

Responses to editor and reviewer comments for

“Water storage and redistribution effect evaporation, retention, and infiltration of forest floor sites”

Responses in blue Added text in purple

Response to the Editor

The similarity index is rather high (14%). I checked the report and noticed that this high number mostly derives from your previous paper and your EGU-abstract. For now that is OK, but please consider to reduce the similarity during the possible revision phase.

Please also know that your current data availability statement is not in line with HESS its policy (https://www.hydrology-and-earth-system-sciences.net/policies/data_policy.html). HESS highly promotes open data sharing and only in exceptional cases embargo is possible. Please clearly state why the data is not open. Or better: make it open.

Dear Dr. Coenders-Gerrits,

Thank you for handling our manuscript and for the opportunity to submit a revised version.

We will try to reduce similarity in a revision and we will make sure to publish the used data on the open data sharing repository FreiData. Below we provide point-by-point replies to all comments made by you the editor and the reviewers.

With kind regards,

Heinke Paulsen (on behalf of all co-authors)

Response to Reviewer #1

This manuscript reports on experiments to quantify storage, drainage, and evaporation from the forest floor, using innovative instrumentation to obtain rare observations of these processes. The experiment and analysis are simple and the interpretations are clear, providing a remarkably clean story about the controls on water storage and release from this important store. I have only a few suggestions to improve the manuscript.

Reply We thank the reviewer for their supportive feedback and the constructive suggestions to improve our manuscript. We are glad the reviewer found the interpretations clear and the data valuable.

In the revised manuscript, we will implement the suggested changes, specifically regarding the consistency of our storage metrics, the addition of detailed material descriptions in the Supplementary Material, and the refinement of our discussion on wetting mechanisms and spatial variability. Detailed responses to each individual comment are provided in the following sections

There is an inconsistency in how C_{max} is defined. In Fig 8 and many other places in the results, C_{max} is the same as or less than C_{min} . I expected C_{max} to be defined as in Fig 2 and L150, where it is defined as an absolute amount of water, not as excess above C_{min} as it seems to be in most of the Results. I hope you will choose to define and report it as in Fig 2 and L150.

Reply In the presented analysis C_{min} is always smaller than C_{max} . But yes, the difference is not as high as one would expect from classical soil physics. This is caused by the generally wet conditions in the observed period. We discussed this in Section 3.1 and 4.

The methods and discussion unfortunately gloss over how much mineral soil is in these lysimeters. Are there data available on organic volume or mass vs mineral volume or mass? Though it might be difficult to compare across humus forms, this kind of analysis would help in extrapolating results to other settings.

Reply To provide better transparency regarding the composition of our measurement system, we will add a detailed description figure in the Supplementary Material. This figure explicitly quantifies the volumes and masses of organic versus mineral material within each lysimeter, allowing for better extrapolation of our results to other forest settings.

In Fig 9A, why are there large hourly oscillations between high evaporation and zero? Can we be sure this is due to sunflecks and not to an error in the data logger or analysis? Even if there are no data errors, it may be worth explaining the limitations of Fig 9 when estimating evaporation at the ha or catchment scale; i.e., sunflecks affect these installations differently than they affect catchments.

Reply We have verified the data and can confirm that the hourly oscillations result from "sunflecks" and the high temporal dynamics of evaporation at the forest floor interface, rather than logger errors. In a revised manuscript, we will add a discussion on the limitations of these high-resolution measurements when scaling up to a hectare or catchment level, noting that such variability is typically smoothed out at larger scales.

Changes Our high-resolution data further reveal hourly oscillations between peak evaporation and near-zero values. These fluctuations are attributed to the influence of direct radiation through gaps in the canopy reaching the forest floor ("sunflecks"), which trigger rapid, transient increases in evaporation. It should be noted, however, that while these oscillations are prominent at the lysimeter scale, they may be smoothed out when scaling these results to a hectare or catchment level, as the local photo-environment varies significantly across the landscape.

I think Fig 10 is one of the most important figures in the manuscript. We almost never get to see spatially resolved estimates of infiltration, and I found the variation to be strikingly high. It's interesting to see that the forest floor appears to be a stronger source of variability than redistribution by the canopy. Is it possible to add some discussion on how these magnitudes compare to spatial redistributions caused by other sources of heterogeneity?

Reply We agree that the high spatial variability is a key finding. In the revised discussion, we will expand the analysis to explain how the structural heterogeneity of the forest floor (e.g., preferential flow paths) acts as a primary driver of this variability, often exerting a strong influence just like the redistribution of water caused by the vegetation.

Changes The high variation in spatially resolved infiltration suggests that the structural heterogeneity of the FF, specifically the presence of preferential flow paths, acts as a primary driver of this variability. This internal

redistribution within the FF appears to exert a strong influence on infiltration patterns just like the spatial redistribution caused by incoming canopy throughfall and stemflow processes. However, further investigation is needed to fully quantify these effects and compare them to the throughfall variability, particularly under dry and potentially hydrophobic conditions, which were not sufficiently represented in the period of study

I appreciate the discussion comparing intensity effects on storage capacity in the canopy vs forest floor, but I think there is a missing element, namely the contrasting mechanisms in how those two stores become wetted. Unlike in canopies, the role of drop velocity in wetting the forest floor is low, and wetness depends more on low-velocity flow into detrital pores. This difference (momentum vs van der Waals dominance) seems to me a more obvious explanation than the laboratory vs. field distinction emphasized in the manuscript. I do agree that the duration is

Reply This is an important theoretical distinction. In the revised manuscript, we will expand the discussion to contrast the wetting mechanisms of the canopy and the forest floor. We will clarify that while canopy interception is heavily influenced by drop velocity and momentum, the wetting of the forest floor is primarily governed by low-velocity flow and capillary forces within detrital pores.

Changes This suggests that the storage capacity is determined not only by the volume of available pores but also by the specific mechanisms that allow water to enter them. This is evident when contrasting the wetting mechanisms of the canopy and the forest floor: whereas canopy interception is heavily influenced by the momentum and velocity of raindrops (Nanko et al., 2022), the wetting of the forest floor is governed by low-velocity flow and capillary forces within detrital pores. Consequently, storage in the FF is less a result of drop impact and more a function of the hydraulic conductivity and pore structure of the organic-mineral interface.

Figure 3 given that evaporation is always near zero and the details are also in Table 3, I suggest replacing that uninformative line with a time-varying estimate of C, as in Figure 2.

Reply That is a very nice idea, we will try to incorporate this.

Table 5 last rows should be mm per ... h?

Reply Yes, thank you. We will correct this.

Line 293: what does “saturated” mean here? This word must be chosen carefully.

Reply The term "saturated" has been replaced with "moist" as we recognize that the litter layer often starts draining before reaching full saturation.

Response to Reviewer #2 Seyed Mohammad Moein Sadeghi

General comments:

This manuscript presents an interesting and valuable field-based investigation of forest floor (FF) water storage, retention, evaporation, and redistribution, using a novel grid-lysimeter approach across three sites in the Black Forest. The dataset is rich (1,570 events), and the in situ measurement of FF processes is a clear strength, addressing a well-known gap in the literature that has long been dominated by laboratory-based studies.

Importantly, the study also brings much-needed attention to forest floor hydrology—an aspect of the water balance that is often overlooked or, in some cases, not explicitly considered in ecohydrological analyses. From my perspective as someone who works extensively in forest systems, I strongly disagree with this tendency. The forest floor is not a passive layer; it is a dynamic interface that regulates water storage, delays runoff, and mediates energy exchange. In that sense, I particularly appreciate the authors' effort to quantify these processes under field conditions and to elevate the role of the forest floor within the broader hydrological framework.

The manuscript would benefit from clearer terminology regarding the measured compartment. At present, “forest floor,” “litter layer,” and “organic layer” are used somewhat interchangeably in the Introduction and Methods, whereas the lysimeter setup appears to include not only the organic horizons but also a portion of the upper mineral soil. Because this distinction can influence the interpretation of storage, drainage, and evaporation dynamics, I encourage the authors to clarify exactly what component is being measured throughout the manuscript and to use the terminology consistently. In particular, please clarify whether the reported C_{min} and C_{max} values should be interpreted as properties of the litter layer, the full forest floor, or the combined organic–mineral surface layer represented in the lysimeter design.

The manuscript compares hydrological responses across the three study sites (e.g., Figure 7); however, these sites differ not only in altitude but also in precipitation regimes, temperature, and environmental conditions, which are key drivers of the observed fluxes. While this comparison is informative, it would be helpful to clarify whether the results are intended to be descriptive (process-based) or statistical comparisons of site characteristics. This distinction is particularly important because differences among sites may reflect not only intrinsic forest floor properties, but also variations in forcing conditions (e.g., rainfall amount, intensity, and duration). This point becomes especially relevant in Figure 7, where statistical differences in C_{min} and C_{max} among sites are presented using t-tests. Given the differing precipitation regimes, it would be useful to clarify how these comparisons should be interpreted, and whether the observed differences are driven primarily by site properties or by differences in precipitation forcing. A brief clarification or discussion of this aspect (e.g., whether comparisons are normalized, event-based, or influenced by precipitation variability) would strengthen the interpretation of the results.

Overall, this is a strong and well-executed study with a rich dataset and clear relevance to forest ecohydrology. The manuscript would benefit from some targeted clarifications and modest improvements in conceptual framing and methodological explanation, but these are minor and can be addressed without substantial reanalysis. I therefore recommend major revision.

Reply Dear Seyed Mohammad Moein Sadeghi,

We would like to thank you for this insightful review. It is rare to receive feedback that is both so detailed in its technical critique and so supportive of the overarching ecological importance of the research. We especially appreciate the reviewer's shared passion for forest floor hydrology and the guidance on how to better represent this "dynamic interface" in our writing.

Following the reviewer's guidance, we have revised the manuscript. Key improvements include a clarified definition of the measured compartment to avoid confusion between the litter and mineral layers, a more careful treatment of site-to-site comparisons, and a refined title that highlights the importance of mixed temperate stands.

We provide a detailed point-by-point response to all comments below and thank the reviewer once again for the invaluable contribution to this work!

Specific comments (line-by-line / section-based)

**** Title:**

1. *The title reflects the general scope of the study; however, I found it somewhat difficult to follow on first reading. The phrasing is not fully clear (e.g., “effect” vs. “affect,” and “forest floor sites”), and the sentence structure makes it harder to quickly grasp the main message. Beyond identifying issues, I believe it is helpful to suggest possible improvements. I therefore encourage the authors to revise the title for clarity and readability, using a more direct structure that clearly highlights the key processes and the role of the forest floor. For example, a formulation such as: “Forest floor water storage and redistribution affect evaporation, retention, and infiltration in mixed temperate forests” may improve clarity and better reflect the ecological context of the study. That said, this is only a suggestion—the authors may choose an alternative phrasing. From a forest management perspective, explicitly indicating mixed stands (or mixed temperate forests) could be valuable, as hydrological processes in mixed systems are often more complex and less studied than in pure stands.*

Reply We really like your remark since and consider to change the tile accordingly.

**** Abstract:**

2. *First, I suggest introducing the ecological context more explicitly at the beginning of the abstract. In particular, mentioning that the study was conducted in mixed *Picea abies* and *Fagus sylvatica* stands would provide important context and improve clarity from the outset. Second, the description of the study sites as “three sites with different altitudes located in the Black Forest” is somewhat misleading. Based on Table 1, the sites differ not only in altitude, but also in temperature and precipitation regimes, which are critical drivers of hydrological processes. I recommend revising this statement to better reflect the full range of environmental gradients represented in the study. Third, there are a few minor language issues (e.g., “effect” should be “affect”), and some sentences could be streamlined for better readability. Finally, I do not fully agree with the last sentence of the abstract. It is currently too general and does not reflect a specific contribution of this study. In its present form, it reads more like a broad statement that could be made independently of the results. I recommend revising this sentence to more clearly emphasize the novel findings or specific insights derived from your dataset and experimental approach.*

Reply We will include the fact that the sites were mixed beech-dominated forest stands. We will clarify that the different altitudes of the sites result in changed climatic conditions. We will clarify the specific knowledge gains developed from our study in the last sentence.

Changes We present the results of our network at three mixed temperate forest sites with different altitudes, and therefore diverging climatic conditions, located in the Black Forest, southwest Germany.

Changes The findings of the lysimeter network provide a comprehensive understanding of how not only the thickness of the FF but rather characteristics like the share of organic fine material define the water cycle within forest ecosystems.

**** Intro:**

3. *Several aspects would benefit from clarification and refinement to strengthen the conceptual framing and improve the overall impact. The manuscript would benefit from clearer and more consistent use of terminology at the beginning of the Introduction. Terms such as “forest floor,” “litter layer,” and “organic layer” are used closely and at times interchangeably, although they can represent different components of the soil profile. I recommend that the authors explicitly define these terms early in the Introduction, and clearly state which component is the focus of the study. This is particularly important because the hydrological behavior of the litter*

layer can differ from that of the entire forest floor, and later in the Methods, the measurement system appears to include not only organic material but also part of the upper mineral soil.

Reply In a revised text we will revise the first paragraph to make the different components clearer. We will define the FF as the combination of organic litter layer and top mineral soil. This is also what the lysimeters are filled with.

Changes Understanding the partitioning and movement of water within temperate forest ecosystems is essential for predicting hydrological responses to changing environmental conditions. At the interface between the atmosphere and the lithosphere, the forest floor (FF), comprising both the organic litter layer and the underlying, organic rich top mineral soil, serves as a critical mediator of water fluxes. This system regulates the movement of water through the interaction between its organic and mineral components, directly influencing runoff generation, soil moisture recharge, and evaporation. The organic layers interaction with the mineral soil dictates the fate of precipitation before it infiltrates deeper into the profile, thereby affecting water availability and overall ecosystem functioning (Ilek et al., 2021).

4. L 30: While the Introduction summarizes relevant literature effectively, the research gap is not yet clearly articulated. The manuscript notes that field-based data are limited, but it would strengthen the paper to explicitly state: what specific limitations exist in previous studies (e.g., lack of in situ, event-scale, or spatially distributed measurements), and how the present study addresses these limitations.

Reply We will add a paragraph to the Intro to clearly articulate the research gaps and lead in to our hypotheses.

Changes Despite the theoretical understanding of FF dynamics, much of the existing data has been derived from controlled laboratory experiments or grab-sample studies, which often fail to capture the complex environmental variability of the field. There remains a critical shortage of high-resolution, in-situ data that accounts for the interaction between litter properties and thickness, varying pre-wetness conditions, and the lateral movement of water. Specifically, the relationship between organic layer depth and its actual impact on throughfall retention and mineral soil evaporation in a natural setting remains poorly quantified. To address these gaps and move beyond laboratory approximations, we established an extensive FF lysimeter network, installed at three different sites throughout the Black Forest, SW Germany, to test the following hypotheses:

5. L 43-44: The statement that “broadleaf litter has a larger interception capacity than needle litter due to its higher surface area-to-weight ratio” is supported by laboratory-based studies at the leaf scale. However, I encourage the authors to be more cautious in presenting this as a general conclusion. At the forest floor scale, interception and storage are not only controlled by leaf morphology, but also by litter accumulation, decomposition rates, and structural properties of the litter layer. For example, needle litter often decomposes more slowly and can form thicker and more porous layers, which may result in greater overall water storage at the ecosystem scale, even if the per-mass storage capacity is lower. I suggest clarifying the scale of interpretation (e.g., per unit mass vs. per unit area or whole-layer storage) or slightly rephrasing this statement to avoid overgeneralization.

Reply We will rephrase the statement and add information on the different scales and underlying processes.

Changes At single leaf scale it has been observed that broadleaf litter has a larger interception capacity than needle litter, attributed to its higher surface area-to-weight ratio (Walsh and Voigt, 1977; Zhao et al., 2022). At the FF

scale this interception can be influenced by additional factors like litter accumulation and decomposition rates, e.g. needle litter is characterized by the formation of more porous material due to slower decomposition rates (Sato et al., 2004).

6. L 59-62: *The hypotheses are clearly stated and relevant. However: Hypotheses (1) and (2) are somewhat overlapping (both related to forest floor thickness and storage), and could be streamlined. Hypothesis (4) ("barrier for evaporation") would benefit from a slightly stronger mechanistic framing (e.g., reference to capillary barrier effects or vapor transport limitations).*

Reply We will revise the hypotheses as you suggest.

Changes ...to test the following hypotheses: (1) thicker FF results in higher total water storage capacities and higher initial throughfall retention, depending on initial wetness conditions, (2) infiltration is highly heterogeneous on a small spatial scale, influenced not only by the spatially variable incoming canopy throughfall but enhanced by lateral redistribution of water in the FF, and (3) thicker FFs function as a barrier for evaporation from the mineral soil by increasing the diffusion path for water vapor and reducing capillary connectivity between the soil and the atmosphere.

7. L 64: *I recommend removing it to improve the structure of the manuscript.*

Reply We will restructure the paragraph with the hypothesis.

**** Methods:**

8. Table 1: *Table 1 provides useful background information; however, it would benefit from additional site descriptors to improve clarity and reproducibility. I recommend including the geographic coordinates (latitude and longitude) for each site, ideally representing the center point of the study area. In addition, including topographic information, such as the mean slope (or slope range), would be valuable, as slope is an important control on hydrological processes (e.g., runoff, infiltration, and lateral redistribution).*

Reply We will add information on site characteristics to the table.

9. *The manuscript refers to the study sites as mixed stands, but the stand composition is not sufficiently described. Please clarify the proportion of each species (e.g., percentage of *Fagus sylvatica* vs. *Picea abies*) at each site. This could be reported based on tree frequency, basal area, or another standard forest metric. This information is important because different mixtures (e.g., beech-dominated vs. spruce-dominated stands) can result in distinct canopy structures and, consequently, different hydrological responses.*

Reply Our research sites are located in beech-dominated forests, we will try to make this clearer in the text. Specific forest metrics were not applied.

Changes This study utilizes a network of novel forest floor lysimeters (Paulsen and Weiler, 2025) and soil moisture probes deployed across three mixed beech-dominated (*Fagus sylvatica*) forest sites with patchy groups of spruce (*Picea abies*) trees.

10. *In addition to stand composition, I recommend including information on canopy structure, such as canopy cover and/or leaf area index (LAI), if available. These variables are particularly important because they strongly*

influence understory rainfall partitioning (e.g., throughfall and stemflow) and are also indirectly related to forest floor development and litter accumulation, which are key controls on interception and storage processes.

Reply We will check within the Research unit if LAI is available.

11. L 70: The manuscript refers to humus forms such as F-Mull and Moder, which are well-established classifications in European forest ecology. However, these terms may not be familiar to all international readers, particularly those working in North America or other regions where different classification systems are more commonly used. I recommend briefly clarifying the key characteristics of these humus forms in the main text (e.g., differences in litter decomposition, thickness, and organic matter structure), so that readers can better interpret their hydrological implications. In addition, the authors may consider providing a more detailed description or classification framework in the Supplementary Material, for readers who are less familiar with this system.

Reply We will add a definition of the classified humus forms in the text/site description.

Changes In this classification the Mull is characterized by rapid decomposition rates of litter, resulting in thin organic and thick A-horizons, an Oh-layer (humified material) is never present. In contrast, Moder is characterized by moderately-to-poorly decomposable litter. Therefore, the A-horizon is thinner and the organic horizon is composed of fresh and fragmented litter. Oi and Of horizons are present all year while sometimes a thin or patchy Oh layer can occur (AG Boden, 2024).

12. L 81: The manuscript states that the lysimeters contain the forest floor organic layer together with the upper mineral soil (top ~15 cm). While this is clearly described, it introduces some ambiguity in the interpretation of results. Please clarify whether the system is intended to represent: the litter layer, the entire forest floor, or a combined organic–mineral surface layer.

Reply We will clarify in the results that our system represents the forest floor, so that it includes the interplay of organic and mineral layers.

13. L 75-90: Please clarify how was disturbance minimized during installation?

Reply We will add that the filling was not possible using a FF monolith, but we reconstructed the different layers. For a detailed description we refer to our technical note, since the editor already remarked too much similarity.

14. Table 2 is useful but could be moved to the Supplementary Material, as it serves mainly as supporting information. This would improve the flow of the manuscript.

Reply We will move it to the supplementary material.

15. L 106: There is a typo error: May is correct.

Reply We will correct it.

16. L 115: The assumption that negative storage change ($\Delta S < 0$) in the absence of drainage represents evaporation is reasonable, but I recommend briefly acknowledging potential sources of uncertainty, such as: load cell measurement noise, OR temperature effects on sensor readings OR possible small, unrecorded drainage fluxes.

Reply This point was discussed in detail in our previous technical note describing the development of the Forest Floor Grid Lysimeter. To avoid redundancy in accordance with the editor's request, we have referred the reader to that publication rather than expanding the discussion here.

17. L 139: The criteria used to define precipitation events (threshold and 6-hour separation) are clearly described; however, it would be helpful to clarify whether these thresholds were selected based on prior studies or tested within this dataset. In addition, under closed canopy conditions, the forest floor can remain wet for several days after rainfall, suggesting that a 6-hour dry period may not fully represent hydrologically independent events, particularly with respect to storage and pre-event moisture.

Reply The inter event time was based on other studies using this specific time, we will include the references in the text. It allows to include dripping from the canopy to be included to the event and not be separated into a new one. But of course, it is too short for the forest floor to be completely dried out.

18. Figure 2: I suggest adding a short sentence in the text explicitly linking this conceptual framework to how these phases are quantified in the dataset (e.g., using storage change, drainage, and timing thresholds).

Reply We will add this to the first sentence.

Changes Fehler! Verweisquelle konnte nicht gefunden werden. shows a schematic representation of an exemplary precipitation event. The event can be divided into several typical phases, calculated using the storage change (ΔS) and drainage (D) at 10-minute resolution.

19. The manuscript includes several statistical comparisons (e.g., t-tests); however, a dedicated Statistical Analysis subsection is currently missing. I recommend adding a short subsection at the end of the Methods to clearly describe: the statistical tests used, assumptions (e.g., normality, independence), significance levels, and the software or packages applied.

Reply We will add a small paragraph before describing the different processes.

**** Results:**

20. I recommend briefly explaining why this event was selected (e.g., representative, average, or specific hydrological characteristics). Without this clarification, it is unclear how representative this example is relative to the full dataset (1,570 events).

Reply This event was selected since it was captured in parallel from all twelve lysimeters. We will add this information to the text.

21. Table 3 (and so on): If abbreviations are used in figures or tables (e.g., PTF, CE, CM), please ensure that they are clearly defined in each figure or table caption (or as a footnote). Figures and tables should be able to stand alone, and readers should not need to refer back to the main text to interpret abbreviations.

Reply We will carefully check the abbreviations in figures and tables.

22. In Figure 3, the blue bars are labeled as "Precipitation," whereas in the text and Table 3 the corresponding variable is referred to as canopy throughfall (PTF). This creates some ambiguity, as precipitation and throughfall represent different hydrological inputs. Given that the measurements are conducted below the canopy, it appears that the figure is showing throughfall rather than gross precipitation.

Reply You are right, the blue bars should be referred to as canopy throughfall rather than precipitation, we will correct this and adapt the text to be more consistent.

23. *Figure 5 (and whole parts of manuscript): In Section 3.2 (and Figure 5), the term “event size” is used to describe rainfall characteristics. I recommend revising this terminology to “rainfall amount” or “event magnitude”, which are more standard and precise in hydrological studies. The term “size” may be ambiguous and could be misinterpreted (e.g., as referring to drop size or spatial extent), whereas “rainfall amount” more clearly reflects the total precipitation depth (mm) associated with each event.*

Reply This is a good remark, we will change it to throughfall amount.

24. *L 203 (and table 1): The manuscript reports retention values in g/g, but in the text it is also written as “g/g mm,” which is unclear and potentially inconsistent.*

Reply Thank you for finding this. It’s an artefact from a previous version of the text. We will make sure to check, that the retention is given as mass fraction.

25. *Table 4: Please use one consistent term (“water retention” or “retention”) throughout.*

Reply We decided to use water retention in a more general context, when we don’t explicitly refer to our data. When referring to our data we will use the term throughfall retention, since our input to the forest floor is throughfall influenced by the canopy. We will delete the term rainfall retention in our manuscript

26. *L 206: Please indicate which figure or table supports the statement regarding increased retention with higher precipitation totals.*

Reply We don’t show this in a table or figure yet. But we will add a table to the supplementary material, with more detailed numbers.

27. *L 210: The statement “Given the absence of significant differences in retention amounts and durations among the three sites” requires clarification. The three study sites differ in key climatic drivers, particularly precipitation amount, intensity, and event characteristics (e.g., duration and intra-event variability). These factors directly influence retention dynamics. Therefore, direct statistical comparison of retention metrics across sites may not be fully appropriate without accounting for these differences in forcing conditions. For example, even if the number of events is similar across sites, the events themselves may differ substantially in total rainfall, intensity (including short-term intensities), and duration, which can strongly affect retention behavior. I recommend clarifying how these differences were addressed in the analysis (e.g., normalization, event-based comparison, or inclusion of precipitation characteristics as covariates), or revising the statement to reflect that comparisons are descriptive rather than strictly comparable across sites.*

Reply The comparisons for the throughfall retention were made rather descriptive because of the points mentioned by you. We will rephrase this section to make it clearer.

28. *In Figure 6, the x-axis labels (“yes” and “no”) are not sufficiently self-explanatory. Please clearly define the meaning of these categories in the figure caption (e.g., whether they refer to the occurrence or absence of initial rainfall retention). Given that many readers may come from different disciplines or may be students, the figure should be fully interpretable without requiring prior familiarity with the terminology or referring back to the main text.*

Reply We will adapt the figure caption.

Changes Figure 1: Tendency for throughfall retention to occur in relation to soil moisture at 5 and 15 cm depth. No = There is no retention occurring in the first interval (10 minutes) of an event, yes = retention occurs during the first interval of the event. We found a significant difference in soil moisture leading to initial throughfall retention or not in Waldkirch and at Kandel. Retention usually occurs with lower soil moisture contents.

29. Fig 6: The x-axis label in Figure 6 is currently phrased as a question (“Does retention of initial rainfall occur?”), which is unconventional for figure labeling. I recommend revising it to a more concise and descriptive format, such as: “Initial rainfall retention” with categories “Yes” and “No”, or “Retention occurrence” (Yes/No). In addition, please ensure that the x-axis category labels are consistently formatted, with the first letter capitalized (i.e., “Yes” and “No” instead of “yes” and “no”).

Reply See above.

30. L 222-223: In the text, $C_{\max}$ is reported as: “1.5 g/g at Waldkirch to 4.4 mm at Kandel” This mixes g/g and mm, which is inconsistent. Please ensure that units are consistent when reporting ranges (either all in g/g or all in mm).

Reply As in point 24, this was an artefact from the previous version, we will correct it to mass fraction.

31. L 229: Figure 7 caption: Figure 7 refers to: “different precipitation durations” However, panel (a) shows site comparison, not duration classes.

Reply This is a mistake, we will correct the caption.

32. Table 5: Please ensure that intensity units are consistently expressed (mm/h) across figures and tables, as “mm” alone is ambiguous.

Reply We will adapt the Table caption. This was maybe irritating because here we show the total amount of stored water in the lysimeters in mm in the first two columns of the table, additionally to the previously used mass fraction.

Changes Table 5: Mean values and standard deviation (brackets) for minimum storage capacity ($C_{\min}$) and maximum storage capacity ($C_{\max}$) in total amounts (mm) and as mass fraction (g/g) at the three different sites and for distinct precipitation durations and intensities.

33. L 255: Figure10 OR 9?

Reply Thank you! It has to be Figure 9, we will correct it.

34. Section 3.5 presents evaporation patterns based on specific short time periods (Spring 2025 and Summer 2024). It is not clear why these limited periods were selected instead of using the full observation period, particularly since other sections of the manuscript appear to rely on more comprehensive datasets. This raises two points that require clarification: the rationale for selecting these specific time windows, and whether the results shown are representative of the overall dataset In addition, the temporal scope of the analysis appears inconsistent across sections, which makes it difficult to compare results directly.

Reply These two periods were selected for exemplary purposes only. The table shows the data for the full observation periods. In the figure we wanted to show the diurnal pattern, and hint for the differences in the seasons e.g. caused by the canopy. We will clarify this in the according text passages.

35. Table 6: Column headings: “Days without rainfall” “First day following rainfall” Consider standardizing wording slightly (e.g., “Rain-free days” vs “Post-rainfall day”).

Reply We like the remark and will change it.

36. L 264: “proportion of precipitation between the left and right sides” However, based on the study design, measurements are taken below the canopy, meaning this likely represents throughfall, not gross precipitation.

Reply Yes, we will adapt it to throughfall.

37. L 273-290: The manuscript refers to: “left and right sides” but these are not explicitly defined. Please clearly define what is meant by “left” and “right” (e.g., orientation, relative to tree stem, or lysimeter setup), either in the text or figure caption.

Reply We will add the information to the text and figure caption.

38. L 289-290: the caption does not fully explain: how significance is determined what statistical test is used.

Reply We will add the information.

**** Discussion:**

39. Section 4.1: The discussion provides a reasonable interpretation of the observed relationship between antecedent moisture conditions and retention capacity; however, several aspects would benefit from clearer mechanistic justification and stronger linkage to the results. First, the statement that a saturated forest floor lacks the capacity to retain additional water is broadly consistent with hydrological theory, but the manuscript does not explicitly distinguish between capillary storage, gravitational storage, and dynamic storage thresholds (e.g., C_{min} vs. C_{max}). Clarifying which storage domain is being referred to would improve precision and avoid oversimplification. Second, the attribution of outliers in the “no retention” class to hydrophobicity is presented as an assumption (“we assume...”), but no direct measurements or supporting indicators (e.g., water drop penetration time, soil organic composition, or temperature history) are provided. Given that hydrophobicity is a process-dependent and temporally variable phenomenon, this explanation should either be: supported with additional evidence, or more clearly framed as a hypothesis rather than a conclusion. Third, the discussion would benefit from a clearer connection to the results presented in Section 3.3 and Figure 6. For example, the observed relationship between lower pre-event soil moisture and higher likelihood of retention is consistent with the argument presented here, but this link is not explicitly stated. Finally, the statement regarding 2024 being an unusually wet year is important, but its implications are not fully developed. In particular, the limited occurrence of dry conditions may: constrain the range of observed FF states, and reduce the ability to robustly assess processes such as hydrophobicity or extreme dry-state behavior. A brief acknowledgment of this limitation would strengthen the interpretation (I would like to see section 4.5 as a new section for research limitation, and future directions. So you can move this limitation to a new subsection).

Reply 1. We will change the wording saturated to moist. We recognize that it’s misleading, since litter usually starts draining already before being fully saturated.

Reply 2. We will rephrase the sentence so that it is more hypothesizing.

Changes We hypothesize that severe dryness induces hydrophobicity within the litter layer, which in turn reduces water retention and promotes rapid percolation. But this hypothesis needs further testing.

Reply 3. We will revise the section. To include the mentioned implications.

Changes The unusually wet year 2024 precluded the development of these hydrophobic conditions necessary for observation. As a result, the specific processes associated with extreme dry-state behaviour remained dormant and could not be analyzed during the study period.

40. Sectikon 4.2: L 300: Please define what is meant by “heavier” forest floor (e.g., mass per unit area, bulk density, organic matter content, or thickness).

Reply This was regarding Mull vs. Moder FF. We will adapt the text.

41. L 301: The conclusion: “storage is influenced by the proportion of organic fine material (OFM) rather than just FF thickness” is interesting but not clearly supported by the results shown. Please clarify: how OFM was quantified, and whether a direct analysis (e.g., correlation or regression) supports this statement Otherwise, this should be framed more cautiously.

Reply We conclude this, since we presented our storage data (C_{max} , C_{min}) as mass fraction. The difference between Mull and Moder FF-form lies in the amount of fragmented material (which is OFM in huge parts). Therefore, we can interpret that the higher storage capacity has to be controlled by higher share of OFM. But we will revise the text to be more cautious in the wording for the discussion.

41. L 304: The discussion notes: storage was often not empty at the onset of events. This is a key methodological point. However, the manuscript does not clearly show: how initial storage conditions were quantified, or how strongly they influenced C_{min} and C_{max} .

Reply We mentioned this in the methods. C_{min} and C_{max} are the complete storage, not only the net gain per event. So basically, the storage was never empty. In a revised text we will sweep the word “often”

42. L 320-321: The discussion compares with: Sato et al. (2004) Keim et al. (2006) However, the explanation: “differences are due to lab vs field conditions” is somewhat broad. I suggest refining this by explicitly distinguishing: rainfall intensity distribution event duration structure water input magnitude rather than grouping all differences under “lab vs field”.

Reply We will adapt the sentence to better explain the differences.

Changes These discrepancies may be attributed to differences between field observations and laboratory simulations. In field studies, water storage is influenced by naturally varying pre-moisture conditions and precipitation characteristics, including intensity, duration, and magnitude. In contrast, these parameters are strictly controlled in laboratory settings, allowing for precise manipulation.

43. L 320: The sentence: “The most parsimonious explanation...” is quite strong. Consider softening to: “A likely explanation...” or “One possible explanation...” unless this is quantitatively demonstrated.

Reply We will change it.

44. Section 4.3: The text refers to Penman–Monteith variables (radiation, humidity, wind), but: no full Penman–Monteith implementation or comparison is presented.

Reply A full comparison with Penman-Monteith was not intended for our study. This would be a study itself and might be done in the future.

45. L 336-337: The discussion states that: temperature is not the primary controlling factor, while radiation, humidity, and wind dominate. However, this conclusion is not directly supported by a formal sensitivity analysis or multivariate decomposition in the manuscript.

Reply No, you are right this is more a hypothesis. We will use imply instead of indicate.

46. L 339: The statement: higher evaporation in spring due to higher radiation under less dense canopy is plausible, but: canopy density is not quantified seasonally in this section spring vs summer also differs in VPD, soil moisture, and phenology.

Reply See above (Point 44 and 45)

47. Section 4.4: The text refers to “Figure 11” as supporting evidence for infiltration-related findings. However, there is no Figure 11 present in the manuscript.

Reply Thank you, it needs to be Figure 10. We will correct it.

48. The Discussion section would benefit from a clearer separation of content into distinct sub-sections. I recommend adding: Section 4.5 – Limitations and future directions: to explicitly address methodological constraints (e.g., temporal coverage, climatic variability, and limited observation of extreme dry conditions) and to outline clear avenues for future research. Section 4.6 – Management implications for forest ecosystems and water resources to translate the findings into applied insights for forest management, particularly regarding forest floor dynamics, water storage regulation, and hydrological functioning under different stand conditions.

Reply We like the idea of adding a separate paragraph for limitations. We don’t think management applications would fit into the manuscript.

**** Conclusion:**

50. The Conclusion currently repeats several detailed findings already discussed in the Results and Discussion (e.g., differences in storage capacity, precipitation controls, and evaporation patterns). I recommend restructuring it as a synthesis-oriented section, focusing only on the main take-home messages rather than restating results. Example improvement: Instead of: “We found significant differences in storage capacities among the three sites...” A more appropriate conclusion-style synthesis would be: “Storage capacity varied among sites, primarily driven by forest floor composition and moisture conditions.” Some expressions in the Conclusion are not sufficiently precise, such as “nearly all lysimeters” or “significant differences.” In a Conclusion section, such vague wording reduces scientific clarity. Please replace such statements with explicit quantitative summaries where possible or remove them if they are not essential.

Reply In a revised version of our manuscript we will reframe the conclusion to be more synthesizing.