu/images/Documents/Model%20Evaluation%20Case%20Studies_web.pdf, 2015.

Benavides, J., Snyder, M., Guevara, M., Soret, A., Pérez García-Pando, C., Amato, F., Querol, X., and Jorba, O.: CALIOPE-Urban v1.0: coupling R-LINE with a mesoscale air quality modelling system for urban air quality forecasts over Barcelona city (Spain), Geoscientific Model Development, 12, 2811–2835, doi:10.5194/gmd-12-2811-2019, 2019.

Birmili, W., Rehn, J., Vogel, A., Boehlke, C., Weber, K., and Rasch, F.: Micro-scale variability of urban particle number and mass concentrations in Leipzig, Germany, Meteorologische Zeitschrift, 22, 155–165, doi = 10.1127/0941-2948/2013/0394, 2013.

Brandt, A. and Livne, O. E.: Multigrid Techniques, Society for Industrial and Applied Mathematics, doi:10.1137/1.9781611970753, 2011.

Brown, M.: QUIC: A fast, high-resolution 3D building-aware urban transport and dispersion modeling system, AWMA Environmental Manager April issue, pp. 28–31, doi:10.2172/1134791, 2014.

Bröker, O. and Grote, M.: Sparse approximate inverse smoothers for geometric and algebraic multigrid, Applied Numerical Mathematics, 675   41, 61–80, doi:10.1016/S0168-9274(01)00110-6, 2001.

Calhoun, D. and LeVeque, R. J.: A Cartesian grid finite-volume method for the advection-diffusion equation in irregular geometries, Journal of Computational Physics, 157, 143 – 180, doi:10.1006/jcph.1999.6369, 2000.

Carlino, G., Pallavidino, L., Prandi, R., A., A., Matteucci, G. L., F., R., Bajardi, P., and L., B.: Micro-scale modelling of urban air quality to forecast NO2 critical levels in traffic hot-spots, 2016.

680   Chang, S. Y.: High Resolution Air Quality Modeling for Improved Characterization of Exposures and Health Risk to Traffic-Related Air Pollutants (TRAPs) in Urban Areas [doctoral-degree thesis], University of North Carolina, Chapel Hill, USA, https://cdr.lib.unc.edu/concern/dissertations/q811kk98b (last access: 08.10.2020), 2016.

Chorin, A.: Numerical solution of the Navier-Stokes equations, Mathematics of Computations, 22, 745–762, doi:10.1090/S0025-5718-1968-0242392-2, 1968.

685   Croitoru, C. and Nastase, I.: A state of the art regarding urban air quality prediction models, E3S Web of Conferences, 32, 01 010, doi:10.1051/e3sconf/20183201010, 2018.

Deardorff, J. W.: A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers, Journal of Fluid Mechanics, 41, 453–480, doi:10.1017/S0022112070000691, 1970.

Doms, G., Förstner, J., Heise, E., Herzog, H., Mironov, D., Raschendorfer, M., Reinhardt, T., Ritter, B., Schrodin, R., Schulz, J.-P., and Vogel, 690   G.: A description of the nonhydrostatic regional COSMO model. Part II: Physical parameterization. Deutscher Wetterdienst, Offenbach, http://www.cosmo-model.org (last access: 28.12.2020), 2013.

Drew, D.: Mathematical modeling of two-phase flow, Annual Reviews on Fluid Mechanics, 15, 261–291, doi:10.1146/annurev.fl.15.010183.001401, 1983.

Efstathiou, G. A., Beare, R. J., Osborne, S., and Lock, A. P.: Grey zone simulations of the morning convective boundary layer development, 695   Journal of Geophysical Research: Atmospheres, 121, 4769–4782, doi:10.1002/2016JD024860, 2016.

Fallah-Shorshani, M., Shekarrizfard, M., and Hatzopoulou, M.: Integrating a street-canyon model with a regional Gaussian dispersion model for improved characterisation of near-road air pollution, Atmospheric Environment, 153, 21–31, doi:10.1016/j.atmosenv.2017.01.006, 2017.

Fernández, G., Rezzano Tizze, N., D'Angelo, M., and Mendina, M.: Numerical simulation of different pollution sources in an urban envi-700   ronment, E3S Web of Conferences, 128, 10 005, doi:10.1051/e3sconf/201912810005, 2019.

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, Geoscientific Model Development, 8, 975–1003, doi:10.5194/gmd-8-975-2015, 2015.

705   Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmospheric Environment, 39, 6957–6975, doi:10.1016/j.atmosenv.2005.04.027, 2005.

Haga, C. J. B.: Numerical methods for basin-scale poroelastic modelling [doctoral-degree thesis], University of Oslo, Oslo, Norway, http://urn.nb.no/URN:NBN:no-29117 (last access: 08.10.2020), 2011.

Hanna, S. and Chang, J.: Acceptance criteria for urban dispersion model evaluation, Meteorology and Atmospheric Physics, 116, 133–146, doi:10.1007/s00703-011-0177-1, 2012.

Hanna, S. R., Brown, M. J., Camelli, F. E., Chan, S. T., Coirier, W. J., Hansen, O. R., Huber, A. H., Kim, S., and Reynolds, R. M.: Detailed simulations of atmospheric flow and dispersion in urban downtown areas by computational fluid dynamics (CFD) models - an application of five CFD models to Manhattan, Bulletin of the American Meteorological Society, 87, 1713–1726, doi:10.1175/BAMS-87-12-1713, 2006.

Harrison, R.: Urban atmospheric chemistry: a very special case for study, npj Climate and Atmospheric Science, 1, 20 175–20 180, doi:10.1038/s41612-017-0010-8, 2018.

Haupt, S. E., Kosovic, B., Shaw, W., Berg, L. K., Churchfield, M., Cline, J., Draxl, C., Ennis, B., Koo, E., Kotamarthi, R., Mazzaro, L., Mirocha, J., Moriarty, P., Muñoz-Esparza, D., Quon, E., Rai, R. K., Robinson, M., and Sever, G.: On bridging a modeling scale gap: Mesoscale to microscale coupling for wind energy, Bulletin of the American Meteorological Society, 100, 2533–2550, doi:10.1175/BAMS-D-18-0033.1, 2019.

Hicken, J., Ham, F., Militzer, J., and Koksal, M.: A shift transformation for fully conservative methods: turbulence simulation on complex, unstructured grids, Journal of Computational Physics, 208, 704 – 734, doi:10.1016/j.jcp.2005.03.002, 2005.

Jensen, S. S., Ketzel, M., Becker, T., Christensen, J., Brandt, J., Plejdrup, M., Winther, M., Nielsen, O.-K., Hertel, O., and Ellermann, T.: High resolution multi-scale air quality modelling for all streets in Denmark, Transportation Research Part D: Transport and Environment, 52, 322 – 339, doi:10.1016/j.trd.2017.02.019, 2017.

Jähn, M., Knoth, O., König, M., and Vogelsberg, U.: ASAM v2.7: A compressible atmospheric model with a Cartesian cut cell approach, Geoscientific Model Development, 8, 317–340, doi:10.5194/gmd-8-317-2015, 2015.

Kadaverugu, R., Sharma, A., Matli, C., and Biniwale, R.: High resolution urban air quality modeling by coupling CFD and mesoscale models: a review, Asia-Pacific Journal of Atmospheric Sciences, 55, 539–556, doi:10.1007/s13143-019-00110-3, 2019.

Kanda, M., Inagaki, A., Miyamoto, T., Gryschka, M., and Raasch, S.: A new aerodynamic parametrization for real urban surfaces, Boundary-Layer Meteorology, 148, 357–377, doi:10.1007/s10546-013-9818-x, 2013.

Karam, M., Sutherland, J., Hansen, M., and Saad, T.: A Framework for Analyzing the Temporal Accuracy of Pressure Projection Methods, AIAA Aviation 2019 Forum, doi:10.2514/6.2019-3634, 2019.

Kemm, F., Gaburro, E., Thein, F., and Dumbser, M.: A simple diffuse interface approach for compressible flows around moving solids of arbitrary shape based on a reduced Baer-Nunziato model, Computers Fluids, 204, 104 536–104 561, doi:10.1016/j.compfluid.2020.104536, 2020.

Kim, Y., Sartelet, K., Raut, J.-C., and Chazette, P.: Influence of an urban canopy model and PBL schemes on vertical mixing for air quality modeling over Greater Paris, Atmospheric Environment, 107, 289 – 306, doi:https://doi.org/10.1016/j.atmosenv.2015.02.011, 2015.

Korhonen, A., Lehtomäki, H., Rumrich, I., Karvosenoja, N., Paunu, V., Kupiainen, K., Sofiev, M., Palamarchuk, Y., Kukkonen, J., Kangas, L., Karppinen, A., and Hänninen, O.: Influence of spatial resolution on population PM2.5 exposure and health impacts, Air Quality, Atmosphere Health, 12, 705–718, doi:10.1007/s11869-019-00690-z, 2019.

Kurppa, M., Roldin, P., Strömberg, J., Balling, A., Karttunen, S., Kuuluvainen, H., Niemi, J. V., Pirjola, L., Rönkkö, T., Timonen, H., Hellsten, A., and Järvi, L.: Sensitivity of spatial aerosol particle distributions to the boundary conditions in the PALM model system 6.0, Geoscientific Model Development Discussions, 2020, 1–33, doi:10.5194/gmd-2020-163, 2020.

Larsson, J., Lien, F. S., and Yee, E.: Conditional semicoarsening multigrid algorithm for the Poisson equation on anisotropic grids, Journal of Computational Physics, 208, 368–383, doi:10.1016/j.jcp.2005.02.020, 2005.

Lee, M., Leitl, B., and Patnaik, G.: Model and application-specific validation data for LES-based transport and diffusion models, Proceedings of the Eighth Conference on Coastal Atmospheric and Oceanic Prediction and Processes, Phoenix, Arizona, 2009.

Llorente, I. M. and Melson, N. D.: Behavior of plane relaxation methods as multigrid smoothers, Electronic Transactions on Numerical
750   Analysis, 10, 92–114, 2000.

Louis, J. A.: A parametric model of vertical eddy fluxes in the atmosphere, Boundary-Layer Meteorology, 17, 187–202, doi:10.1007/BF00117978, 1979.

Maronga, B., Gross, G., Raasch, S., Banzhaf, S., Forkel, R., Heldens, W., Kanani-Sühring, F., Matzarakis, A., Mauder, M., Pavlik, D., Pfafferott, J., Schubert, S., Seckmeyer, G., Sieker, H., and Winderlich, K.: Development of a new urban climate model based on the model
755   PALM–Project overview, planned work, and first achievements, Meteorologische Zeitschrift, 28, 105–119, doi:10.1127/metz/2019/0909, 2019.

Martilli, A., Clappier, A., and Rotach, M.: An urban surface exchange parameterisation for mesoscale models, Boundary-Layer Meteorology, 104, 261–304, doi:10.1023/A:1016099921195, 2002.

Miller, M. J. and Thorpe, A. J.: Radiation conditions for the lateral boundaries of limited-area numerical models, Quarterly Journal of the
760   Royal Meteorological Society, 107, 615–628, doi:10.1002/qj.49710745310, 1981.

Mittal, R. and Iaccarino, G.: Immersed boundary methods, Annual Review of Fluid Mechanics, 37, 239–261, doi:10.1146/annurev.fluid.37.061903.175743, 2005.

Mohr, M. and Wienands, R.: Cell-centred multigrid revisited, Computing and Visualization in Science, 7, 129–140, doi:10.1007/S00791-004-0137-0, 2004.

765   Noh, Y., C. W. H. S. and Raasch, S.: Improvement of the K-profile Model for the Planetary Boundary Layer based on Large Eddy Simulation Data, Boundary-Layer Meteorology, 107, 401–427, doi:10.1023/A:1022146015946, 2003.

Paas, B., Schmidt, T., Markova, S., Maras, I., Ziefle, M., and Schneider, C.: Small-scale variability of particulate matter and perception of air quality in an inner-city recreational area in Aachen, Germany, Meteorologische Zeitschrift, 25, 305–317, doi:10.1127/metz/2016/0704, 2016.

770   Rakai, A. and Gergely, K.: Microscale obstacle resolving air quality model evaluation with the Michelstadt case, The Scientific World Journal, 2013, 781 748, doi:10.1155/2013/781748, 2013.

Rieger, D., Bangert, M., Bischoff-Gauss, I., Förstner, J., Lundgren, K., Reinert, D., Schröter, J., Vogel, H., Zängl, G., Ruhnke, R., and Vogel, B.: ICON–ART 1.0 – a new online-coupled model system from the global to regional scale, Geoscientific Model Development, 8, 1659–1676, doi:10.5194/gmd-8-1659-2015, 2015.

775   Schatzmann, M., Leitl, B., Harms, F., and Hertwig, D.: Field data versus wind tunnel data: The art of validating urban flow and dispersion models, Proceedings of the 9th Asia-Pacific Conference on Wind Energy, Auckland, New Zealand, doi:10.17608/k6.auckland.5630923.v1, 2017.

Schubert, S., Grossman-Clarke, S., and Martilli, A.: A double-canyon radiation scheme for multi-layer urban canopy models, Boundary-Layer Meteorology, 145, 439–468, doi:10.1007/s10546-012-9728-3, 2012.

780   Schumann, U.: Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli, Journal of Computational Physics, 18, 376 – 404, doi:10.1016/0021-9991(75)90093-5, 1975.

Sedlacek, M.: Sparse approximate inverses for preconditioning, smoothing, and regularization [doctoral-degree thesis], Technical University of Munich, Munich, Germany, https://d-nb.info/1029819246/34 (last access: 08.10.2020), 2012.

Shu, C.-W.: Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws, Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, 1998.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, 1988.

Sweby, P. K.: High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws, SIAM Journal on Numerical Analysis, 21, 995–1011, doi:10.2307/2156939, 1984.

Tang, W.-p. and Wan, J.: Sparse approximate inverse smoother for multigrid, SIAM Journal on Matrix Analysis and Applications, 21, doi:10.1137/S0895479899339342, 2000.

Wang, J., Mao, J., Zhang, Y., Cheng, T., Yu, Q., Tan, J., and Ma, W.: Simulating the effects of urban parameterizations on the passage of a cold front during a pollution episode in megacity Shanghai, Atmosphere, 10, 79, doi:10.3390/atmos10020079, 2019.

Wicker, L. J. and Skamarock, W. C.: A time-splitting scheme for the elastic equations incorporating second-order Runge–Kutta time differencing, Monthly Weather Review, 126, 1992–1999, doi:10.1175/1520-0493(1998)126<1992:ATSSFT>2.0.CO;2, 1998.

Wolf, T., Pettersson, L. H., and Esau, I.: A very high-resolution assessment and modelling of urban air quality, Atmospheric Chemistry and Physics, 20, 625–647, doi:10.5194/acp-20-625-2020, 2020.

Wolke, R., Knoth, O., Hellmuth, O., Schröder, W., and Renner, E.: The parallel model system LM-MUSCAT for chemistry-transport simulations: Coupling scheme, parallelization and applications, 13, 363 – 369, doi:10.1016/S0927-5452(04)80048-0, 2004.

Wolke, R., Schröder, W., Schrödner, R., and Renner, E.: Influence of grid resolution and meteorological forcing on simulated European air quality: A sensitivity study with the modeling system COSMO-MUSCAT, Atmospheric Environment, 53, 110 – 130, doi:doi.org/10.1016/j.atmosenv.2012.02.085, 2012.

Wu, X.: Inflow turbulence generation methods, Annual Review of Fluid Mechanics, 49, 23–49, doi:10.1146/annurev-fluid-010816-060322, 2017.

Xie, Z. and Castro, I. P.: LES and RANS for turbulent flow over arrays of wall-mounted obstacles, Flow, Turbulence and Combustion, 76, 291, doi:doi.org/10.1007/s10494-006-9018-6, 2006.

Yavneh, I.: On Red-Black SOR Smoothing in Multigrid, SIAM Journal on Scientific Computing, 17, 180–192, doi:10.1137/0917013, 1996.