

Anonymous Referee #1, 04 May 2026

Report #1

Submitted on 04 May 2026 Anonymous referee #1	
Notes for the submission of interactive comments	
Anonymous: Yes No	
Formal manuscript rating and recommendation to the editor (visible to authors and reviewers only)	
1) Scientific significance Does the manuscript represent a substantial contribution to modelling science within the scope of this journal (substantial new concepts, ideas, or methods)?	Excellent Good Fair Poor
2) Scientific quality Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)? Do the models, technical advances and/or experiments described have the potential to perform calculations leading to significant scientific results?	Excellent Good Fair Poor
3) Scientific reproducibility To what extent is the modelling science reproducible? Is the description sufficiently complete and precise to allow reproduction of the science by fellow scientists (traceability of results)?	Excellent Good Fair Poor
4) Presentation quality Are the scientific results and conclusions presented in a clear, concise, and well structured way (number and quality of figures/tables, appropriate use of English language)?	Excellent Good Fair Poor
For final publication, the manuscript should be accepted as is accepted subject to technical corrections accepted subject to minor revisions reconsidered after major revisions rejected Please note that this rating only refers to this version of the manuscript!	
Were a revised manuscript to be sent for another round of reviews: I would be willing to review the revised manuscript. I would not be willing to review the revised manuscript.	
Referee Comment: 'Comment on egusphere-2026-551'	

General comments:

This study describes a new one-tile urban scheme in the ORCHIDEE land surface model, which helped improve energy flux simulations and better capture the urban signature in water fluxes at 20 Urban-PLUMBER sites. The manuscript is clearly written and presented, and represents continued efforts by the global community towards improving urban representation in global-scale models. My main concern lies in the significant (up to 200 W/m^2) overestimation of sensible heat flux during the day in summer. As the authors mentioned a few times, the lack of anthropogenic heat fluxes may explain some of the underestimation in sensible heat fluxes in the winter, but similar argument can apply in the summer as well, which means the inclusion of anthropogenic heat flux could further worsen the overestimation of sensible heat flux in summer. Could the authors explain what could be contributing to the significant overestimation of sensible heat in summer during the day? As this bias persists in the original (baresoil), natural vegetation, as well as the urban experiments, could there be reasons linking to model structure or other more fundamental assumptions in ORCHIDEE? Does this overestimation show up in rural/natural vegetation settings as well? Having a detailed description on the potential reasons contributing to the bias could help readers better contextualize and add to the trustworthiness of the model development.

We thank the reviewer for this constructive comment and for their positive assessment of the manuscript.

During the revision, we identified an issue in the experimental setup used for the submitted manuscript: all simulations had been performed using the default ORCHIDEE soil texture rather than the site-specific soil-texture information provided by Urban-PLUMBER. We therefore reran every model configuration (Baresoil, Lowveget, Urban0, Urban1, and Urban2) at all 20 sites using the corresponding site-specific soil texture. These new simulations replace the original simulations throughout the revised manuscript. Consequently, all simulation-based result figures, including the boxplots, diurnal cycles, MAE comparisons, and water-budget analysis, as well as the associated numerical values and conclusions, were recomputed. The changes in the distributions, including the higher minimum latent heat flux values, therefore result from the revised simulations.

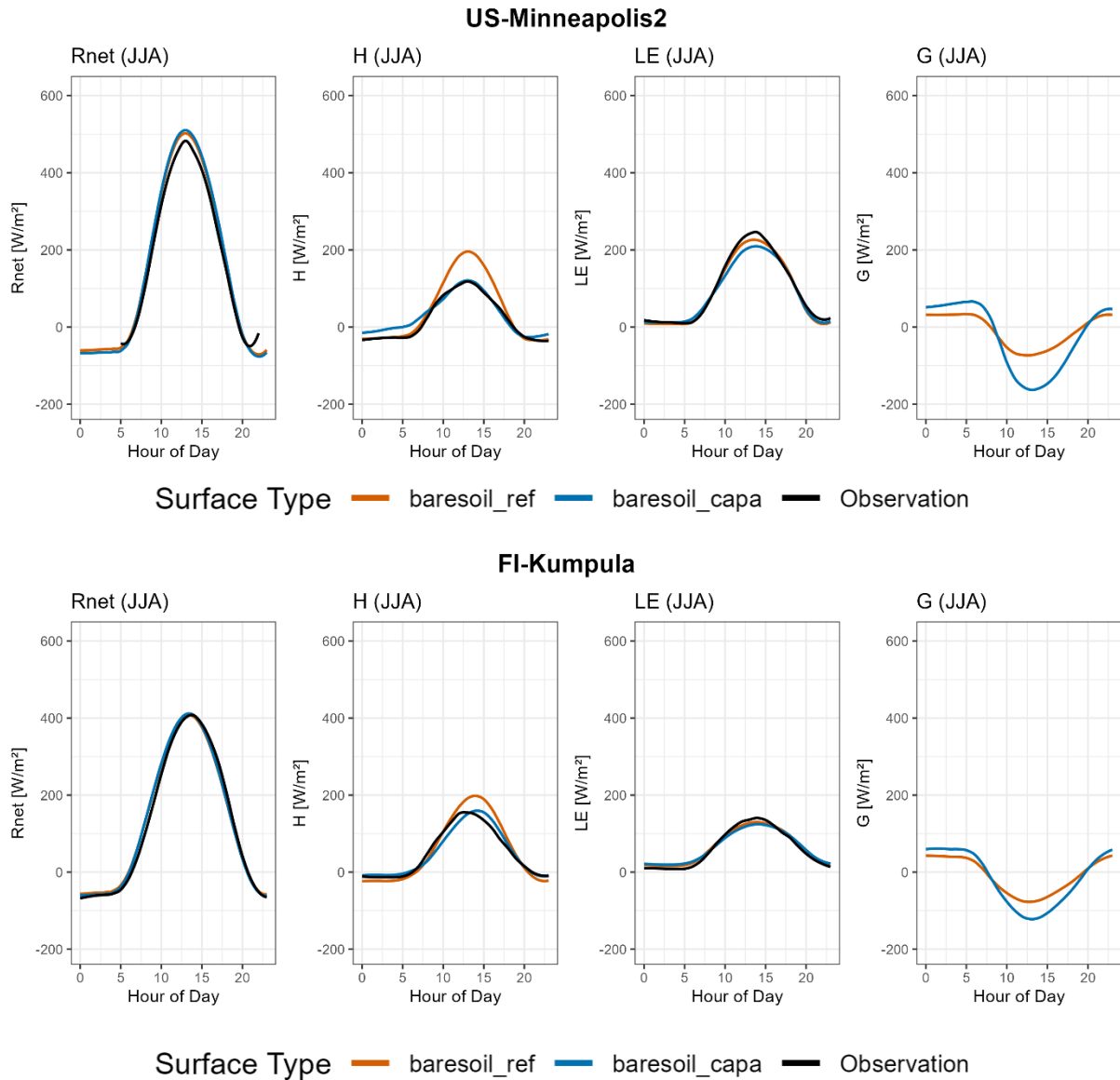
The effect of the soil-texture correction is site dependent: it reduces the sensible heat flux bias at some sites but increases it at others. The remaining large biases therefore cannot be attributed solely to the use of the default soil texture. To investigate these persistent biases, we performed a separate sensitivity experiment in which higher urban thermal conductivity and volumetric heat capacity were prescribed in the Baresoil configuration. This thermal-property experiment is used only as a diagnostic: it is not included in the final model configuration or in the results reported in the revised manuscript. It identifies soil thermal parameterization as an important avenue for future ORCHIDEE development.

Detailed analysis explaining how we figured out the soil texture issue and the identified underestimation of heat capacity and conductivity at semi-urban stations:

The largest biases in the summer mean daily maximum sensible heat for Baresoil occurred at US-Minneapolis2 (228 W m^{-2}), FI-Kumpula (204 W m^{-2}), US-Minneapolis1 (188 W m^{-2}), AU-Preston (183.8), GR-HECKOR (162 W m^{-2}), KR-Ochang (129 W m^{-2}), and US-WestPhoenix (113 W m^{-2}). Importantly, US-Minneapolis1 and US-Minneapolis2 have relatively low urban fractions and are therefore closer to semi-rural sites than to densely urbanized sites (respectively 21% and 5% urban). Moreover, the bias is of a similar magnitude in the bare-soil, natural-vegetation, and urban configurations (except for AU-Preston, where in the Urban configuration, bias decrease to 68 W m^{-2}). This indicates that it is not primarily introduced by the new urban parameterization, but as the reviewer suggested, instead reflects a more general limitation in the representation of surface energy partitioning in ORCHIDEE at these sites.

To investigate the source of the bias, we examined the summer mean diurnal cycles, and we show the results for two sites, US-Minneapolis2 and FI-Kumpula. We focus on the

baresoil simulation because the magnitude and timing of the bias are similar across the different model configurations. The corresponding figures are provided below.



The reference simulations, shown by the orange curves, indicate that the excessive daytime sensible heat cannot be attributed to an equivalent positive bias in net radiation or to the small underestimation of latent heat flux. Instead, the energy partitioning suggests that the simulated daytime ground heat uptake is too weak, leaving an excessive fraction of the available energy to be released as sensible heat. Although uncertainties related to observational energy-balance closure must also be considered, this result points towards the representation of soil heat storage and ground heat flux as an important contributor to the bias.

As a sensitivity test, we replaced the default soil thermal conductivity and volumetric heat capacity with the values used for urban surfaces in the new scheme (in the standard urban configuration, these values are applied only when the urban fraction exceeds 50%, this criterion is not met at US-Minneapolis2 or FI-Kumpula). For the sensitivity

experiment, the urban thermal properties were instead prescribed throughout the bare-soil simulation. The resulting simulations are shown by the blue curves.

This modification substantially reduces the daytime summer QH bias at most of the affected sites. For example, the bias decreases from 228.3 to 103.8 W m⁻² at US-Minneapolis2, from 188.1 to 68.0 W m⁻² at US-Minneapolis1, from 161.7 to 38.7 W m⁻² at GR-HECKOR, and from 128.9 to 5.6 W m⁻² at KR-Ochang. The complete results are shown below.

Site	Bias in summer daily max Qh (W m ⁻²)		
	Baresoil_ref (1st version result)	Baresoil_capa (change of soil capacity)	Revised with soil texture (new results, in revised manuscript)
AU-Preston	183.8	115.9	169.7
AU-SurreyHills	54.3	33.2	47.6
CA-Sunset	9.2	-109.6	25.5
FI-Kumpula	203.5	141.6	77.0
FI-Torni	82.2	41.5	-26.4
FR-Capitole	81.5	-57.2	78.6
GR-HECKOR	161.7	38.7	171.0
JP-Yoyogi	57.3	-29.1	25.0
KR-Jungnang	-10.8	-114.5	15.1
KR-Ochang	128.9	5.6	113.3
MX-Escandon	88.8	6.8	90.6
NL-Amsterdam	23.3	-16.2	43.5
PL-Lipowa	105.0	9.2	141.4
PL-Narutowicza	79.7	-22.0	100.0
SG-TelokKurau06	54.8	1.7	64.0
UK-KingsCollege	19.5	-78.7	38.2
UK-Swindon	8.7	-73.2	9.7
US-Baltimore	21.6	-64.2	29.8
US-Minneapolis1	188.1	68.0	144.4
US-Minneapolis2	228.3	103.8	168.9
US-WestPhoenix	112.6	-44.3	109.7

These sensitivity experiments strongly implicate the representation of subsurface thermal properties, particularly soil thermal conductivity, as an important source of the positive sensible heat bias. In the standard model, thermal conductivity and heat capacity are calculated from soil texture class and state variables, including porosity, soil moisture, liquid water content, and quartz fraction. The use of a limited number of broad soil texture classes, together with uncertainty in these state-dependent relationships, can produce substantial errors in simulated ground heat storage. This limitation is not specific to ORCHIDEE and is increasingly recognized within the land-surface modelling

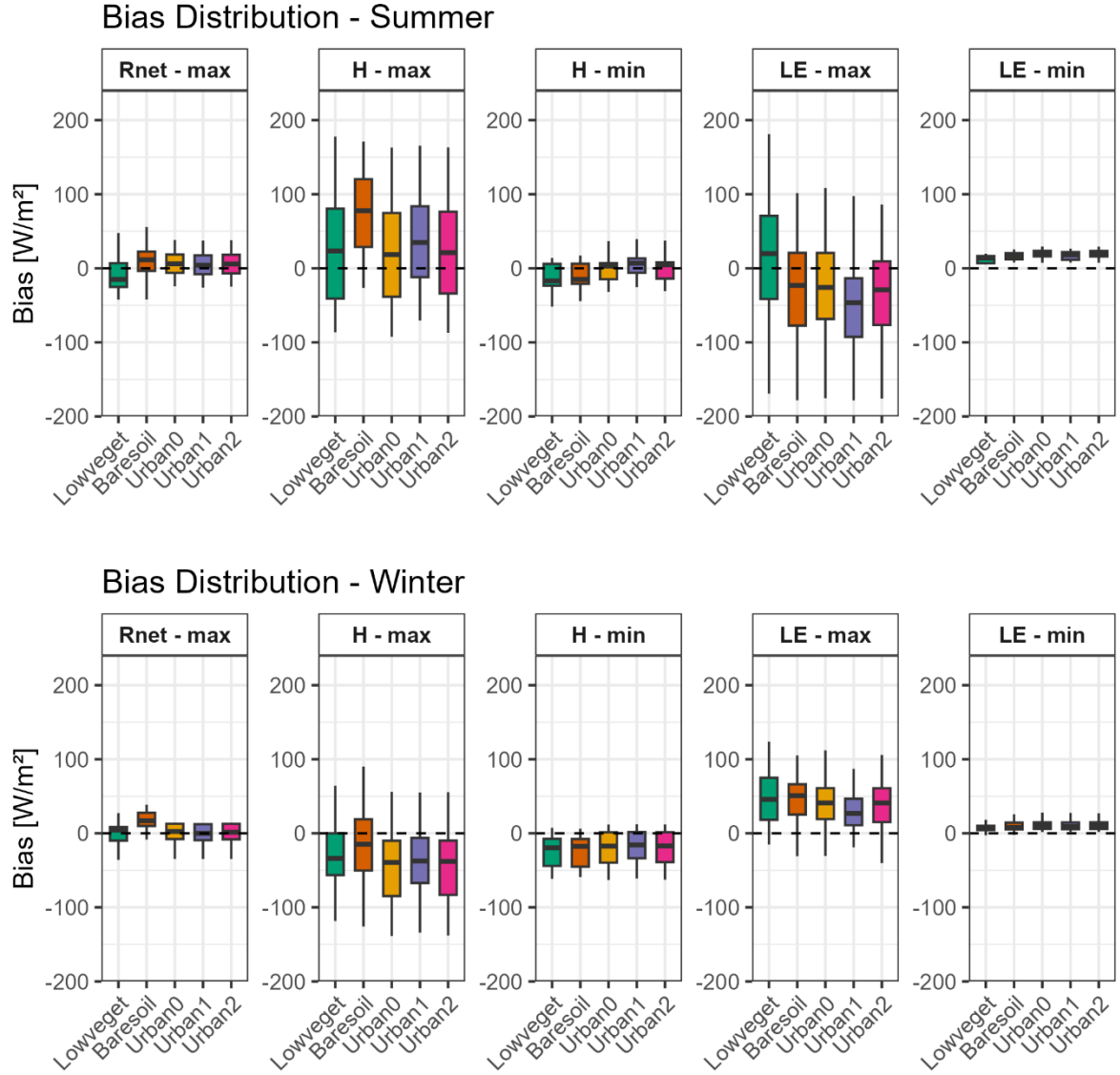
community, as discussed by Verhoef et al. (2026). A comprehensive revision of soil thermal parameterization is beyond the scope of the present urban-scheme development, but these results identify it as an important priority for future ORCHIDEE development.

We have added a discussion of this limitation to the revised manuscript:

“Still, all simulations tend to overestimate sensible heat, the bias exceeding 150 W m^{-2} at some stations. The largest summer sensible heat flux biases are common across the baresoil, vegetation, and urban simulations. Additional sensitivity experiments indicate that they are partly related to underestimated ground heat uptake associated with the representation of soil thermal properties: prescribing higher thermal conductivity and heat capacity substantially reduces the biases at most affected sites, all of which have urban fractions below 50%, highlighting soil thermal parameterization as an important target for future ORCHIDEE development (Verhoef et al., 2026).”

Finally, we revised the manuscript describing the results of our new simulations which now account for soil texture information from the Urban-plumber data (indicated in Table 1 of the manuscript), which changes some of the model outputs and performances:

New figure 3:



We adapted the description and analysis of the figure in consequence:

“The different parameterizations used in the model configurations lead to the largest differences for the maximum daily sensible heat flux. During the summer, the Baresoil simulations have a more pronounced overestimation compared to the urban simulations, showing a median bias close to 77 W m^{-2} . In contrast, the Urban0 and Urban2 simulations have a bias distribution centered closer zero (respectively having a median bias of 14 and 18 W m^{-2}), and 50% of biases between 66 and -36 W m^{-2} . The substantial change in simulated sensible heat is due to increased thermal conductivity in urban simulations, as already seen in Fig. 2. However, many simulations do not match the maximum daily observed sensible heat as closely as for AU-Preston, although the biases are significantly reduced. The Urban2 simulation introduces only minor changes compared to the Urban0 simulation for sensible heat. Urban1, on the other hand, increases the bias compared to Urban0 and Urban2 during summer, but still brings improvement compared to Baresoil.

Still, all simulations tend to overestimate sensible heat, the bias exceeding 150 W m^{-2} at some stations. The largest summer sensible heat flux biases are common across the baresoil, vegetation, and urban simulations. Additional sensitivity experiments indicate that they are partly related to underestimated ground heat uptake associated with the representation of soil thermal properties: prescribing higher thermal conductivity and heat capacity substantially reduces the biases at most affected sites, all of which have urban fractions below 50%, highlighting soil thermal parameterization as an important target for future ORCHIDEE development (Verhoef et al., 2026). In winter, maximum daily sensible heat is generally underestimated across the simulations. Baresoil has the least negative bias, although its bias distribution remains relatively broad, while Lowveget and the Urban simulations show similarly negative biases. This similarity suggests that the winter underestimation of sensible heat cannot be primarily attributed to the increased thermal conductivity introduced in the urban configurations. At the same time, all simulations show a positive bias in maximum latent heat flux, indicating that excessive partitioning of the available energy toward latent heat may contribute to the negative sensible heat bias. The absence of anthropogenic heat fluxes from the simulations may additionally contribute to the underestimation at some highly urbanized sites. For example, Urban-PLUMBER reports mean anthropogenic heat fluxes of 78.5 W m^{-2} at UK-KingsCollege and 92.7 W m^{-2} at KR-Jungnang.

Minimum sensible heat flux, which corresponds to the nocturnal sensible heat flux, is generally higher in cities than in natural areas due to the release of stored heat accumulated during the day and the persistence of heat transfer by convection at night. As expected, both in winter and summer, Baresoil and Lowveget simulations exhibit a negative bias, meaning they simulate nocturnal sensible heat that is too low. The Urban0, Urban1, and Urban2 simulations lead to improvements in the simulation of nocturnal sensible heat flux. Furthermore, the Urban1 simulation, which includes stronger imperviousness, further improves the representation of nocturnal sensible heat during both seasons. Still, Urban0, Urban1, and Urban2 show negative bias in minimum sensible heat fluxes in winter. This underestimation is unlikely to be primarily explained by latent heat flux, for which nighttime biases remain comparatively small. It may instead partly reflect an insufficient release of stored heat at night, as well as missing anthropogenic heat fluxes at highly urbanized sites.

Looking at maximum latent heat flux, Fig. 3 shows a clear seasonal contrast. In winter, biases are positive for all simulations, indicating an overestimation of daytime latent heat flux. In summer, biases are instead generally negative, indicating an underestimation of maximum latent heat flux, although the distributions remain broad and span both positive and negative values across stations. The summer underestimation may partly result from an insufficient representation of active vegetation and from the absence of irrigation in the model. This interpretation is supported by the Lowveget simulation, whose median summer bias is close to zero and which also shows a maximum sensible heat flux bias

centered around zero. However, the distribution of Lowveget latent heat biases remains very broad across stations, suggesting that highly local factors such as vegetation characteristics, vegetation management, and irrigation, which are not explicitly represented by the model configuration, strongly influence latent heat flux. The Urban0 and Urban2 simulations show only limited differences compared with Baresoil. In contrast, Urban1, with stronger imperviousness and therefore lower water availability for evapotranspiration, further increases the summer underestimation while reducing the winter overestimation. Biases in minimum latent heat flux are comparatively small but consistently positive in both seasons across all simulations, indicating an overestimation of nocturnal latent heat flux.

Overall, the urban configurations improve the representation of the surface energy balance through both a better simulation of net radiation and changes in energy partitioning. In summer, the large overestimation of maximum sensible heat flux in Baresoil appears to arise from combined biases: net radiation is imperfectly represented, while both daytime heat storage and latent heat flux are underestimated, leaving too much of the available energy to be transferred as sensible heat. The Urban configurations improve net radiation and substantially improve the representation of daytime heat storage, which strongly reduces the sensible heat bias. Maximum latent heat flux, however, remains generally underestimated, suggesting that evapotranspiration, including the representation of local vegetation characteristics and irrigation, is an important remaining source of error. In contrast, the Urban configurations do not improve sensible heat in winter, they exhibit a systematic underestimation of maximum sensible heat, consistent with an overestimation of maximum latent heat flux. Despite this limitation, the urban configurations systematically improve nocturnal sensible heat in both seasons relative to Baresoil and Lowveget.”

To summarize, we have rerun the simulations with soil texture information, which changes the performance of the model. It improves the energy balance over some stations and deteriorates it over other stations, but it highlights better the overestimation of sensible heat in summer resulting from an underestimation of latent heat, and the underestimation of sensible heat in winter resulting from a too high latent heat. The rest of the manuscript results (MAE, water balance), are also re computed based on these new simulations.

Specific comments/Technical corrections:

L19: perhaps missing a noun after “high-resolution convection-permitting”?
“Applications” or “simulations”?

Done

L81: Define PFTs here as it is the first time it appears in the text.

Done

L85: Missing “of” between “the geometry” and “buildings and streets”.

Done

L87: “cities buildings” should be “buildings in cities”.

Done

L92: The comma is not needed.

Done

L196: λ_s and λ_w should be given units as well. Also, the comma after “solid soil” is not needed.

The units are ($\text{W m}^{-1} \text{K}^{-1}$), we added them again after λ_s and λ_w to make it more evident as suggested by the reviewer.

L239 - 241: This sentence seems to be grammatically incorrect. Perhaps rewrite “ $f=2\text{m}$ ” into “ f is set to 2m ”?

We updated the text as:

“ f is set to 2 m^{-1} ”

Eq. (17): “ β ” should be “ βET ”

Done

L261: and becomes closer to zero as soil moisture gets limiting

Done

L262: the different ET components

Done

L385: refer readers to Figure 2 here.

Done

Figure 3: describe in the caption what each element of the boxes represents (e.g., 95th percentile, 75th percentile, etc.).

We added in the caption:

“For each boxplot, the central horizontal line indicates the median, the lower and upper edges of the box indicate the 25th and 75th percentiles, respectively, and the whiskers extend to the most extreme values within 1.5 times the interquartile range (IQR). The dashed horizontal line indicates zero bias.”

Anonymous Referee #2, 13 May 2026

Report #2

Submitted on 13 May 2026 Anonymous referee #2	
Notes for the submission of interactive comments	
Anonymous: Yes No	
Formal manuscript rating and recommendation to the editor (visible to authors and reviewers only)	
1) Scientific significance Does the manuscript represent a substantial contribution to modelling science within the scope of this journal (substantial new concepts, ideas, or methods)?	Excellent Good Fair Poor
2) Scientific quality Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)? Do the models, technical advances and/or experiments described have the potential to perform calculations leading to significant scientific results?	Excellent Good Fair Poor
3) Scientific reproducibility To what extent is the modelling science reproducible? Is the description sufficiently complete and precise to allow reproduction of the science by fellow scientists (traceability of results)?	Excellent Good Fair Poor
4) Presentation quality Are the scientific results and conclusions presented in a clear, concise, and well structured way (number and quality of figures/tables, appropriate use of English language)?	Excellent Good Fair Poor
For final publication, the manuscript should be accepted as is accepted subject to technical corrections accepted subject to minor revisions reconsidered after major revisions rejected Please note that this rating only refers to this version of the manuscript!	
Were a revised manuscript to be sent for another round of reviews: I would be willing to review the revised manuscript. I would not be willing to review the revised manuscript.	
Referee Comment: 'Comment on egusphere-2026-551'	

This study presents the implementation of an “urban” tile in the ORCHIDEE land cover and vegetation model for future applications in climate modeling at various spatial scales. The evaluation of energy fluxes is based on the Urban-Plumber database, and shows a nearly systematic improvement in modeling of sensible and latent heat fluxes which demonstrates the added value of the new parameterization. However, some biases persist in ORCHIDEE: it would be relevant to suggest some ways to improve the model and to assess the model’s sensitivity to certain parameters (e.g. roughness and drag coefficient). For the hydrological component, one limitation of the study is the lack of data needed to conduct an objective evaluation. Consequently, this section presents only a sensitivity analysis comparing the different ORCHIDEE configurations, which makes it difficult to critically analyze the results. We would expect to gain a better understanding of the specific characteristics of certain sites that result in a “non-linear” response from the Urban2 version comparing to Urban1.

We thank the reviewer for these suggestions. We agree that the remaining biases require further investigation and that the lack of hydrological observations limits the interpretation of the water-budget results.

For the energy-budget evaluation, we examined the sensitivity of the simulated sensible heat flux to the aerodynamic parameters, including roughness length and the drag coefficient. These diagnostics indicate that the remaining sensible heat biases cannot

be explained primarily by the aerodynamic formulation. Further analysis, detailed in our response to Reviewer 1, instead points to two main limitations: the prescribed soil thermal conductivity and heat capacity at sites with relatively low urban fractions, and an underestimation of latent heat flux at many sites. We have revised the manuscript accordingly and now identify the refinement of the soil thermal properties and evapotranspiration parameterizations as priorities for future model development.

We also agree that the hydrological analysis remains limited because runoff, soil moisture, and the other water-budget components are not observed at the Urban-PLUMBER sites. The results should therefore be interpreted as a sensitivity analysis of the different ORCHIDEE configurations rather than as an objective hydrological validation, but we plan to explore performance at the catchment scale in future work. We have revised the discussion of the apparently non-monotonic response in Urban2 to clarify this limitation and identify possible mechanisms to investigate in future work.

We replaced:

“The Urban2 simulation, which includes a lower level of imperviousness, also leads to an increase in surface runoff (with a corresponding decrease in subsurface runoff), but the magnitude of change is much smaller. Interestingly, at a few stations (e.g., KR-Ochang and SG-TelokKurai), Urban2 even shows a slight increase in subsurface runoff, suggesting a non-linear response of the water balance to moderate imperviousness. Urban2 also results in a decrease in surface runoff at some stations (e.g., FI-Torni and FI-Kumpula).

with:

“The Urban2 simulation, which represents a more moderate reduction in saturated hydraulic conductivity, generally increases surface runoff and decreases subsurface runoff, although the magnitude of these changes is much smaller than in Urban1. However, the response is not consistent across all sites: Urban2 slightly increases subsurface runoff at KR-Ochang and SG-TelokKurai and decreases surface runoff at FI-Torni and FI-Kumpula. These site-dependent changes indicate that the annual water-budget response to a moderate conductivity reduction is not necessarily monotonic. Although saturated hydraulic conductivity is scaled linearly with imperviousness, the resulting water-budget response may be nonlinear because runoff generation and vertical water transfer depend on infiltration thresholds, rainfall intensity, antecedent soil moisture, soil texture, evapotranspiration, and changes in soil-water storage.”

P2, L42

"urban bulk schemes" : add references

Done

P2, L61

"Chancibault et al., 2014": The reference listed here does not appear to be correct; I would suggest Stavropoulos-Laffaille et al. (2018) instead. > DOI: 10.5194/gmd-11-4175-2018

Done

P3, L66-68

You should add a point about the lack of data for process documentation and model evaluation.

We agree that the limited availability of observations is an important constraint on both process understanding and model evaluation. The sentence highlighted by the reviewer was removed when the urban hydrology section was restructured; however, this point is now addressed explicitly at the end of the revised paragraph:

"Progress is also constrained by the limited availability of observations, as urban sites measuring turbulent fluxes rarely provide all water-balance components required to evaluate runoff, irrigation, drainage, and storage changes consistently at the same spatial scale (Jongen et al., 2024)."

P3, 79

Please specify the model name using ORCHIDEE here.

The full name of the ORCHIDEE (Organising Carbon and Hydrology In Dynamic Ecosystems) model has now been specified at its first occurrence.

P3, L85

"...by having a lower albedo"

It is important to note that we are referring here to a relative albedo that takes into account the geometry of the urban canopy, not a surface albedo.

Yes, we changed the sentence in : *"First, cities modify the absorbed shortwave radiation through changes in the effective urban canopy albedo relative to rural areas because of the materials used in cities and the geometry of buildings and streets (Kotopouleas et al., 2021; Qin, 2015)."*

P3, 87

"This lower albedo reflects less solar radiation leading to an increase in net radiation (Rnet), and ultimately in higher surface temperatures"

Some caution is warranted regarding this cause-and-effect relationship: (1) net radiation also increases because of radiative trapping of infrared emissions from surfaces inside the urban canopy, and (2) surface temperatures also rise due to thermal properties of materials.

We have revised the text to clarify that urban net radiation is influenced by both enhanced shortwave absorption and longwave radiative trapping within the urban canopy, while higher surface temperatures also depend on the thermal properties of urban materials: *“Urban geometry also affects longwave exchanges by trapping infrared radiation within the canopy. Together with the high heat capacity and thermal conductivity of urban materials, these radiative effects contribute to increased heat storage and often higher urban surface temperatures.”*

P3, L94

"This is the result of the higher thermal capacity and conductivity of urban materials compared to natural areas..."

The amount of surface area available for heat storage—which is linked to urban geometry—is also a key factor.

Yes, we modified the sentence into:

“This larger heat storage results from both the higher heat capacity and thermal conductivity of urban materials and the greater surface area available for heat exchange and storage within the urban canopy, including roofs, walls, and roads”

P4, L120

"version 2.2 of ORCHIDEE (rev. 8133)."

What does “rev. 8133” mean? Could you clarify and provide the reference?

“Rev. 8133” refers to the specific source-code revision of ORCHIDEE version 2.2 used in this study. This identifier defines the exact code snapshot and is therefore provided to ensure reproducibility. The corresponding repository information and access details are given in the Code and data availability section.

We have clarified the text as follows:

“This study uses ORCHIDEE version 2.2, revision 8133; further details and access information are provided in the Code and data availability section.”

P4, L124

"Hydrological and energy fluxes associated with vegetation ..."

Change by "Water and energy fluxes associated with vegetation and natural soils ..."

"interception" > please clarify (intercepted by vegetation?)

"evaporation" > please clarify (soil evaporation?)

Yes the sentence is now: “Water and energy fluxes associated with vegetation and natural soils, including transpiration, evaporation of water intercepted by the vegetation canopy, and soil evaporation, ...”

P5, L125

It is necessary to clarify what you mean by “flux-aggregation approach”

We added: “(calculated separately for each PFT and then aggregated to the grid-cell scale as area-weighted fluxes)” after flux-aggregation approach

Also it is essential to include a figure illustrating the discretization/description of the grid and the soil column for modeling hydrology and thermal processes.

We now include this new figure to present the model:

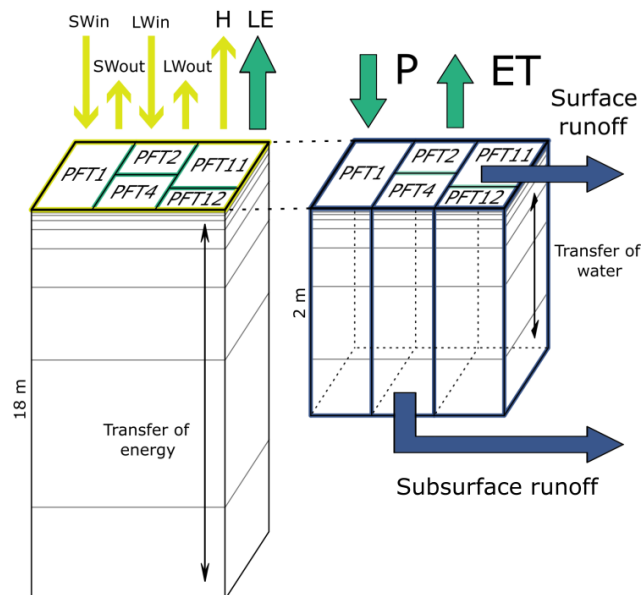


Figure 1: Schematic representation of the energy- and water-budget calculations in ORCHIDEE for a single grid cell, its plant functional types (PFTs), and the three associated soil columns representing bare soil, high vegetation, and low vegetation. Yellow arrows denote fluxes calculated directly at the grid-cell scale, green arrows denote fluxes calculated separately for each PFT and subsequently aggregated to the grid-cell scale, and dark-blue arrows denote fluxes calculated independently for each soil column. For example, sensible heat flux (H) is calculated at the grid-cell scale, evapotranspiration (ET) at the PFT scale, and surface runoff and drainage at the soil-column scale. Each soil column is discretized into 11 hydrological layers extending to a depth of 2 m, while the thermal discretization extends to 18 m. Abbreviations: P, precipitation; ET, evapotranspiration; H, sensible heat flux; SW, shortwave radiation; LW, longwave radiation; PFT, plant functional type.

P5, 144

"First, the model estimates change in surface temperature based on the downwelling and upwelling radiative terms ..."

Requires clarification on how the surface temperature (or the temperature of the first soil layer) is calculated

We have clarified that T_{surf} is the grid-cell surface temperature, distinct from the temperature of the uppermost soil layer. It is calculated by implicitly solving the

complete surface energy balance, including radiative, turbulent, and conductive heat exchanges. The resulting surface temperature is then used as the upper boundary condition for updating the soil-temperature profile.

We modified the text as:

“At each time step, ORCHIDEE implicitly solves a single grid-cell surface temperature, T_{surf} , from the surface energy balance, accounting for radiative, turbulent, and conductive heat exchanges. T_{surf} is distinct from the temperature of the uppermost soil layer and provides the upper boundary condition for the implicit soil heat-diffusion scheme. The net radiation is then calculated as:”

P5, 147

" ϵ is the emissivity (1, unitless)"

Why is emissivity equal to 1 ?

In the standard ORCHIDEE configuration used in this study, surface longwave emissivity is fixed to 1 unless a site-specific value is prescribed. We have clarified that this is a model assumption corresponding to a blackbody approximation, rather than a general physical property of land surfaces.

We changed the text in:

“ ϵ is the surface longwave emissivity, fixed to 1 in these simulations, corresponding to a blackbody approximation in the thermal infrared,”

P5, L149-152

"Each vegetation PFT has its own value of albedo (α)"

The albedo should be noted for a leaf consistency with the with the rest of the text and the equations ?

The text on soil albedo needs to be reworded, ex. "For bare soil, the specification of albedo must take into account the spatial variability of reflectance, resulting from composition, moisture, and texture heterogeneities".

We thank the reviewer for pointing this out. We have corrected the notation to α_{leaf} for consistency and replaced the description of bare-soil albedo with:

The use of MODIS-derived bare-soil albedo fields accounts for the spatial variability of soil reflectance resulting from heterogeneities in soil composition, moisture, and texture.

P6, L156

"... weighted by the fractional areas of the PFTs of the grid cell (Equation 3)."

Change by "... weighted by the fractional areas of the PFTs of the grid cell :"

Done

P6, L159

"... of each vegetated PFTs"

Change by "... of each vegetated PFT"

Done

P6, L164

"The latent and sensible heat fluxes are calculated respectively thanks to Eq. (4) and Eq. (5)."

Change by "The sensible and latent heat fluxes are calculated respectively thanks to Eq. (4) and Eq. (5):" to be consistent with the order of equations

Done

P6, 172

"The calculation of C_d (also expressed ..."

If I understand correctly, you should add "for a given PFT"

Done

P7, L183

"... the soil thermal conductivity and the soil heat capacity, and both vary with soil texture and with the water content of each soil layer."

You should remove "and with the water content" because this point is discussed later.

Done

P7, L191-193

"Finally, the soil thermal properties also change with the presence of water inside the soil layers. Eq. (9) shows the representation of the soil heat capacity as a function of the fraction of water inside the layer. Eq. (10) represents the evolution of soil thermal conductivity as a function of the same fraction:"

I would put it simply as "Finally, the soil thermal properties also change according to the soil water content of each soil layer:"

You mention that thermal properties depend on the soil moisture content, but the soil column used for modeling water exchange and heat transfer is not the same depth. How do you deal with this?

We have simplified the introductory sentence as suggested and clarified how soil moisture information is transferred between the hydrological and thermal discretizations. Over the common upper 2 m, volumetric soil moisture calculated by the hydrological scheme is interpolated from hydrological nodes to thermal-layer interfaces, while the saturation degree and layer-integrated water content are transferred directly. Below the hydrological domain, the saturation degree and volumetric water contents are held equal to those of the deepest hydrological layer, while the layer-integrated water content is adjusted according to the thickness of each thermal layer.

We added to the text:

“The hydrological and thermal discretizations share the upper 2 m of the soil column. Within this common domain, hydrological soil-moisture variables are transferred to the thermal grid, including interpolation of volumetric water content to the thermal-layer interfaces. Below 2 m, the saturation degree and volumetric water contents are held equal to those of the deepest hydrological layer, while the total water content is adjusted for the thickness of each thermal layer.”

P7, L196

"where, λ_s and λ_w are the heat conductivities ..."

Subscripts are capitalized in equations and lowercase in the text.

Done

P8, L215

"Where $K(\theta)$..."

Change by "where $K(\theta)$..."

Done

P9, L225

"Reference saturated hydraulic conductivity (K_{sref}) at the depth $z_{lim} = 0.30m$, derived from the dominant USDA ..."

What does this “ z_{lim} ” depth refer to?

Explain USDA

U.S. Department of Agriculture

P9, L230

"... represented by a factor $FK_{roots}(z,c)$, applied only"

Could you clarify what is "c" ?

Here, (c) denotes the soil-column index. The factor $F_{Kroots}(z, c)$ depends on the type of soil column: it represents the root-induced enhancement of saturated hydraulic conductivity in vegetated soil columns, whereas it is set to 1 for the bare-soil column. We have clarified this notation in the revised manuscript.

P9, L239-240

Check unit of "f" in m-1

We corrected "f = 2 m" into: "f is set to 2 m⁻¹"

P9, L242

"Infiltration is computed using a Green–Ampt type approach"

Add a reference

Done

P10, L250-256

Add the units to the terms E_{total} and βET

Done

P10, 262

"Once the latent heat flux is calculated with Eq. (5) and Eq. (17), ORCHIDEE calculates the different ET, as:"

I suggest "Once the latent heat flux is calculated with Eq. (5) and Eq. (17), ORCHIDEE calculates the different terms contributing to total evapotranspiration:"

Done

Section 2.2.1

This section requires more details to fully understand how the various parameters are defined and calculated/aggregated: Is the albedo a composite (or relative) albedo? How is roughness calculated? ... And here again, a figure would be helpful for understanding.

P11, L288 (and Eq 22)

"the vegetation root factor $F_{Kroots}(z,c)$ with a constant urban factor $K_{facturban}$..."

To ensure greater consistency in the notation, shouldn't we replace $K_{facturban}$ with F_{Kurban} ?

Yes, we replaced it by F_{urb}

P12, Eq 24

The “max” is unnecessary because, by definition, $1 - 0.9f_{imp}$ is always greater than or equal to 0.1

Done

P12, L302-304

Please provide references.

We modified the sentence into:

“When site-specific information is unavailable, the modeler may prescribe a representative value. For example, the Copernicus Urban Atlas distinguishes discontinuous medium-density urban fabric ($0.30 < f_{imp} \leq 0.50$), discontinuous dense urban fabric ($0.50 < f_{imp} \leq 0.80$), and continuous urban fabric ($f_{imp} > 0.80$; European Environment Agency, 2020).”

European Environment Agency. (2020). Copernicus Land Monitoring Service (CLMS): Urban Atlas Land Cover/Land Use and Street Tree Layer 2012 and 2018—Product User Manual, version 6.3 (6.3.1).

(https://library.land.copernicus.eu/products/Urban_Atlas_Land_Cover-Land_Use_and_Street_Tree_Layer_2012_and_2018_PUM_v6.pdf)

P12, Fig1

Add unit

Done

P12, L313

"... for the grid cells where urban areas cover more than 50% of the grid cell"

I am not sure to understand what is done if the fraction is less than 50%?

For grid cells with an urban fraction of 50% or less, the thermal conductivity and volumetric heat capacity are not modified and retain the values calculated by the standard ORCHIDEE formulation as functions of soil texture and soil moisture. No weighted averaging between natural and urban thermal properties is applied in the current implementation. We have clarified this point in the revised manuscript:

“In grid cells where the urban fraction is below 50%, the standard ORCHIDEE values calculated from soil texture and soil moisture are retained.”

P12, L321

"The thermal conductivity is set to $3.24 \text{ W m}^{-1} \text{ K}^{-1}$ and the heat capacity to $1890 \text{ kJ m}^{-3} \text{ K}^{-1}$ "

In slab-based approaches (as used in SURF, Wouters et al.), thermal properties are adjusted to account for surface density (and can therefore vary depending on the urban typology). How are these properties defined and selected in this study?

In the present implementation, thermal conductivity and volumetric heat capacity are prescribed as fixed effective urban values and do not vary with urban morphology or typology. The values of $3.24 \text{ W m}^{-1} \text{ K}^{-1}$ and $1890 \text{ kJ m}^{-3} \text{ K}^{-1}$ were adopted from the Noah land-surface model. The prescribed thermal conductivity is relatively high, including compared with the effective value of $1.55 \text{ W m}^{-1} \text{ K}^{-1}$ used in SURF, and therefore provides a first-order bulk representation of enhanced heat conduction in urban areas. However, unlike SURF, these properties are not explicitly derived from surface-area density or other morphological parameters.

We added: *“These values are not adjusted according to urban typology, building morphology, or surface-area density.”*

P13, L343

"The benchmark simulations are the Baresoil and the Lowveget simulations...."

The definition of land cover and land use characteristics is not very clear. Is it only the urban portion of the site that is classified as “lowveg” or “baresoil”? For example, in AU-PRESTON, 23% of the area is covered by trees—are these trees taken into account in the different configurations? Table S1 should be clarified and completed with all characteristics.

We have clarified the treatment of land cover in the different simulations. Only the fraction identified as urban in the Urban-PLUMBER site characteristics is reassigned between experiments. The other observed land-cover fractions, including trees, low vegetation, bare soil, and water, are retained in all configurations. Thus, for example, the 22.5% tree fraction at AU-Preston is represented in every simulation, while its 62% urban fraction is represented as bare soil in Baresoil, as low vegetation in Lowveget, and using the new urban PFT in Urban0, Urban1, and Urban2.

We modified the text to:

“The benchmark simulations are Baresoil and Lowveget. In all simulations, the observed non-urban land-cover fractions of each site, including trees, low vegetation, bare soil, and water, are retained. Only the fraction classified as urban is represented differently among the experiments. In Baresoil, the urban fraction is assigned bare-soil properties, corresponding to the default treatment of urban areas in ORCHIDEE. In Lowveget, the urban fraction is assigned the properties of the dominant low-vegetation type in the surrounding area, either grassland or cropland, providing a non-urban baseline.”

The complete land-cover and land-use fractions for all sites are already provided in Table 1. We have therefore not detailed these values for each simulation in Table S1, as doing so would considerably increase the size of the table, and make it difficult to read.

P13, L351

"For each site, the model is first spun up for 40 years, using repeated cycles of 10 years... "

The data from the Urban-Plumber sites does not cover such long periods of time. How are atmospheric forcings built over long time periods for each site (using which data) ?

As stated in the manuscript, the 40-year spin-up is generated by repeating the 10-year meteorological forcing available for each Urban-PLUMBER site four times. No additional atmospheric forcing dataset is used. We have slightly rephrased the sentence to make this procedure more explicit.

We changed:

"For each site, the model is first spun up for 40 years by repeating four times the 10-year meteorological forcing, until soil moisture and temperature reach a stable seasonal cycle."

Into:

"For each site, the model is first spun up for 40 years by repeating four times the 10-year meteorological forcing provided by Urban-PLUMBER, until soil moisture and temperature reach a stable seasonal cycle."

P15, Table2

Could you add the roughness ?

Roughness is not prescribed as an independent parameter in these simulations. For the urban PFT, the aerodynamic roughness length is diagnosed from the site-specific building height reported in Table 1 using the standard ORCHIDEE formulation described in Eq. (6), also expressed as $1/(rah \cdot u)$ where u is the wind speed and rah the aerodynamic resistance, with the ratio between roughness length and PFT height, z_0 , fixed to 0.0625. The resulting drag coefficients are subsequently aggregated at the grid-cell scale according to the PFT fractions. Consequently, roughness varies among sites and cannot be represented by a single value for each simulation. The standard ORCHIDEE aerodynamic treatment is retained for the Baresoil and Lowveget configurations.

We modified:

“Turbulent transfer of the fluxes between the surface and the atmosphere, is represented thanks to the drag coefficient C_d . The calculation of C_d (also expressed as $1/(r_{ah} \cdot u)$ where u is the wind speed and r_{ah} the aerodynamic resistance) for a given PFT in the standalone version of ORCHIDEE, is calculated with the following equation:

$$C_d = \frac{k^2}{\ln\left(\frac{z_{lev}}{h_{PFT} \cdot z_0}\right)^2} \quad (6)$$

where k is the Von Karman constant; z_{lev} is the height of the first atmospheric layer (10m in standalone simulations); h_{PFT} is the height of the canopy of the considered PFT (fixed) and z_0 is a factor that estimates the roughness height above the canopy and is fixed to 0.0625 (1/16, default value in ORCHIDEE for all PFTs in standalone simulations).”

To:

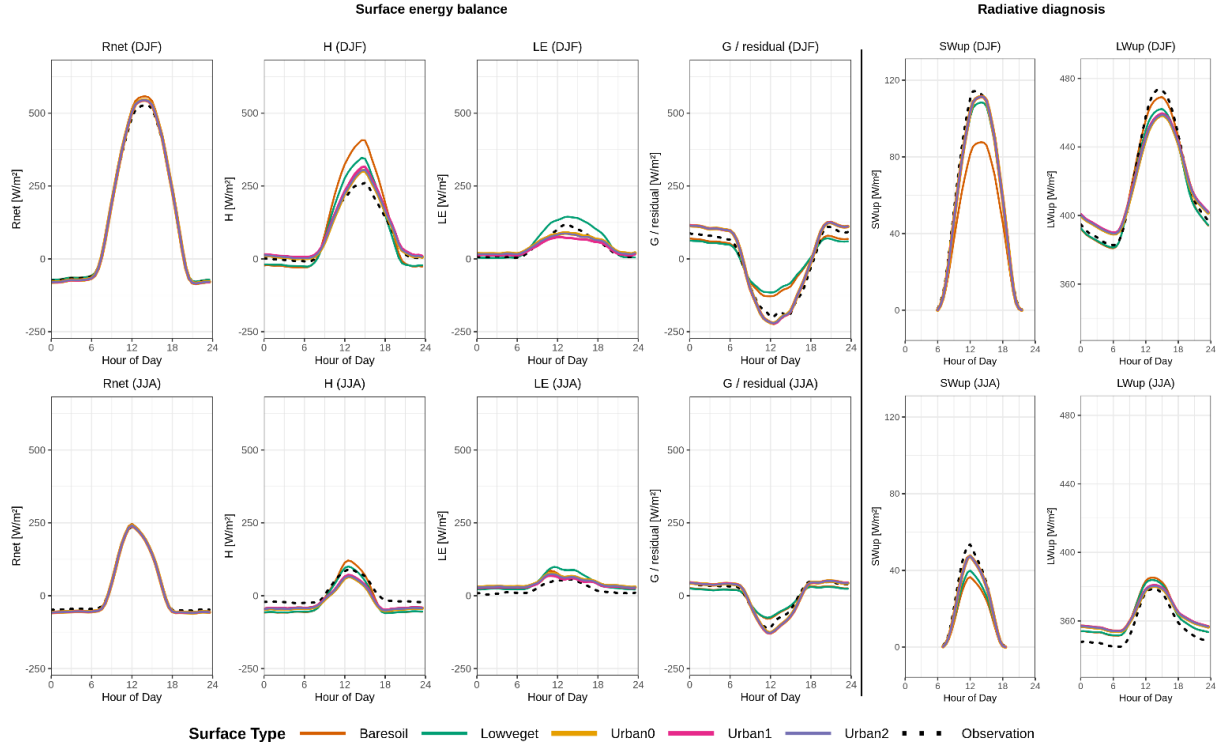
“Turbulent transfer of the fluxes between the surface and the atmosphere, is represented thanks to the drag coefficient C_d . For each PFT in the standalone version of ORCHIDEE, C_d which can also be expressed as $1/(r_{ah} \cdot u)$, where u is the wind speed and r_{ah} the aerodynamic resistance, is calculated as follows:

$$C_d = \frac{k^2}{\ln\left(\frac{z_{lev}}{h_{PFT} \cdot z_0}\right)^2} \quad (6)$$

where k is the Von Karman constant; z_{lev} is the height of the first atmospheric layer (10m in standalone simulations); h_{PFT} is the height of the canopy of the considered PFT (fixed) and z_0 is the dimensionless roughness-length-to-height ratio and is fixed to 0.0625 (1/16, default value in ORCHIDEE for all PFTs in standalone simulations). Thus, the product $h_{PFT} \cdot z_0$ in Eq. (6) corresponds to the aerodynamic roughness length. The aerodynamic roughness length is therefore not prescribed as an additional independent parameter but is diagnosed from the PFT height. For the urban PFT, h_{PFT} corresponds to the prescribed site-specific building height; consequently, its aerodynamic roughness varies among sites. Because ORCHIDEE calculates a single energy budget per grid cell, the PFT-level drag coefficients are subsequently aggregated to the grid-cell scale using the corresponding land-cover fractions.”

Section 4.1

- It seems to me that the Urban-Plumber database provide incoming and upwelling solar and IR radiation terms. It would be useful to compare observed and simulated upwelling S and L to understand the errors noted for Rnet.



Thank you for this suggestion. We have added the observed and simulated upwelling shortwave and longwave radiation components to Fig. 2. The incoming shortwave and longwave radiation terms are prescribed from the Urban-PLUMBER atmospheric forcing and are therefore identical across simulations, so they are not shown.

The figure displays the new simulations accounting for the soil texture information.

The radiative decomposition shows that Baresoil substantially underestimates upwelling shortwave radiation, while the urban configurations are closer to the observations because of their modified albedo. The urban configurations nevertheless slightly underestimate upwelling longwave radiation around midday. The remaining positive bias in Rnet therefore results rather from an underestimation of emitted longwave radiation.

We changed the text in consequence:

“However, looking at upwelling shortwave (SWup) and longwave (LWup) components reveals compensating differences among the configurations. The grid-cell albedo increases from 0.110 in Baresoil and 0.131 in Lowveget to 0.141 in Urban0, Urban1, and Urban2. Consequently, the urban configurations reflect more shortwave radiation than the benchmark simulations: Baresoil produces the lowest SWup, Lowveget intermediate values, and the three urban configurations the highest values. The urban configurations also modify the diurnal cycle of upwelling longwave radiation. Relative to Baresoil and Lowveget, they emit less longwave radiation around midday and more during the night, particularly in summer. This pattern indicates a reduced diurnal amplitude of surface

radiative temperature and is consistent with the greater effective thermal inertia resulting from the changes in heat capacity and thermal conductivity.”

- Why is the albedo prescribed to 0.141 when it is informed equal to 0.151 in the Urban-Plumber database (Table 1) ?

The value of 0.151 reported in Table 1 corresponds to the albedo assigned to the urban PFT. However, R_{net} is calculated using the grid-cell albedo, obtained by weighting the albedos of all PFTs by their respective land-cover fractions. At AU-Preston, the grid cell also contains tree, low-vegetation, and bare-soil fractions of 0.225, 0.15, and 0.005, respectively. The resulting grid-cell albedo is therefore 0.141. We have revised the text to clarify that 0.141 refers to the grid-cell albedo, whereas 0.151 is the albedo prescribed to the urban PFT:

“The grid-cell albedo increases from 0.110 in Baresoil and 0.131 in Lowveget to 0.141 in Urban0, Urban1, and Urban2. The albedo values discussed here correspond to the effective grid-cell albedo used to calculate net radiation, rather than to the albedo assigned to an individual PFT. The grid-cell albedo is calculated as the land-cover-fraction-weighted mean of the PFT-level albedos.”

- Even though the heat storage flux G is not available in the Urban-Plumber database, it would be possible to compare the “residual term” of the energy budget (i.e., $R_{net} - H - LE$), given that there may be an additional anthropogenic effect (Fig 2). Is G calculated as the residual term in ORCHIDEE ?

Thank you for this suggestion. We have added the observational energy-balance residual to Fig. 2 for a qualitative comparison with the ground heat flux simulated by ORCHIDEE. Because the ORCHIDEE sign convention defines G as negative when heat is transferred into the substrate, we plot the observational residual as $H+LE-R_{net}$, i.e. the negative of the conventional residual $R_{net}-H-LE$.

In ORCHIDEE, G is not calculated as a residual of the surface energy budget. It is explicitly computed from the temperature gradient between the surface and the upper soil or urban-substrate layer, as described in Eq. (7), and provides the upper boundary condition for the soil heat-diffusion equation. We have clarified this distinction in the revised manuscript and discuss the comparison as follows:

“During summer, Baresoil and Lowveget underestimate the magnitude of daytime heat storage, as their simulated G values are less negative than the observational energy-balance residual. The urban configurations reproduce the observed daytime magnitude more closely, although their slightly more negative values indicate an overestimation of heat storage. At night, the urban configurations also agree more closely with the observational residual, but their more positive values suggest that nocturnal heat release is overestimated. This comparison remains qualitative because the colored curves represent the ground heat flux explicitly calculated by ORCHIDEE (eq. 7),

whereas the black curve is an energy-balance residual rather than a direct observation of G. The observational residual may additionally include contributions from unaccounted anthropogenic heat and net horizontal advective transport, as well as errors associated with measurement uncertainty.”

- You suggest that the negative bias in winter for H could be related to the anthropogenic fluxes that are not considered in ORCHIDEE: while this may certainly play a role, in this case the biases appear to be equivalent between H (negative) and LE (positive), which suggests that it is the competition between the two fluxes that is not modeled correctly.

We change the sentence into:

“The comparable negative bias in H and positive bias in LE suggests that the main issue lies in the simulated partitioning of available energy between sensible and latent heat fluxes.”

P19, L441

"...due to the release of stored heat accumulated during the day."

You could add "and the persistence of heat transfer by convection at night"

Done.

P19, Fig3

Could you clarify whether the data in the boxplots represent the daily bias for all modeled days and all sites, or an average bias per site (i.e., 20 values per boxplot)?

Thank you for requesting this clarification. The boxplots do not pool the daily biases from all sites. For each site, simulation, season, and flux, we first calculate the observed and simulated daily maximum or minimum using only matching model–observation timestamps. The daily simulated-minus-observed differences are then averaged over all valid days at that site, yielding one site-level mean bias. Each boxplot therefore contains 21 values (There are 20 sites but Minneapolis has two datasets depending on the wind directions), one per site.

We have clarified this procedure in the text and in the caption of Fig. 3. We retained the site-level representation because it gives equal weight to each site, irrespective of the duration of its observational record. Pooling all daily values would give greater weight to sites with longer records and would represent variability among site-days rather than consistency of model performance across sites. In addition, consecutive daily biases from the same site are not statistically independent.

We added:

“For each site, simulation, season, and flux, the daily simulated-minus-observed differences are calculated using only matching model–observation timestamps and are

then averaged over all valid days. Each site therefore contributes one value to each boxplot, so that all sites are weighted equally regardless of the length of their observational record. This approach evaluates model performance during both the daytime and nighttime phases of the diurnal cycle.”

L21, L485

"...except for CA-Sunset, UK-Swindon, and FI-Kumpula, where performance is slightly degraded"

We notice in Fig5 a significant drop in H score for the Phoenix site, which you don't mention in the text. Could you explain ?

Thank you for pointing this out. US-WestPhoenix indeed exhibits a more pronounced degradation in sensible-heat-flux performance than the other sites mentioned in the text. The H MAE increases from 28.7 Wm⁻² in Baresoil to 37.2 Wm⁻² in Urban1. This degradation is common to all three urban configurations, with MAEs of 38.1, 37.2, and 37.8 Wm⁻² for Urban0, Urban1, and Urban2, respectively. It is therefore not primarily caused by the Urban1 imperviousness formulation.

Urban configurations substantially reduce the mean negative R_{net} bias at this site, from -18.9 in Baresoil to -1.0 W m⁻² in Urban1. This improvement results from reductions in both upwelling radiative components. Based on the mean summer daytime cycle, Urban1 decreases SW_{up} by approximately 9 W m⁻² relative to Baresoil, consistent with the decrease in grid-cell albedo from approximately 0.206 to 0.190. It also decreases LW_{up} by approximately 11 W m⁻², consistent with its lower daytime radiative surface temperature. Together, these changes increase daytime R_{net} by approximately 20.9 W m⁻², which most of this additional available energy is transferred to sensible heat.

US-WestPhoenix has an urban fraction of 48%, slightly below the 50% threshold used to activate the prescribed urban thermal conductivity and heat capacity. The urban configurations therefore modify the radiative and aerodynamic properties without activating the corresponding increase in substrate thermal inertia. This may contribute to the insufficient daytime heat storage and excessive sensible heat simulated at this site. We have added US-WestPhoenix to the text and clarified this site-specific behaviour.

We modified the text to:

“For sensible heat, improvements are also seen at nearly all sites, except for CA-Sunset, US-WestPhoenix, UK-Swindon, and FI-Kumpula, where performance is slightly degraded. Detailed analysis suggests that these degradations arise from site-specific changes in energy partitioning. At FI-Kumpula and US-WestPhoenix, improvements in net radiation are not accompanied by sufficient partitioning toward latent heat and/or heat storage, resulting in increased sensible heat flux. Their urban fractions (46% and

48%, respectively) lie just below the 50% threshold used to activate the prescribed urban thermal conductivity and heat capacity. Consequently, the urban configurations modify the radiative and aerodynamic properties at these sites without activating the corresponding increase in substrate thermal inertia, which may contribute to insufficient daytime heat storage and the degradation in sensible-heat-flux performance.”

Section 4.3

- Fig6 : Wouldn't it make more sense to present the different components (drainage, runoff, evap) as percentages of total precipitation (with the total amount of precipitation indicated in the fig)? It seems to me that this would make comparison easier, and it would allow for a single bar (combining the three components) for each experiment (and each site).

Thank you for this helpful suggestion. We agree that expressing the water-balance components relative to total precipitation facilitates comparison among experiments and sites. We have therefore revised Fig. 6 to show evapotranspiration, surface runoff, and subsurface runoff as percentages of annual precipitation. The three components are now combined into a single stacked bar for each experiment at each site, and the mean annual precipitation amount is indicated above each panel. The figure caption and corresponding discussion in Section 4.3 have been updated accordingly.



- The fact that this section is based solely on a sensitivity analysis comparing the ORCHIDEE experiments, without any objective evaluation, remains a significant limitation of this study.

We agree that the absence of an objective evaluation of the simulated hydrological fluxes is an important limitation of this study. The Urban-PLUMBER dataset provides latent heat flux observations, allowing the evapotranspiration response to be evaluated in the preceding section, but it does not provide runoff, drainage, or soil-moisture observations at the investigated sites. Consequently, the analysis presented in this section should be interpreted as a sensitivity assessment of how the different ORCHIDEE configurations partition the water balance, rather than as a validation of the simulated runoff and drainage.

This observational limitation is not specific to our study. Jongen et al. (2024), in their evaluation of the water balance in 19 Urban-PLUMBER models, were also limited by the unavailability of these observations, and where still able to show large inter-model spread in runoff, showing the values produced by ORCHIDEE and its different configuration allow us to place our results in this spread.

We have revised the manuscript to state this limitation more explicitly and to avoid interpreting the simulated changes as evidence of hydrological performance. An objective evaluation would require spatially distributed simulations over instrumented urban catchments, including river routing, followed by comparison with observed streamflow. Producing such simulations also requires appropriate high-resolution meteorological forcing (because urban catchments are of small size), which remains challenging and is beyond the scope of the present study, but is a priority for future work.

In the introduction:

“Despite these advances, recent intercomparison work shows that urban water-balance representation remains uncertain in ULSMs, especially regarding water-balance closure, storage dynamics, and runoff parameterization (Jongen et al., 2024). Progress is also constrained by observations, because urban sites providing turbulent fluxes do not provide all water-balance terms needed to evaluate runoff, irrigation, drainage, and storage changes consistently at the same spatial scale (Jongen et al., 2024).”

In the 4.3 Annual water budget:

“Because runoff and subsurface-runoff observations are unavailable at these sites, this analysis should be interpreted as a sensitivity assessment rather than a validation of hydrological performance. Evaluating the simulated runoff would require spatially distributed simulations over instrumented urban catchments and comparison with observed streamflow, which is beyond the scope of this study.”

- Couldn't a surface water retention basin be easily defined in ORCHIDEE “urbanized”?

A surface water retention reservoir could, in principle, be introduced into the urban version of ORCHIDEE. However, its implementation would require additional assumptions regarding, for example, storage capacity, the fraction of impervious area connected to the reservoir, infiltration and evaporation losses, and the rate at which stored water is released. The available observations at the Urban-PLUMBER sites do not allow these parameters to be constrained or the resulting model behaviour to be evaluated. For this first urban implementation in ORCHIDEE, we therefore chose to remain parsimonious and to focus on representing the effects of imperviousness on infiltration, runoff, and drainage. A surface water retention reservoir could be considered in future developments if catchment-scale evaluations indicate that such a process is necessary to reproduce observed urban hydrological behavior.

We added a mention of rainfall interception by urban structures to the conclusion.

“Additional developments, such as spatially and temporally varying urban parameters (e.g., from WUDAPT), interception of rainfall on urban structures, refinement of thermal

properties, and the inclusion of anthropogenic heat fluxes, will further enhance the performances of this scheme.”

Discussion

- The points raised in the “Discussion” section are interesting but should be more clearly contextualized in relation to the scientific literature and the approaches used in other climate models applied to similar applications (as it is done for thermal properties).

We thank the reviewer for pointing this out. We added to the manuscript:

«Despite the implemented urban developments, the simulations exhibit persistent seasonal biases. Most notably, during summer, urban simulations show an overestimation of daily sensible heat flux and an underestimated latent heat flux, indicating that too much of the available energy is partitioned into sensible rather than latent heat. This high-sensible-heat, low-latent-heat bias is a recognized limitation of urban land-surface models and has been associated with the omission or insufficient representation of urban vegetation and water availability (Best and Grimmond, 2015). Indeed, latent heat flux remains underestimated by many current urban models (Lipson et al., 2024). In winter, the daytime bias reverses, with an underestimated sensible heat and an overestimated latent heat. The absence of anthropogenic heat in the current scheme may contribute to the winter underestimation of sensible heat, as its representation has been identified as important for simulating wintertime urban energy budgets (Hertwig et al., 2020; Jin et al., 2021; Karsisto et al., 2016). Future work will test whether prescribing anthropogenic heat fluxes from observation-based gridded datasets can reduce this bias in kilometer-scale simulations. On the contrary, the winter overestimated latent heat is not consistently reported across other models (Lipson et al., 2024) and may therefore reflect ORCHIDEE-specific seasonal controls on evaporation and water availability. At night, sensible heat also remains underestimated in winter, while latent heat exhibits smaller positive biases in both seasons. The bias pattern displayed by this new urban representation therefore suggests that evaporation is insufficient in summer, when urban vegetation and irrigation are most active, but insufficiently constrained in winter and at night. A potential development for ORCHIDEE would therefore be to embed an active vegetated fraction, including seasonal phenology, directly within the urban PFT. This would allow urban vegetation, and potentially irrigation, to enhance transpiration during the summer while limiting vegetation-related evaporation during winter.»

And:

«A more refined representation of urban hydrology could also help address the above biases. Urban land surface models use a range of runoff strategies, from routing a fixed fraction of rainfall directly to runoff to more process-based formulations accounting for

infiltration capacity, soil saturation, surface-water storage, and transfers between pervious and impervious tiles. The simpler formulations may be insensitive to rainfall intensity, antecedent soil moisture, and site characteristics (Jongen et al., 2024). The present study introduces a novel physically based representation of imperviousness by linking it to saturated hydraulic conductivity, thereby affecting infiltration, subsurface water transfer, and runoff generation. This development directly addresses the weak sensitivity of runoff to imperviousness found in 7 of the 18 models evaluated by Jongen et al. (2026). Nevertheless, Jongen et al. (2026) also found that 10 models omitted at least one major runoff-generation mechanism. Future developments should therefore represent surface ponding and interception storage, irrigation, drainage connectivity, and water transfers between urban and vegetated tiles. These processes should ultimately be evaluated against hydrological observations, because the present sensitivity analysis alone cannot establish whether the simulated water partitioning is realistic. Importantly, advancing the hydrological realism of the model should not be guided by latent heat performance alone, but by a comprehensive evaluation of the full water balance (Jongen et al., 2024). However, such evaluation is currently limited by a lack of observational data for key hydrological variables (e.g., runoff and soil moisture) at the urban flux tower sites used in this study (Jongen et al., 2026, 2024). This underscores the need for coordinated efforts to collect hydrological data in urban environments to support model development and validation.»

P25, L552-557

"Our current thermal conductivity value, derived from the NOAH land surface model (He et al., 2023) and relatively high compared to values used in CLM (Lawrence et al., 2019) for instance, may require further refinement. Our current thermal conductivity value ($3.24 \text{ W m}^{-1} \text{ K}^{-1}$), taken from the NOAH land-surface model, is substantially higher than values commonly used in CLM for example, $0.767 \text{ W m}^{-1} \text{ K}^{-1}$ in its original urban implementation (Loridan and Grimmond, 2012) or $1.55 \text{ W m}^{-1} \text{ K}^{-1}$ in the later SURY configuration (Wouters et al., 2016), and may therefore require further refinement."

There's a problem with repeated sentences here. You could simplify: "Our current thermal conductivity value ($3.24 \text{ W m}^{-1} \text{ K}^{-1}$), taken from the NOAH land-surface model (He et al., 2023), is substantially higher than values commonly used in CLM for example, $0.767 \text{ W m}^{-1} \text{ K}^{-1}$ in its original urban implementation (Loridan and Grimmond, 2012) or $1.55 \text{ W m}^{-1} \text{ K}^{-1}$ in the later SURY configuration (Wouters et al., 2016), and may therefore require further refinement."

Done, thank you.

Anonymous Referee #3, 01 Jun 2026

Report #3

Submitted on 01 Jun 2026
Anonymous referee #3

[Notes for the submission of interactive comments](#)

Anonymous: Yes No

Formal manuscript rating and recommendation to the editor

(visible to authors and reviewers only)

1) Scientific significance Does the manuscript represent a substantial contribution to modelling science within the scope of this journal (substantial new concepts, ideas, or methods)?	Excellent Good Fair Poor
2) Scientific quality Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)? Do the models, technical advances and/or experiments described have the potential to perform calculations leading to significant scientific results?	Excellent Good Fair Poor
3) Scientific reproducibility To what extent is the modelling science reproducible? Is the description sufficiently complete and precise to allow reproduction of the science by fellow scientists (traceability of results)?	Excellent Good Fair Poor
4) Presentation quality Are the scientific results and conclusions presented in a clear, concise, and well structured way (number and quality of figures/tables, appropriate use of English language)?	Excellent Good Fair Poor

For final publication, the manuscript should be

accepted as is
accepted subject to technical corrections
accepted subject to minor revisions
reconsidered after major revisions
rejected
Please note that this rating only refers to this version of the manuscript!

Were a revised manuscript to be sent for another round of reviews:

I would be willing to review the revised manuscript.

I would not be willing to review the revised manuscript.

Referee Comment: '[Comment on egusphere-2026-551](#)'

General Comments

Overall, this study provides a solid and well-organized description of a new one-tile “slab” urban module implemented in the ORCHIDEE v2.2 land surface model. The authors have done a commendable job situating their work within the broader context of urban land surface modeling and presenting the model evaluation in a coherent manner. The manuscript is generally well-written and the figures and tables are informative. However, I have several major concerns that I believe should be addressed before the manuscript is suitable for publication.

- Limited discussion of urban hydrology in the introduction. While the introduction dedicates space to the evolution of urban energy balance parameterizations, the treatment of urban hydrology is comparatively brief and could be strengthened. A more thorough review of how existing urban schemes represent urban hydrology, and what the key modeling gaps are, would help readers better appreciate the novelty of the approach proposed here. This concern is reinforced by specific issues throughout the manuscript: the definition of “imperviousness” itself (L271) is never made explicit—does it refer to the impervious surface fraction, a hydrological connectivity metric, or something else? The authors should clarify this early in the manuscript and use it consistently.

We thank the reviewer for this comment. We agree that the previous version of the introduction did not sufficiently describe how urban hydrology is represented in existing

ULSMs, nor did it clearly define how the term “imperviousness” is used in our model. We have therefore expanded the introduction to better situate our approach within previous developments in urban hydrology parameterizations and clarified the definition of imperviousness in the model description.

In the revised introduction, we now discuss how urbanization modifies surface and subsurface hydrological pathways, including infiltration, runoff generation, soil water availability, groundwater recharge, and drainage-related processes. We also clarify that “imperviousness” is not uniquely defined in urban hydrology, as it may refer to the total impervious surface fraction, to the fraction effectively connected to the drainage network, or to an effective hydrological control on infiltration and runoff generation.

We then review how urban hydrology has been introduced in ULSMs through several modelling choices, including simplified impervious representations, pervious/impervious surface distinctions, water storage on artificial surfaces, infiltration through artificial surfaces, coupling with LSM soil hydrology, and more detailed approaches including urban subsoil processes, lateral water transfer, and sewer drainage. This expanded discussion better explains the context for our approach and highlights that runoff parameterization, storage dynamics, water-balance closure, and the lack of complete water-balance observations remain important challenges.

We also clarified the definition of imperviousness in the model description. In this study, imperviousness does not represent effective connected impervious area or hydrological connectivity to the drainage network, and it is not used to define separate pervious and impervious hydrological sub-tiles. Instead, it is used as an effective hydrological parameter that modifies saturated hydraulic conductivity as a function of depth, thereby affecting infiltration, runoff generation, soil water storage, and vertical water transfer within the ORCHIDEE soil hydrology.

Finally, we revised the novelty statement to clarify that our contribution is not a detailed urban drainage or multi-tile hydrological scheme, but a parsimonious, physically motivated representation of imperviousness designed to remain consistent with the one-tile structure and existing hydrology of ORCHIDEE.

We updated this part in the introduction:

“Similarly, urban hydrology has progressed from simplified impervious representations (Kusaka et al., 2001; Masson, 2000) to more complete schemes incorporating pervious/impervious distinctions, water reservoirs, irrigation, and subsurface routing (Chancibault et al., 2014; Lemonsu et al., 2007; Oleson et al., 2008; Wouters et al., 2015). Urban hydrology introduces additional challenges due to soil sealing, compaction, preferential flow pathways, and anthropogenic drainage infrastructure. Imperviousness affects both surface and subsurface processes and is not uniquely defined; it may refer to sealed surfaces, hydrologically disconnected areas, or the

effective fraction contributing to runoff (Saadi, 2020). These processes influence infiltration, soil moisture, and groundwater recharge, yet remain poorly understood (Bhaskar et al., 2016; Fletcher et al., 2013). This complexity highlights the need for parameterizations in Urban LSMs that explicitly link imperviousness to hydrological properties within the soil column.”

To:

“On top of vegetation, hydrology has also been shown as a key element to capture correctly latent heat fluxes. Urbanization also modifies surface and subsurface hydrological pathways by increasing soil sealing, compaction, changing preferential flow pathways, and through water use infrastructures (Fletcher et al., 2013; O’Driscoll et al., 2010; Bhaskar et al., 2016). These changes increase surface runoff and shorten hydrological response times, while their effects on soil moisture, groundwater recharge, low flows, and evapotranspiration depend on local soil properties, vegetation, water use, drainage connectivity, and sewer or stormwater networks (Fletcher et al., 2013; Bhaskar et al., 2016; Oudin et al., 2018; Saadi et al., 2020). This complexity is partly reflected in the use of the term “imperviousness”, which is not unique in urban hydrology. Depending on the modeling and spatial context, it can refer to the total impervious surface fraction, the fraction effectively connected to the drainage network, or an effective hydrological control on infiltration and runoff generation (Arnold and Gibbons, 1996; Shuster et al., 2005; Jacobson, 2011; Saadi et al., 2020). This distinction is important for ULSMs because urban hydrology influences not only runoff, but also soil water availability and evapotranspiration, which directly links the water and energy balances.

In ULSMs, urban hydrology has been introduced progressively through several modelling choices. Early urban canopy schemes primarily focused on radiative, thermal, and aerodynamic exchanges, while hydrological processes were absent or simplified, with artificial surfaces commonly represented as impervious and water exchange mainly limited to surface interception, evaporation from wet sealed surfaces, and runoff (Masson, 2000; Kusaka et al., 2001; Martilli et al., 2002). Later developments introduced more explicit hydrological distinctions between urban surface types. One direction was to distinguish hydrological behaviour between pervious and impervious urban fractions, for example through pervious and impervious canyon-floor fractions in CLM-U (Oleson et al., 2008). TEB-related developments add infiltration through artificial surfaces, and connected rainfall interception on roofs and roads toward a soil component (Lemonsu et al., 2007). Other schemes refined the representation of water stored on impervious surfaces and its evaporation after rainfall, as in TERRA-URB (Wouters et al., 2015). Hydro-microclimate approaches further extended urban hydrology to include subsoil beneath built surfaces, lateral water transfer between pervious and impervious tiles, sewer drainage, and irrigation, as in TEB-Hydro (Chancibault et al., 2014; Stavropoulos-Laffaille et al., 2018, 2021). Despite these advances, recent intercomparison work shows that urban water-balance representation remains uncertain in ULSMs, especially regarding

water-balance closure, storage dynamics, and runoff parameterization (Jongen et al., 2024). Progress is also constrained by observations, because urban sites providing turbulent fluxes do not provide all water-balance terms needed to evaluate runoff, irrigation, drainage, and storage changes consistently at the same spatial scale (Jongen et al., 2024).”

We also updated the following part:

“Second, we introduce a physically based representation of imperviousness that affects both surface and subsurface hydrological processes. Unlike most one-tile schemes that treat imperviousness as a surface-only property, we allow imperviousness to influence water storage and transfer throughout the soil column by modifying saturated hydraulic conductivity as a function of depth.”

Into:

“Second, we introduce a parsimonious, physically motivated, representation of imperviousness. Rather than resolving separate water budgets for pervious and impervious urban sub-tiles, the proposed approach preserves the existing ORCHIDEE soil hydrology with an urban soil tile where imperviousness is represented through its effect on the hydraulic properties of the soil column. Specifically, imperviousness modifies saturated hydraulic conductivity as a function of depth, allowing it to influence infiltration, runoff generation, soil water storage, and vertical water transfer. This formulation is designed for one-tile, kilometer-scale and coarser LSM applications, where the objective is to represent their aggregate effect on the water budget and surface fluxes feedback to the atmosphere.”

We modified Line 271: *“For the water budget, the key parameter is imperviousness, which controls both infiltration and runoff generation.”*

Into:

“For the water budget, the key input parameter is the impervious surface fraction, ranging from 0 to 1. In this study, this fraction is not used to define separate pervious and impervious hydrological sub-tiles, nor does it represent the fraction directly connected to the drainage network. Instead, it is converted into an effective hydrological parameter that modifies saturated hydraulic conductivity as a function of depth, thereby affecting infiltration, runoff generation, soil-water storage, and vertical water transfer within the ORCHIDEE soil hydrology.”

Finally in the discussion we added:

“Urban land surface models use a range of runoff strategies, from routing a fixed fraction of rainfall directly to runoff to more process-based formulations accounting for infiltration capacity, soil saturation, surface-water storage, and transfers between pervious and impervious tiles. The simpler formulations may be insensitive to rainfall intensity,

antecedent soil moisture, and site characteristics (Jongen et al., 2024). The present study introduces a novel physically based representation of imperviousness by linking it to saturated hydraulic conductivity, thereby affecting infiltration, subsurface water transfer, and runoff generation. This development directly addresses the weak sensitivity of runoff to imperviousness found in 7 of the 18 models evaluated by Jongen et al. (2026). Nevertheless, Jongen et al. (2026) also found that 10 models omitted at least one major runoff-generation mechanism. Future developments should therefore represent surface ponding and interception storage, irrigation, drainage connectivity, and water transfers between urban and vegetated tiles. These processes should ultimately be evaluated against hydrological observations, because the present sensitivity analysis alone cannot establish whether the simulated water partitioning is realistic. Importantly, advancing the hydrological realism of the model should not be guided by latent heat performance alone, but by a comprehensive evaluation of the full water balance (Jongen et al., 2024). However, such evaluation is currently limited by a lack of observational data for key hydrological variables (e.g., runoff and soil moisture) at the urban flux tower sites used in this study (Jongen et al., 2026, 2024). This underscores the need for coordinated efforts to collect hydrological data in urban environments to support model development and validation.”

- Equations require polishing for consistency and completeness. Several equations contain inconsistencies or missing terms that need to be corrected. First, Equations 1 and 2 are inconsistent: Equation 1 includes a heat storage term (dW/dt), but this term is absent from Equation 2. The authors should clarify whether these terms are explicitly resolved, implicit in other terms, or assumed negligible, and this justification should be provided in the text (see also specific comment at L162–163). Second, the same symbol K_s is used in both Equation 15 (saturated hydraulic conductivity for non-urban PFTs) and Equation 22 (for the urban PFT), creating ambiguity. A distinct notation should be adopted for the urban case to avoid confusion.

We thank the reviewer for this comment, which helped us improve the consistency and clarity of the model description. We agree that the previous presentation of the equations was ambiguous. Equation 1 was intended as a general surface energy-budget framework, whereas Equation 2 only defined the net radiation term used by ORCHIDEE. However, this distinction was not sufficiently clear in the text, and the use of a separate heat-storage term in Equation 1 could be confusing because ORCHIDEE does not diagnose an additional storage flux independently from the conductive ground heat flux.

We therefore revised Section 2.1.1, “Representation of the energy fluxes and surface energy budget”, to distinguish more clearly between the conceptual surface energy balance and the individual fluxes computed by ORCHIDEE. In the revised formulation, the surface energy balance is written as:

“

$$R_{net} = H + L_E + G$$

where R_{net} is the net radiation, representing the total balance between incoming shortwave (SW) and longwave (LW) radiation and their respective outgoing components, H is the sensible heat flux, L_E is the latent heat flux, and G is the ground heat flux. The heat stored and released by the soil or urban substrate is not neglected, but is represented through (G), which is used as the upper boundary condition of the one-dimensional heat diffusion equation for the soil temperature profile. Thus, no additional storage term is introduced in the surface energy budget.”

We further clarified in the manuscript how these terms are connected in the ORCHIDEE computation. At each time step, ORCHIDEE first uses the prescribed or computed surface properties, including albedo, emissivity, roughness length, and soil or substrate thermal properties. The model then computes net radiation from incoming shortwave and longwave radiation, surface albedo, emissivity, and surface temperature. This corresponds to the radiative term of the surface energy budget.

The turbulent fluxes are computed consistently with this same surface temperature. Sensible heat depends on the surface–air temperature gradient and on the aerodynamic exchange coefficient, which is affected by wind speed and surface roughness. Latent heat depends on the surface–air humidity gradient, the saturation humidity at the surface temperature, the same turbulent exchange coefficient, and the water-availability limitations imposed by the hydrological state of the soil and surface. Thus, Equations 2, 4, and 5 are linked through the surface temperature and through the turbulent exchange formulation: Equation 2 defines the net radiation term, while Equations 4 and 5 define the latent and sensible heat fluxes that appear in the surface energy budget.

The conductive ground heat flux is computed from the thermal gradient between the surface and the upper soil or substrate layer, together with the thermal conductivity and heat capacity. It provides the upper boundary condition for the one-dimensional heat diffusion equation, which updates the soil or substrate temperature profile. Heat storage and release are therefore represented through the prognostic evolution of this temperature profile, rather than through a separate diagnosed storage term. The updated surface and soil or substrate temperatures are then used as thermal state variables at the next time step.

Finally, the water budget is updated through precipitation partitioning, infiltration, surface runoff, drainage, soil water redistribution, and evapotranspiration. The updated soil moisture then feeds back on latent heat fluxes and on soil thermal properties at the following time steps. We added this clarification to make explicit that the radiative,

turbulent, conductive, and hydrological terms are coupled in ORCHIDEE, even though they are introduced through separate equations in the model description.

We also revised the notation for saturated hydraulic conductivity to avoid ambiguity between the standard ORCHIDEE profile and the urban-modified profile. In the revised manuscript, $K_s(z)$ denotes the standard ORCHIDEE saturated hydraulic conductivity profile defined in Eq. (15), while $K_s^{urb}(z)$ denotes the saturated hydraulic conductivity profile used for the urban soil column. The urban profile is now described as :

$$K_s^{urb}(z) = K_s^{ref} F_K(z) F_{urb}$$

where F_{urb} is the dimensionless reduction factor, new notation of K_{fact}^{urban} . This notation clearly distinguishes the reference ORCHIDEE formulation from the imperviousness-modified urban formulation.

- The novelty of this urban scheme relative to existing one-tile schemes is insufficiently articulated. The manuscript describes the new ORCHIDEE urban scheme as a one-tile “slab” approach and positions it within the Lipson et al. (2023) classification, but it does not clearly delineate what distinguishes this scheme from other existing one-tile urban parameterizations in the literature. The authors highlight the physically based treatment of imperviousness through saturated hydraulic conductivity as a key novel contribution, but this is not placed in the context of other schemes’ imperviousness representations. A concise comparison, even qualitative, would help readers assess the incremental scientific contribution and appreciate what is genuinely new here.

We thank the reviewer for this comment. We agree that the previous version did not sufficiently articulate the novelty of the proposed ORCHIDEE urban scheme relative to existing one-tile and bulk urban parameterizations. We added in the manuscript:

“On top of vegetation, hydrology has also been shown as a key element to capture correctly latent heat fluxes. Urbanization modifies surface and subsurface hydrological pathways by increasing soil sealing, compaction, changing preferential flow pathways, and through water use infrastructures (Bhaskar et al., 2016; Fletcher et al., 2013). These changes generally increase surface runoff and shorten hydrological response times, while their effects on soil moisture, groundwater recharge, low flows, and evapotranspiration depend on local soil properties, vegetation, water use, drainage connectivity, and sewer or stormwater networks (Bhaskar et al., 2016; Fletcher et al., 2013; Oudin et al., 2018; Saadi et al., 2020). This complexity is partly reflected in the use of the term “imperviousness”, which is not unique in urban hydrology. Depending on the modeling and spatial context, it can refer to the total impervious surface fraction, the fraction effectively connected to the drainage network, or an effective hydrological control on infiltration and surface runoff generation (Jacobson, 2011; Saadi et al., 2020).

This distinction is important for ULSMs because urban hydrology influences not only surface runoff, but also soil water availability and evapotranspiration, which directly links the water and energy budgets.

In ULSMs, urban hydrology has been introduced progressively through several modelling choices. Early urban canopy schemes primarily focused on radiative, thermal, and aerodynamic exchanges, while hydrological processes were absent or simplified, with artificial surfaces commonly represented as impervious and water exchange mainly limited to surface interception, evaporation from wet sealed surfaces, and runoff (Kusaka et al., 2001; Martilli et al., 2002; Masson, 2000). Later developments introduced more explicit hydrological distinctions between urban surface types. One direction was to distinguish hydrological behaviour between pervious and impervious urban fractions, for example through pervious and impervious canyon-floor fractions in CLM-U (Oleson et al., 2008). TEB-related developments add infiltration through artificial surfaces, and connected rainfall interception on roofs and roads toward a soil component (Lemonsu et al., 2007). Other schemes refined the representation of water stored on impervious surfaces and its evaporation after rainfall, as in TERRA-URB (Wouters et al., 2015). Hydro-microclimate approaches further extended urban hydrology to include subsoil beneath built surfaces, lateral water transfer between pervious and impervious tiles, sewer drainage, and irrigation, as in TEB-Hydro (Stavropoulos-Laffaille et al., 2021, 2018). Despite these advances, recent intercomparison work shows that urban water-balance representation remains uncertain in ULSMs, especially regarding water-balance closure, storage dynamics, and runoff parameterization (Jongen et al., 2024). Progress is also constrained by observations, because urban sites providing turbulent fluxes do not provide all water-balance terms needed to evaluate runoff, irrigation, drainage, and storage changes consistently at the same spatial scale (Jongen et al., 2024).”

And:

“Second, we introduce a parsimonious, physically motivated, representation of imperviousness. Rather than resolving separate water budgets for pervious and impervious urban sub-tiles, the proposed approach preserves the existing ORCHIDEE soil hydrology with an urban soil tile where imperviousness is represented through its effect on the hydraulic properties of the soil column. Specifically, imperviousness modifies saturated hydraulic conductivity as a function of depth, allowing it to influence infiltration, surface runoff generation, soil water storage, and vertical water transfer. This formulation is designed for one-tile, kilometer-scale and coarser LSM applications, where the objective is to represent their aggregate effect on the water budget and surface fluxes feedback to the atmosphere.”

And:

«A more refined representation of urban hydrology could also help address the above biases. Urban land surface models use a range of runoff strategies, from routing a fixed

fraction of rainfall directly to runoff to more process-based formulations accounting for infiltration capacity, soil saturation, surface-water storage, and transfers between pervious and impervious tiles. The simpler formulations may be insensitive to rainfall intensity, antecedent soil moisture, and site characteristics (Jongen et al., 2024). The present study introduces a novel physically based representation of imperviousness by linking it to saturated hydraulic conductivity, thereby affecting infiltration, subsurface water transfer, and runoff generation. This development directly addresses the weak sensitivity of runoff to imperviousness found in 7 of the 18 models evaluated by Jongen et al. (2026). Nevertheless, Jongen et al. (2026) also found that 10 models omitted at least one major runoff-generation mechanism. Future developments should therefore represent surface ponding and interception storage, irrigation, drainage connectivity, and water transfers between urban and vegetated tiles. These processes should ultimately be evaluated against hydrological observations, because the present sensitivity analysis alone cannot establish whether the simulated water partitioning is realistic. Importantly, advancing the hydrological realism of the model should not be guided by latent heat performance alone, but by a comprehensive evaluation of the full water balance (Jongen et al., 2024). However, such evaluation is currently limited by a lack of observational data for key hydrological variables (e.g., runoff and soil moisture) at the urban flux tower sites used in this study (Jongen et al., 2026, 2024). This underscores the need for coordinated efforts to collect hydrological data in urban environments to support model development and validation.»

- The term “drainage” should be replaced with “subsurface runoff” or other terms throughout. The manuscript uses “drainage” (e.g., L132, L245–247, L516–527) to refer to gravitational outflow at the bottom of the 2 m soil column. However, the term “drainage” in an urban hydrology context typically refers to engineered drainage infrastructure (sewers, storm drains), which is explicitly noted as absent from this model. This conflation of terms is potentially misleading, particularly for readers from the urban hydrology community and in the context of the stated objective of advancing urban hydrological representation in LSMs. The authors themselves use “subsurface runoff” as a synonym in places (e.g., L132). Please use “subsurface runoff” or other terms consistently throughout the manuscript to avoid ambiguity.

We thank the reviewer for this important clarification. We agree that “drainage” may be misleading in an urban hydrology context, as it can refer to engineered infrastructure. We have therefore replaced it with “subsurface runoff” throughout the manuscript.

Specific Comments

- L27–29: Would the authors provide some examples of which land surface models they were referring to?

Yes, we now include 3 examples for each:

“Historically, their development followed two pathways shaped by the application scale. GCM-oriented LSMs, such as ORCHIDEE, CLM, and JSBACH, operate on coarse grids (tens to hundreds of kilometers; e.g. Fisher and Koven, 2020) and focus on global water and carbon cycles (Krinner et al., 2005; Lawrence et al., 2019; Raddatz et al., 2007). By contrast, LSMs used in mesoscale and limited-area models, such as Noah-MP, TERRA, and ISBA-SURFEX, have long used finer resolutions, favoring more accurate landscape representation and parameterizations relevant at kilometer scale over restricted domains or shorter periods (Doms et al., 2011; Masson et al., 2013; Niu et al., 2011), but neglecting groundwater processes (Barlage et al., 2021).”

- L57: What does “leading to significant advances” refer to here? Please clarify or provide specific references.

We meant improvement of fluxes simulation; we removed this last part of the sentence.

- L81: Please spell out “PFTs” at first use.

Done

- L83–100: This is a coherent paragraph, but the authors could be more explicit about which key surface properties they are referring to. Listing them explicitly (e.g., (1) albedo, (2) roughness, (3) thermal properties ...) would improve clarity, as the current prose is slightly convoluted.

We added: “(1) radiative properties, particularly albedo and longwave-radiation trapping; (2) aerodynamic properties, including surface roughness and urban geometry; (3) thermal properties, especially heat capacity and thermal conductivity; (4) vegetation cover, imperviousness, and water availability; and (5) anthropogenic heat emissions.”

- L125–127: Do the authors mean “energy fluxes associated with radiative and aerodynamic parameters are aggregated for individual grid cells as PFT-weighted parameters”?

Yes, we modified as suggested.

- L128–133: This paragraph raises two points of confusion. First, what distinguishes the soil columns for bare soil versus high and low vegetation? Do they differ in soil properties, depth, or only rooting effects? Maybe this detail embedded in ORCHIDEE, but it is important for readers who are not familiar with ORCHIDEE. Second, the model has a 2 m hydrology column and an 18 m temperature column, both discretized into 12 layers—does this imply a mismatch in layer thicknesses between hydrological and thermal calculations? Please clarify.

Thank you for highlighting this issue, we added a new figure 1 to describe ORCHIDEE structure, and we modified the paragraph into:

“Each grid cell contains three soil columns, one for bare soil, one for high vegetation, and one for low vegetation. These columns have the same depth, vertical discretization, soil texture, and texture-dependent hydraulic properties. They differ in their vegetation-related processes: the high- and low-vegetation columns have distinct rooting profiles and root-water uptake, whereas the bare-soil column has no root extraction. Their soil-moisture states and water budgets therefore evolve independently. Each soil column is vertically discretized (12 layers, down to a depth of 2 m for hydrology, and 18 m for soil temperature computation), allowing the model to simulate temperature and moisture transfer through the soil profile. For each soil column, ORCHIDEE calculates the surface runoff, and subsurface runoff, representing gravitational outflow at the bottom of the soil column.”

Later in the manuscript we also now specify:

“The hydrological and thermal discretizations share the upper 2 m of the soil column. Within this common domain, hydrological soil-moisture variables are transferred to the thermal grid, including interpolation of volumetric water content to the thermal-layer interfaces. Below 2 m, the saturation degree and volumetric water contents are held equal to those of the deepest hydrological layer, while the total water content is adjusted for the thickness of each thermal layer.”

- L140: Consider using the plural “theses” here, as multiple Ph.D. and HDR theses are cited.

Done

- L142–144: The text states that snow-covered conditions and floodplains are “not considered in this description.” Does ORCHIDEE represent them at all, and they are simply excluded from the description for brevity, or are they absent from the model? Please clarify.

We clarified that snow-covered conditions and floodplains are represented in ORCHIDEE but are omitted from this methodological description for brevity, as they are not directly relevant to the urban one-tile developments presented in this study:

“For clarity, snow-covered conditions and floodplain processes, although represented in ORCHIDEE, are omitted from the description below because they are not relevant to the urban one-tile developments presented here.”

- Equation 2: Where are the ground heat flux (G) and heat storage (dW/dt) terms? Does the model assume these are negligible, or are they resolved elsewhere? This should be explicitly stated, as they appear in Equation 1.

Thank you for highlighting this point. Equation (2) defines only the net-radiation component of the surface energy budget and is therefore not intended to include the turbulent or conductive fluxes. In ORCHIDEE, the sensible and latent heat fluxes are

calculated separately in Eqs. (4) and (5), while the ground heat flux G is calculated from the temperature gradient between the surface and the upper substrate layer in Eq. (7). This flux provides the upper boundary condition for the one-dimensional heat-diffusion equation used to update the soil-temperature profile. Heat storage and release within the soil or urban substrate are therefore represented prognostically through G and the evolution of the thermal profile; they are not neglected or introduced as an additional independent dW/dt term. We revised the text and Eq. (1) to clarify this formulation and avoid double counting heat storage.

- L162–163: The list of latent heat flux components appears incomplete. Evaporation from canopy interception is missing as in L250, ORCHIDEE does include evaporation from vegetation interception (canopy interception). Please double check.

Thank you. ORCHIDEE includes evaporation of water intercepted by the vegetation canopy, also referred to as interception loss. This component was inadvertently omitted from the list, although it was included later in the methodological description. We have revised the sentence to list all four components of the latent heat flux: bare-soil evaporation, plant transpiration, canopy-interception evaporation, and snow sublimation.

- L237–241 and L239: This paragraph is slightly confusing and could benefit from clearer structure. Also, please provide the unit for F_{K_max} .

Thank you for this comment. We reorganized and simplified the paragraph to distinguish more clearly between the reference conductivity, its depth-dependent reduction, and the root-induced enhancement:

“In version 2.2 of ORCHIDEE, the vertical profile of saturated hydraulic conductivity $K_s(z)$ is prescribed as a function of depth. Its value at any depth z results from three components:

- (i) Reference saturated hydraulic conductivity (K_s^{ref}) at the depth $z_{lim} = 0.30\text{m}$, derived from the dominant USDA (U.S. Department of Agriculture) soil texture of the grid cell (12-class texture classification). Each grid cell's dominant soil texture influences soil transfer (Ducharne et al., 2018; Tafasca et al., 2020).
- (ii) An exponential decrease with depth, controlled by the dimensionless factor $F_K(z)$. This decrease applies from z_{lim} to the bottom of the soil column (2 m).
- (iii) A root-induced enhancement in the upper soil layers, represented by a factor $F_{Kroots}(z, c)$, where c denotes the soil-column index. This enhancement is applied only to vegetated soil columns, whereas $F_{Kroots}(z, c)$ is set to 1 for the

bare soil column, which therefore retains a constant K_s over the first 30 cm (d'Orgeval et al., 2008; De Rosnay et al., 2002).

The resulting expression for the saturated hydraulic conductivity profile is:

$$K_s(z) = K_s^{ref} F_K(z) F_{Kroots}(z, c) \quad (15)$$

where $F_K(z)$ controls the exponential decrease with depth:

$$F_K(z) = \min\left(\max\left(\exp\left(-f(z - z_{lim})\right), \frac{1}{F_K^{max}}\right), 1\right) \quad (16)$$

Apart from the texture-dependent K_s^{ref} , the other three parameters (z_{lim} , f , F_K^{max}) are constants, i.e. they do not depend on the PFT, the soiltile, and the soil texture. The depth z_{lim} is set to 0.30 m, the decay factor f is set to 2 m^{-1} , and the dimensionless parameter F_K^{max} is set to 10. Therefore, $F_K(z)$ cannot be smaller than $1/F_K^{max}=0.1$.

- L255: Please use “potential evapotranspiration”.

Done

- L271: The clear definition of “imperviousness” can be helpful here or when this terminology was first introduced. Does it refer to the impervious surface fraction (the interpretation implied by Eq. 24) or a more general hydrological concept?

Thank you for this comment. As detailed in our response to the first major comment, we now define imperviousness explicitly in both the introduction and the model description. In Eq. (24), imperviousness corresponds to the prescribed impervious surface fraction, ranging from 0 to 1. However, it is not used to partition the urban tile into separate pervious and impervious fractions, nor does it represent the fraction connected to the drainage network. Instead, it is translated into an effective hydrological control by scaling saturated hydraulic conductivity throughout the urban soil column. We have clarified this distinction at its first introduction and at Line 271.

- Equation 22: Please use a different notation for the saturated hydraulic conductivity of the urban PFT. The symbol K_s is already used in Equation 15 for non-urban PFTs, and reusing it creates unnecessary ambiguity.

Yes, thank you for this observation, we have change $K_s(z)$ to $K_s^{urb}(z)$.

- L320: According to He et al. (2023), the reference model is Noah-MP, not NOAH. Please verify and correct this throughout the manuscript.

Done

- L426: Since maximum daily sensible heat fluxes naturally take the largest absolute values, it is not surprising that they also show the largest absolute

biases. It would be more informative to evaluate model performance in relative terms (e.g., normalized bias or percentage error) to allow fair comparison across variables.

We agree that absolute bias magnitudes should not be interpreted as a direct ranking of relative performance across flux variables, because their characteristic magnitudes differ. Figure 3 is intended primarily to compare model configurations within each flux diagnostic while retaining the sign and physically interpretable magnitude of the biases. We have revised the text to clarify this scope and no longer interpret the larger absolute sensible heat biases as evidence of poorer relative performance than for the other fluxes.

We did not introduce normalized or percentage errors because a single normalization applicable to all diagnostics would not be robust. In particular, the observed nighttime minima frequently approach zero and may change sign, making normalization by the observed mean or individual values unstable or undefined. We have instead explicitly acknowledged that the small absolute biases in minimum latent heat flux do not necessarily indicate high relative accuracy, while their consistently positive sign provides evidence of systematic nocturnal overestimation:

"Because the characteristic magnitudes differ among fluxes and between daytime maxima and nighttime minima, the absolute bias magnitudes should not be interpreted as a direct ranking of relative model performance across panels. The comparisons below focus primarily on differences among model configurations within each diagnostic. "

- L451: Similarly, minimum daily latent heat flux values are inherently small, so small absolute biases here are expected. A relative comparison would be more meaningful.

We agree that the small absolute biases partly reflect the inherently small magnitude of nocturnal latent heat flux and should not, by themselves, be interpreted as indicating high relative accuracy. As explained in our response to the previous comment, we retained the absolute bias but clarified its interpretation in the revised manuscript.

- Figures 4 & 5: The text in the first x-axis tick label is cut off. Please fix.

We updated the figures 4 and 5 of the manuscript with a complete x-axis that is not cut off.