Referee #2

Overall evaluation:

This study develops an EnKF-like image assimilation system for the Common Land Model and validate the assimilation results. It is a well-motivated work with good novelty.

Response:

We sincerely thank you for your valuable time and positive assessment of our manuscript. We appreciate your careful reading and encouraging feedback on the motivation and novelty of this study.

Below, we provide point-by-point responses to your comments. All corresponding revisions have been marked in blue and red in the revised manuscript, with blue indicating added text and red indicating deleted text.

Minor Comments:

1. The content of section 3 includes the model description, but the title does not correspond. It is recommended to modify it.

Response:

Thanks for your suggestions. We revised the title of Section 3 to "Methodology and experimental design" to cover the model description, assimilation system construction, and experimental design more concisely.

2. Line 160, coarsest-scale spatial features, Line 165, large-scale, Line 170, continental-scale. In this section, the author provides various descriptions of the scale names. A brief explanation of the basis for this classification is needed.

Response:

Thank you for your comments. Following the commonly used scale classification in meteorology, we have revised these descriptions and now describe the spatial features extracted by the curvelet multiscale decomposition as planetary-scale, large-scale, and mesoscale features from coarse to fine levels. We also added a brief

explanation of the basis for this classification in the revised manuscript (**Lines 189–191**), as follows:

The curvelet-derived spatial features are grouped according to the meteorological scale convention, from planetary-scale to large-scale and mesoscale features. These categories correspond respectively to broad background patterns, regional structures, and finer variations associated with local heterogeneity.

3. Line 210. Is X^f stand for the variable or spectral coefficients after curvelet transformation? Please explicitly state it.

Response:

Thank you for your comments. We rewrote the relevant section using more explicit notation, with x representing variables in physical space and c representing curvelet coefficients in spectral space. We also provided additional details on the generation of perturbed ensemble members, the application of the curvelet transform, and the estimation of covariance matrices in spectral space. Furthermore, a new schematic flowchart has been added as Figure 2 to clearly illustrate the overall structure and workflow of the EnKF-like image assimilation system. The revised manuscript now includes the following text in **Lines 249–286:**

Leveraging the mathematical exactness of curvelet analysis, we can transform not only the structure of variables in physical space into spectral coefficients, but also the corresponding errors into coefficients in spectral space. In this EnKF-like assimilation framework, following the error estimation strategy of the EnKF, we first estimate the background error of soil moisture at each grid point in physical space, random perturbations with zero mean and a standard deviation corresponding to the physical-space error are added to the background soil moisture field, thereby generating an ensemble of perturbed members. Let the i-th ensemble member in physical space be denoted as $\boldsymbol{x}_i$ ($i = 1, 2, \ldots, N$), where N is the ensemble size. In this study, the ensemble size is set to $N = 50$. By applying the curvelet transform operator C, these members are mapped one by one into spectral space, yielding the corresponding curvelet coefficients with errors:

$$\boldsymbol{c}_i = C(\boldsymbol{x}_i) \tag{7}$$

The background error covariance matrix **B** in spectral space is dynamically estimated directly from the ensemble samples in spectral space:

$$\mathbf{B} = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{c_i} - \bar{\boldsymbol{c}})(\boldsymbol{c_i} - \bar{\boldsymbol{c}})^{\mathrm{T}} \tag{8}$$

where $\bar{\boldsymbol{c}}$ denotes the ensemble mean of the curvelet coefficients, and the superscript T indicates matrix transpose. The observation error covariance matrix **R** is transformed in a similar manner. The standard deviations of the random perturbations are prescribed as 0.15 m^3/m^3 for the background soil moisture field and 0.1 m^3/m^3 for the observation field. In fact, owing to the orthogonality of different basis functions, the background error covariance **B** and observation error covariance **R** are diagonal in the spectral space. Thus, the unrealistic divergence of error impacts can be avoided, making the error localization and inflation procedure unnecessary.

After **B** and **R** are constructed, the assimilation system performs the EnKF analysis step in spectral space to compute the updated curvelet coefficients $\boldsymbol{c}^a$. Finally, the inverse curvelet transform operator C^{-1} is applied to map the updated curvelet coefficients back to physical space, thereby obtaining the final analyzed field:

$$\boldsymbol{x}^a = C^{-1}(\boldsymbol{c}^a) \tag{9}$$

To provide a clearer illustration of the overall architecture of the curvelet-transform-based EnKF-like image assimilation system, we present a schematic flowchart in Figure 2. As shown in Figure 2, one assimilation cycle of the framework consists mainly of four key steps. First, in the physical space, the initial ensemble is generated by adding random perturbations to both the background field and the observation field. Second, the forward curvelet transform operator is applied to map the ensemble members from physical space to spectral space. Third, in spectral space, the background and observation error covariances are dynamically estimated from the ensemble samples, and the Kalman filter update is performed to obtain the analysis field of curvelet coefficients. Finally, the inverse curvelet transform is applied to convert the updated curvelet coefficients back to the physical space, thereby optimizing the spatial structure of the model state variables as a whole.

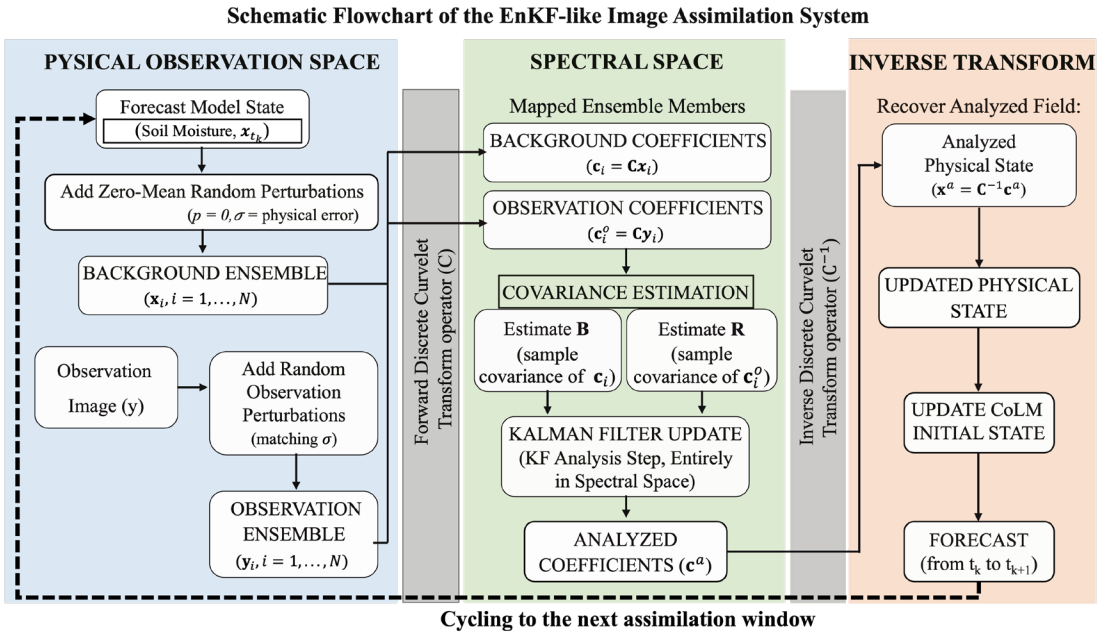

Figure 2: Schematic Flowchart of the EnKF-like Image Assimilation System.

4. The proposed image-based method is reasonable for land assimilation. However, the authors do not show the comparison with point-based method. It is recommended to conduct a simple comparative analysis, or to address the limitations in the conclusion.

Response:

Thank you for your comments. Following your suggestions, we have added a discussion in the revised manuscript to acknowledge that a direct comparison with point-based assimilation methods was not included in this study and to clarify this limitation. The corresponding revision has been made in **Lines 607–615**. The added text is as follows:

Although the above analysis demonstrates the potential advantages of image assimilation in preserving spatial structures, a direct comparison with point-based assimilation methods would further clarify its relative strengths and limitations. However, a mature point-based assimilation system specifically applicable to CoLM is not yet available, making it difficult to objectively compare the performance differences between image-based and point-based assimilation at this stage. From a methodological perspective, the two are highly complementary. Image assimilation offers clear advantages in capturing spatial patterns and maintaining structural continuity, whereas

point-based assimilation is more effective for assimilating high-accuracy in-situ observations and handling localized extreme anomalies. Therefore, future work will focus on developing hybrid assimilation strategies that apply scale-appropriate techniques at different spatial scales, fully leveraging the strengths of each method.