We thank reviewer #1 for their constructive comments which have made the paper more accessible and rigorous, especially in regards to the validation practices. The main changes made in this revision are:

- Further explanation on the technical elements of the paper;
- Detailed validation metrics for the DAPPLE case;
- Substantial improvement of the East-side demo:
  - Reran the experiment with the use of a driver simulation rather than a periodic simulation in order to prevent heating up of the air;
  - This allows the simulation time to be 8 hours rather than one hour.
- Fluxes are now presented conforming to atmospheric convention.
- A thorough check of all simulation cases in the repository, with updated input and run-options.

We respond to the specific points made by the reviewer below.

**Comment 1:** P1, l17: "... since these structures channel high-momentum air downward to ground level". This keeps popping up in the context of urban boundary layer (UBL) flows, yet I still have not seen proper evidence for it, let alone an explanation what is the physical mechanisms for high-rise buildings creating high-momentum flow structures towards the ground. The downwash in front of the building or the wake turbulence behind it should not warrant such a simplistic characterization of wind conditions near the ground. I checked (Blocken et al. 2015) and it appears that the introduction, at least, provides only anecdotal evidence. If the evidence lies in some of the references within Blocken 2015, please point to it directly or rephrase the original statement so that the readers will not continue to spread this claim as a fact.

**Authors reply:** We thank the reviewer for pointing this out. By no means we want to trivialise the flow structures in complex urban environments. The flow patterns at ground level of high-rise buildings will always be determined by the surrounding geometry, topography and meteorological conditions, indeed this is the reason why we bring up high-rise buildings as example for the use of CFD studies.
The study of wind conditions at pedestrian level and observation that high-rise buildings generally increase mean wind speed and gustiness around the corners of tall buildings goes back to the 1960s, with a number of important publications following the International Conference on Wind Effects on Buildings and Structures in Heathrow, London (e.g. Lawson and Penwarden, 1975, Isymov and Davenport, 1975).

As the wind speed in the boundary layer increases with height, tall obstacles tend to intercept the stronger winds high up and redirect them down to ground level, causing accelerated flow zones in front and around the corners of the building (ASCE, 2004). The amplification of ground level wind speed is often measured in terms of an amplification factor, that compares the existing geometry with/without the tall building (Blocken et al., 2003). While improved building design can alleviate some of these effects (for example by using a podium shaped base or a floor plan area that decreases with height), the building height relative to the height of the surrounding buildings is always a decisive factor (ASCE, 2011).

Publications by the American Society of Civil Engineers (ASCE 2004, 2011) were added as references since these provide an excellent overview of the most important works, research progress and detailed explanation of the physical mechanisms influencing pedestrian level wind comfort. We also added “can” to the sentence to account for the fact that ground level flow always depends on the individual geometries.

**Comment 2:** P5, l122: The pressure field is solved via FFT (which, by default, requires horizontally periodic boundary conditions), but some of the lateral boundary conditions described in Section 2.3.2 and Figure 1 are not horizontally periodic. Please specify how the pressure field is solved in those bc configurations.

**Authors reply:** When using inflow/outflow boundary conditions, the y direction is periodic so is still solved using FFT, but x and z are solved using cyclic reduction. The inflow boundary condition for pressure is Neumann and the outflow is Dirichlet. The bottom and top are both Neumann boundary conditions. We have clarified this in section ‘2.2 Immersed boundary method’.

**Comment 3:** P7, l173: Please clarify are the dimensions (Lx x Ly x Lz)? Is the order of grid resolutions now correct? It’s best to document with clarity so you don’t leave the reader guessing.

**Authors reply:** To avoid any confusion we now clearly state the size of the model domain in number of cells & size of cells and total extent in metres

**Comment 4:** P7, l183: You claim to present quantitative agreement, but do not provide any quantity. Eye-balling the temperature contours from Figure 2 is not a satisfying validation strategy. Please provide a transparent validation metric for the comparison which other researchers in the field can compare their results to or reduce the comparison to a qualitative one.

**Authors reply:** We thank the reviewer for noticing this – we intended to state the agreement was qualitative. We reproduce figures of Cai without having the original data and cannot carry out a quantitative comparison. The text in ‘2.4.1 Wall-functions for momentum and temperature’ was changed to make clear that the comparison is qualitative.

**Comment 5:** P12, l295: Section 3.1 needs to be improved. The description of how surfaces within the discretized model are handled as facets is not clear. It is very difficult to understand how the LES grid relates to the facets used by the UEB model. Given the IBM implementation, the surfaces of all building blocks are made up of potentially a large number of cell faces, so are they grouped to form facets? It appears so but it’s hard to get a clear picture of the method. What happens when the buildings are diagonal with respect
to the LES grid? The Figure 4 provides some help but not much. Figure 11a is much better so you should make better use of it. Now it comes at the very end of the section, which is a struggle to reach with the current description.

**Authors reply**: The process is to generate ‘blocks’ first. These blocks represent e.g. entire buildings and have to conform to the LES grid. In case a block does not touch any other blocks, the facets would correspond to the 5 faces of said block (the face touching the ground is ignored). However, if two blocks touch, one has to subdivide one or both blocks to ensure that faces are entirely exposed to the fluid (external) or entirely within the touching blocks (internal).

Diagonal buildings cannot be represented directly as the blocks have to conform to the grid. These are represented by stair-case-like patterns, which indeed requires many blocks and therefore many facets. We will be working to remove this limitation in one of the next versions of uDALES.

We extended the introductory part of ‘3.1 Urban facets’ to clarify the process and terminology.

**Comment 6**: P22, l537: Section 4.1 on Validation. Again, the standards for performing validation against measurements are not sufficient. Comparing vertical velocity and Reynolds stress profiles at one location visually does not constitute validation. And the scalar concentration along a line juxtaposed with measurements alone does not either. Validation metrics are derived for this purpose. What if the authors choose to further improve the numerics of uDALES and perform the same DAPPLE test case again? Will they judge the level of improvement by eye-balling the curves? Of course not. It is true that Xie and Castro (2009) do not carry out validation either, but their paper appears to focus more on establishing sufficient LES modelling criteria for urban flows.

**Authors reply**: We appreciate the comments on the shortcomings in terms of validation. We now define clear validation metrics: the root-mean-square error (RMSE), fractional bias (FB) and factor of two (FAC2). We have calculated these for the mean profiles of velocity magnitude for the location shown (Figure 9) and 8 other locations, as well as for the mean scalar concentration profile (Figure 10). The validation metrics are shown in Table 2. The RMSEs range from 0.06 to 0.15 (0.04 for scalar concentration). The FBs are typically slightly below zero, indicating that the simulation had a small tendency to under-predict the velocities. The velocities generally yield high FAC2 scores, with some locations where all simulated velocities lie within a factor two of the experiment data (i.e., FAC2 = 1).

These results are now reported in section ‘4.1.2 Results’ (note the change in notation: measurement point R -> P1; axis for scalar concentration x_r -> x_s). The section ‘4.1 Validation’ and ‘4.1.1. Simulation set-up’ were also further expanded to include a more detailed description of the study case and wind-tunnel experiment.

**Comment 7**: One of their conclusions is that the resolution requirement is ~1 m for this case. Here, validation is attempted and claimed, yet 2 m resolution is used without justification. Thus, please improve the validation section by providing appropriate validation metrics and justifications for the numerical setup.

**Authors reply**: Indeed Xie and Castro say that ~1m resolution is sufficient for the problem. However, it is not necessarily a requirement. We analysed the subgrid fluxes and they do not exceed 4.5% of the total flux anywhere but in the lowermost cells indicating that the resolution is sufficient to resolve the majority of the energetic turbulent scales of the urban flow field. As always there is a computational cost associated with finer resolutions and we found 2m to be adequate for the problem.

We now mention that the subgrid-scale fluxes are small in ‘4.1.1 Simulation set-up’

**Comment 8**: The wordings relating to the validation in Concluding remarks should also be
made more scientific so that we do not have to engage in discussions trying to determine the degree of “soundness” in agreements. Verification is a different process and I have no objections with Section 4.2.

**Authors reply:** We have updated the validation against measurements in section 4.1 and reflect this now in section 5. The text in ‘5 Concluding remarks’ has been changed.

**Comment 9:** P17, l432: Only a suggestion: The lower case $\sigma_{L,i}$ does not go well with upper case L in the subscript. If this is not a universally accepted nomenclature, consider using another greek symbol.

**Authors reply:** We used the fact that we assumed ‘emissivity = absorptivity’ in the longwave and eliminated the symbol altogether. In addition, the sign of some of the fluxes in the surface energy balance have been adjusted to better represent the normal conventions in the literature.

**Comment 10:** P28, l617: “… in this dissertation” This manuscript is likely part of a dissertation but perhaps the wording should be changed here.

**Authors reply:** Thanks for pointing this out. Changed.