This is a preliminary point-by-point response prepared for the discussion stage to summarize our current position and planned revisions. A complete and finalized response, including the corresponding manuscript changes and finalized additional analyses, will be provided together with the revised manuscript.

General response

We thank the referee for the careful review and for identifying several aspects of the manuscript that require clearer framing and stronger supporting analyses. We agree with a number of the comments concerning terminology, representation of previous literature, the distinction between model-derived and observed quantities, and the need for a more self-contained description and evaluation of the reconstruction. These points will be addressed directly in the revised manuscript.

At the same time, several comments appear to rely on an assumption that a dynamic water-deficit trajectory and the prognostic active-storage state analyzed here are essentially equivalent, and that the manuscript must therefore demonstrate superiority over a deficit calculation to establish scientific novelty. We respectfully disagree with this interpretation. We will add a direct deficit benchmark to determine empirically where the two representations agree, where they differ, and whether the manuscript's scientific findings depend on the chosen representation.

Major comments

1. Novelty, necessity and relation to the deficit literature

The distinction between aS_rz and existing deficit approaches is overstated. Deficit methods generate a time-varying state; a capacity estimate is a subsequent summary of its extremes. Recharge, drawdown, persistence, and multi-year carryover therefore do not fall outside that framework. Storage and deficit have already been described explicitly as complementary (Gao et al., 2024. https://doi.org/10.5194/hess-28-4477-2024). Non-resetting deficits have also been used as dynamic states to explore bedrock water use and post-drought runoff (Ehlert et al., 2024. https://doi.org/10.1029/2023WR036719; Lapides et al., 2022. https://doi.org/10.1029/2022GL100505). MOREDO adds learned partitioning, multiple conceptual reservoirs, and basin-to-grid parameter transfer. Those are model developments: subtracting the record minimum from a modeled storage trajectory is not a new observation of active storage. The manuscript should state the incremental methodological contribution and test whether it adds information beyond the established deficit water-balance calculation (Gao et al., 2014. https://doi.org/10.1002/2014GL061668; Wang-Erlandsson et al., 2016. https://doi.org/10.5194/hess-20-1459-2016; Stocker et al., 2023. https://doi.org/10.1038/s41561-023-01125-2). Let me put it another way. Comparable minimum-referenced storage trajectories could be extracted from many hydrological or land-surface models containing vegetation-accessible reservoirs. The essential question is therefore what independent evidence demonstrates that the MOREDO state is better identified, or contains more hydrological information, than existing modeled storage states or a plain deficit calculation. An important virtue of the stand-alone deficit approach is its transparent bookkeeping based on externally estimated precipitation and ET fluxes, which limits dependence on the internal architecture and latent states of a particular hydrological model. These flux products have their own uncertainties, but those uncertainties are more readily exposed and propagated than the structural assumptions governing an unobserved internal reservoir.

Response: We agree that our original description of deficit approaches was too restrictive. We will correct this point.

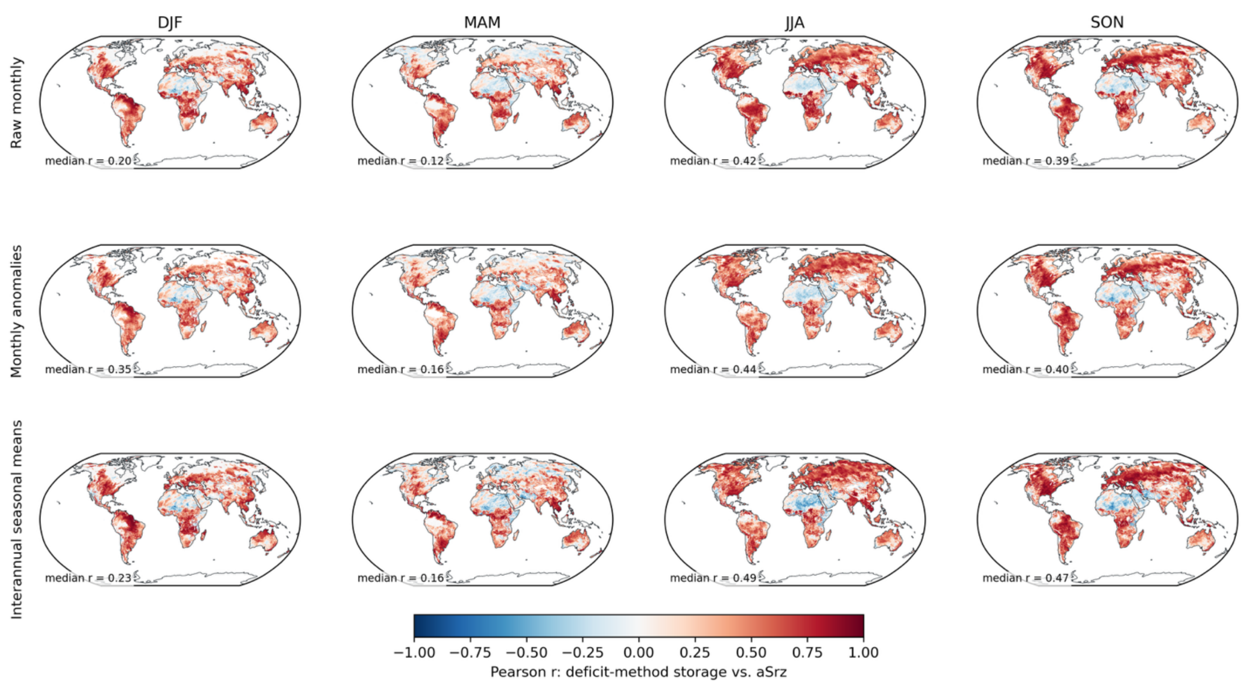
We respectfully disagree, however, with treating aSrz and a conventional deficit-derived state as effectively interchangeable solely because they share a first-order water-balance basis. The distinction we intend to make is not static versus dynamic, but how the ecosystem-accessed storage trajectory is constructed. In a deficit formulation, the storage state follows directly from the cumulative imbalance between prescribed liquid-water input and ET. In our framework, the water input to the ecosystem-accessed reservoir is not prescribed a priori. Instead, incoming liquid water is dynamically partitioned between the ecosystem-accessed reservoir and runoff or drainage pathways, and the resulting reservoir evolves prognostically as it is replenished by the inferred accessible input and depleted through ET. The observational constraints help restrict this partitioning and the resulting storage trajectories. Thus, the two approaches share a water-balance basis, but the ecosystem-accessed storage trajectory is not constructed from the same flux balance in the two formulations. We will use the deficit calculation as a benchmark to quantify where the two representations agree and where they differ.

Moreover, while deficit-based approaches can retain dynamic information, the major global applications most relevant here have primarily used deficit trajectories to infer rooting-zone storage capacity from extremes or specified return periods (Wang-Erlandsson et al., 2016; Stocker et al., 2023). Stocker et al. (2023), for example, explicitly describes its global S0 estimate as a snapshot and does not consider temporal changes in S0. Gao et al. (2024) similarly presents the cumulative deficit trajectory as the basis from which SRmax can be inferred, while also emphasizing that storage and deficit are complementary representations. The dynamic deficit studies cited by the referee address important complementary questions, including multi-year bedrock-water use and post-drought runoff at regional scales. Ehlert et al. (2024), for example, explicitly develops a non-resetting deficit formulation to account for multi-year deficit accrual. We will therefore revise the novelty statement so that it does not imply that dynamic plant-accessible storage is itself a new concept. The contribution of the present study lies instead in the global analysis of the reconstructed active-storage state, including its spatial organization, recent changes, turnover, and its relationship with bulk terrestrial water storage across timescales.

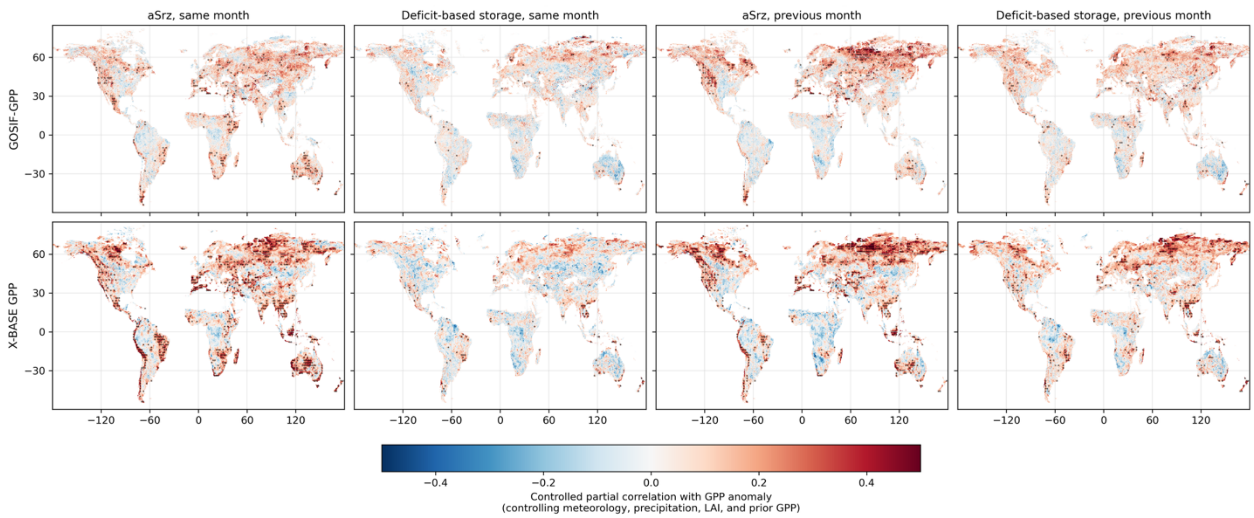
In the revision, we will

1. rewrite the relevant Introduction paragraphs to represent dynamic and non-resetting deficit approaches explicitly;
2. distinguish root-zone storage capacity from the time-varying state analyzed here without implying that deficit methods lack dynamics;
3. broaden the literature;
4. add a direct deficit-based benchmark calculated using consistent water-input and ET data, clearly distinguishing the two methods.
5. evaluate the two storage representations against GPP products to assess whether their differences are also reflected in ecosystem functioning.

The preliminary benchmark does not support treating the two representations as interchangeable for the analyses considered here. At the monthly timescale, their temporal correspondence is strongly regime dependent, with particularly weak agreement in cold and dry regions and a global median pixel-wise correlation of r=0.39. Their 2001–2020 climatological means also show markedly different spatial organization, with a spatial correlation of r=−0.31 across common valid pixels. We will examine where these similarities and differences emerge across hydroclimatic regimes.



We are also adding a functional comparison using GPP products that were not used as model training targets. After controlling for meteorological conditions, precipitation, LAI, and previous-month GPP, preliminary analyses show more spatially coherent relationships between aSrz and GPP anomalies than between deficit-derived storage and GPP anomalies. This provides an additional line of evidence that differences between the two storage representations are also reflected in their relationships with ecosystem functioning.



2. dependence on an unreviewed model

Although Appendix A summarizes the conceptual structure and Section 2.1 describes the principal training data and design, the detailed model architecture and the decisive sensitivity, identifiability, and product-evaluation evidence are deferred to a non-peer-reviewed companion manuscript (Blougouras et al., 2026. https://doi.org/10.22541/essoar.15005507/v1). Neither S_a nor aS_rz is a training target. ET constrains the accessible bucket only because the architecture assigns ET to that bucket; runoff constrains the complementary pathway; and GRACE constrains the sum of modeled stores, not their decomposition. Reproducing these targets does not establish a unique internal partition.

The present manuscript must contain enough information to evaluate the reconstruction independently. It should compare aS_{rz} with a transparent deficit calculation using identical precipitation, ET, and snow inputs/targets.

Response: As described in our response to Comment 1, we will add a direct benchmark against a deficit-derived state. Here, however, the referee raises a separate question: whether the model-derived state is sufficiently constrained and evaluated to support the scientific interpretations made in this manuscript. We agree that this requires a more self-contained presentation.

We also agree that neither the ecosystem-accessed storage state nor aSrz is directly observed or used as a training target. Reproducing ET, runoff, SWE, and TWSA does not establish a mathematically unique decomposition of subsurface storage, and we will avoid language implying such uniqueness. Uniqueness, however, is not required for a latent model state to be scientifically informative. The relevant question is whether the inferred state is sufficiently constrained by the available observations, robust to reasonable perturbations of the reconstruction, and consistent with information not used in its estimation.

In our framework, ET, runoff, SWE, and TWSA jointly restrict the set of water-balance trajectories compatible with the model structure. The interpretation of aSrz therefore does not rely on claiming that it represents the unique physical decomposition of subsurface water. Rather, it relies on establishing that the reconstructed state is sufficiently constrained and ecohydrologically meaningful for the scientific questions addressed here.

The revised manuscript will make this basis much more self-contained by providing the governing water balance, state definitions, parameterization, cross-validation design, ensemble uncertainty, and key product-level evaluation needed to interpret aSrz. We will also add targeted tests that directly address the referee's concern. As described in our response to Comment 1, a deficit-based benchmark will test whether aSrz simply reproduces a transparent cumulative water-balance calculation. We will additionally evaluate its relationship with GPP products that were not used as model training targets, and include withholding experiments to assess the sensitivity of the reconstructed state and subsequent analyses to individual observational constraints.

The scope of these additions is to establish the interpretive basis needed for the scientific analyses presented here. The primary purpose of the manuscript remains the analysis of the reconstructed active-storage state, including its global spatial organization, recent changes, turnover, and relationship with bulk terrestrial water storage across timescales.

3. catchment runoff vs grid grid cells

This is more of a comment about Moredo than this manuscript, but it matters for this manuscript. Pixel-scale deficit formulations generally do not include observed runoff directly: excess water input after a deficit has been replenished is treated implicitly as drainage or runoff, while discharge observations are generally available only at the integrated catchment scale. In MOREDO, each gauged basin is treated during training as one lumped unit: forcings and attributes are aggregated over the basin, and simulated runoff is compared with outlet discharge. during global inference, each 0.25-degree cell is run independently, without routing, upstream accumulation, lateral exchange, or a cell-scale runoff observation. Outlet discharge therefore does not directly constrain the partitioning or water balance of individual grid cells. It constrains an attribute-to-parameter relationship at catchment scales, whose validity at grid-cell support is assumed.

This matters directly for aS_rz because the learned partition between accessible water and runoff controls its magnitude. Effective catchment parameters may absorb subcatchment heterogeneity and partitioning of precip into runoff vs storage avaailble to plants.Because both runoff generation and storage dynamics are nonlinear, parameters inferred by applying the model to basin-mean forcing need not reproduce the area-average behavior obtained by applying it to heterogeneous grid-cell forcings

Response: We agree with the referee that outlet discharge is not a direct observation of runoff generation or storage partitioning at the 0.25° grid-cell scale. We also agree that this distinction matters for aSrz, because the inferred accessible-storage trajectory depends on the same water partitioning that governs the runoff-producing pathway. We will revise the manuscript wherever our wording may have implied a direct grid-cell runoff constraint.

The role of runoff in the reconstruction, however, is not to assign observed outlet discharge to individual grid cells. During training, each gauged catchment is treated as a water-balance control volume, and outlet discharge provides an integrated constraint on the runoff-producing pathway and, together with ET, SWE, and TWSA, restricts the range of water-partitioning behavior compatible with the observed catchment water balance. In the global application, the same prognostic water-balance equations are solved for each 0.25° model control volume using local forcing and attributes. The grid-cell reconstruction therefore represents a transfer of the learned water-balance behavior to a different spatial support, rather than a direct disaggregation of catchment runoff observations.

We agree with the referee that this change in spatial support is not mathematically neutral. Because input partitioning, runoff generation, and storage dynamics are nonlinear, and because subcatchment heterogeneity and lateral exchanges are not explicitly resolved in the grid-cell application, the response obtained from basin-mean forcing need not be identical to the area-average response obtained from heterogeneous grid-cell forcing. This is a genuine structural uncertainty in the global transfer and will be stated explicitly in the revised manuscript.

To address this point directly, we will evaluate scale consistency on held-out catchments by applying the trained model at the 0.25° grid scale and aggregating the resulting fluxes and storage dynamics back to the catchment support. We will compare these distributed-to-catchment results with the corresponding lumped-catchment simulations and with available catchment-scale observations. This provides a direct test of whether the water-balance behavior learned at catchment scale is preserved when the model is applied at grid-cell support. Because the issue here is water-balance partitioning instead of river-routing dynamics, this comparison will focus on temporally aggregated runoff volumes, ET, and storage behavior rather than on the timing of the outlet hydrograph.

We will therefore clarify in the revision that runoff provides an integrated catchment-scale constraint, that it does not constitute a grid-cell runoff observation, and that application to 0.25° model units involves an explicit scale-transfer assumption whose robustness can be evaluated empirically.

Line-level comments lines 10-13: the stated observational gap is too broad. existing observations include profiles of soil moisture and water potential, lysimeters, boreholes, geophysics, rock-moisture measurements, groundwater wells, sap flow, eddy covariance, isotopes, and catchment water balances. the defensible gap is the absence of a direct, continuous, globally gridded observation of the entire plant-accessible domain.

Response: Agreed. The original wording was too broad. The intended gap is not the absence of observations of subsurface or plant-water processes, but the absence of a direct, continuous, globally gridded observation covering the full domain of water accessed by ecosystems. We will revise the statement accordingly.

line 30: "unsaturated stores," "weathered substrate," and "shallow groundwater connections" are overlapping categories of hydrologic state, material, and pathway. In other words, you're contrasting apples and oranges

Response: Agreed. We will separate these concepts, for example by referring to water stored in soils and weathered bedrock, potential connections to shallow groundwater, and recharge or lateral-flow pathways separately.

lines 42-64: the discussion of observations is restricted almost entirely to satellite products and land-surface models. field and critical-zone observations need to be represented, including work that distinguishes plant-accessible capacity, available water, and observed drawdown (Klos et al., 2018. https://doi.org/10.1002/wat2.1277).

Response: Agreed. The Introduction currently emphasizes globally available satellite products because the manuscript addresses global reconstruction. We will broaden this discussion and cite field-based work on plant-accessible storage, observed drawdown, weathered-bedrock water, and groundwater access.

lines 65-70: deficit approaches do retain recharge, drawdown, persistence, and non-resetting multi-year behavior. revise this paragraph around the actual difference in model structure.

Response: Agreed. We will remove the statement that recharge, drawdown, persistence, and multi-year dynamics fall outside the deficit framework. The revised paragraph will distinguish the approaches based on how the storage state is constructed and how incoming water is partitioned.

lines 70-78: the companion study introduced this notation and model formulation not the general concept of dynamic ecosystem-accessible storage. apparent or lower-bound accessible storage was already articulated directly (e.g. Hahm et al., 2024. https://doi.org/10.1029/2023WR035362).

Response: Agreed. We do not intend to claim that the general concept of dynamic ecosystem-accessible subsurface storage was introduced by the companion study. We will explicitly acknowledge earlier concepts of apparent, lower-bound, and dynamic plant-accessible storage.

figure 1: this relates to the claims of novelty. compare the time series with e.g. the deficit time series and bucket schematic in lapides et al. (Lapides et al., 2022. https://doi.org/10.1029/2022GL100505). the two represent the same first-order cumulative balance (moredo just constrains input differently).

Response: We agree that Figure 1 should not be interpreted as claiming novelty for the general bucket or cumulative-water-balance concept. The purpose of the figure is to define and explain the storage state used in this manuscript for the readers who are not familiar with the concept. We will revise the caption and accompanying text to acknowledge the shared first-order water-balance basis with deficit approaches. However, as shown by the direct benchmark described above, sharing this first-order balance does not lead to equivalent storage trajectories or spatial patterns.

lines 90-100: the first study question substantially overlaps the global analysis of climate, vegetation, topography, and rooting-zone storage in other papers, e.g. stocker et al. (Stocker et al., 2023. https://doi.org/10.1038/s41561-023-01125-2). clarify which questions are genuinely new.

Response: We agree that climate, vegetation, and topographic controls on rooting-zone storage capacity have already been studied globally, particularly by Wang-Erlandsson et al. (2016) and Stocker et al. (2023). We will not present the existence of these controls as new. The distinction is that our first analysis concerns the realized active-storage state reconstructed over 2001–2020 rather than a rooting-zone storage capacity inferred from deficit extremes. The subsequent questions concerning recent changes, storage turnover, and timescale-dependent coupling with TWSA are not addressed by the global capacity studies cited above.

In the revision, we will sharpen Question 1 so that it does not imply novelty in identifying climate, vegetation, or topographic controls per se, but instead asks how these established controls are expressed in the reconstructed active-storage state.

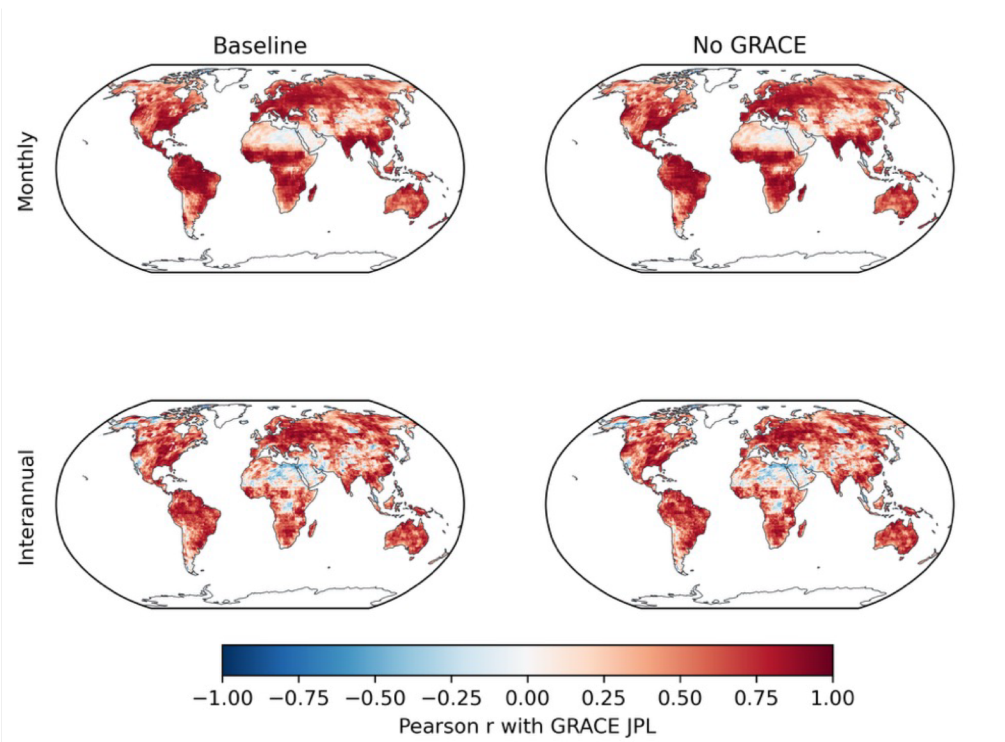
lines 115-118: convergence of biome-mean values with increasing window length does not test sensitivity to the timing of an extreme multi-year drought.

Response: Agreed. Convergence with increasing window length demonstrates sensitivity to record length, but it does not by itself isolate sensitivity to the timing of a particular extreme multi-year drought. We will revise this interpretation and no longer present the window-length experiment as a test of sensitivity to extreme-drought timing. The limitation associated with the particular hydroclimatic sequence sampled during 2001–2020 will be stated explicitly.

Because GRACE TWSA enters the training loss, the subsequent aS_rz–TWSA coupling analysis is partly circular(?) and cannot independently validate the inferred relationship. This limitation should be acknowledged or tested through a GRACE-withheld experiment

Response: We agree that because GRACE TWSA enters the baseline training loss, the original aSrz-TWSA analysis cannot be regarded as an independent validation. Our purpose was a coupling analysis, but we agree that a withholding experiment provides a stronger test of whether the reported coupling structure is imposed by the GRACE constraint.

We have performed the suggested GRACE-withholding experiment using a matched ensemble. Preliminary results show that the broad spatial structure of both monthly and interannual aSrz–GRACE correlations remains very similar when GRACE is excluded from training. We will quantify the differences between baseline and no-GRACE experiments and include the result in the revised manuscript. The text will also clearly distinguish coupling analysis from independent validation.



Please maintain a consistent distinction between an observation-constrained model output and an observed quantity; several statements in the abstract and conclusions blur this distinction

Response: Agreed. We will consistently use terms such as “observation-constrained reconstruction”, “model-derived storage state”, or “inferred active storage” and remove language that could imply that aSrz is directly observed.

Throughout this was super confusing: the prefixed a in aS_rz and subscript a in S_a are difficult to distinguish while denoting "active" and "accessible." use clearly distinct symbols

Response: We will retain aSrz for the reported diagnostic and rename the underlying prognostic accessible-storage state throughout the manuscript and Appendix.

References:

Ehlert, R. S., Hahm, W. J., Dralle, D. N., Rempe, D. M., & Allen, D. M. (2024). Bedrock controls on water and energy partitioning. *Water Resources Research*, *60*(8), e2023WR036719.

Gao, H., Hrachowitz, M., Wang-Erlandsson, L., Fenicia, F., Xi, Q., Xia, J., ... & Savenije, H. H. (2024). Root zone in the Earth system. *Hydrology and Earth System Sciences*, *28*(19), 4477-4499.

Lapides, D. A., Hahm, W. J., Rempe, D. M., Whiting, J., & Dralle, D. N. (2022). Causes of missing snowmelt following drought. *Geophysical Research Letters*, *49*(19), e2022GL100505.

Stocker, B. D., Tumber-Dávila, S. J., Konings, A. G., Anderson, M. C., Hain, C., & Jackson, R. B. (2023). Global patterns of water storage in the rooting zones of vegetation. *Nature geoscience*, *16*(3), 250-256.

Wang-Erlandsson, L., Bastiaanssen, W. G., Gao, H., Jägermeyr, J., Senay, G. B., Van Dijk, A. I., ... & Savenije, H. H. (2016). Global root zone storage capacity from satellite-based evaporation. *Hydrology and Earth System Sciences*, *20*(4), 1459-1481.

Bachofen, C., Tumber-Dávila, S. J., Mackay, D. S., McDowell, N. G., Carminati, A., Klein, T., ... & Grossiord, C. (2024). Tree water uptake patterns across the globe. *New Phytologist*, *242*(5), 1891-1910.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences*, *114*(40), 10572-10577.

McCormick, E. L., Dralle, D. N., Hahm, W. J., Tune, A. K., Schmidt, L. M., Chadwick, K. D., & Rempe, D. M. (2021). Widespread woody plant use of water stored in bedrock. *Nature*, *597*(7875), 225-229.