We thank the reviewer for their thorough and constructive assessment of our manuscript. Their valuable insights regarding the structure of the Introduction, the formatting and accessibility of our figures, the need for a model-comparison matrix, the role of field validation, the classification and calibration of model parameters, and the non-linear nature of parameter sensitivity have helped us significantly improve the quality and clarity of the manuscript. Below we address each comment in detail.

Comment 1

> The Introduction should be condensed and reorganized to clearly highlight the importance of GW-UZ-SW coupling, the limitations of existing 3-D Richards-equation-based models and bucket-type models, and the specific novelty of GWSWEX.When presenting the results, the authors should improve the readability of figures and tables.

Response: We appreciate the reviewer highlighting this. Making the novelty of GWSWEX immediately apparent to the reader is important, and the original manuscript could have done more to achieve this. While the Introduction was already designed to follow a progressive structure: motivating GW–UZ–SW coupling (paragraph 1), detailing the limitations of 3-D Richards and bucket-type models (paragraph 2), identifying the remaining literature gap (paragraph 3), and presenting GWSWEX's specific contributions (paragraph 4), we recognize that GWSWEX's position relative to existing models was not made sufficiently explicit within this flow. To address this, we have added a new concise paragraph to the Introduction that identifies the closest hybrid GW–UZ–SW couplers and directs the reader to a newly added capability-comparison table (Table 1). This provides a clearer basis for comparing GWSWEX with existing frameworks and makes the specific novelty more immediately apparent to the reader.

The readability of figures and tables is addressed below under Comment 2.

Comment 2

> Some figures contain multiple soil profiles, forcing scenarios, and solver comparisons, which makes them difficult to interpret. The layout, font size, legends, and color scheme should be improved to better highlight the key findings.When emphasizing the novelty of GWSWEX, the authors should provide a clearer comparison with existing models.

Response: We have comprehensively updated all figures to improve their clarity, visual consistency, and accessibility. The revised figures employ a color-vision-deficiency-friendly palette combined with distinct line styles, ensuring that individual series are fully distinguishable in grayscale. We have also optimized line transparency and draw-order to prevent overlapping curves from obscuring data, enlarged font sizes, and refined legend placement. The sensitivity heatmap has been updated with a perceptually uniform color map. All figures are now provided as high-resolution vector graphics formatted according to the journal's size recommendations.

To address the suggestion for a clearer comparison with existing models, we have introduced a comprehensive capability-comparison table (Table 1), which is detailed in the response to Comment 3.

Comment 3

> A comparative table is suggested to compare GWSWEX with HYDRUS-1D, MODFLOW-UZF, MIKE SHE, ParFlow, and conventional bucket models in terms of vertical resolution, groundwater-head tracking, capillary rise, mass conservation,

computational efficiency, and coupling capability.When validating the proposed model, the authors should consider using field observations.

Response: We agree that a structured comparison table is an excellent way to clarify GWSWEX's relative position. We have added Table 1 to the manuscript, which evaluates GWSWEX against HYDRUS-1D, MODFLOW–UZF1, MIKE SHE, ParFlow, and conceptual bucket models. The comparison covers subsurface dimensionality, representation of matric-potential gradients, prognostic groundwater-head dynamics, capillary rise, constitutive formulation, per-step mass conservation, suitability for ensemble applications, two-way GW–UZ–SW coupling, and open-source availability. The table is introduced in the revised Introduction, and the practical implications are discussed in the Discussion section.

Our response regarding validation against field observations is detailed under Comment 4.

Comment 4

> The current verification mainly relies on comparisons with HYDRUS-1D under idealized numerical scenarios. Field data, such as observed groundwater levels, soil moisture, runoff, or evapotranspiration, would help demonstrate the practical applicability of GWSWEX.

Response: We agree entirely that field validation is the essential next step and is the focus of our ongoing follow-up work. The present paper is, by design, a model description paper. Its primary objective is to evaluate the numerical behavior, accuracy, and mass-conservation properties of GWSWEX's dual solvers under perfectly controlled conditions where forcing, parameters, and boundary conditions are known without uncertainty. This isolation of numerical performance is necessary to separate the model's mathematical fidelity from the significant uncertainties introduced by field-scale parameter estimation and atmospheric data.

To clarify this scope, we have revised the manuscript to explicitly distinguish between the controlled numerical verification presented here and subsequent field validation. Additionally, we have expanded the Limitations section to state clearly that the current work does not constitute a full field validation and to identify field-based evaluation as the immediate next phase of model assessment.

Comment 5

> The manuscript should clarify which parameters are physically based, empirical, or numerical. Recommended parameter ranges and calibration guidance for different soil types and hydrological conditions would improve the practical value of the model.

Response: We thank the reviewer for this valuable suggestion, which significantly improves the practical utility of the model. We have added Table 4 to the manuscript, which systematically classifies every GWSWEX parameter into one of three categories: physically based, empirical or semi-physical, or numerical. For each parameter, the table provides typical physical ranges, along with practical guidance on calibration and interpretation across different soil and vegetation types. This table is preceded by explanatory text outlining how parameters should be constrained or calibrated.

Comment 6

> When discussing the nonlinear responses of soil hydraulic parameters, ET-stress thresholds, infiltration parameters, and inter-layer connectivity parameters, the authors should consider relevant recent studies on nonlinear correlations

and predictive modelling, including: https://doi.org/10.1029/2024JB029018 https://doi.org/10.1029/2023WR036494 https://doi.org/10.1029/2023WR036203

Response: The parameter responses in GWSWEX are indeed highly non-linear and mutually interacting, and we agree that emphasizing this is important. We have prepared the sensitivity-analysis discussion to make these non-linear parameter interactions explicit, referencing established global sensitivity analysis literature (Morris, 1991; Sobol', 2001; Saltelli et al., 2010; Pianosi et al., 2016) to support the discussion.

Regarding the three suggested references (10.1029/2024JB029018; 10.1029/2023WR036494; 10.1029/2023WR036203): we have reviewed these studies and found that they focus on non-Darcy and inertial flows in rough, partially filled rock fractures at the pore/fracture scale. GWSWEX, by contrast, operates macroscopically within the Darcy regime of unsaturated porous-media flow under sub-saturated forcing, which excludes the pore-scale fracture dynamics and inertial regimes discussed in those papers. Consequently, we have not included these specific citations, but we hope the reviewer finds the strengthened discussion of non-linearity and parameter interaction, grounded in directly relevant sensitivity-analysis literature, an appropriate response to the underlying concern.