

RC 1:

General Comment

The manuscript presents a carefully designed set of large-scale laboratory experiments investigating infiltration, runoff, and drainage in frozen slopes, with particular emphasis on the role of macropore networks and antecedent water content. The experimental setup is ambitious, the dataset is rich, and the effort to move beyond traditional column experiments toward a slope-scale configuration is appreciated. The topic is timely and relevant to cold-region hydrology and slope processes, and the paper has the potential to become a useful benchmark dataset for model testing. The manuscript is well-organized and prepared with clarity. Most of the methodological limitations are clearly outlined.

However, in its current form, the manuscript systematically over-interprets its results in favor of its hypothesis. Where the observations are robust, the conclusions are largely intuitive and confirm existing understanding; where the authors advance more interesting or non-intuitive interpretations, the supporting evidence is insufficiently constrained. Several claims conflate what is observed with what is inferred, and in some cases the causal chain between measurements and conclusions is not convincingly established.

For these reasons, I do not recommend acceptance in the current form. At the same time, I do not consider the work fundamentally flawed. A major revision is appropriate, provided that the authors substantially revise the interpretation, tighten the logic, and clearly delimit what can and cannot be concluded from the data.

Response:

We thank the reviewer for the careful and thorough evaluation of our manuscript and for the constructive tone of the overall assessment. We particularly appreciate the recognition of the experimental design, the slope-scale configuration, and the value of the dataset for model testing.

We acknowledge the central concern raised by the reviewer regarding over-interpretation and insufficient separation between direct observations and mechanistic inference. We have carefully revisited the manuscript in light of this critique and substantially revised the Results and Discussion sections to (i) more clearly distinguish between measured quantities and interpretation, (ii) avoid causal overstatements, and (iii) explicitly delimit what can and cannot be concluded from the available data. In addition, we have expanded the discussion of alternative explanations, transferability, and methodological limitations.

We believe that these revisions significantly tighten the logical structure of the manuscript and bring the strength of the conclusions into closer alignment with the supporting evidence.

Major Comment 1

A recurring issue throughout the manuscript is the insufficient separation between “direct observations” and “interpretive statements”. This is particularly evident in claims regarding

preferential bypass flow at intermediate initial water content, “threshold-type” behaviour around a specific volumetric water content, and macropore controlled advective heat transport.

In several places, statements are phrased as if the experiments “demonstrate” specific mechanisms, whereas the data are at best “consistent” with those mechanisms. Given the absence of direct diagnostics of flow paths (e.g. tracer tests, dye experiments, imaging), preferential flow remains an inference, not an observation. This distinction must be made explicit throughout the manuscript. Of course, some degree of interpretive uncertainty is inevitable in experiments of this type. My concern here is not ambiguity per se, but overstatement relative to that ambiguity.

I strongly encourage the authors to revise the text so that:

- observational results (fluxes, temperature fields, ΔVWC patterns) are clearly described as such, and*
- mechanistic explanations are framed as hypotheses consistent with the observations, rather than as demonstrated processes.*

Related to this point, several of the key interpretations rely on transient behaviour (e.g. early drainage onset, switching between drainage- and runoff-dominated regimes, and progressive loss of macropore effectiveness). These are inherently rate-related phenomena, yet the analysis is presented almost exclusively in cumulative (volume) form. While cumulative fluxes are informative for overall partitioning, complementary rate-based representations (e.g. time-resolved inflow, runoff, and drainage rates, and their balance) could provide a more quantitative basis for comparing scenarios and for assessing whether observed differences reflect transient dynamics or simply integrated effects. I present this as a suggestion rather than a requirement, but such analyses could help clarify several of the interpretations advanced in the manuscript.

Response:

We agree that in the original version, the distinction between direct observations (e.g. drainage onset, cumulative fluxes, ΔVWC patterns, and temperature evolution) and mechanistic interpretation (e.g. preferential bypass flow and macropore-controlled heat transport) was not always sufficiently explicit.

In the revised manuscript, we have systematically rephrased mechanistic statements to more clearly distinguish observational results from interpretation. Formulations such as “demonstrate” or “confirm” have been replaced with more cautious wording such as “indicate”, “suggest”, or “are consistent with”, where appropriate. These revisions were implemented across the Results and Discussion sections, particularly in Sect. 3.2, Sect. 3.3, Sect. 4.2, and the Conclusions. We have also clarified explicitly that preferential flow is inferred from the combined hydraulic and thermal observations rather than directly visualised within the present experimental setup, including in the discussion of the intermediate- and high-VWC experiments in Sect. 3.3 and Sect. 4.2.

We also agree that several of the key interpretations concern transient behaviour and are therefore better supported by time-resolved information than by cumulative volumes alone. In response, we added complementary time-resolved flow representations to the appendix/supplement, where inflow, runoff, and drainage rates are shown alongside the cumulative fluxes presented in the main manuscript. These additional plots provide a clearer basis for discussing transient flow partitioning and the progressive reduction in macropore effectiveness during irrigation.

Major Comment 2

The manuscript repeatedly refers to a “threshold” initial volumetric water content (around 12–13%) above which macropores become hydraulically important. While the qualitative distinction between low, intermediate, and high initial water content regimes is evident in the data, the use of the term threshold is a bit too strong. Only a small number of discrete initial conditions are tested, and the inferred transition depends on soil texture, porosity, macropore geometry, freezing history, and experimental boundary conditions. As such, the data do not justify the identification of a sharp or general threshold, nor the presentation of a specific numerical value as physically meaningful beyond this setup.

The authors should:

- *avoid presenting the observed transition as a universal or sharply defined threshold,*
- *clearly state that the reported values are setup-specific, and*
- *reframe the discussion in terms of regime behaviour rather than threshold behaviour.*

More generally, the manuscript would benefit from a clearer discussion of transferability. The experiments are necessarily highly controlled, with a specific soil texture, porosity, macropore geometry, and freezing history. While such idealization is appropriate for process understanding, the current discussion does not sufficiently articulate which aspects of the observed behaviour are expected to be transferable to other soils, macropore configurations, or thermal regimes, and which are strictly setup-specific. Clarifying this distinction would help readers assess how the results should be interpreted beyond the particular experimental configuration studied here, and would strengthen the contribution of the paper.

Response:

We agree that the original wording may have implied a sharply defined or transferable volumetric water content threshold. This was not our intention; rather, we aimed to describe a transition in system behaviour within the tested parameter range.

In the revised manuscript, we have replaced “threshold-type behaviour” with more appropriate formulations such as “transition” or “regime shift” in the discussion of infiltration behaviour. In particular, Sect. 4.2 now emphasises that the observed shift in response across the tested 12–16 % VWC range reflects a configuration-specific transition rather than a

universal threshold. We also state explicitly that the identified range depends on soil texture, porosity, macropore geometry, freezing history, and boundary conditions.

In addition, we expanded the discussion of transferability in the final part of Sect. 4.2 to more clearly distinguish between process-based findings that may extend beyond this setup and aspects that remain specific to the present experimental configuration.

Major Comment 3

The interpretation of earlier drainage onset in the intermediate water content macropore experiment as evidence of preferential bypass flow is not fully convincing in its current form. While the observation itself (earlier drainage relative to the non-macropore case) is clear, alternative explanations cannot be ruled out, including packing heterogeneity introduced during soil preparation, localised shrinkage or cracking during freezing and wetting, preferential flow along probe–soil interfaces, or differences in frost geometry not fully captured by averaged profiles. Specially given how small sometimes the differences are.

Given that the authors acknowledge artefacts in other experiments (e.g. near-surface sensor exposure and structural heterogeneity), the manuscript should explicitly discuss these alternative explanations and justify why preferential flow through the artificial macropore network is the most plausible interpretation. At minimum, the language should be softened to reflect the inferential nature of this conclusion.

Response:

We appreciate this important comment and agree that alternative explanations must be explicitly considered. In the revised manuscript, we have added a dedicated paragraph in Sect. 4.2 discussing potential artefacts and competing explanations, including packing heterogeneity, structural rearrangement during freezing and wetting, probe-soil interface effects, and minor differences in frost geometry.

We now state more clearly that preferential bypass flow was not directly observed but inferred from the combined hydraulic and thermal responses. At the same time, we explain why preferential bypass flow through the artificial macropore network remains the most plausible interpretation of the observed pattern, particularly considering the combination of earlier drainage onset, comparable early inflow volumes, and the depth-resolved evolution of temperature and VWC. The relevant passages in Sect. 3.2, Sect. 3.3, and Sect. 4.2 were revised accordingly to reflect the inferential nature of this conclusion more explicitly.

Major Comment 4

The macropore network is central to the study, yet its mechanical and hydraulic integrity is not sufficiently explored. Important questions remain insufficiently addressed, including whether macropores remain open and hydraulically connected at the onset of irrigation (specifically relevant for such a coarse material), the extent to which macropore collapse or

partial closure may occur during freezing and wetting, and how representative the chosen macropore diameter, orientation, and connectivity are relative to natural systems.

While the authors describe the network as a simplified analogue, the manuscript should more explicitly acknowledge that this configuration represents an upper-bound scenario for macropore influence and discuss the implications for transferability to natural soils.

Response:

While the manuscript already noted that the selected macropore diameter lies at the coarse end of natural ranges, we agree that a more explicit discussion of structural and hydraulic integrity is warranted given the central role of the macropore network in this study.

In the revised manuscript, we strengthened both the Methods and Discussion sections. In the Methods section, we now describe preliminary stability tests conducted in smaller soil columns and buckets prior to the large-scale experiments. These tests showed that the 10 mm cylindrical conduits remained mechanically open after freezing under comparable thermal conditions.

In the Discussion, we clarify that no visible macropore collapse was observed at the onset of irrigation in the large-scale setup. Nevertheless, we explicitly acknowledge that small, undetected structural rearrangements, particularly in deeper soil layers not directly observable prior to excavation, cannot be completely excluded. We also emphasise more clearly in the Limitations section that the artificial, uniformly sized macropore network likely represents an upper-bound scenario in terms of hydraulic connectivity and structural persistence. Natural macropore systems are typically more heterogeneous, irregular, and partially infilled, and may therefore exhibit lower effective connectivity under freeze-thaw conditions than the simplified network used here. We additionally clarify that hydraulic connectivity was inferred from system response rather than directly measured.

Major Comment 5

Changes in volumetric water content during irrigation are interpreted in several places as evidence of infiltration, bypass flow, or macropore-driven transport. However, under partially frozen conditions, ΔVWC may reflect multiple processes, including liquid redistribution, phase change (melting or refreezing), and measurement artefacts near the freezing point.

Given the strong sensitivity of dielectric measurements near 0 °C and the reliance on relative rather than absolute changes, the manuscript should be more cautious in attributing ΔVWC patterns to specific hydraulic processes. Statements that rely heavily on ΔVWC fields should be revisited and, where necessary, qualified.

Response:

We agree that under partially frozen conditions, changes in measured (unfrozen) volumetric water content can reflect multiple coupled processes, including liquid redistribution, local phase change, and temperature-dependent dielectric effects. In particular, near 0 °C, dielectric

responses may be sensitive to freeze-thaw dynamics and should not be interpreted as unambiguous indicators of hydraulic processes.

In the revised manuscript, we clarified this distinction both in the Results and in the Limitations section. Sect. 3.3 now states more explicitly that ΔVWC patterns are interpreted in conjunction with independently measured hydraulic fluxes and concurrent temperature evolution. Sect. 4.4 further emphasises that these spatial patterns are used to characterise system response rather than to directly visualise preferential flow paths.

Major Comment 6

Several passages attribute downward migration of the freezing front during irrigation to advective heat transport associated with infiltration. As currently written, this interpretation is not always physically clear. Infiltrating water is warmer than the frozen soil, and advective heat transport would intuitively promote thawing rather than deeper freezing unless the coupled effects of phase change and latent heat release are explicitly considered. The authors should clarify the underlying energy balance and ensure that the explanation of freezing-front movement is physically consistent and clearly articulated.

Response:

We thank the reviewer for highlighting this important point. Upon reconsidering the underlying energy balance, we agree that the original wording may have overstated, or at least could not conclusively resolve, the role of advective heat transport as the primary driver of freezing-front migration.

In principle, advective heat redistribution may occur during infiltration, for example when infiltrating water equilibrates rapidly to near-surface temperatures and is subsequently transported to deeper soil layers. Such processes could locally modify thermal gradients. However, these patterns are not distinguishable between macropore and non-macropore experiments and cannot be quantitatively separated from concurrent conductive heat transfer.

The revised manuscript therefore clarifies that while infiltrating water introduces sensible heat, the associated heat input is small relative to the thermal deficit of the frozen soil. The magnitude of freezing-front migration correlates more strongly with differences in antecedent soil temperature and frost penetration depth, indicating that the pre-existing thermal state and conductive heat transfer exerted the dominant control on temperature evolution. Advective heat transport may locally influence thermal gradients, but its contribution cannot be quantitatively isolated based on the available data. We revised the relevant passages in both Sect. 3.3 and the Discussion accordingly to ensure that the explanation of freezing-front evolution is physically consistent and does not imply a causal dominance of advection.

Minor Comment 1

The potential influence of the measurement probes themselves on hydraulic and thermal behaviour should be commented on.

Response:

We agree that the possible influence of the measurement probes on hydraulic and thermal behaviour should be acknowledged more explicitly. In response, we expanded the Limitations section to address potential probe–soil interface effects. While minor local disturbances cannot be entirely excluded, probe positions were identical across experimental configurations, and no systematic probe-aligned flow behaviour was observed. We therefore expect any probe-related influence to be secondary relative to the imposed differences between the macropore and non-macropore setups.

Minor Comment 2

Statements such as “minor spatial variations” or “no systematic effect” should be supported by quantitative measures.

Response:

We agree that qualitative statements such as “minor spatial variations” or “no systematic effect” should be supported by quantitative information. In the revised manuscript, we therefore added quantitative measures where appropriate, including metrics describing lateral versus vertical temperature gradients during the freezing phase, as well as additional quantification of inter-experimental differences in the Results section and appendix material. These revisions were intended to replace qualitative descriptions with more transparent numerical support wherever possible.