Response to comments by Anonymous Referee # 2

We thank the reviewer for the careful and thoughtful assessment of our manuscript. We appreciate the constructive and detailed comments, which will help us to improve the clarity and scope of the paper.

Please note that, in the following point-by-point response, the reviewer's comments are reproduced in red letters, followed by our replies in black text.

Major Comments

1. The scheme relies on fixed characteristic connectivities between tile types. It is unclear how these should be specified for real landscapes. In particular, equation 1 for the weighted connectivity c(t,s) is hard to understand from the text alone. The authors should consider clarifying this concept, as it is vital to the methods that follow. It would strengthen the paper to include an example or add an appendix including it.

We fully agree with the reviewer that – in addition to clarifying the descriptive text – examples constitute a straightforward way to make the concept of the connectivities more readily accessible for the reader. Here, our idea would be to provide the connectivity matrices for the two example applications that are already included in the manuscript, detailing how they were derived for the specific tile-definition. With respect to the question how this would be specified in real landscapes, please see our response to the third comment.

2. The spatial relationship module assumes circular patch shapes to compute center-to-center distances, as shown in equation 3. This is a substantial simplification. How robust are the results to this assumption? Many natural patterns are not circular. I would encourage the authors to discuss this decision more extensively. Additionally, the authors mention that only two-level nesting is allowed. Please discuss the implications of this decision. At a minimum, I consider that the limitations of the circular-patch assumption should be stated clearly.

Here, the reviewer is absolutely right that the circular shape is a substantial simplification with important implications for the simulated lateral exchange. The diffusive fluxes, for example, scale linearly with the contact length for two connected patches and are inversely proportional to the centre-to-centre distance (for a given temperature- or moisture difference between the pair of tiles). And both of these properties depend on the assumed shape of the patches. In extreme cases, in which the represented surface features have an elongated shape, e.g. as a plausible assumption for the floodplains along rivers, the pathways of the diffusive fluxes could easily be an order of magnitude smaller than the centre- to-centre distance based on assumed circles. And as shown by our second example case, this can make the difference between the fluxes being able to largely harmonize the states of two adjacent tiles or not. The circular geometry was chosen for a number of reasons, first and foremost because it is possibly the least bad option for many landscape features such as the patterned ground targeted in our second example or for smaller lakes and ponds, with many physical and biogeochemical lake models making this assumption. For a broader applicability of the scheme and as future development steps, the assumptions could be made more flexible, to better represent non-circular geometries, using (tile- and grid cell specific) shape-factors (S_1; S_2) in the key ratios of perimeter (C) and radius (r): $C = r \cdot 2 \cdot pi \cdot S1$ and radius and area (A): $r = (A \cdot pi)^{0.5} \cdot S2$. Or, as an additional option, the relevant lengths could be provided directly by the user. Concerning the manuscript, we would be happy to include the above as an outlook in the discussion section of a revised version.

3. Both examples use highly idealized tiling configurations (i.e., a two-tile upland/depression split, and uniform concentric circles everywhere in the Arctic). Even though the rationale for selecting these 2 configurations is clear (i.e., testing the effects of different lateral transport processes introduced by T-REX), the results may depend strongly on these assumptions. How sensitive are the findings to these choices? It would be helpful to discuss how one would apply T REX to a more complex tiling. Can the authors comment generally on their conclusions beyond the toy cases? Perhaps an additional test or sensitivity (even qualitative) could be suggested.

One of the fundamental limitations of the scheme, in its present form, is that it can only be applied to highly idealized tiling configurations, in the sense that the definition of the tiles must provide a set of characteristic connectivities that is globally valid. Thus, the scheme can be used to approximate real world landscapes only to that extent that these can be constructed from such idealized tiles. The four-tile upland/lowland, depression/elevation split, for example, could possibly be extended to include a floodplain tile, since it appears to be a reasonable assumption that large floodplains would predominantly be connected to the lowland tiles. In principle, these topography-based tiles could then further be subdivided based on a second property, say soil types. However, for many factors it will be very difficult to establish a globally valid logic that provides the connectivities. E.g. to which degree would lowland depressions with a coarse-grained soil material be connected to floodplain patches with loamy soils and to which degree to floodplain patches with sandy soils? For a tiling for which there is no such a globally valid logic, the only option (at present) is to assume that the tiles are randomly distributed within the landscape. In this case, a given tile could be assumed to have the same characteristic connectivity with any of the other tiles, with the actual connectivity being largely determined by the cover fractions of the tiles within a specific grid cell. This could be a permissible assumption for many land-surface characteristics such, as vegetation cover, but constructing real-world landscapes with a higher degree of realism would at least require the ability of the model to work with grid-cell specific connectivity matrices, which could be included in a future model development.

We expect the choice of the tile-configuration to have profound consequences for the results not only because it determines the geometry of the patches represented in a setup, but also the properties of the tiles. Unfortunately, we cannot really provide any robust conclusions beyond the toy cases since the application of the scheme to a less idealized setting is still work in progress. Here, our assumption is that the two key factors are hydrology and land cover, with our next goal being to (be able to) run the model with the four-tile upland/lowland, depression/elevation split and allow the model to establish the vegetation composition for the different tiles individually. Such a setup could greatly enhance the inter-tile temperature variability – shown in Fig. 5 – if, for example, the tiles with the better ability to retain water also feature a vegetation mix that transpires more readily. However, it could also reduce the variability, if the wetter tiles feature vegetation with a lower albedo than the dryer tiles, e.g. trees instead of grass.

4. In the new tiling structure, many associated parameters are introduced (e.g., characteristic elevations, areas, contact fractions, connectivities, slopes, etc.). The manuscript should explicitly list the variables. Some guidance, or even just a table, with parameter definitions would improve clarity. This is especially important since the formulations presented in equations 3 to 5 can be confusing without context on the variables used.

We would be happy to provide such an overview table in a revised version of the manuscript.

5. One aspect that remains unclear is whether the snow redistribution scheme accounts for wind direction and speed. These are critical drivers of snow transport in natural systems, particularly in open tundra or alpine environments. Yet, the manuscript does not indicate whether wind forcing influences the redistribution pattern or magnitude. Additionally, the effects of vegetation, such as snow interception by shrubs or reduced wind exposure under forest canopies, are not addressed, despite their role in limiting wind-driven snow transport. These simplifications may be reasonable for a first implementation, but should be explicitly acknowledged. Furthermore, the manuscript lacks dedicated test cases or diagnostic outputs focused solely on snow redistribution. Without a targeted evaluation, it is difficult to assess the realism or the impact of this component relative to the other lateral processes included in T-REX. The authors should consider highlighting this as a current limitation and outlining whether validation or further development is planned.

Unfortunately, the scheme does not account for wind direction or -speed and we also do not see a straightforward way to include these in future development steps. The main reason for this is that ICON's land-surface component, as many other land-surface models, does not "know" about directions. Since, the land surface heterogeneity is never spatially resolved explicitly in the model, there is no information if patches of surface type A are predominately located north, east, west or south of the patches of surface type B. Here, it is conceivable to derive such information in the pre-processing and then use direction-specific connectivity matrices in the model. However, while we agree with the reviewer (that the simplicity of the snow transport is an important limitation) and would be happy to highlight the issue in a revised version of the manuscript, we fear that we do not have the resources to work on this shortcoming in the foreseeable future. As a result, the applicability of the snow redistribution will be limited to cases where it can be assumed that the patches of tiles between which the snow is shifted are randomly distributed and well interspersed so that it is permissible to ignore the wind-direction.

6. The study uses prescribed meteorological forcings (i.e., uncoupled simulation). The authors conclude that the grid-mean surface fluxes are hardly affected by the implementation of T-REX. However, even small changes in moisture or temperature patterns can alter atmospheric response in a fully coupled simulation. The paper should acknowledge this limitation. It would strengthen the conclusions to note that "while our offline results show negligible change in mean FLUX/STATE, coupled simulations would be needed to assess any feedback on climate fully."

We fully agree with the reviewer and will include such a statement in a revised manuscript.

Minor fixes

In the figure 5 caption, correct "sugrid-scale" and "subgrig-scale" to subgrid-scale, and remove the stray comma in "sum of, surface runoff".

Will be corrected.

In the figure 4 caption the phrase "layers fn the vertical grid" should be "layers in the vertical grid"..

Will be corrected.