This study presents a novel EnKF-like image assimilation system that integrates curvelet transform into the Common Land Model (CoLM) to improve the spatial structure of soil moisture analyses. By shifting the assimilation process from point-based magnitude corrections to multi-scale structural adjustments in the spectral domain, the proposed method addresses a critical limitation of conventional land data assimilation systems. The experimental design is sound, the use of GLDAS and in-situ observations for validation is appropriate, and the results clearly demonstrate the advantages of the approach in enhancing spatial correlation and reducing errors across different soil layers and vegetation types. The manuscript is well-written and the scientific contribution is significant. I recommend publication after major revisions.

Major Comments

1. Methodological Clarifications

Flowchart/Schematic (Section 3): Section 3 would benefit from the inclusion of a flowchart or schematic diagram illustrating the overall structure of the EnKF-like image assimilation system. Such a visual aid would help readers better understand the integration of curvelet transform, ensemble-based error estimation, and the Kalman filter update within the spectral space. It would also clarify the sequential flow of data from observation space to image space and back.

Curvelet mode interpretation (Lines 170–175): The description of reconstruction using different modes is informative but somewhat abstract. It would be helpful to briefly explain what is meant by "first mode," "first two modes," etc., in terms of spatial scale (e.g., largest scale features correspond to low-frequency modes). This would improve accessibility for readers unfamiliar with curvelet decomposition.

Error covariance transformation (Lines 210–215): The explanation of how background and observation error covariances are transformed from observation space to spectral space is critical but brief. A more explicit description—perhaps with a step-by-step outline or a mathematical summary—would help readers understand how ensemble perturbations are propagated through the curvelet transform.

Sensitivity to GLDAS biases (Section 3.4): The decision to omit bias correction to preserve spatial structure is reasonable given the study's focus. However, the potential sensitivity of results to systematic biases in GLDAS is not addressed. A brief discussion or a sensitivity test (even in a supplement) would strengthen the robustness of the conclusions.

2. Data and Validation

Interpolation effects (Lines 100–105): Bilinear interpolation of GLDAS data from 0.25° to 1.4° is mentioned, but the potential impact of this resampling on spatial structure—especially in heterogeneous regions—is not discussed. A short comment on how this might affect the comparison would be valuable.

Scale mismatch in in-situ validation (Lines 380–385): The independent validation using in-situ data is a strong point. However, station density in some regions (e.g., western China) is low, and the representativeness of point-scale measurements for grid-scale (1.4°) soil moisture is not addressed. A brief note on this scale mismatch and its implications for interpreting correlation and ubRMSE values would improve transparency.

3. Figures and Visual Presentation

Figure 2 caption (Line 225): As noted in the initial review, the caption should be corrected to read: "Figure 2: (a) Soil moisture background field and (b) difference between the background field and the reconstructed field after adding background error perturbations and applying first-mode curvelet inverse transformation."

4. Terminology and Consistency

uRMSE vs. ubRMSE (Line 370): The term "uRMSE" appears here, while "ubRMSE" is used elsewhere. Please ensure consistent terminology throughout (preferably "ubRMSE" as defined earlier).

5. Interpretation and Discussion

Vertical propagation mechanisms (Lines 325–335): The discussion of how assimilation increments propagate under different vegetation types is insightful. It could be strengthened by linking these patterns to known hydrological processes (e.g., root water uptake, preferential flow) and noting whether the model's parameterizations are capable of representing such processes.

Hybrid assimilation outlook (Lines 475–490): The conclusion states that image assimilation and point-based assimilation are complementary. This is an important point, but the manuscript does not explore how such a hybrid system might be implemented. A brief outlook on potential hybrid approaches (e.g., scale-dependent assimilation weights) would add value.

Acknowledgement of ChatGPT: The acknowledgment of ChatGPT for language polishing (Lines 490–495) is appreciated. However, it might be more appropriate to place this in the Acknowledgements section rather than the main text.

6. Minor Technical Issues

Reference formatting:

Line 625: The reference for Shen et al. (2023) is listed as "Numerical methods," which appears incomplete. Please update with the full journal or conference name.

Summary

The manuscript presents a well-conceived and rigorously evaluated assimilation system. The points raised above are minor and do not detract from the scientific contribution. Addressing them would, however, improve the clarity, reproducibility, and overall impact of the work. I recommend minor revisions based on the comments provided.