Response to Reviewer

Manuscript Title: Cold Climates, Complex Hydrology: Can A Land Surface Model Accurately Simulate Deep Percolation?

Authors: Alireza Amani, Marie-Amélie Boucher, Alexandre R. Cabral, Vincent Vionnet, and Étienne Gaborit

Reviewer Comments and Author Responses

- **The reviewer's comments in orange**
- **The author's response in blue**
- **Excerpts from the manuscript in *green***

We would like to thank the reviewer for their time and constructive feedback. The responses are provided in separate sections, similar to the document submitted by the reviewer.

Comment: The title is not appropriate and does not correspond to the conclusions of the study
Response: Our conclusion will explicitly answer the question posed in the title. The answer is no, SVS cannot accurately simulate deep percolation. This is primarily because the current version of SVS does not include processes such as frozen soil infiltration and water movement through preferential flow pathways. We have shown that these limitations can make the model unreliable for simulating deep percolation in cold regions.

The introduction needs to be more improved

Comment:. I find the introduction disorganized and incomplete. Preferential flows need to be mentioned, and also add more details on LSMs and lysimeters. To facilitate the reader's understanding of the impact of your study, I suggest an organization similar to the below: i) context : groundwater, complications and challenges of cold climates, characteristic and properties of landfill soils (this last part is really missing); ii) models can be used, including Lsm (give more details: their history, why they were developed, aren't they mainly used on a regional or global scale? What's the point of looking at an LSM on the scale of a lysimeter?); iii) Talking about lysimeters: this part is too limited and needs more details.
Response: We appreciate the feedback and the suggested organization concerning the introduction. We will rethink and revise the introduction section to better communicate the importance of the topic to readership. The revised introduction will include a concise exploration of LSMs and lysimeters.
The soils (cover material + topsoil) used in filling the lysimeters and building the experimental plot described in our study are locally available materials and do not belong to any special category such as 'landfill soils'. The soil description is given in the Methods section.
That being said, our field investigation included other experimental plots filled with contaminated soils. Those plots were not part of this study.

Comment: I find the literature search interesting: however, it is also necessary to give some values in the text (and not only in the Figure). Also, if the values correspond to frequencies, then why does the sum not equal 100? And I also don't understand why there are 57 cases in fig 1a while there are 28 in fig 1b; Be coherent.

Response: We will provide statistics within the revised text. We will also revise the figures, using percentages instead of frequencies.
Both figures (Figure 1a and Figure 1b) include 57 cases. In Figure 1b, the " Total" bar may have been interpreted as representing the cumulative count of cases. This was not the intended interpretation. In our figure, this bar represents the number of case studies where the total (annual or cumulative) amount of deep percolation is compared between model and simulation. We will choose an unambiguous label (e.g. Annual / Cumulative) for the revised figure.

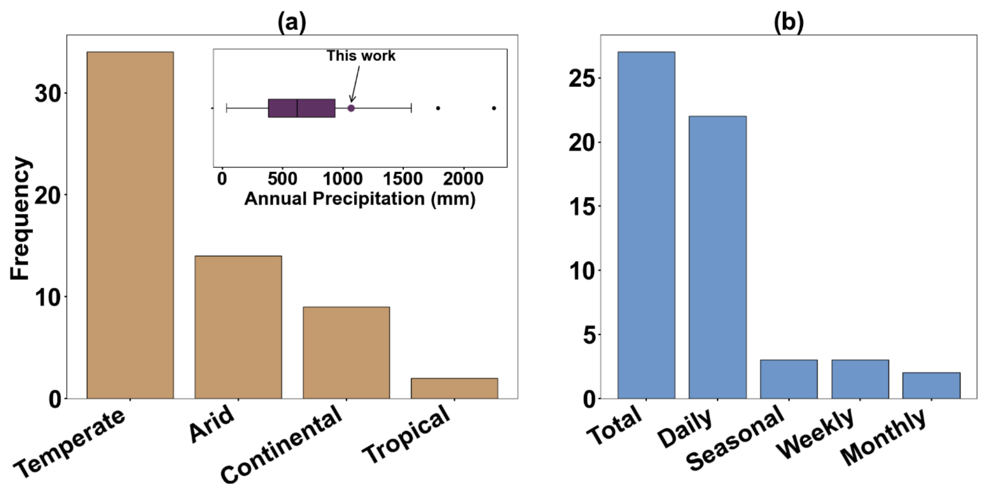


Figure 1

Comment (page 1 line 15)**:** Why are they important?
Response: Currently, the importance of numerical (land surface) models is explained using examples starting from line 19:
LSMs can explore 'what if' scenarios, enhance understanding of physical processes, and provide process-based predictions with explanatory power (Willcox et al., 2021). For example, they help quantify how changing precipitation patterns impact groundwater recharge (Trenberth, 2011; Wu et al., 2020). This is particularly important in regions where fluctuations of groundwater resources heavily influence ecosystem water availability (Huang et al., 2019; Orellana et al., 2012). Beyond natural systems, LSMs can guide the design of landfill final covers. Designers can optimize the covers by simulating how prospective covers interact with local hydrometeorological conditions (Ho et al., 2004) to minimize leachate production. This pre-construction analysis saves time and resources.
We will improve and augment this aspect of the introduction to communicate the importance of land surface models better.

Comment (page 2 line 34)**:** There are many other limitations (Vereecken et al., 2019).
Response: We acknowledge that there are other limitations associated with numerical models that are not mentioned within the text or addressed by our model evaluation effort. Our focus was limited to identifying potential process representation within the SVS model that needs to be improved to better simulate deep percolation in cold regions.

In the conclusion of the revised manuscript, we will highlight the need for future research to explore the additional limitations identified by Vereecken et al. (2019) and others, particularly those related to the numerical solution of the Richards equation.

Comment (page 2 line 53)**:** “This study aims to improve deep percolation simulation”. I cannot see the improvement, as there is no comparison with the initial version...
Response: We will revise the last paragraph of the introduction to communicate our message better: process-based model evaluation is a prerequisite and contributes to improving deep percolation simulation. (We addressed the comment regarding comparing different model versions later in this document.)

Comment (page 2 line 58)**:** Not soil matric potential?
Response: The experimental plot (inside and outside the lysimeters) were also equipped with soil matric potential sensors at various depths. However, these measurements were not included in any of our analyses concerning this manuscript, therefore we decided not to include them in the figures and not mention them in the text.
In the section where we introduce the experimental plot, in the revised manuscript, we will acknowledge the presence of soil matric potential sensors.

Comment (page 3 line 60)**:** Line 60 : Maybe for SVS, but not for LSMs (Boone et al., 2001 ; Decharme et al., 2016)
Response: In this sentence:
Additionally, we implemented a simple soil-freezing scheme for SVS, which is assessed for the first time in the scientific literature.
We meant only for SVS. We will replace it with:
Additionally, we implemented a simple soil-freezing scheme, which is assessed for the first time in the scientific literature for SVS.

Description of lysimeters, soils, datas and parameters are incomplete

Comment: The description of the soils and lysimeters is really incomplete, which makes it difficult to interpret the results. What are the characteristics of these "landfill" soils? What makes them special? And is the soil covered with vegetation? Please comment texture values when you describe soil. 0% of Clay?
Response: We will provide more information (in Section 2.1) on the soils (e.g. soil water retention curves, and classification) in the revised text. We will also refer to soil texture in the text, as suggested.

- As addressed above, no special type of soil is used in filling the L1 and L2 lysimeters.
- The soil is covered with grass, as Figure 2b shows. However, we have not mentioned this in the text. We will include this information within the revised text.

(b)

Figure 2

Comment: There is a wide diversity of lysimetric devices and methods. What are the dimensions, the installation procedures...? I later learn that there are runoff measurements... Is there an imposed slope on these lysimeters? Explain the differences between the lysimeters used in this study, etc...
Response: The installation procedures are detailed in a published article from our group.
Kahale, T., Ouédraogo, O., Duarte Neto, M., Simard, V. and Cabral, A.R., 2022. Field-based assessment of the design of lysimeters for landfill final cover seepage control. *Journal of the Air & Waste Management Association*, *72*(12), pp.1477-1488.
We will briefly provide the information within this reference in the revised text (in Section 2.1)

- The lysimeters are 2% sloped. This was considered in the simulations (as an SVS model's input parameter). We will provide a new figure that visualizes this slope in the revised text.

- Two surface water (runoff) collection systems were designed and constructed in the field. Despite these efforts, neither system yielded reliable measurements. Therefore, they were not used in our model evaluation effort.
- The only difference between the L1 and L2 lysimeters is their height. This was the focus of Kahale et al. (2022) and they have shown that the L2 lysimeter (shorter walls) collects practically the same percolation volume as L1.

Comment: The scheme in fig. 2c could be improved and made bigger with more detail. For me, figs. 2a and 2b are not visible and not necessary.
Response: We will improve Figure 2 (all of its sub-figures) in the revised text, as suggested.

Comment: Some suggestions for Table 1 :
- Add the heading "Observation" for the 1st column and "Model" for the 2nd column
- Illustrate the variability of soil (textures+parameters) for the two lysimeters
- What is N ?

Response: We will include additional headers, as suggested. The lysimeters are filled with the same soil types. The variability associated with the physical properties of these soils is reflected in the wide range of their laboratory-estimated values. The 'N' refers to the number of tests in the laboratory. We will include this definition in the caption of the revised table.

Comment: As you mentioned in your introduction, model parameters are extremely important for the quality of simulations. However, no comment is given in the text on the values obtained. Are these values comparable with these soils ? Are they the same as those generally estimated by LSMs? A value of b equal to 1 is surprising, especially in comparison with pedotransfer functions used by LSMs. Please comment, especially given its importance in the dynamics of percolation simulation.
Response: We will provide a table in a supplementary document (supplementary material) to compare parameter values as estimated by SVS pedo transfer functions and estimated in our laboratory.

Comment: Moreover, the parameters are estimated for the period between April to September 2019. Some studies recognised that changes to the soil and its structure can be significant in lysimeters (Weihermüller et al., 2007; Séré et al., 2012; Seneviratne et al., 2012). Have you performed this protocol at other periods? If so, are the parameters measured different? If not, this could be something to consider, especially in view of the low scores obtained after 2020, and would be a point for discussion.
Response: We have not performed inverse modeling in any other period. However, we acknowledge that the soil properties within the lysimeters might have changed over time. This could potentially explain the low model performance scores. We will include this element in the discussion section of the revised manuscript.

Comment: If there is vegetation, please explain how you set the parameters.
Response: The experimental plot is fully covered with short grass. SVS would classify this vegetation type as "low vegetation." The parameters for low vegetation, such as leaf area index (LAI), roughness length, heat capacity, albedo, stomatal resistance, and root depth, are derived from lookup tables containing values for 21 vegetation classes. In this case, the lookup table would be used to select the appropriate parameters for short grass. The details are provided in:
Alavi, N., Bélair, S., Fortin, V., Zhang, S., Husain, S.Z., Carrera, M.L. and Abrahamowicz, M., 2016. Warm season evaluation of soil moisture prediction in the Soil, Vegetation, and Snow (SVS) scheme. *Journal of Hydrometeorology*, *17*(8), pp.2315-2332.
This information will be added to Section 3.2 of the manuscript.

Comment: Comment on the quality of your data, the gaps, and how you integrate it into your analysis.
Response: The potential influence of these gaps on the analysis will be discussed in Section 4 (Results) of the revised manuscript. When calculating performance metrics, time steps with missing measurements were not considered.

Comment: For me, the "Methods" section is not simply a Methods section, in particular the model description. Find a more appropriate name for this section or a better structure.
Response: We will choose a more descriptive name for this section (e.g. 'Methodology and Model Description').

There is no comparison with the initial version

Comment: As I said earlier, in order to assess the contribution of the gel to this model, we need to compare it with the original version. This procedure is essential to highlight this work.
Response: The primary objective of this manuscript was to conduct a process-based model evaluation of the SVS model. Assessing the improvement associated with adding the freezing module was not part of our study.

In the revised manuscript, we will briefly discuss the differences in model performance with and without the soil freezing module in Section 4.2. We will refer the reader to the supplementary material for a comprehensive comparison between the SVS model with and without the soil freezing module activated in the supplementary material.

Analysis between variables (periods and criterias) are confusing

Comment: It's very complicated to compare Figs 3, 4, 5, 7 and 8, because they don't have the same evaluation periods. I can understand Fig 3, but not for the others. You have to be consistent. It is very confusing for me. In addition, evaluation metrics are not always identical from one analysis to another, especially correlation. This is a criterion to be included in each analysis.
Response: We will revise the figures and we will ensure the suggested consistency is respected.

Results not presented in enough detail

Comment: You never mention two essential concepts: ETP and water balances. However, these are still key aspects of hydrology and can be estimated using lysimeters. An assessment of the model's capacity to reproduce these two elements would be advisable.
Response: During our field investigation, direct measurements of evapotranspiration (ETP) were not conducted due to the construction of pan lysimeters instead of weighing lysimeters. As previously mentioned, we tried to measure runoff from the surface of the experimental plot by designing and constructing two systems; however, neither system yielded reliable measurements. Consequently, indirect estimation of ETP is also not feasible in our case.
Evaluation of SVS's ability to calculate surface energy balance (including ETP) was the focus of another study:
Leonardini, G., Anctil, F., Abrahamowicz, M., Gaborit, É., Vionnet, V., Nadeau, D.F. and Fortin, V., 2020. Evaluation of the Soil, Vegetation, and Snow (SVS) land surface model for the simulation of surface energy fluxes and soil moisture under snow-free conditions. *Atmosphere*, *11*(3), p.278.
Leonardini et al. (2020) assessed SVS's ETP simulation capabilities across six FLUXNET sites (2004-2015) and found it generally performed well. The reference to the ETP study by Leonardini et al. (2020) will be included in section 2.1, where the limitations in measuring runoff are discussed.

Comment: Then, I find there is very poor comment on the observations, yet observations are also a result in themselves.
Response: A comprehensive process-based evaluation of field observations is the topic of a separate manuscript (written by the same authors as the current manuscript) presently being reviewed for publication in Hydrological Processes. We expect this manuscript to be published soon. In that case, we will cite the published article in the revised manuscript.

Comment: Also, in the observations, it would be good to highlight the variability between lysimeters, as you did for Fig 8. That would make it possible to assess heterogeneity, the influence of the size of the lysimeters, changes in the soil over time, etc.

Response: This is a very good suggestion. We will modify and reproduce the relevant figures accordingly.

Comment: I also think it would be interesting to have a seasonal cycle, in addition to Fig 6. to have a temporal dimension in the differences between observation and model.
Response: The caveat in displaying the comparison between the model and the observations seasonally is the duration (months of data) with absent measurements for lysimeters. This may result in a prejudiced comparison. With this reservation, we will provide a seasonal cycle in the supplementary material.

Comment: In the Figs 3, 4, 5, 7 and 8, it would also be good to include the statistical values over the complete period (although this is included in the text).
Response: We will modify and reproduce the figures to include the evaluation metric over the entire period.

Snow Depth

Comment: Where is the year 2022?
Response: Figure 3-c shows the comparison (SVS vs measurement) for the third winter, including 2022.

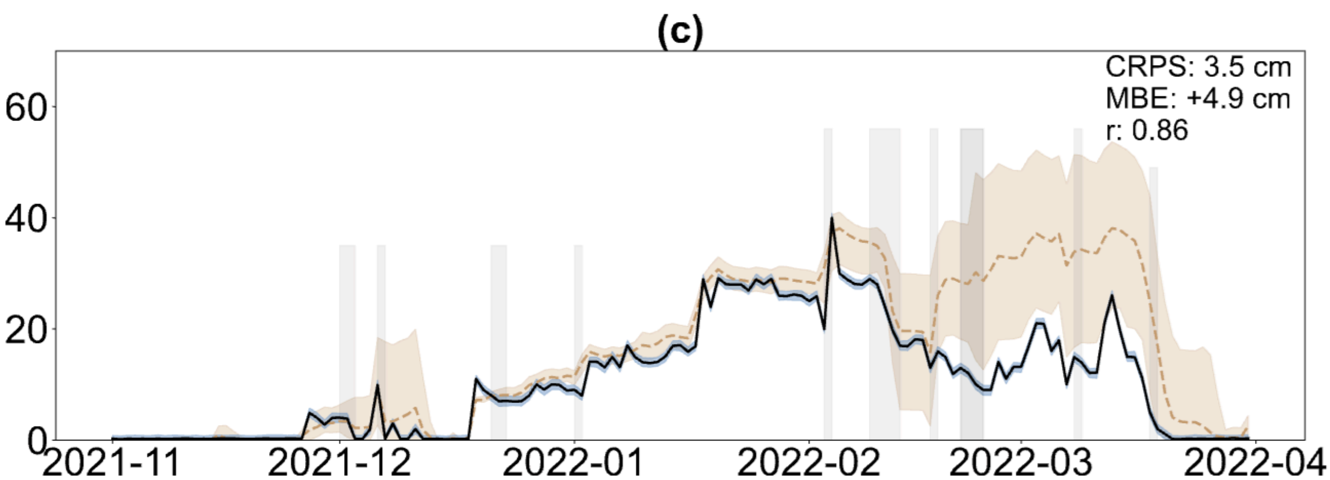


Figure 3

Comment: Comment the very similar dynamics (sensitivity) of the model at the end of winter. For example, the maximum peak is almost always at the same time of year, with a value very close to 50 cm.
Response: We will investigate the results (data + model output) to identify potential explanations. For instance, this might be due to a similar volume of snowfall ocurring during the three winter periods.

Near-surface Soil Temperature

Comment: Do you have an estimation of the depth of frost?
Response: There are soil temperature sensors 600 mm below the surface in the E1 enclosure. This may help to have an upper bound for the first depth. We will include this in Section 5.1, where we discuss the cold bias issue during the snow-free condition.

Near-surface Soil moisture

Comment: Please add the correlation in your analyse. I guess the correlation is not very good. Comment. And notably in summer when the observation is very low until model is stable. This may illustrate the fact that the model does not represent ETP very well. Is there any vegetation that can increase ETP?
Response: The correlation (for warm/cold months) is depicted in Figure 6b. This statistic will also be included in time-series figures. We will provide commentary on the potential for the model to experience difficulty with ET as a possible factor contributing to the discrepancy between the model and observations during the summer months.
The plot is fully vegetated, and this is taken into account by the model. As stated earlier, we will add information about the vegetation in the revised text.

Deep percolation

Comment: Please comment the observation! What do the periods without observation lines mean (e.g. 2020-11 to 2021-03); Is this a period without drainage? or a period without observation? How is this taken into account in your statistical evaluation?
Response: To clarify Figure 8, we will modify it to distinguish between times when no volume was collected and times when observations were not possible due to equipment malfunctions.

Comment: I don't agree with your comment when you say the model correctly reproduces the observations. Perhaps the CRPS is good, but the dynamics are not good, with significant dephasing and a model that reacts too quickly and too intensely.
Response: This is a fair argument. We will modify the text to better reflect the model's struggles concerning deep percolation.

Comment: Is it possible that the low scores after 2021 are due to a transformation of the environment within the lysimeters over time? in structure and hydrodynamic parameters?
Response: It is possible. Nonetheless, our analyses showed that the model's inability to capture mid-winter events in 2022 is the primary factor contributing to the low scores. This failure is due to the model's inability to simulate frozen soil infiltration and the movement of water through preferential flow pathways.

Discussion should be improved

Comment: I rather like the discussion, which is well constructed and well developed.
I think an analysis of the sensitivity of the parameters could be useful for this study. For example, if we only perturbed one parameter, or only the meteorological forcing, how will the simulations be affected?
Response: We also believe such analysis would be an interesting dimension to explore. However, it was outside the scope of the current study. This can be pursued in the future using the publicly shared dataset from our field

investigation using SVS or other land surface models. We will mention this as a suggestion in the revised Conclusion.

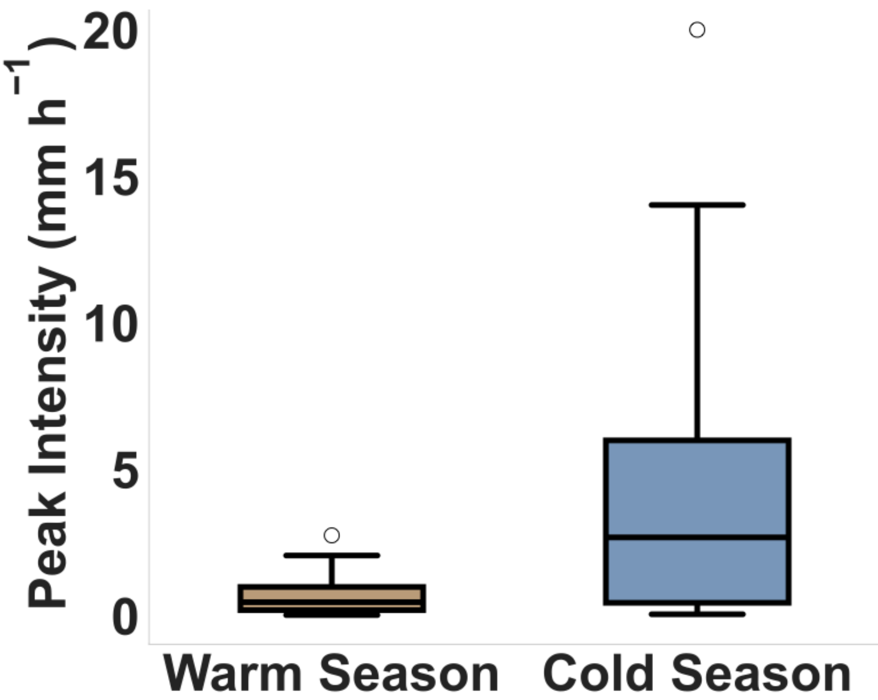
Comment: Figures should be improved with the use of hovmoller diagram. These diagrams are often used on lysimeters to follow the movement of water in the soil and to help comprehension (depth vs time) (Abdou et al., 2004, Decharme et al., 2016, Sobaga et al., 2023).
Response: We agree that such diagrams would help to better visualize the water movement across the soil column. However, as only a few sensors exist, the Hovmoller diagram might not be as informative as expected. Additionally, the sensors are not placed at regular intervals, which would make the interpretation of the diagram even more challenging. Therefore, we decided not to include Hovmoller diagrams in our current manuscript.

Comment: For the example you are using in this section, please use a symbol (a star?) on the previous chronicles (Fig 3 - 8).
Response: We will use symbols as suggested.

Limitations of SVS in simulating soil moisture and deep percolation in winter

Comment: Can you examine how many events the model does not react to while the observations do react? And give a statistic for the cases where preferential flows are observed?
Response: We will examine the results and provide statistics for when the model fails to react and when it is likely that we observed preferential flows.

Comment: Furthermore, don't you think that these preferential flows would not be due (or at least reinforced) by the lysimeters? Either by edge effects? or by a modification of the soil?
Response: In another manuscript–submitted before this one and currently under review after revision–, where we analyzed deep percolation observations from four lysimeters (including the L1 and L2 lysimeters), we showed that preferential flow mostly happened during cold months, reflected by distinctly high peak intensities for deep percolation events starting in the cold season. This is shown in the below figure. While we acknowledge that edge effects within the lysimeters may have reinforced preferential flow to some extent, the strong seasonal pattern of preferential flow events suggests that lysimeter artifacts are not the sole cause.



We will discuss in Section 5.2 the possibility of the influence of lysimeters on preferential flows.

Influence of Capillary Barrier on Simulated Subsurface Soil Moisture

Comment: Do you try to change the thickness of the last layers?
Response: Changing the thickness of the last layers from 2.5 cm to 2.0 cm (lower value will cause numerical stability in our version of SVS) does not change the simulation results and the observed behavior.

Comment: Are these results also verified on the Lysi L2? If not, why not?
Response: We will investigate the L2 observations (soil volumetric liquid water content at 1850 / 1750 mm) and comment on them in Section 5.3.

Comment: Please comment the Fig 10.a,
Response: we will provide more comments concerning Figure 10a in the revised manuscript.

Comment: A simple simulation with the same parameters and a free LBC would be welcome to confirm these observations, and to check if this is only the effect of the LBC.
Response: This is a good idea, however, an option to select a free drainage LBC is not implemented in the SVS model. We will mention in the conclusion that working on implementing this option is an interesting improvement for SVS.

Other Comments

Comment: Lsm are generally used on a regional or global scale. In your opinion, what would be the main challenges of the change of scale for these cold climates? And how would you integrate preferential flows on a larger scale?

Response: While LSMs are used at regional or global scales, we believe their application at smaller scales is straightforward. Working at smaller scales can be advantageous, as we have reduced uncertainty in parameter values and meteorological data. This allows for better model calibration and validation, ultimately leading to more reliable simulations of the hydrological processes specific to cold environments.

We will include in the revised manuscript (introduction), with relevant references (see below for examples), the fact that LSMs are also employed in point-scale.

- Zhang, Li, et al. "Evaluation of the Community Land Model simulated carbon and water fluxes against observations over ChinaFLUX sites." Agricultural and Forest Meteorology 226 (2016): 174-185.
- Denager, Tanja, et al. "Point-scale multi-objective calibration of the Community Land Model (version 5.0) using in situ observations of water and energy fluxes and variables." Hydrology and Earth System Sciences 27.14 (2023): 2827-2845.
- Chadburn, S., et al. "An improved representation of physical permafrost dynamics in the JULES land-surface model." Geoscientific Model Development 8.5 (2015): 1493-1508.
- Harper, Anna B., et al. "Improvement of modelling plant responses to low soil moisture in JULESvn4. 9 and evaluation against flux tower measurements." Geoscientific Model Development Discussions 2020 (2020): 1-42.

Comment: When you use LBC seepage for hydraulic conditions, how do you define LBC for Temperature?

Response: This is described in the A2 section of the Appendix.

Additional notes: We will address and incorporate all the suggestions designated as ’Minor Corrections’ in the revised manuscript.