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.

Reviewer 2 – General Assessment

The authors present an excellent experimental study on the effects of macropores on the partitioning of rainfall between infiltration, runoff and drainage in frozen hillslopes. The hillslope-scale experiments are a significant advancement of past 1D soil column experiments, and allows investigation of more realistic multi-dimensional hillslope behavior. The results are not nearly as novel as the authors claim, almost all of the findings presented here have been shown in other studies (see below). However the results still provide valuable insight into the importance of antecedent soil moisture on preferential flow, as well as enhanced infiltration and bypass flow in frozen soils due to preferential flow, refreezing of infiltrated water in macropores and its effect on the temporal evolution of runoff generation in drainage in frozen soils.

A novel aspect is the direct observational evidence of freezing of water in macropores initiating from the pore-walls, corroborating the findings of Watanabe and Kugasaki (2017). We need more than one study to confirm this behavior, and this study does a great job of showing this in hillslopes, also corroborating the conceptual model put forward in Mohammed et al. (2018). The other major advancement in this study is providing a well-controlled dataset that will be extremely valuable for testing emerging dual-permeability models of water flow in frozen soils. I very much look forward to future modeling work with this dataset. This is a solid contribution and should be published. However, the authors need to walk back some of the strong language used in the manuscript, as this experimental work also suffers from many of the limitations of field studies, in that many of the preferential flow processes discussed are inferred rather than directly observed.

Preferential flow in frozen soil is still a pretty niche field compared to more mature fields in cold regions hydrology, but interest is growing due to increased consensus that it has significant control on water partitioning in frozen soils. There has been previous work that has shown many, if not all, of the findings in this work, yet the authors have not acknowledged many of these important studies. They appropriately mention the seminal experimental work of Watanabe and Kugisaki (2017) and Pittman et al. (2020), yet Holten et al. (2019) and Grant et al. (2019) have also shown many of these infiltration, preferential flow and refreezing dynamics. The authors need to acknowledge these works as their work builds upon these previous studies. Similarly, there have been pioneering field and modeling studies on hillslopes that have also shown almost all the findings of this work that the authors don't acknowledge either, including Mohammed at al. (2019), Rey et al. (2021), Mohammed et al. (2021b), Hyman-Rabeler and Loheide (2023), and Sanchez-Rodriguez et al. (2025). They refer to a review paper like Walvoord and Kurylyk (2016) that has almost nothing to do with the work presented here and yet don't mention Ireson et al. (2013). One might think that the authors have purposely left out references to these studies in this manuscript to seemingly enhance the perceived novelty of the work presented. The authors need to give credit where credit is due.

I believe with minor revisions, this will be an excellent contribution to the field of cold regions hydrology.

Response:

We thank Reviewer 2 for the positive and encouraging assessment of our work. We particularly appreciate the recognition of the slope-scale experimental design and the value of the dataset for testing dual-permeability and freeze-thaw modelling approaches.

We acknowledge the reviewer's concern regarding the strength of certain formulations and the framing of novelty. In the revised manuscript, we have moderated causal language, clarified the inferential nature of preferential flow interpretations, and strengthened the positioning of our study within the existing body of literature. We agree that some statements in the original version were formulated too strongly and that similar mechanisms have been observed and discussed in previous studies.

We also appreciate the reviewer's suggestions regarding additional literature. We recognise the importance of thoroughly situating our work within the established research landscape on preferential flow in frozen soils. In response, we have carefully incorporated the recommended references and expanded both the Introduction and Discussion to more explicitly acknowledge prior contributions. The inclusion of these studies has strengthened the contextualisation of our results and clarified how our work builds upon existing knowledge.

We emphasise that the novelty of our contribution does not lie in identifying preferential flow processes per se, which have indeed been documented previously, but in investigating these processes under controlled slope-scale conditions using a reproducible three-dimensional macropore network and dense spatial instrumentation. This intermediate hillslope-scale experimental configuration allows us to observe process interactions that bridge traditional one-dimensional column experiments and field-scale studies.

Minor Comments

RC2: L43: Should cite Larsbo et al. (2019) and Mohammed et al. (2021a) here, as these two were the first to include differences in freezing between macropores and the soil matrix in numerical models, using the conceptual model put forward in Mohammed et al. (2018).

AC: We have incorporated the suggested references into the revised manuscript and integrated them more explicitly into the Introduction to better position our study within the existing literature.

RC2: L52: This sentence should reference works like Mohammed et al. (2021b) and Sanchez-Rodriguez et al., (2025).

AC: We added the citations to this sentence.

RC2: L67-68: Should also cite Mohammed et al. (2019) and Sanchez-Rodriguez et al., (2025) here as well.

AC: We agree and added the suggested literature to the sentence.

RC2: L100: Please report on the stability of the temperature range of the compressor used to main temperatures, i.e. show that the temperature control is actually stable during frozen conditions.

AC: We agree that the original wording may have implied that the climate chamber maintained a perfectly stable air temperature at the nominal set point, which is not the case. The cooling unit operates in periodic compressor cycles, resulting in controlled temperature oscillations around a target value. In the revised manuscript, we have clarified the description of the temperature regulation and adjusted the wording to avoid any misleading implication of constant air temperature. We now explicitly state that this is the possible operating range of the control unit. In addition, we have added a representative temperature record of the compressor cycling behaviour to the Appendix (E) for transparency. We thank Reviewer 2 for highlighting this point, which has helped us improve the clarity and accuracy of the methodological description.

*RC2: L102-104: What was the insulation material surrounding the soil box? The authors state that insulation "ensured that freezing initiates **exclusively** at the air-soil interface, akin to natural conditions". It is not sufficient to just state this, you have to show it. What you should have done is place temperature sensors at the edges of the box and compare the thermal profiles at the center and edges of the box. You should use the thermal profiles of the probes S1-S5 to show that ambient temperature interference was actually minimized, and lateral thermal gradients are at least an order of magnitude smaller than vertical gradients. Nagare et al. (2012) and Mohammed et al. (2014) showed that passive insulation like that used here will always fail at some point, so you need to show that the set-up actually mimicked top-down thaw, as the thawing period is more important that the freezing period for these experiments. While I don't necessarily think this is a major issue for this type of experiment, statements like this made with such certainty need to be backed up by data. Your figures 5-7 clearly show lateral thermal gradients with those curved lines showing frost depths, especially FN10.*

AC: We agree that a clearer quantification of lateral versus vertical temperature gradients strengthens the methodological transparency of the experimental setup. In the revised manuscript, we have moderated the original wording in the Methods section to avoid implying strictly one-dimensional freezing conditions. We now explicitly acknowledge that lateral thermal influences are present. In addition, we have included a chapter in the appendices (F) showing the temporal evolution of the relative magnitude of lateral and vertical temperature gradients.
We agree that lateral influences are detectable even during freezing, and we now state this explicitly in the manuscript. Future experimental configurations may include additional boundary-focused temperature measurements to further refine the assessment of lateral heat exchange.

RC2: L246-247: Again, here you mention The temperature and VWC distributions were found to be largely uniform across the soil body, with only minor spatial variations and weak boundary influences near the walls and bottom You're asking the reader to take your word for this. See my suggestion above about comparing lateral versus vertical thermal gradients. This could be in the appendices. Figure 3: Somewhat related to the previous comments above. I understand why the authors have averaged the VWC and temperature probe readings, I suspect that there are plans in the pipeline to model this experiment with dual-permeability model similar to Khanahmadi et al. (2026) but modified for freeze-thaw. However, it would be nice to see the profiles from the individual probe profiles. This, again, could go into the appendices.

AC: We agree that the original wording was too qualitative and did not sufficiently quantify the reported spatial variability. In the revised manuscript, we have added explicit quantitative measures of lateral and vertical temperature gradients in the appendices (see also answer to L. 102-104) and changed the original wording. We also added profiles from individual probes for two representative experiments in the appendices (G). We did not include all profiles, as including all plots in the manuscript would substantially increase its length. The experiments FN15 and FM15 were selected because they clearly illustrate the main experimental patterns discussed in the manuscript. These experiments were not chosen because they exhibit the strongest or most variable responses, but because they provide the most influential results.

RC2: L263: You state that 'Concurrent VWC decreases in the upper 10-20 cm (> 0.5 %) indicate early redistribution of unfrozen water'… so where did this water go? All depths show either decreases or no change. Cryo-suction redistributes water to the freezing front, so was water redistributed to the shallow soil where no sensors were present? If you don't have measurements to confirm, another reason could be that there was vapor loss as well, so you can't say for sure that this loss was due to redistribution... especially in the coarse-grained soil used here.

L265-266: Downward redistribution during freezing goes against our current understanding of the effects of freezing on water migration (i.e. cryo-suction). This does not make sense, especially at drier conditions in coarse grained soil as the authors have used here.

AC: We agree that classical cryo-suction mechanisms promote upward water migration toward the freezing front, particularly in fine-grained soils. The slight decreases in VWC observed in the upper 10-20 cm occur during the early cooling phase, prior to the onset of significant ice formation, and are therefore not interpreted as cryo-suction-driven redistribution. In the high initial water content experiments, small water losses into the drainage layer were detected before freezing commenced. These early changes are most plausibly explained by gravitational drainage during cooling and minor structural settling of the soil matrix. Given the coarse texture and low capillarity of the experimental soil, substantial cryo-suction effects are not expected. Temperature-dependent sensor effects associated with soil cooling cannot be entirely excluded (even though the applied three- and four-phase dielectric mixing approach explicitly incorporates temperature effects). However, the magnitude of the observed variations is small and consistently detected across experiments. Regarding the reviewer's question "where did the water go?", we note that during the early cooling phase, part of the observed VWC decrease in the upper layer corresponds to measurable drainage losses (in the high VWC experiments).

At later stages during freezing, the reduction in measured VWC primarily reflects phase change rather than mass redistribution. As described in the methods section, the sensors measure liquid volumetric water content. Consequently, a decrease in VWC during freezing represents conversion of liquid water to ice rather than physical water loss from the system.

To avoid confusion, we have revised Sect. 3.1 to clarify (i) the distinction between early gravitational redistribution prior to freezing and (ii) later reductions in liquid VWC due to phase change. We have also explicitly reiterated that the reported VWC values represent liquid water content. Furthermore, we have moderated the wording describing early VWC decreases in the upper layers during the initial cooling phase to reflect the associated uncertainty in their precise physical origin.

RC2: L279: Again, do you have data to support this statement?

AC: We revised the corresponding statement, as we agree that it was formulated too definitively.

RC2: L287-289: These cumulative plots are great, but it would be nice to see some temporal fluxes in addition to the cumulative plots.

AC: We agree that time-resolved flux representations provide valuable additional insight into the transient dynamics of the system. Especially at those experiments, where fluxes are changing dynamically. In the revised manuscript, we have therefore added a dedicated section in the appendices (H) presenting temporal inflow, drainage, and surface runoff rates of FM15 and FM16 in addition to the cumulative fluxes shown in the main text.

RC2: L317: 'as evidence of' should be replaced by 'suggests'. The similarity in infiltration behavior, and differences in drainage volumes help support this inference of bypass flow as well, since it suggests that more water is stored in the matrix in the FN12 versus FM13.

AC: We have revised the original sentence with more cautious wording ("is consistent with") to better reflect the inferential nature of the interpretation. We also clarified the argument by explicitly linking the earlier drainage response in FM13 to the combination of similar infiltration behaviour and differences in cumulative drainage volumes, thank you for this suggestion (see also response to Reviewer 1, Major Comment 1).

RC2: L320: Not sure your data shows any differences or 'acceleration' in infiltration rates… it does suggest bypass flow though.

AC: We agree that the term "acceleration" was imprecise and potentially misleading in this context. Our intention was to describe that the onset of drainage and the effective transmission of infiltrating water through the system occurred earlier and more rapidly in the macropore experiments compared to the non-macropore case. In the revised manuscript, we have avoided the term "acceleration" to avoid implying a quantified change in infiltration rate beyond what is supported by the data.

RC2: L335-336: 'which enabled water to bypass frozen regions and resulted in earlier and more pronounced drainage'…I'm wondering if you could also show the VWC profiles at the onset of drainage in FM15 and FM16. My reasoning is that if the probes show no changes prior to the onset of drainage, that will significantly strengthen your argument of bypass flow through macropores.

AC: We thank the reviewer for this helpful suggestion. In the revised manuscript, we have included the VWC profiles at the onset of drainage for FM15 and FM16 (see revised Fig. 8 and revised Results 3.3.) These profiles illustrate the spatial distribution of liquid water content immediately at the observed drainage response.

RC2: L337: Replace 'reflecting' with 'suggesting'. You don't have direct evidence of this.

AC: We agree that the term "reflecting" was too strong in this context. The wording has been revised ("consistent with") to better reflect the inferential nature of the interpretation (see also response to Reviewer 1, Major Comment 1).

RC2: Figures 5, 6, and 7 clearly show edge effects and that your set-up did not produce exclusively top-down thaw. That being said, I don't think this affects the interpretation of your results. I suggest the authors walk back their strong statements about the set-up's ability to reproduce one-dimensional vertical freeze-thaw. This is another reason where showing the vertical versus lateral thermal gradients would be helpful… although from these figures I doubt the vertical gradients are at least an order of magnitude greater than the lateral thermal gradients.

AC: The original wording may have overstated the degree to which the experimental setup reproduced strictly one-dimensional, top-down freeze-thaw conditions. In the revised manuscript, we have moderated this language to avoid implying exclusively vertical thermal propagation. In addition, we have quantified the relative magnitude of vertical and lateral temperature gradients during the freezing phase and included these metrics in the appendices as the reviewer suggested above. We therefore no longer describe the setup as producing purely one-dimensional freeze-thaw behaviour, but rather as predominantly top-down freezing under controlled boundary conditions.

RC2: L368-369: Soil shrinking or soil compaction/settling? I'm having a hard time seeing how the very coarse soils used in experiments would have that much shrinking? I would expect this to be more prevalent in fine grains soils.

AC: We agree that the original wording may have been misleading. Given the coarse grain size distribution and the absence of clay in our artificial soil mixture, classical shrinkage effects are unlikely. A more plausible explanation is mechanical settling or compaction of the soil matrix during wetting, particularly under low initial water content conditions. The effect appears to depend on initial water content, suggesting that increased pore-water lubrication may have facilitated structural rearrangement and densification of the granular matrix. This interpretation is consistent with the relatively high bulk density of the soil and its broad grain size distribution. However, we acknowledge that the precise mechanism cannot be resolved definitively within the present experimental framework. In the revised manuscript, we have therefore replaced the term "soil shrinking" with wetting-induced settlement.

RC2: L375: Are you sure this is due to advective heat transport? I don't think your data allows you to make such a strong statement. See previous comments about edge effects. Soften your language please.

AC: We agree that the original wording was too strong and could be interpreted as implying a causal dominance of advective heat transport. In the revised manuscript, we have substantially revised the corresponding passages in both the Results and Discussion sections to clarify the underlying energy balance. The observed downward migration of the freezing front is now described as a coupled thermo-hydraulic response of the system, primarily controlled by initial thermal conditions and conductive heat transfer. While advective heat redistribution associated with infiltration may contribute locally in the macropore experiments, its relative importance cannot be quantified or isolated based on the available data. We have therefore removed language suggesting that advection was the primary driver and reformulated the interpretation accordingly (see revised L. 370, L. 395, L. 399, and L. 439 (see response to Reviewer 1, Major Comment 6).

RC2: L402: Your figures are units of degrees Celsius, yet you discuss in Kelvins. While I know the changes are equivalent, you should be consistent, I suggest using degrees Celsius.

AC: We agree and changed all units of temperature in degrees Celsius in the revised manuscript.

RC2: L427: Replace 'demonstrate' with 'suggests'.

AC: The sentence has been revised accordingly to use more cautious wording ("suggesting"), in line with the broader revisions addressing inferential language throughout the manuscript (see also response to Reviewer 1, Major Comment 1).

RC2: L434-435: The main reason why these results may differ from Pittman et al. (2020) is that soils used in these experiments had significant amounts of smectite and thus swelled significantly at high saturation which likely sealed many of the macropores. Similar observations were seen at the field site where these cores were taken from in Mohammed et al. (2019), at the site named Triple G (figure 5) where recharge through was frozen ground was observed in in MW2, but at 80 cm the soil slowly became saturated while under the zero-degree curtain and no further infiltration and groundwater recharge was observed until the soil profile thawed.

AC: We thank the reviewer for this insightful comment and for highlighting the role of soil mineralogy, particularly smectite swelling, in influencing macropore functionality under frozen conditions. We agree that this mechanism provides an important context for interpreting differences between our results and those reported by Pittman et al. (2020). In the revised manuscript, we have incorporated this explanation into the Discussion (Sect. 4.2), explicitly noting that the clay-rich soils used by Pittman et al. likely experienced swelling and macropore sealing at high saturation, thereby limiting preferential flow. We also now refer to the field observations reported in Mohammed et al. (2019), where saturation under the zero-degree curtain inhibited further infiltration until thaw occurred.

RC2: L461-463: This is also a very novel, direct observation of pore-blockage due to sediment deposition and promoting freezing in macropores. Very nice!

AC: We appreciate the reviewer's positive assessment of this aspect of the study. The direct observation of pore blockage due to sediment deposition and associated freezing within macropores represents an important finding.