Reviewer #3

The paper by Radovic et al. (2026) presents an evaluation of the PALM model's radiation module RTM against measurements from four sites in terms of shortwave radiation. PALM is forced with radiation from two different WRF set-ups for 16 days and run in spin-up mode, i.e. without the resource-consuming calculation of air quantities. PALM is a complex model and has mostly been evaluated with full, realistic set-ups and quantities that require different components of PALM. This paper focusses on a systematic evaluation of one component and is thus highly welcome. In most parts of the paper, in particular in the title and the abstract, this, however, does not become clear. In addition, since the model WRF is used as input, it is actually an evaluation of the coupled WRF/PALM-RTM set-up, which introduces additional uncertainties. I would like to ask the authors to make this clearer in the paper and to clarify the relationship with the stations Karlov and Libus (details below). I thus recommend consideration for publication after major revision.
Major issues:

1. **This paper mostly evaluates the RTM module of PALM. While the land-surface and the building surface module as well as the mesoscale nesting module are also used, the former only supplies surface albedo and the latter facilitates only the radiation input to my understanding. In particular, I assume that only WRF *radiation* data is used and not other fields; the latter is implied by "dynamic meteorological forcing" (L159). Please clarify this in the paper, in particular in the abstract. Are the surface albedos constant in time or (partly) sun-angle-dependent?**

Answer: *During the spinup mode, PALM uses only the incoming shortwave and longwave radiation from the WRF mesoscale forcing to drive the RTM module and performs surface energy balance calculations. Other meteorological variables are not dynamically coupled to the PALM simulation per se; they are used from WRF to create the "complete" dynamic driver set. The wind speed, humidity, and temperature are prescribed through the configuration p3d file. In the present setup, temperature follows an idealised daily sinusoidal cycle, while the remaining atmospheric variables are kept constant in time, and present only in the dynamic driver.*
We will clarify in the manuscript that mesoscale nesting is used solely to provide radiative forcing fields rather than full dynamic meteorological forcing, and that the remaining fields are downscaled into the dynamic driver file but not used. Please see Lines 204-206 in the revised manuscript.
Surface albedo (reflectance) values are prescribed for each individual surface grid element and remain constant in time; no explicit incidence-angle-dependent albedo parameterisation is applied. On the other hand, the total effective albedo (ratio of reflected to incoming radiation) for larger areas does change in time as a result of changing solar geometry when the 3-D radiation with multiple reflections is simulated by RTM. We have revised the manuscript to address the comments regarding albedo. Please refer to the newly included section 2.1.1, "Spin-up

simulations," in which all albedo-related concerns are addressed. Additionally, we would also like to point to our reply to Reviewer #1's comment #3 regarding albedo, in which we evaluated the time series of observed albedo and discussed the additional analysis presented.

2. **The authors also imply that PALM-RTM could partly correct errors in the radiation input (L12, L290, L357, L376). I think that this is not the case. I consider the radiation fluxes of WRF to be above its (unresolved) canopy and exactly like this, it is considered as input in RTM. RTM mostly distributes it geometrically within the canopy without doing any atmospheric adjustments. This is why using WRF as input actually results in an evaluation of the coupled WRF/PALM-RTM system. Thus, forcing with radiation measurements above the canopy, for example from rooftops, would have removed the uncertainty from the evaluation. Please discuss this. What about the stations Karlov and Libus? Their data is used in the paper but both stations are not introduced at all. Please describe these stations as well. Could their data be used as forcing?**

Answer: *We agree that PALM-RTM does not explicitly correct biases in the incoming radiation forcing from WRF. The RTM primarily redistributes incoming shortwave and longwave radiation within the urban canopy through geometric shading, absorption, reflection, and multiple scattering. Therefore, uncertainties in the WRF radiation forcing may propagate into the PALM simulations.*
Using radiation measurements above the canopy as forcing would indeed reduce uncertainties associated with mesoscale radiation forcing and would allow a more isolated evaluation of RTM performance. However, the objective of the present study was to assess the applicability of PALM-RTM under realistic mesoscale-forced conditions representative of operational urban climate simulations.
We will additionally expand the description of the Karlov and Libuš stations in the manuscript, including their locations, measurement characteristics, and observational datasets used in the evaluation. While these stations could potentially provide observational radiation forcing data, in the present study, they were used exclusively as out-of-domain reference evaluation stations and not as direct model forcing. Please see Lines 254-262 in the revised manuscript.

Minor issues:

3. **L12: Does PALM, in particular RTM, compensate for any errors in the radiation input? My understanding is that RTM distributes the radiation within the canopy received as input at the top of the canopy. This input is expected to be correct.**

Answer: *The incoming radiation input coming from WRF (and further dynamic driver) can contain errors (e.g., cloud representation and timing of cloud cover, aerosols). So, the radiation arriving at the canopy top of the PALM domain is not guaranteed to be necessarily "correct" or close to reality. The RTM uses whatever radiation it receives as boundary/input forcing and, per se, is not designed to diagnose or correct biases in the forcing; it just redistributes and* *physically processes the given data within the canopy. Hence, the quality of the RTM results still depends on the accuracy of the incoming radiation forcing.*

4. **L31: "recognised by THE World Meteorological Organization"**

Answer: *We thank the reviewer for the correction. The text has been updated to include the missing article (please see Line 32).*

5. **L46: As the authors write in L45, MRT cannot be derived from shortwave radiation alone, but longwave radiation needs to be considered as well.**

Answer: *We thank the reviewer for pointing out this inconsistency. The sentence is revised to clarify the issue. Now it states: "Outdoor MRT combines the impacts of short-wave and long-wave radiation fluxes in outdoor environments. While long-wave radiation is an important component of the overall radiative balance, especially in indoor or shaded environments, short-wave radiation is often the dominant driver of elevated MRT values outdoors, particularly at high solar altitudes and under clear-sky conditions. Although MRT derivation strictly requires integrating both short-wave and long-wave radiation flux densities, the variability of outdoor MRT during the daytime is heavily influenced by short-wave radiation flux \citep{lee2014}." For the corrections, please see Lines 45-50.*

6. **L106: Without parentheses around the citation.**

Answer: *Done. The parentheses around the citation have been removed. For the corrected manuscript, please see Line 106.*

7. **L144: Are there any differences in the results of RTM 4.1 compared to RTM 4.0 or RTM 3.0 described in Krc et al. (2021) when only a 2.5D geometry is used? The description mentions only numerical advancements.**

Answer: *The differences are indeed mostly computational enhancements (such a, localized raytracing, improved MPI exchanges, etc.), with the single exception of the full 3D geometry as opposed to the 2.5D geometry in RTM 3.0. However, the actual number of overhanging structures in the simulated scenario is very small, and none of them are close to the measurement locations, therefore, we do not expect this to have an impact on the results.*

8. **Section 2.2: Please include more details of the WRF simulations:**

   *** How is WRF forced? Only the discussion section mentions ERA5.**

Answer: *We apologise for the omission in the Methods section. The WRF model was indeed driven by ERA5 reanalysis data. We have updated section 2.2 "Initial and boundary conditions" and Lines 206-207 to explicitly state this and provide the necessary technical details regarding the WRF's forcing data.*

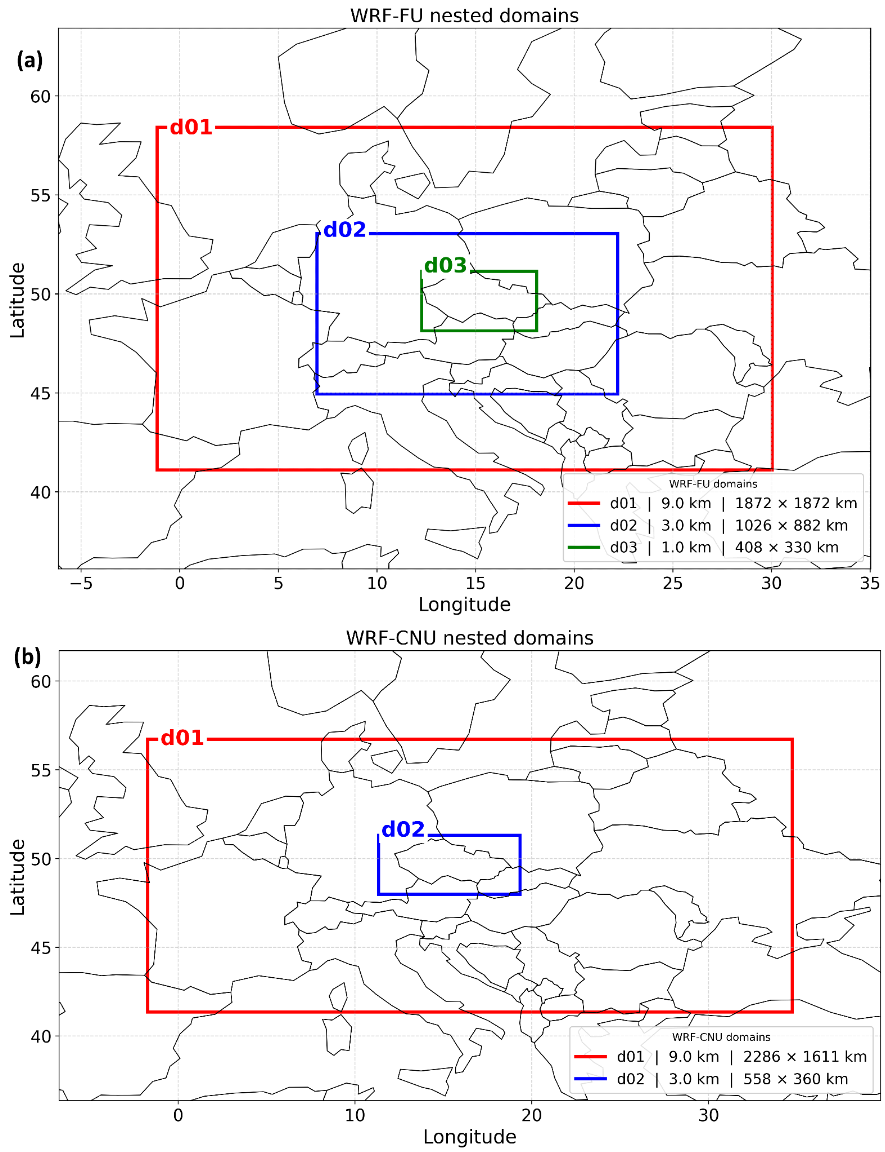
   *** What are the domain sizes, and are the FU simulations nested into CNU? Or are there any other nesting steps in between? If not, is the difference in**

resolution between the forcing of WRF and WRF itself (in particular for the FU simulations) not a problem?

Answer: *We thank the referee for pointing out the omission of the description of WRF domains and nesting. We added the information to section 2.2, “Initial and boundary conditions,” and to Lines 200-202. In addition, we provided a detailed Figure S14 in the supplementary materials to support our answer, as well as here as the continuation of our answer.*

To clarify, the CNU and FU WRF simulations were performed in two independent domain configurations. The CNU simulations were performed as a two-domain setup, with an inner domain at 3 km resolution and an outer domain at 9 km resolution. The FU simulations have a triple-nesting chain with resolutions of 9 km, 3 km and 1km. We tried to make the WRF setup standard. We think that ERA5 (with 0.25° resolution) forcing WRF with 9km horizontal resolution is a frequent choice, as well as further nesting of WRF domains to 3 km and 1 km resolution. From today’s operational forecasting perspective, the CNU setup is rather minimalistic. The triple nesting in FU with domain sizes as depicted in the picture is (in our opinion) comparable to similar cases reported in the literature.

Figure: WRF model simulation domain configurations and nested grid hierarchies for the two experimental setups: (a) the three-domain WRF-FU setup (red: d01 at 9.0 km, blue: d02 at 3.0 km, and green: d03 at 1.0 km horizontal resolution) and (b) the two-domain WRF-CNU setup (red: d01 at 9.0 km and blue: d02 at 3.0 km horizontal resolution). The legends denote the horizontal grid spacing and total geographic coverage area for each parent and nested domain.



9. L170: Is the diffuse shortwave radiation also stored?

Answer: *Yes, the diffuse short-wave radiation is stored in the WRF model's 10 min-resolution output files and passed to PALM-meteo for further processing into the dynamic driver. In our setup, the* ***SWDDIF*** *variable ("Shortwave surface downward diffuse irradiance") from the WRF output is mapped to the* ***rad_sw_in_dif*** *variable within the PALM dynamic driver. This allows PALM to use the meso-scale model's specific partitioning of direct and diffuse radiation as a* *boundary condition, rather than relying on an internal parameterisation for the split into direct and diffuse components, which would be applied if the external diffuse component were unavailable.*

10. **Table 1: Which urban canyon parametrization is used? Probably BEP (Martilli et al. 2002), however, SLUCM+BEM (Takane et al. 2024) is also available.**

Answer: *The WRF simulations utilised the building environment parameterisation (Martilli et al., 2002), in combination with a building energy model, so BEP+BEM by Salamanca and Martilli (2009), corresponding to* **`sf_urban_physics = 3`** *in WRF. We have clarified this in the revised manuscript in Table 1. We also thank the reviewer for pointing out the recent work by Takane et al. (2024) on the SLUCM+BEM implementation, and we believe it could be explored and tested in future studies.*

11. **Table 4: Columns Category and CNU/FU are redundant.**

Answer: *We agree with the reviewer's observation regarding the redundancy in Table 4. We have removed the redundant columns in Table 4.*

12. **L220: Please explain exactly the output quantities: Are the values taken from the bottom surface (height 0m) at the locations of the measurements? For completeness: what does SWin include, only direct and diffuse radiation from the sky or also reflections from the surroundings?**

Answer: *We clarify that the SWin variable is not limited to the direct and diffuse solar radiation coming down from the sky, but it accounts for secondary reflections from the surrounding environment (such as neighbouring walls, ground, and vegetation). The model simulates reflection steps among surfaces that share a mutual view factor. At each step, surfaces and canopy elements absorb a portion of the radiation, while the remainder is re-reflected according to surface albedo, whereas the facets with no mutual view to other structures (e.g., unobstructed roof surfaces) receive only the direct and diffuse sky radiation, meaning their incident shortwave irradiance does not receive reflected environmental components. The values are taken from the grid cell nearest to the vertical position of the pyranometers, that is, 1 m above ground.*
For the revised manuscript, please look at section 2.1.1 "Spin-up simulations" and Lines 170-274 for the first part of your comment.
For the second part, please see our response to Reviewer 1, which addresses the optional suggestion to include a diagram. We have implemented this suggestion in Appendix B of the manuscript.

13. **Figures 3 and 4: According to Table 4, the common episodes are e1 to e6. Why do the captions say it is (e3, e5, e6, e8, e9, e16)?**

Answer: *We apologise for this inconsistency. The discrepancies in the episode numbering in the captions of Figures 3 and 4 were a clerical error resulting from a later episode's re-indexing* *during the final manuscript preparation. We have updated the captions to correctly match the numbering (e1–e6) used in Table 4. Please see Figures 3 and 4.*

14. **Figure 5, in particular (a): Please discuss why the ratio of PALM-CNU In to WRF-CNU In is so different from the ratio of PALM-FU In to WRF-FU In. Is this related to the the relationship of diffuse and direct radiation? This would highlight that not only the total shortwave input needs to be correct but also the distinction between diffuse and direct.**

Answer: *We thank the reviewer for highlighting this important topic. To investigate this observation, we examined the diffuse shortwave radiation supplied to PALM via the dynamic driver, as well as the WRF input data for both simulation setups. The qualitative analysis reveals substantial differences between the PALM-CNU and PALM-FU simulations in the diffuse shortwave radiation forcing itself, with PALM-CNU exhibiting considerably larger diffuse radiation values throughout the daytime period (see Figure 1). This finding supports the reviewer's interpretation that the differing relationships between WRF and PALM incoming shortwave radiation are likely related to differences in the partitioning between direct and diffuse radiation components. Since PALM further processes the incoming radiation through urban radiative transfer mechanisms such as shading, obstruction, and reflections, the diffuse-to-direct ratio becomes highly important for the resulting near-surface radiation fields. Consequently, even when total incoming shortwave radiation is comparable (as shown in Figure 2), differences in diffuse radiation forcing can produce substantially different urban radiative responses within PALM. These results, therefore, highlight that accurate representation of the diffuse and direct shortwave radiation components is crucial for reproducing total radiation fluxes, and that PALM simulations are sensitive to the diffuse/direct partitioning inherited from the mesoscale forcing model.*

An additional discussion addressing this valuable point has been added to Section 4.2, "Driving data impact: WRF-CNU vs. WRF-FU configuration performance" (Lines 249–255).

Figure 1. *Comparison of diffuse short-wave radiation supplied to PALM-CNU (solid blue) and PALM-FU (dashed orange) setups through the dynamic driver (panel (a); variable rad_sw_in_dif) with the corresponding forcing data from WRF-CNU (solid blue) and WRF-FU (dashed orange) (panel (b); variable SWDDIF) for the clear-sky episode e6 at the FLE station.*

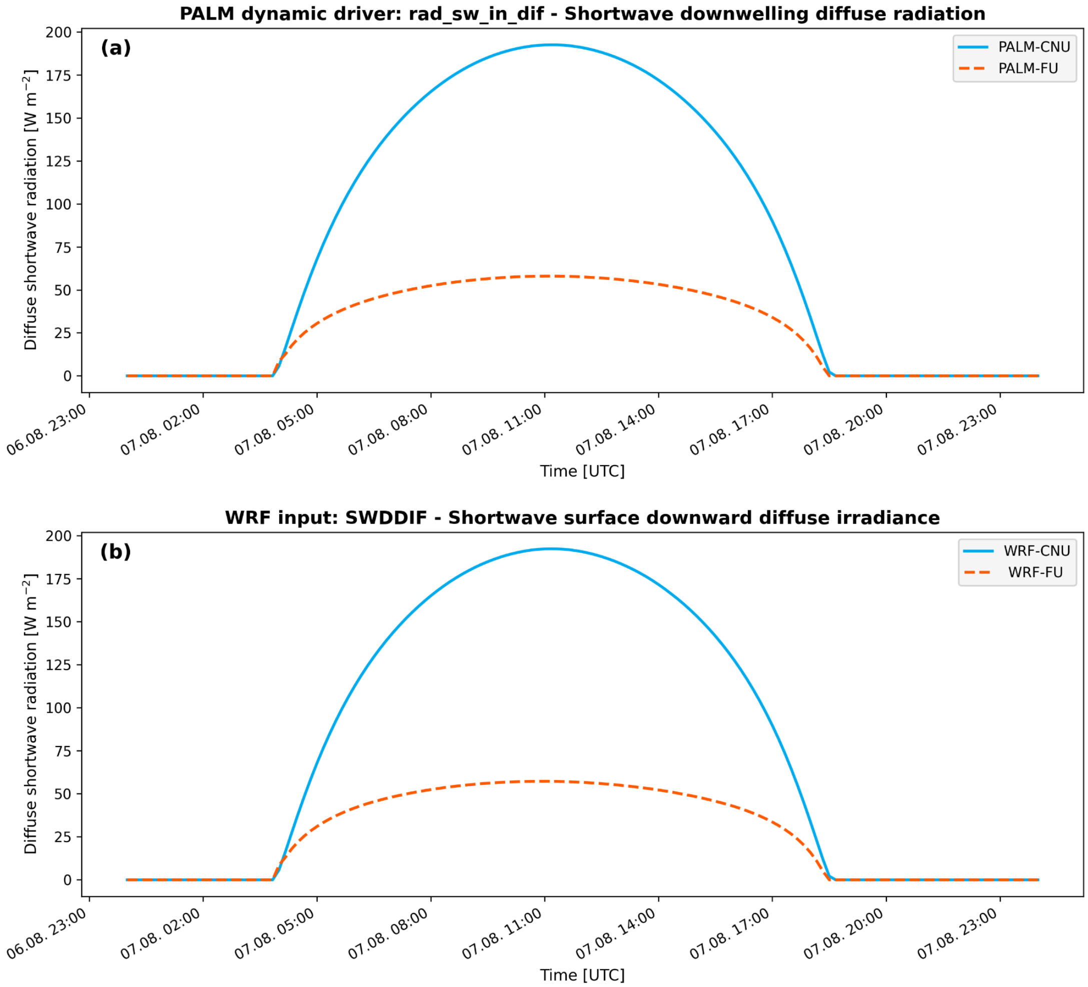
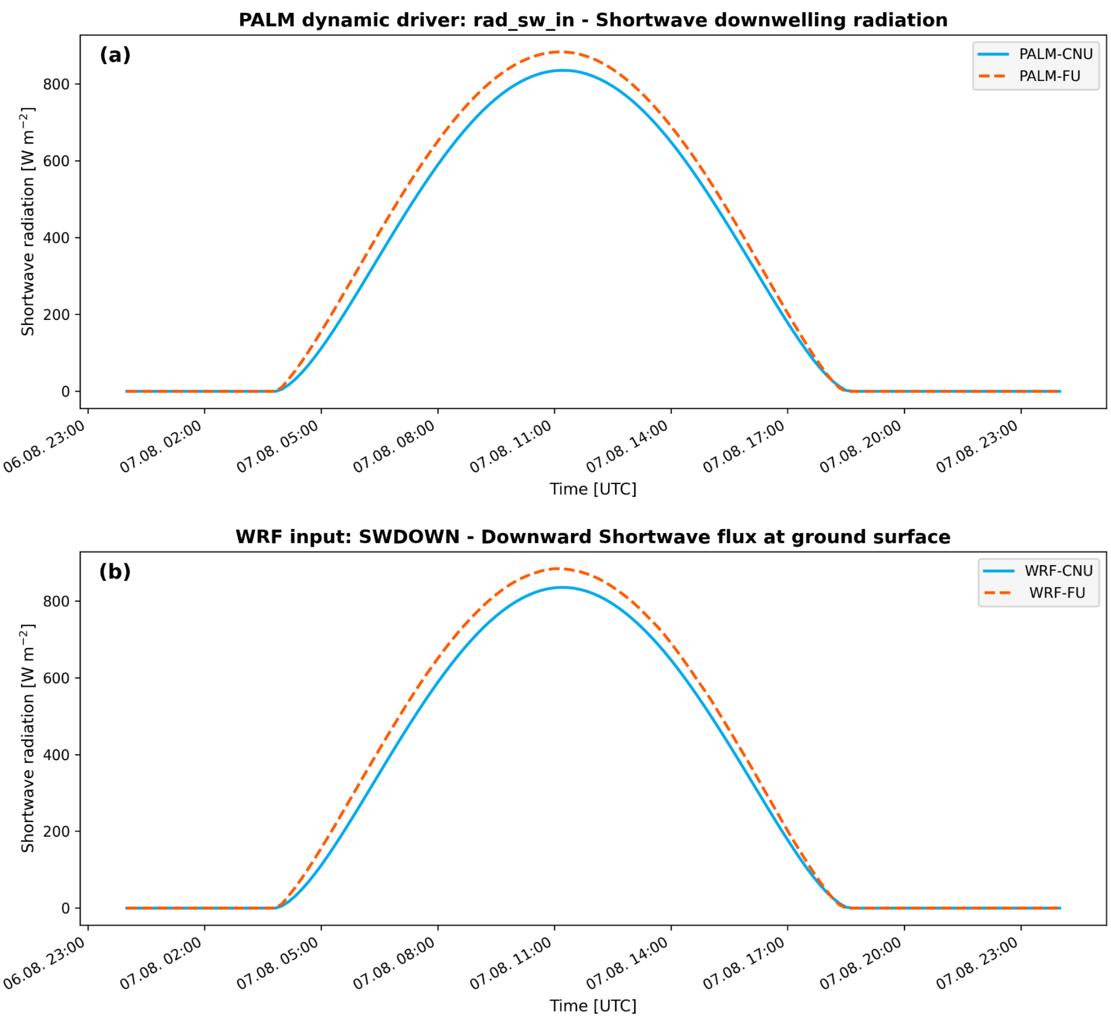


Figure 2. *Comparison of the direct short-wave radiation supplied to PALM-CNU (solid blue) and PALM-FU (dashed orange) setups through the dynamic driver (panel (a); variable rad_sw_in) with the corresponding forcing data from WRF-CNU (solid blue) and WRF-FU (dashed orange) (panel (b); variable SWDOWN) for the clear-sky episode e6 at the FLE station.*



15. L331: episode e5 while Figure 6 says e9.

Answer: *Once again, we apologise for this inconsistency. The discrepancies in the episode numbering are due to the reasons explained in our answer to your previous comment. We have updated the caption to correctly match the numbering used in Table 4. Please see Figure 6.*