Reviewer 1

This paper addresses the role of the driving model in an urban-area LES model that covers an 8 km square. In this study multiple configurations of the driving model with different biases were used to provide initial and boundary conditions to PALM. the main finding was that the biases do propagate into the PALM results, but PALM can also alleviate some of them. The main verification was wind and temperature for soundings and the surface, but the results were also compared to the spread of the driving model. Many results were presented in tables and supplementary plots.

Overall the paper is worth publishing because this type of downscaling is becoming a common tool for urban modeling that addresses some important problems affected by local details in built environments. It is useful to emphasize to what extent the driving model matters, and this is quantified by the methods adopted here. I will only have some minor comments to address, but I think some may be useful additional thoughts for consideration.

Minor Comments

1. Table 1. It would have been useful to put the month with the episode number here.

Answer: Thank you for the suggestion, a new row has been added to the table with the dates of the episodes. Please see Table 1.

2. Line 135. Does the urban model include anthropogenic heating and a building energy model?

Answer: The simulation configuration employs the Building Surface Module (BSM; Resler et al., 2017, Maronga et al., 2020) with prescribed inner building temperature. The inner building temperature is prescribed according to the season, and heat exchange through walls/roofs is calculated accordingly. The exchange is modeled as a heat transfer between the inner building environment with a realistic inner temperature and the outer environment based on properties of walls and roofs (insulation, thermal capacity, thickness, etc.), and thus the heat energy produced inside the building (mainly heating or cooling) is considered. We use this method because we have more precise information about building construction than about installed heating/cooling systems and their operational conditions. We thus expect our approach to be less prone to large errors in our test cases than the Building indoor model (Pfafferott et al., 2021) since we do not have sufficient information for its complex calculation and due to that, we could bring significant uncertainties into the model energy balance.

1. Resler, J., Krč, P., Belda, M., Juruš, P., Benešová, N., Lopata, J., Vlček, O., Damašková, D., Eben, K., Derbek, P., Maronga, B., and Kanani-Sühring, F.: PALM-USM v1.0: A new urban surface model integrated into the PALM large-eddy simulation model, Geosci. Model Dev., 10, 3635–3659, https://doi.org/10.5194/gmd-10-3635-2017, 2017.

2. Maronga, B., Banzhaf, S., Burmeister, C., Esch, T., Forkel, R., Fröhlich, D., Fuka, V., Gehrke, K. F., Geletič, J., Giersch, S., Gronemeier, T., Groß, G., Heldens, W., Hellsten, A., Hoffmann, F., Inagaki, A., Kadasch, E., Kanani-Sühring, F., Ketelsen, K., Khan, B. A., Knigge, C., Knoop, H., Krč, P., Kurppa, M., Maamari, H., Matzarakis, A., Mauder, M., Pallasch, M., Pavlik, D., Pfafferott, J., Resler, J., Rissmann, S., Russo, E., Salim, M., Schrempf, M., Schwenkel, J., Seckmeyer, G., Schubert, S., Sühring, M., von Tils, R., Vollmer, L., Ward, S., Witha, B., Wurps, H., Zeidler, J., and Raasch, S.: Overview of the PALM model system 6.0, Geosci. Model Dev., 13, 1335–1372, https://doi.org/10.5194/gmd-13-1335-2020, 2020.

3. Pfafferott, J., Rißmann, S., Sühring, M., Kanani-Sühring, F., and Maronga, B.: Building indoor model in PALM-4U: indoor climate, energy demand, and the interaction between buildings and the urban microclimate, Geosci. Model Dev., 14, 3511–3519, https://doi.org/10.5194/gmd-14-3511-2021, 2021.

3. Line 225. Define MRT, PET and UTCI.

Answer: The following paragraph has been reformulated, with the definitions included, and additional information was added about the reasons for evaluating the thermal indices as suggested by Reviewer#2 in his comment on L225-276. For the new formulation please take a look at L263-L276.

4. Figure 2. Is it one point for each time and profile-average and each simulation, or each day? It is not clear how many points to expect to see here?

Answer: The additional information you've asked for in this comment we included in the manuscript. Please take a look at the paragraph from L300 to L307.

5. Supplementary Figures. It is noticed that the night time stable layer development is missed by both models, nor does PALM seem to show any improvement when in calm conditions it should be able to develop cooling if it was able. I am not sure that the driving model could be solely blamed for missing these and more likely both models are missing something. This should be discussed more even if it is just seen in Supplementary plots.

Answer: Thank you for pointing this out. We do agree with your opinion and have considered mentioning this issue in the manuscript. This is one of the first sets of experiments of such kind, and we would like to focus more on this issue in future research. This remains to be researched and it is yet to be seen if this behavior by PALM is systematic or is it not. In this instance, we can't "blame" either model based only on the data and simulations we have, thus, we would need more simulations. However, we've concluded that this effect is related to the PALM's dynamic core.
We included additional information about this comment in the manuscript. Please see L361-L367.

6. Line 318. "fringe-like pattern". Is this because the flow is from the west and the variation is along the west boundary?

Answer: We checked the data and concluded that the prevailing flow is from west to east (200-300 deg). We also checked the PALM original data (not differences which are presented in Figure 3b) and this pattern is present there as well. Our hypothesis for the formation of these waves is a consequence of the inflow boundary conditions, and for the details, we refer to the second reviewer's comment on Figure 3b.

7. Figure 3. We see a strong local signal which is presumably an entirely urban area, but it is hard to know because no maps show the WRF urban area. Could an outline be added for WRF urban grid cells?

Answer:
According to your suggestion, we've added WRF urban grid cells to Figures 3, and 4, and corresponding supplementary figures. For this purpose, we used the LU_INDEX WRF model variable which shows the landuse information. If the square is outlined with a black color the majority of the land use in that square is of urban type. If the square is not outlined with the black line, then the majority of the grid box area is not of urban type, albeit some of it can be.

Please see Figures 3, and 4 and supplementary figures with the description on them, and an additional description on L372-L373.

8. Figure 5b and others. It is interesting to note that the more urban cells have less difference due to driving data. I think this implies that the buildings are deterministic in some way in the local climate. This could be an interesting aspect of the results to discuss. What characteristics of urban density do these areas have?

Answer: In connection to your comment #11 we added a short comment in the manuscript on this feature. Please see the L470-L473.

In more detail, in Fig05b are differences between members 05 and 07, which means BouLac BEP urban physics (07) and with YSU BEM (12). It seems that both members have a similar parametrization for heterogeneous built-up areas (prevailing terrain is flat, there are only minor differences in building heights or density). The highest differences are located in forested areas of urban greenery (see classes 1410 and 3100 on Fig05a) in complex terrain, mostly on north-oriented valley sides. It means that both parameterizations have important differences in the parametrization of urban greenery and/or terrain orientation (probably the effect of different irradiation).

9. Figure 6. Some of these rows are redundant being just the reverse of other rows. It can be reduced to 6 rows without loss of information.

Answer: The unnecessary rows have been removed. In addition, as suggested by Reviewer#2 the font size has been increased, the displayed hours have been reduced to every three hours, and an additional x-axis has been added. Figure 7, and supplementary figures have been adjusted accordingly.

10. Line 405. As above, I think it is not completely clear that the local errors in both models can be blamed on the mesoscale model in cases where both miss the night-time stability. There is no indication that PALM has tried to correct this bias even in conditions of light wind where local effects should dominate.

Answer: *This was more of a general comment, but we see how it might be confused with a discussion of our experiment. As per the suggestion of the second reviewer, the general text was moved to the introduction and only the discussion of the actual results was kept here. Please see L48-L56.*

11. Line 412. Also related to this as stated earlier, dense urban areas appear less affected by the driving data, so in these areas the parameters may matter more.

Answer: Thank you for this observation, we added a comment about this in the manuscript (see also the answer to comment #8). Please see the L470-L473.

12. Line 444. Yes, WRF has recently added the local climate zone map which should improve urban morphology. LCZ could be mentioned here.

Answer: We have added information about the LCZ., please see L509-L510.

13. General comment. I think the large number of tables is not critical to the bulk of the paper while some soundings in the Supplementary data are. I think a better balance could be achieved. Or maybe some way could be found to condense the table information more.

Answer: We moved two tables summarizing the technical details from the manuscript text to the appendix. One is the table describing WRF model parameterizations, and the other one summarizes selected WRF ensemble members as the best/worst. They are now Appendix A and B respectively.