Review of

Validation of the Dynamic Core of the PALM Model System 6.0 in Urban Environments: LES and Wind-tunnel Experiments

By Gronemeier et al

The manuscript compares a modelled urban flow (obtained with a new PALM model version) to wind-tunnel measurements (representing a part of the city of Hamburg). The results are solid, but I am not sure if the employed double-blind strategy helps more than it distracts from the ultimate goal of validating the code and demonstrate the model capabilities. Moreover, I am not sure if "validation of the dynamic core" in the manuscript title oversells the results as only one case study with a single simulation is performed.

Major comments

1. In the beginning, I interpreted the title of the manuscript in a way that the dynamic core of the new PALM model version is validated. But I do think this is not achieved by limiting the validation to a single case study. The first sentence of the abstract reflects much better what was achieved here. Please adapt the title accordingly or expand the study to a comprehensive dynamical core validation.
2. Generally, I appreciate the double-blind test strategy. Nevertheless, at some point it would have been nice to see if the model-observation agreement is better when other values of the discussed parameters would have been chosen. In the present evaluation, any disagreement can be explained by a "wrong" setting of some parameter. In the end, the reader does not know, if the model could reproduce the observation if it had used better values of "external" parameters. In the end, it may turn into a philosophical question, how much a priori knowledge you allow to be included in your model setup? When I first read the manuscript, I thought "double-blind" is the optimal approach as it guarantees an unbiased and honest comparison. After thinking more about it, I felt that the double-blind test strategy leaves me unsatisfied as you often blame initial inconsistencies for the observed discrepancies. But what conclusions regarding the validation of the model remain at the end of the day (Q1)? In my impression, the manuscript gives more an answer to the question, how well can I simulate a specific scenario and what uncertainties remain having in mind that I do not know precisely which values to choose for some parameters in the beginning (Q2). Imagine the situation: You are asked to perform a PALM simulation for a different city where you obtain topography data but no measurement stations are installed for comparison. Then the present study helps assessing the faithfulness of your simulation results. It may help, if the manuscript makes clearer what part of the analysis relates more to answering question Q1 or Q2.
   For example, lines 190-194: For the purpose of validation, wouldn't it be better to correct the topography data, eliminate this flaw and redo the simulation? In the present state, this flaw distracts more than it helps, at least when you pursue a model validation.
   Moreover, I would have appreciated to see the sensitivity to the grid resolution (which I suppose can be varied independently of the resolution of the topography data).
   The paper is not long, one could also think of adding the nesting feature to the present manuscript.

3. It is only marginally explained how the scaling works between the wind-tunnel scale and the real-world scale. In the simulation framework, I believe, it would have easily been possible to simulate the flow at the wind-tunnel scale? Why did you choose to simulate it on the real-world scale even though this can introduce ambiguities? Scaling relations for all displayed or discussed quantities should be derived, e.g. in a separate section in the appendix or in a table if it can be done in a compact way.
4. In general, my impression is that the comparison in Section 3.2 is not very elaborate and discusses only few properties. At least, the description of the results is rather subjective and uses often simple phrases.

Minor comments

In general, the figure captions sometimes lack crucial information to understand what it is shown or to make it easier for the reader.

1. I suppose the approaching flow displayed in Fig. 2 is part of the setup and not of the validation. Isn't it possible to prescribe in the model a flow that is closer to measured profile?
2. Figure 5: Which value range does the white colour represent? If the street level is at z=5m, the background colour should be blueish, shouldn't it.
3. Along which direction do streak-like structures appear? Why does it help to shift the flow in cross-stream direction? How does this prevent the formation of streak-like structures?
4. Line 138: I do not understand what you want to describe. "Profiles of nine different quantities…"(?) "The profiles were recorded with a time resolution of 2Hz, i.e. 9 evaluation during one model time step." (?)
5. Issue with staggered grid: Why not additionally evaluate the model flow also at z=3.5m and 10.5m? What's the wind speed deficit?
6. Fig. 8 and 9: You have to explain the meaning of the arrows (in particular, size and why you show differently sized arrows at each station).
7. The caption of Fig. 6 should tell me that the plot uses to different scales on the horizontal axis. Furthermore, you may replace the blue colour by some colour that can be more easily distinguished from the black line.

Technical comments

1. Abstract, line 1: "We **demonstrate"**
2. Line 2: I believe scenario fits better than situation: "The studied **scenario"**
3. In English language there is a difference between "which" and "that" many are not aware of. https://www.diffen.com/difference/That_vs_Which. Please go trough the whole manuscript. Line 35 is one such example.
4. Line 69: length -> period
5. Line 76: ". A detailed description" -> ", which"
6. Line 82: I guess the model DOMAIN was rotated?
7. Line 106: POTENTIAL temperature
8. Line 143: move the variable name in front of the given values (e.g. E=0.8865).
9. Line 153: "magnitude of rotation": Is this the vorticity magnitude?
10. Line 157: remove "to"
11. The paragraph from line 203 on should start with information that holds for all three Figures 10-12. In the next paragraph you should turn the attention the specific results of station 7.
12. Please reformulate line 228.