

Author's Response to Review #1

EGUSPHERE-2025-144 | Development and technical paper

SanDyPALM v1.0: Static and Dynamic Drivers for the PALM Model to Facilitate Urban Microclimate Simulations

Julian Vogel, Sebastian Stadler, Ganesh Chockalingam, Afshin Afshari, Johanna Henning, and Winkler Matthias

Referee #1

General comments

At the end of the manuscript, the authors tried to answer this question, “which dataset most accurately represents the selected district, and how can we further improve data quality?”, but the answer given in the manuscript is that “Different data sources for the same test case can yield varying results, yet the outputs do not diverge excessively.” While the authors showed comparison in time series and averaged fields, the advantage of using an LES model emphasises more on the turbulence-resolving parts. I understand that this is a tool development paper, so more detailed validation or presentation of model results is not required. Can the authors comment more on how the simulated turbulent flow and vertical structure of the boundary layer (e.g. boundary layer height) might be impacted by different static inputs?

Response: Thank you very much for your thorough review and your constructive comments. Regarding the simulated turbulence and especially the vertical structure of the urban boundary layer, we have compared the time-averaged vertical profiles of turbulent kinetic energy, resolved and sub-grid-scale, for the different simulations. The Figures are attached to this response. One reason for observed differences in the profiles is that the OSM and LCZ cases used a global remotely sensed digital elevation model (DEM), while the MOSAIK and Custom cases used a more accurate DEM from the municipality. The difference between the two DEM models can be seen in Figure 10 in the manuscript. Keeping that in mind, we observe that there are more similarities between MOSAIK and Custom, and between OSM and LCZ. Therefore, the DEM may have the biggest impact. The remaining divergence can then be attributed to different building, tree and surface properties. Conclusively, the static input data influence the urban boundary layer turbulence to some extent. However, we hesitate to put these results in the manuscript and open a new discussion, because we feel it is beyond the scope of this research topic.

Specific comments

1. Figure 1, I cannot locate the TUB tower and the DWD station in the figure. Please mark them with more visible markers.

Response: Thank you for this hint. The markers and descriptions of the stations got lost during a previous revision. Similarly, the descriptions of parent and nest domains were missing. We added all overlays again in this revision.

2. L197: “One caveat of OSM data is that it can be quite rich in data outside of cities”. Why rich data can be a caveat? Do you mean there is an imbalance between rural and urban areas?

Response: Thank you for pointing out this mistake. It should mean “One advantage of OSM data ...”. We corrected the text accordingly.

3. LCZ4PALM, the authors developed this novel feature, but the motivation for this is not clear. The authors mentioned that PALM-4U has gap-filling issues in the paragraph in L646-L655, but the LCZ static driver looks the least realistic among the four datasets. (Figures 11-16)

Response: The motivation is as mentioned in the introduction: “LCZ4PALM is particularly useful for parametric studies and for coarser grids, where real building shapes cannot be resolved, making it preferable to have an approximate but meaningful urban representation instead.” The LCZ4PALM method creates a structured morphology of buildings and street canyons for coarse grids. It is true that the LCZ static driver is the least realistic, but the basis for it is just an LCZ map with 100 m resolution. An intention of LCZ4PALM was to prevent the street canyons in the

coarse parent domain from being “lumped together” due to interpolation and gap-filling, which creates additional terrain or building grid cells and results in a modification of the building footprints, see e.g. Figure 19. Even though the street canyons are well resolved in the static driver, it is difficult to prevent the modification due to gap-filling. We are currently investigating if and how much a strong terrain filtering would help, but this is beyond the scope of this paper.

4. L298-299: “There are separate functions that generate the PALM-4U grid and transfer it into a geographic projection”. Please specify.

Response: Thank you for pointing out that the description is not clear enough. The feature that we wanted to highlight is that SanDyPALM offers a simple and independent function to generate the PALM grid positions before performing any time-consuming processing—in the local and projected coordinate systems. The tool plots the PALM domains on a geographic map and the domain extent is given. Vertical levels and thicknesses are printed, which is useful when using vertical grid stretching to quickly try different settings. The according section was adjusted for clarity.

5. From the user-friendly perspective, please describe the file/folder structure of SanDyPALM. This was not in the manuscript and/or the documentation.

Response: We added the file/folder structure in the README.md in the Gitlab Repository, which currently serves as the only documentation outside of the manuscript. We decided against including it in the manuscript to not extend the length further.

6. L324: Please add a link or proper citation for the PALM input data standard (PIDS)

Response: We added a citation that refers to the web address of the PALM input data standard.

7. Dynamic driver generation: does this require a static driver generated by SanDyPALM? Or it is compatible with any static driver provided by the user? Please clarify.

Response: The dynamic driver can be created for other static drivers as well, if the static driver is put in the correct path in the working directory and is renamed as expected by the program code. Besides, the dynamic driver can also be created without a static driver with slightly decreased feature set. We added text to the manuscript that mentions the possibility of standalone dynamic driver generation.

8. Section 3.1, the authors only mentioned figures in the first paragraph and did not refer to the figures later in the text. For example, somewhere in L517-L520 should refer to Figure 10. Same as in Section 3.2

Response: The Figure names were added at all the places where they are discussed.

9. L658: “There is a clear increase in temperature from MOSAIK to Custom to OSM, ...” Why?

Response: We have added an explanation to the according paragraph, where we indicate that the increase of temperatures in this order is most likely due to differences in amount of vegetation in the geographic input data.

Referee #2

General comments

Generally, the manuscript is structured nicely and includes a lot of details. Methodology is sound, and the evaluation provided can be considered trustworthy. The manuscript fits the GMD scope. Before final publication, I would suggest a careful review of the language and sentence structures. Finally, it would be appreciated if the authors would address the following comments before final publication.

Response: Thank you very much for your thorough review and your constructive comments. We are going to revise the language before final publication.

Major comments

1. Connecting the research questions in the introduction with the conclusions. In particular, the conclusion does not provide a clear answer to the questions you posed in the Introduction. Especially questions 2) and 3) on L106-107. Please, try to connect the two (introduction and conclusions) and delve into more detail, if possible, for questions 2 and 3, and try to answer one by one in more detail, concerning your contribution with the presented input data toolkit. In addition to this, the statement on the L720 in the conclusion is too general in terms of “refining these automated tools”. Refining in what way? Can you be more specific about this concerning your toolkit?

Response: Thank you for pointing out this shortcoming of the descriptions. We carefully revised the conclusion to make sure that the stated research questions are all answered completely. We also added explicit examples on how to refine the automated tools.

2. At the beginning of the manuscript, you suggested UHI as a phenomenon enhancing the urban climate problems and evaluated a heat-wave episode as a validation episode. So, what influence do you expect on the UHI evaluation, for example, from microscale simulations with different static drivers in these cases? How much “trust” can we put into our evaluations of such phenomena, depending on the level of detail the input data provides in the case of your datasets? The accuracy of the static driver data has a major influence on the simulated urban climate processes (e.g., radiative processes and consequently the energy balances, etc.). I see that this manuscript is technical and of a development nature, and you provided the validation of the model outcomes to justify the tools. However, one of the PALM’s purposes and superiority over other available tools lies in its details and potential use for urban mitigation strategies and the overall microclimate assessment. Your statement on L124-126 indicates that the proposed toolkit contributes to more reliable urban climate simulations. I would elaborate on this in the discussion/conclusion section a bit more.

Response: Thank you for these important suggestions. A UHI evaluation would be an interesting extension of this work, especially investigating the impact of using different static drivers on the resulting UHI effect. As a result of our current work, we can only give an answer on how the static driver impacts the urban microclimate simulation results in general, as given in section 3.3, where we compared the average, standard deviation, min and max values of near-surface temperature, humidity and wind speed. The answer to the question how trustful the simulations are is rather subjective and our opinion is already stated in the paper. With regards to investigating the UHI effect, we believe that this would be an excellent future extension of this work, but probably beyond the scope of the current manuscript. We now added this as future outlook in the final paragraph of the conclusions.

3. I would suggest including a separate dedicated section addressing the current limitations of the toolkit and potentially comparing the toolkit with the other available sources in terms of their shortcomings and strong sides. In addition, regarding limitations, can you elaborate on how compatible the static data file input from your toolkit is with dynamic drivers created by other tools? Can the static driver created by this toolkit only be used with the one created from the WRF toolkit you described? Can it be utilized and combined with different mesoscale models whose drivers are created by alternate tools for their utilization (e.g., COSMO)?

Response: Our manuscript initially had a section comparing the different available static driver tools. However, in the end we decided to drop this section, because the tools were too different and served different purposes so that it was difficult to compare them. For example, a table we created to show distinct features for several tools was quite sparse, because a feature was often included only in one of the tools. Furthermore, many of the existing tools are still unpublished and in development, which makes it difficult to reference them and track their current state. Regarding the dynamic driver compatibility: in theory, the static driver created in our tool can be combined with dynamic drivers created by other tools. However, one of the advantages of our approach is to combine the generation of the static and dynamic drivers to utilize possible synergies, e.g. adjusting the vertical levels of the mesoscale data during dynamic driver creation to the PALM terrain of the static driver. But generally, the static and dynamic driver creation can be done independently in SanDyPALM. We have clarified this further in our manuscript in section 2.3.

Specific comments

1. Manuscript title: I would suggest removing or modifying the word “realistic” from the title since it can mislead in terms of overstating the accuracy. It is quite subjective, and the word itself needs to be taken with caution.

Response: We removed the word "realistic" from the title.

2. L1-21: At the beginning of the manuscript, the authors discuss the issues related to the urban climate. However, this is quite scarce and not anyhow connected to the rest of the introduction. I would suggest rewriting this part and incorporating it with the rest of the text. The impression I got while reading it is that the authors discuss the general urban climate modeling and jump into the PALM specifics.

Response: We shortened this first paragraph to lessen the focus on the issues related to the urban climate.

3. L17: Please provide a reference for this statement.

Response: We added a reference for this statement.

4. L18-20: I suggest including adequate example studies incorporating/testing the mentioned mitigation/development strategies.

Response: With the response to comment #2, we decided to shorten the first paragraph and drop the examples, because they do not relate significantly to the study at hand.

5. L23: Please provide a reference for this statement.

Response: We revised this part and now include two new references regarding the motivation for microscale modeling.

6. L24: The NS solved equations include the Boussinesq approximation as well, please be more precise in your text (though with an anelastic approximation available too if necessary).

Response: We revised the manuscript to mention that the Navier-Stokes equations are solved using the Boussinesq approximation for buoyancy.

7. L28-30: I would suggest including the chemistry module, as well as the BIO module, in your list of key components, especially since you discuss the urban microclimate and UHI's impact at the beginning of the manuscript. These two play quite an important role when assessing urban microclimates in terms of chemistry and human thermal comfort. I have noticed that you mentioned the PCM module and utilized it in your simulations, so please consider including it in the list as well.

Response: All the recommended models have been included in the manuscripts and citations have been added where possible.

8. L48: Please provide a reference for this statement or rewrite the section. I would also strongly suggest going through the manuscript and filling in all the citation gaps.

Response: The section was revised, and a reference was added. We also checked again for any unreferenced statements.

9. L67: QGIS, explain this abbreviation.

Response: QGIS is not an abbreviation anymore. It was formerly known as Quantum GIS, but now it is just named QGIS.

10. L80-81: Please provide a reference for the WRF model.

Response: A reference for the WRF model was added.

11. L81: Provide the name of the developed tool.

Response: The name of the tool is now provided in the manuscript.

12. L85: WRF model repetitive (see line 81). Provide one full explanation of the abbreviation and use only the abbreviation afterward.

Response: The accidental repetition was removed in the manuscript.

13. Figure 1: Indicate stations clearly in the figure. Use an asterisk or some other symbol for stations. I would also suggest using colors to distinguish between the parent/child domains.

Response: During the review process, the Figure was replaced and the labels accidentally removed. In the revised version, the labels for the stations and the domains are visible again.

14. L162: OSM has already been introduced by the full name at L111, please go thoroughly through the manuscript and check the abbreviations not to repeat them (L186), this applies to other abbreviations, LAD, LAI, etc. Please go through the manuscript carefully and unify all. The same goes for L417, L437, and similar cases. This should all be fixed, and references should be placed there accordingly in cases where they are not provided.

Response: We fixed all the mentioned and other cases, where abbreviations were introduced multiple times or were not used after introducing them.

15. L230-235: What do you consider to be a "coarse grid" in terms of resolution and resolving the urban features? In what aspect can that provide a reasonable (realistic) contribution to microclimate/urban climate studies? I have noticed this term "coarse grid" across the manuscript. Can you maybe add an explanation about the resolution assumed under this term throughout the text (in cases when you don't refer to the 20 m parent domain)?

Response: The first time we mentioned coarse grids, we now specified that we mean grids larger than 10 m.

16. L394: HPC, abbreviation without the full name.

Response: The abbreviation was replaced by its full name.

17. L312-325: Even though this information is known, I would suggest citing either the PALM website or the PALM description paper from Maronga et al., 2020. This part of the text could also be condensed, and the link to the PIDS needs to be provided.

Response: The part was condensed and a reference to the PIDS was provided.

18. L441-453: I would suggest removing the explanation and keeping Table 3 with all the relevant information regarding the domain setup.

Response: We shortened the explanation to only mention the grid spacings and not the extents. This shortens the text but keeps mentioning all grid spacings for completeness.

19. 478: Add reference for the STG, please.

Response: A reference was added.

20. L722: Please check the ZENODO link once more, it does not lead directly to the repository. Also, regarding ZENODO, I would suggest adding a brief description on the ZENODO webpage about its contents.

Response: Clicking the link does not work for some reason, but it works fine when copying it. It is intended that the link does not lead to the repository but to our Zenodo page. The Zenodo page contains the code of version V1.0.0 and also a link to the repository. We agree that the link to the repository is somewhat hidden by default. Therefore, we added the link also in the description. We hesitate to insert another description about the toolbox on Zenodo and prefer to refer to the repository instead. The Zenodo page was adjusted accordingly.

21. Section 2.2.3: I would suggest rewriting and condensing information in this section concerning the quality of data OSM provides. On the one hand, you have L187-188, but on the other, L197-201.

Response: The section was condensed and partially rewritten.

22. L414-417: I would say that the statements you provided here about WRF schemes and resolutions cannot be considered universally true, and that it depends on the study case one is testing. In addition, using a “double” representation of the urbanized areas (in WRF’s more sophisticated schemes and PALM’s high-resolution data) can potentially affect the microscale outputs. Could the authors elaborate on this? Did you try researching this topic or testing sensitivities for different domain setups, cities, etc.?

Response: The description provided at this point is not about the WRF schemes, but only on our own “gap-filling” of data below the first WRF model level height that is done in the SanDyPALM toolbox. We only advise that, when using the single layer urban canopy layer model in WRF and a typical first half level height of about 25 m, a sophisticated “gap-filling” of the WRF data is beneficial, while it may not be necessary when using the multilayer UCM in WRF, where the first vertical half level height is usually 5 m. We could elaborate more on the considerations for specific test cases, but we feel it is beyond the scope of this manuscript. Urban areas are not necessarily represented “double”, because the WRF data is only used for PALM initial and boundary conditions. Therefore, it only represents the temporal data “before” the PALM simulation and the spatial data “outside” of the PALM domain. The approach may not be perfect, but if we leave out the urban representation in WRF, this would mean that “before” the PALM simulation and outside the PALM domain, urban areas would be represented as rural, which is usually not the case.

23. Figures 8 and 9: Please consider moving them to the supplements or appendix sections.

Response: We considered moving the Figures, but we feel that both Figures are important: The terrain height adjustment shown in Figure 8 is a major feature that is best explained by visualization and the three-step nesting setup (Figure 9) is important to get a broad view on our whole simulation approach.

24. L476: Provide reference for the plant canopy model (PCM).

Response: We now mention the PCM in the introduction and added a citation. However, we are not aware of a publication of the PCM within the scope of the PALM model. Therefore, we used the citation (Maronga, 2020), where the PCM has been mentioned and referenced thoroughly.

25. Table 2: Please consider moving this to the supplements or appendix section.

Response: Table 2 was removed, and the text slightly adjusted to contain all necessary information to make the Table obsolete.

26. Table 4 and Table 5: I would suggest moving the tables to the appendix or supplement about the statistical comparison of the static drivers. I would also suggest maybe condensing information about their differences (dominant types, “realness”, overestimating and underestimating certain urban morphological characteristics) into a table if possible.

Response: We considered moving the Tables and creating a new Table to condense information as suggested. However, as the Tables are discussed thoroughly in the text, we think the Tables are needed in place. Otherwise, the whole part would need to be moved to the appendix. Furthermore, we found it difficult to create a Table with condensed information, as it is very subjective, which information is useful for the reader. Instead, we tried to highlight the findings that are most significant to us in the text.

27. L495: I would suggest including a comparison with other available studies (if any) in the discussion section in terms of the static driver data accuracy and selection, as well as its impact on the modeled outputs.

Response: To the best of our knowledge, there are no studies explicitly investigating the accuracy and selection of the static driver.

28. L580: Can you specify what you mean by “keeping the dynamic driver consistent”? Are dynamic drivers synchronized with static drivers while preparing them? Considering you had 4 different static drivers, each of which had different information, the dynamic drivers could not be consistent among the 4 cases. The schemes used in the WRF setup, I suppose yes, but the boundary conditions derived from it would depend on the static driver information that is provided, and would have a further influence on the microscale simulation. Can you comment on this, please, and be clear in the manuscript about your statement on this line?

Response: We clarified this in the text as follows: “The dynamic driver is created using the same mesoscale data, but it may still vary slightly among the cases, because the surface layer model used to fill the near-surface data gaps in the mesoscale input data depends on roughness properties derived from the static driver.”

29. L617: and onwards: Can you describe how you took the PALM data for evaluation in case considering the domain grid box of 20 m?

Response: We have added a description on how the PALM data was evaluated. This is similar for the parent and nest domains, as the vertical grid size is the same at 5 m. The first grid point above ground is then 2.5 m and therefore we used this point directly as a good approximation for the 2m-temperature and 2m-humidity.

30. L691-694: The authors state that: “this research indicates that the outcomes of realistic urban microclimate simulations depend strongly on input data.” However, in the upcoming sentences, they state that the overall features of the microclimate simulations are similar. I suggest clarifying the initial claim on strong dependency on input data or adjusting it.

Response: We have slightly rephrased this part to make a weaker statement. However, we believe that both statements are true: There is a clear dependence of the results on input data, but the overall results are still similar, in that there are no extreme differences when comparing the cases to each other or comparing to measurements.

31. General question about observation measurements: Is the data from the stations assimilated anyhow to the ERA5 data used for driving the WRF model? Is there a bias of such a type? If so, please indicate that in the text.

Response: The DWD measurement station and the TUB tower may have been used in the data assimilation process to create the ERA5 data. But we don't see this as a problematic bias, because it is our intention to recreate the atmospheric processes in our modeling chain from ERA5 reanalysis data through the WRF mesoscale down to the PALM microscale). Our assumption is, if the models would be perfect, they should be able to produce results exactly matching the data obtained from the measurement stations, especially if they were used for data assimilation. But of course the models cannot be perfect, which leads to the deviations.

32. A suggestion for the authors: As far as the PALM/PALM-4U naming, I recommend using “PALM”/ “PALM model” throughout the manuscript since, technically speaking, PALM-4U modules/components are a part of PALM's framework.

Response: Thank you for the suggestion that we now follow throughout the manuscript.

Referee #3

General comments

1. My main issue is that the discussion of the difference of the test cases and the final sentence in the abstract is partly misleading. The authors state that the results of the cases are comparable although they find an spatially and temporally averaged difference of up to 1.4K. How big are local averages? Whether this deviation is acceptable depends on the purpose of the simulation: Is the user interested in spatial averages or in certain microscale effects? This needs to be discussed. This is also connected to the general data availability in the OSM data to assess its practicality. For example, is it possible to give information about the availability of building heights, number of storeys and the complete lack of them?

Response: Thank you very much for your thoughtful comments. Comparing local differences would be indeed an interesting addition allowing a more detailed discussion on the impact of the static driver on simulation results. However, one practical problem is that buildings, vegetation canopies and even terrain vary between simulations, which makes it difficult to compare the cases locally. For example, a grid point in one simulation may be atmospheric and thus contain results while the same grid point in another simulation may be part of a building and hence does not contain any results. Especially for the LCZ case, buildings are not comparable at all to the other cases. For this reason, we decided to only investigate the domain average, std, min and max values. As you pointed out, the

judgement of the simulation results depends on the specific application. In the manuscript, we only wanted to point out two things: 1) all simulation runs by themselves could be seen as validated from their comparison metrics with measurements. 2) At the same time, there are differences between the different cases, as you pointed out, with up to 1.4 K difference in temperature. We believe that the further interpretation of the results is subjective, and we would like to leave it to the readers to make their own judgments based on the provided data. We adjusted the manuscript slightly to weaken our interpretation statements. As for the OSM data, it would be possible to analyze how many of the building footprints have building heights or number of stories provided, and how many buildings needed to be filled by the global raster data WSF3D. However, this data would be specific for the test case we selected for this study and would not be generally indicative of the availability of building data in OSM. Also, the state of OSM is subject to constant change. Instead of giving specific numbers for our domain, we now include a publication that generally investigates building data completeness in OSM.

2. Furthermore, the authors find too high humidity values in the MOSAIK and Custom case, which have probably more realistic vegetation amounts than the OSM and LCZ case. Currently, the water availability applied to resolved vegetation is hard-coded and relatively high. According to the test case description, the days of interest are within a heat wave with like dry soil conditions. What is the soil moisture in WRF? Are measurements for the domain or Berlin available? Could an overestimation of the soil moisture be a reason for the deviations from the measurements?

Response: Unfortunately, we did not find any soil moisture measurements for the stations in our domain. Generally, our experience is that soil moisture is very scarce in measurement stations. We agree that the soil conditions may have a relevant impact on the PALM simulation results. But we believe this is not the major reason for the issue with humidity, since our different PALM simulations using different static drivers showed a highly varying degree of humidity values. Therefore, it is likely that the major factors are either the vegetation amount of the static input data or the way humidity is handled in PALM. The point you make about the water availability is interesting; it would be a valuable investigation to test different settings for water availability and see how it changes our results. But in our opinion this is beyond the scope of the current manuscript, so we added this point in the future outlook in the conclusions.

Minor issues

1. The introduction misses two competing tools. The static driver tool PALM-GEM, while admittedly quite hidden, is mentioned in <https://doi.org/10.1016/j.uclim.2024.102059> and published here: <https://zenodo.org/records/11067859>. The dynamic driver tool PALM-METEO is the successor of the tool used by Resler et al. (2021). It was shipped with PALM 24.4 and supports data input from the models WRF, ICON and Aladin. Please add these.

Response: Both mentioned tools are now included in the introduction and the references.

2. Figure 1 misses the measurement locations

Response: True, the labels for the measurement locations and also for the domains were accidentally removed during the initial submission. The labels will be included again in the next revision.

3. Line 203ff: How is the building height derived from the number of stories? With a constant storey height? What is the "same approach" for tree heights?

Response: We revised this section mentioning that we used a constant story height. We removed the phrase "same approach" and now just mention the default tree height used in case tree height is missing.

4. What kind of tool is SanDyPALM? A commandline tool only, does it have a GUI or is it a library?

Response: SanDyPALM is a command line tool without a GUI. We clarified this in the manuscript.

5. Does SanDyPALM offer the generation of resolved vegetation fields for single trees?

Response: At the moment, SanDyPALM can convert single trees to tree patches and this way the trees are resolved. However, it does not support the generation of 3D trees, yet. We added this in the manuscript and restructured the paragraph for clarity.

6. The current form of Figure 17 and 18 is mostly dominated by the diurnal cycle. I recommend showing only the differences to the measurements to focus on the differences between the simulations.

Response: We considered your suggestion of showing only the differences to measurements, but we found that we would still need to show the absolute values for completeness; otherwise, the reader would not know about the absolute values and the diurnal cycle. And, showing both absolute values and differences would increase the already large number of figures. So, after much consideration, we decided to remain with showing only the absolute values in the graphs. Information about the differences is already included in the error metrics tables, albeit only in an average sense.