

# Response to Reviewer #3 Comments

Yunhao Fu, et al.

April 27, 2026

**General Comments.** The authors proposed an LST assimilation scheme that jointly update soil moisture and soil temperature in the upper soil layers of CoLM. Global experiments show notable improvements in several land surface variables, including soil temperature, snow temperature, snow depth, soil moisture, and surface fluxes. The main contribution is to demonstrate that LST assimilation can be more effective when its information is transferred into coupled water-energy state variables, particularly in freeze-thaw regions and humid areas where vertical water and heat exchanges help preserve and propagate the assimilation signal. The paper is generally well structured and clearly written. However, I have several concerns and suggestions that should be addressed before publication. I therefore recommend MAJOR revision before the manuscript can be considered for publication in GMD.

## **Response:**

We thank the reviewer for the insightful comments. Below is our point-by-point response to each concern.

## Comment 1

### 1. Novelty and positioning of the proposed scheme

The novelty of the manuscript should be clarified more explicitly. LETKF and LST assimilation have been widely explored in previous studies, so the main contribution appears to be the joint update of soil temperature and soil moisture. However, the mechanism by which this joint update improves snow-related variables in freeze-thaw regions is not sufficiently demonstrated. For example, an overestimated LST could be corrected either by reducing soil temperature or by increasing soil moisture, and these two pathways may have different or even competing effects on the surface water-energy balance. The authors should explain how the LETKF balances these increments and why this leads to better snow temperature, snow cover, and snow depth. Moreover, if the proposed mechanism is mainly related to freeze-thaw processes, it is unclear why the strongest soil moisture improvements occur mainly in humid and relatively warm regions rather than in the same freeze-thaw regions. I therefore suggest adding sensitivity experiments that update only soil temperature and only soil moisture, respectively, to demonstrate the added value of the proposed joint-update scheme.

### Response:

Assimilating LST to update only the surface soil temperature yields marginal improvements, as the surface state rapidly relaxes back to the trajectory dictated by atmospheric forcing. Motivated by this finding, we propose to jointly update soil temperature (ST) and moisture (SM) in the top two layers. The proposed scheme leverages the persistent memory of soil moisture and the thermal inertia of soil temperature to prolong the assimilation effect.

Data assimilation inherently distributes the observation innovation to state variables based on the background error covariance rather than physical constraints. This covariance reflects the physical sensitivities and the strength of energy-water coupling within the land surface model. The joint update precisely exploits the covariance between soil temperature and moisture.

This joint-update scheme is proposed to sustain the impact of assimilation, yet it also yields substantial improvements in snow-related variables. As shown in Fig. 1, the joint-update scheme (ST+SM) significantly outperforms the temperature-only update (ST) in capturing snow cover using ERA5-Land. Updating soil moisture modulates the heat capacity and thermal conductivity of the underlying soil. This physically consistent adjustment to the soil's thermal properties regulates the snow energy budget far more effectively than updating temperature alone.

Fig. 2 compares the RMSD of soil moisture and temperature from the joint-update scheme and the temperature-only scheme against in-situ observations from the AmeriFlux networks [4]. This demonstrates the benefits of the proposed scheme, as it consistently improves both water and energy states compared to the temperature-only update scheme.

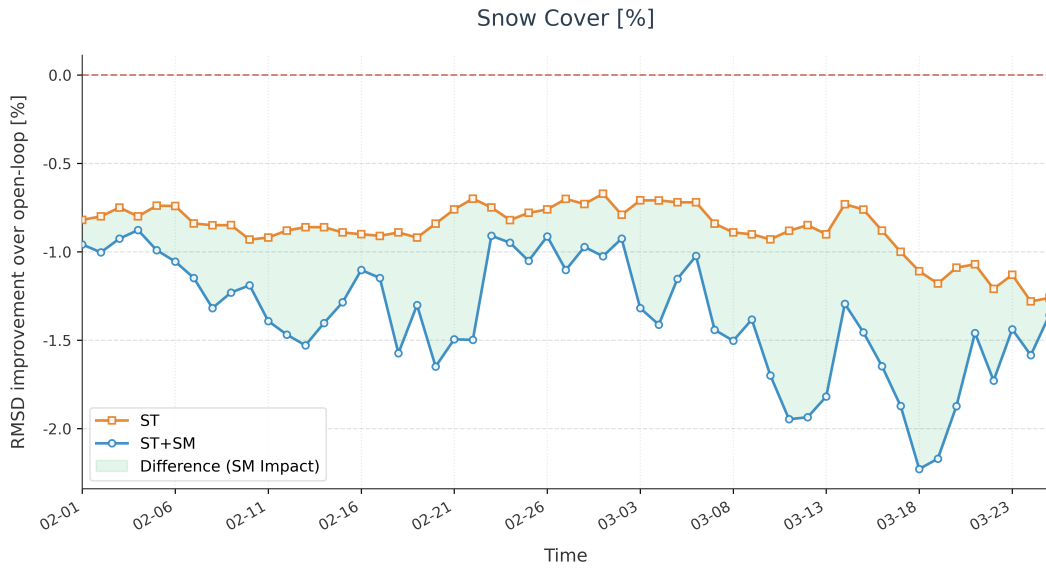


Figure 1: Snow cover  $RMSD_d$  [%] for the joint-update (ST+SM) and temperature-only (ST) assimilation schemes relative to the open-loop experiment.

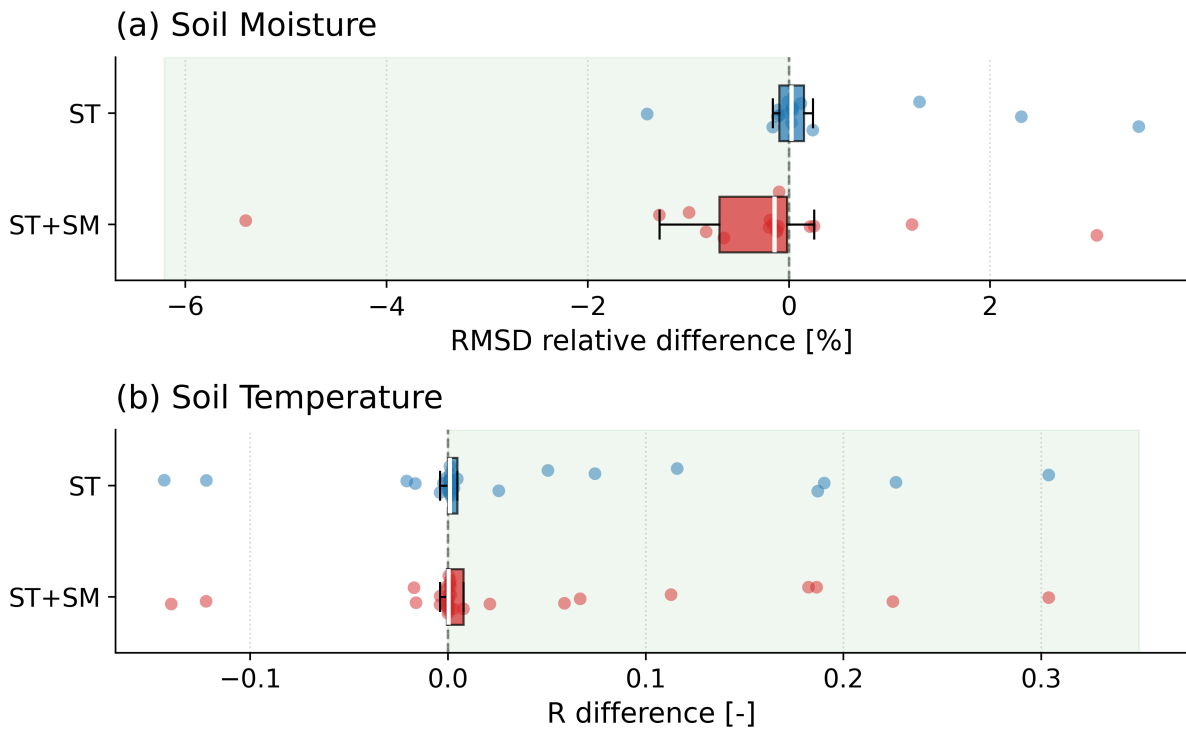


Figure 2: Comparison of assimilation performance between the joint-update (ST+SM) and temperature-only (ST) schemes across AmeriFlux sites. Panel (a) shows the RMSD relative difference (%) for soil moisture. Panel (b) shows the correlation coefficient (R) difference for soil temperature

We have clarified this novelty and the underlying mechanisms in Introduction section of the revised manuscript (L66-74).

## Comment 2

### 2. Some questions on your methods

Several methodological choices are not sufficiently clear, which makes the framework difficult to reproduce and the claimed innovation difficult to assess.

1) The authors generate ensemble spread mainly by perturbing model parameters, rather than atmospheric forcing. This choice needs justification, especially because parameter perturbations may affect water and energy conservation. The authors should clarify whether conservation is maintained or diagnosed. If not, the assimilation experiment should be compared with an ensemble-mean open-loop experiment, rather than a single deterministic open-loop run, to separate the impact of ensemble generation from the impact of LST assimilation.

### Response:

Fig. 3 presents two-month (from January 1 to March 1, 2001) single-member experiments with and without perturbing tunable parameters. The temporal evolution of the global mean energy and water balance error is nearly identical between the two experiments. This demonstrates that perturbing tunable parameters does not compromise the model's intrinsic physical balance or conservation properties, thereby justifying the experimental design adopted in this study. We've added this in L115-117.

## Comment 3

2) The updated state vector should be defined more explicitly. What specific soil layers are included in the "upper soil layers"? Are soil temperature, liquid water, and ice content updated simultaneously? Are snow-layer temperature and water/ice content also updated, or only soil variables?

### Response:

The updated state variables are explicitly defined in Section 2.1. The "upper soil layers" specifically refer to the top two valid layers ( $l_{\text{top}} \leq l \leq l_{\text{top}} + 1$ ). Temperature, liquid water, and solid water/ice content within these valid layers are updated simultaneously. Furthermore, when snow is present, the snow temperature and water/ice content are also updated in the same manner.

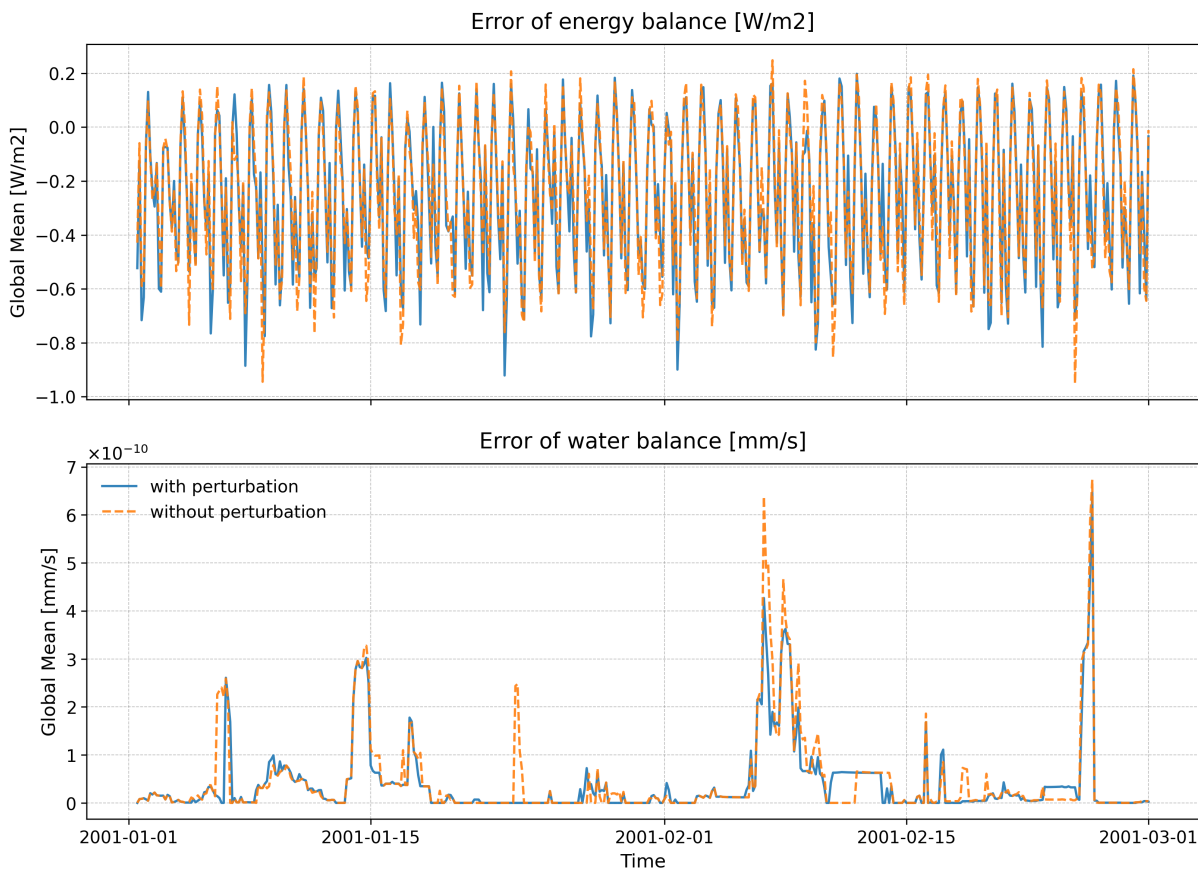


Figure 3: Temporal evolution of the global mean error of energy balance (top panel) and water balance (bottom panel) for simulations with (solid blue line) and without (dashed orange line) perturbations of tunable parameters from January 1 to March 1, 2001.

#### Comment 4

3) Line 210: Where do the observation error values come from? Do they include only MODIS retrieval error, or also representativeness error from aggregating observations to the  $0.5^\circ$  grid? The latitude-dependent weighting/threshold should also be justified. the specification of the observation error needs to be clarified.

#### Response:

As in other typical data assimilation studies, the observation errors are assumed to be independent, so the observation covariance matrix is diagonal. The observation errors are directly derived from the MODIS `lst_uncertainty` field and aggregated to the  $0.5^\circ$  model grid. We did not explicitly introduce the representativeness error, as preliminary tests indicated that the standard deviation of the raw  $0.01^\circ$  LST within a  $0.5^\circ$  grid cell is less than 5% of the aggregated retrieval uncertainty, making the representativeness error negligible.

We use latitude-dependent weighting to account for the varying physical area of the raw MODIS pixels at different latitudes during aggregation, which ensures accurate spatial representativeness. Finally,

the threshold of 1500 valid pixels (60% coverage) is maintained to guarantee statistical reliability. Modifications have been added in L228-233.

#### Comment 5

4) Several key thresholds and parameters require justification or sensitivity tests, including the localization radius and its latitude dependence, the 8 K innovation rejection threshold, and the 3-standard-deviation limit on increments. These choices can strongly affect the analysis increments and the number of assimilated observations.

#### Response:

The localization radius uses a cutoff distance of 730 km at the equator. This radius is smaller than typical atmospheric system scales, and while larger than the minimum scale of land surface heterogeneity, it properly accounts for the spatial representation of aggregated satellite observations. The latitude dependence is simply implemented to avoid unreasonable analysis increments at high latitudes. Regarding the 8 K innovation rejection threshold, our sensitivity tests indicated that tighter thresholds (5 K) overly reject valid observations driven by the LST diurnal cycle, while looser thresholds (9 K) fail to filter gross errors. The 8 K threshold empirically provides a good balance. Finally, the 3-standard-deviation limit is a standard measure to clip the outlier analysis. Modifications have been added in L198-201, L240-241 and L284.

#### Comment 6

5) The temporal treatment of MODIS LST observations is unclear. When assimilating every 3 hours, are all observations within the 3-hour window used at the analysis time, or only observations valid exactly at that time?

#### Response:

For a specific analysis time  $t$ , all observations whose actual scan times fall within the  $[t - 1.5, t + 1.5]$  hours window are aggregated and assimilated. We've added this in L221-224.

#### Comment 7

6) The 52-year spin-up and the "time-lag method" should be described more clearly, especially how the initial ensemble is generated.

**Response:**

To generate the initial ensemble, we first conducted a 52-year spin-up simulation to allow the model to reach a stable equilibrium state. Following this spin-up, the model was run for an additional full year with three-hourly outputs, from which 50 initial members were randomly sampled. This time-lag approach helps the selected initial members maintain the intrinsic energy and water balance of the model while providing a reasonable climatological spread of land surface states. Modifications have been added in L279-282.

**Comment 8**

7) The authors should explain why different variables are evaluated with different metrics and reference datasets, and whether this affects the consistency of the conclusions.

**Response:**

We have clarified the rationale for selecting specific evaluation metrics and reference datasets in the revised manuscript (L302-305 and L311-314). The metrics were chosen based on the typical error structures of the variables. Soil moisture evaluation uses unbiased RMSD to remove systematic biases caused by differing soil layer definitions across products, thereby robustly assessing dynamic variations. Conversely, the other variables (temperatures, snow depth, and fluxes) rely on standard RMSD and BIAS to quantify absolute accuracy. The selection of reference datasets depends on data availability and regional reliability. Specifically, GLDAS lacks snow surface temperature and snow cover fraction data, while MERRA2 lacks bare soil evaporation data. Furthermore, because ERA5-Land soil temperature in permafrost regions of Asia is known to exhibit large biases [1], we excluded it to ensure evaluation accuracy.

**Comment 9**

8) How to process imbalance of water and energy during DA?

**Response:**

Because data assimilation alone lacks explicit physical constraints, we implemented specific processing steps to mitigate potential water and energy imbalances. Specifically, snow mass is updated only when all ensemble members share the same number of snow layers. In addition, a non-negative limit is systematically applied to soil moisture. Furthermore, when the top two soil layers are updated, if the top layer contains solid ice while the second layer has none, a small amount of ice is redistributed from the first layer to the second to ensure stable model integration. We've added this in L288-292.

## AEXP - OEXP

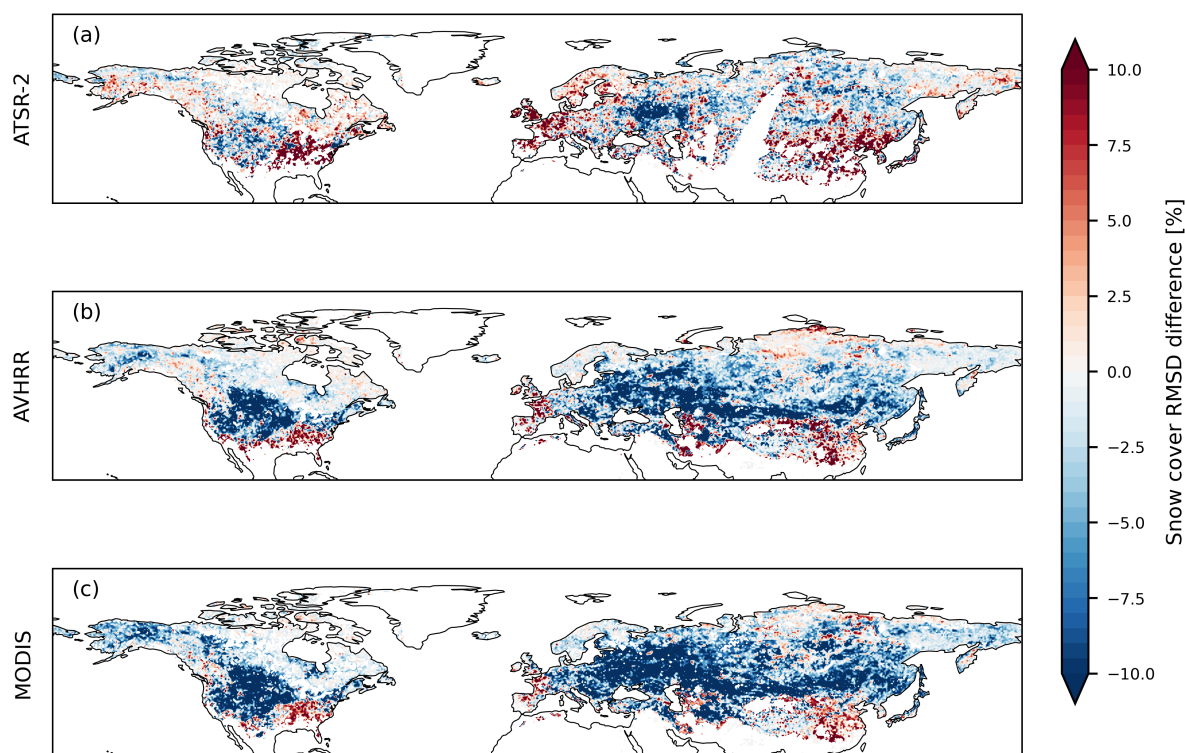


Figure 4: Snow cover [%] yearly (31 January 2001 to 31 December 2001)  $\text{RMSD}_d$  evaluated based on different satellite products: (a) ATSR-2, (b) AVHRR, and (c) MODIS.

### Comment 10

#### 3. Need for more independent validation

The evaluation relies mainly on ERA5-Land, GLDAS, and MERRA2. Although using multiple products is helpful, these datasets are still model-based and are strongly affected by their own forcing, parameterizations, and bias corrections. They do not provide fully independent observations of land surface states. Therefore, I suggest adding more independent validation, such as soil moisture/surface flux from FLUXNET or other in-situ networks, and/or satellite-based products (e.g., MODIS) for relevant land surface variables. This would make the reported improvements more convincing, especially for soil moisture, snow variables, and surface fluxes.

#### Response:

We have introduced independent satellite and in-situ observations into the revised manuscript to robustly evaluate the assimilation performance (L268-272). Using datasets retrieved from multiple satellites, including ATSR-2 [6], AVHRR [2], and MODIS [3], significant improvement on snow cover is seen in Fig. 4. With in-situ observations from the FLUXNET2015 [5] and AmeriFlux [4] networks, improvement on both sensible and latent heat fluxes is evident in Fig. 5. We've added the evaluations in L415-422 and L507-514.

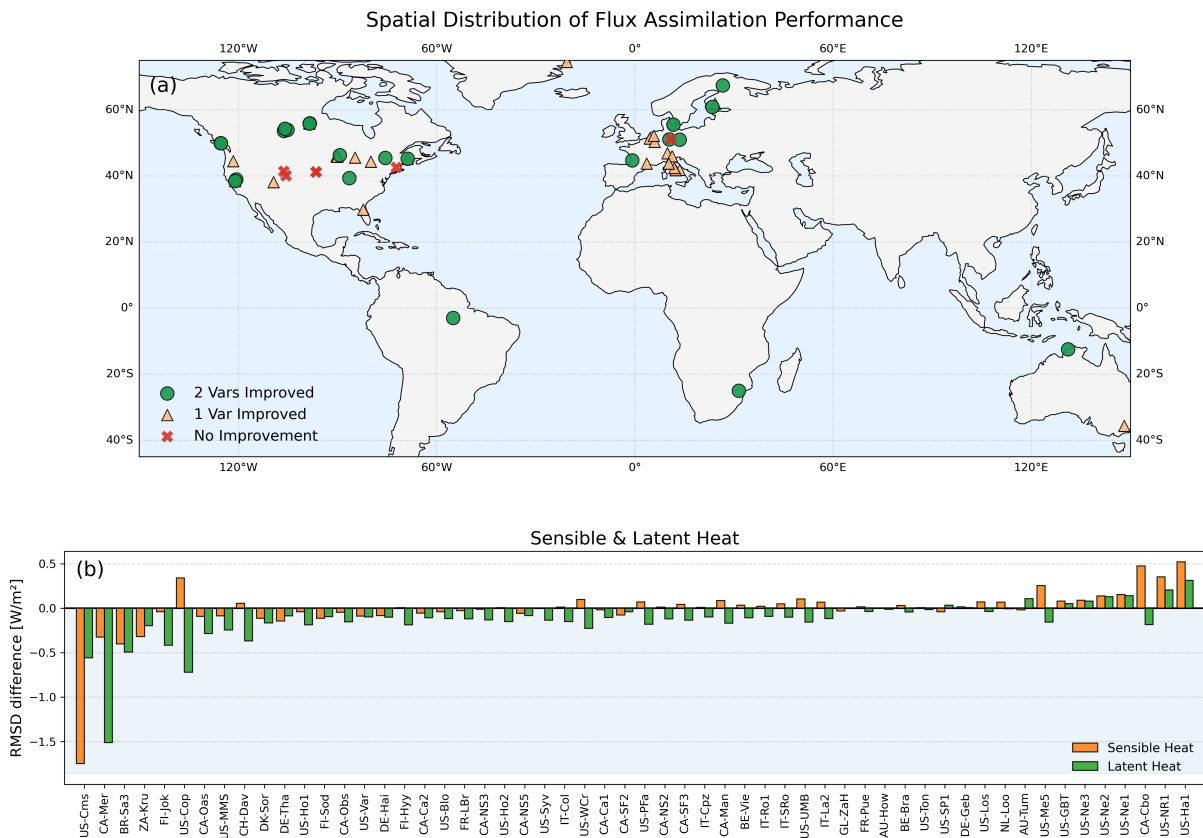


Figure 5: Spatial and quantitative evaluation of the impact of assimilation on surface heat fluxes across global sites. (a) Distribution of assimilation performance based on sensible heat and latent heat. Markers categorize sites by the number of improved flux variables: green circles indicate improvement in both heat fluxes, orange triangles represent improvement in a single variable, and red crosses denote no improvement. (b) Site-specific  $RMSD_d$  for sensible heat (orange bars) and latent heat (green bars) across the year (31 January 2001 to 31 December 2001).

### Comment 11

#### Specific comments

Line 75: I do not agree with this statement. LETKF is also not the core contribution of this study, so I suggest not highlighting it here.

#### Response:

We've removed this sentence (L81-82).

### Comment 12

Line 85: Please specify the version of CoLM used in this study, e.g., CoLM 2014 or CoLM 2024.

#### Response:

The version of CoLM used in this study is based on CoLM 2014 and is operationally employed as the land surface model within the GRAPES (Global/Regional Assimilation and Prediction Enhanced System) global model at the China Meteorological Administration. This information has been specified in the revised manuscript (L96-99).

### Comment 13

Section 2.2: This section mainly describes a physical process already implemented in CoLM and does not appear to be a methodological contribution of this study. If retained, it should be reorganized to clearly explain how soil moisture and soil temperature are linked to LST in the proposed assimilation framework; otherwise, it may not be necessary.

### Response:

We clarify that the observation operator is not claimed as a methodological contribution in this study. However, providing a clear description of the observation operator is a standard and necessary component in data assimilation literature.

Furthermore, this section demonstrates that LST is directly related to soil temperature. The connection between LST and soil moisture is bridged by the temperature-moisture covariance. This cross-variable update mechanism is detailed in Section 2.3. To make this clearer, we have added a sentence at the end of this section pointing to Section 2.3 (L127-128), where this cross-variable update mechanism is explicitly detailed.

### Comment 14

Section 4.2: The improvement in LST is very small. This may be related to the choice of perturbing model parameters rather than atmospheric forcing. Please discuss this possible limitation.

### Response:

We agree that the limited improvement in LST is closely related to the use of deterministic atmospheric forcing. The central goal of our study is to explore how to effectively assimilate LST within an offline framework, even under such forcing conditions. We generated background uncertainty through initial state sampling and parameter perturbations rather than forcing perturbations. Because LST is a fast-varying variable highly sensitive to external atmospheric drivers, the unperturbed forcing strongly constrains the surface energy balance. Consequently, even after LST is corrected at the assimilation step, it quickly relaxes back toward the trajectory dictated by the fixed forcing. Nevertheless, as demonstrated by other

results in the manuscript, the assimilation of LST effectively propagates its impact to deeper soil layers through adjustments in soil temperature and moisture. The modifications have been added in L325-329. To address the forcing-related limitation, we plan to use a land-atmosphere coupled model in future work to account for atmospheric feedbacks, as noted in the manuscript.

#### Comment 15

Line 290: Please clarify how “error” is defined here.

#### Response:

"Error" refers to the observation error, which specifically represents the inherent uncertainty associated with the observational measurements. We have clarified this in the revised text (L332-333).

#### Comment 16

Line 296: “temperature” should be changed to “soil temperature”.

#### Response:

Following the advice of Reviewer #2, we’ve changed the name of this section to "Impacts on Soil and Snow Temperature" (L338).

#### Comment 17

Figure 3: The open-loop experiment appears to perform better in winter. Please explain why this occurs.

#### Response:

In winter, the open-loop experiment exhibits a systematic negative bias in soil temperature. However, assimilating LST alters the snow state, effectively converting "old snow" into "new snow". The enhanced insulation of the new snow then traps excessive heat within the soil layers, shifting the soil temperature toward a positive bias. This reversal from negative to positive bias ultimately leads to the higher RMSD observed in winter.

## Comment 18

Figure 4: In February, performance degrades in many regions. Please explain the possible reasons for these degradations.

### **Response:**

The degradation observed in February is physically linked to the soil temperature biases discussed earlier. Since the atmospheric forcing is prescribed, the thermal exchange at the snow-atmosphere interface is effectively constrained. At the same time, the underlying soil—now anomalously warm due to the enhanced insulation of the updated snowpack—drives a stronger upward ground heat flux into the snow layer. This persistent upward energy transfer artificially elevates the snow temperature, leading to the degradation in RMSD during peak winter conditions.

## References

- [1] Bin Cao et al. “The ERA5-Land Soil Temperature Bias in Permafrost Regions”. In: *The Cryosphere* 14.8 (Aug. 12, 2020), pp. 2581–2595. DOI: 10.5194/tc-14-2581-2020.
- [2] Kathrin Naegeli et al. *ESA Snow Climate Change Initiative (Snow\_cci): Daily Global Snow Cover Fraction - Snow on Ground (SCFG) from AVHRR (1982 - 2018), Version 2.0*. application/xml. NERC EDS Centre for Environmental Data Analysis, 2022. DOI: 10.5285/3F034F4A08854EB59D58E1FA92D207B6.
- [3] Thomas Nagler et al. *ESA Snow Climate Change Initiative (Snow\_cci): Daily Global Snow Cover Fraction - Snow on Ground (SCFG) from MODIS (2000-2020), Version 2.0*. application/xml. NERC EDS Centre for Environmental Data Analysis, 2022. DOI: 10.5285/8847A05EEDA646A29DA58B42BDF2A87C.
- [4] K.A. Novick et al. “The AmeriFlux Network: A Coalition of the Willing”. In: *Agricultural and Forest Meteorology* 249 (Feb. 2018), pp. 444–456. DOI: 10.1016/j.agrformet.2017.10.009.
- [5] Gilberto Pastorello et al. “The FLUXNET2015 Dataset and the ONEFlux Processing Pipeline for Eddy Covariance Data”. In: *Scientific Data* 7.1 (July 9, 2020), p. 225. DOI: 10.1038/s41597-020-0534-3.
- [6] Rune Solberg et al. *ESA Snow Climate Change Initiative (Snow\_cci): Daily Global Snow Cover Fraction - Snow on Ground (SCFG) from ATSR-2 (1995 - 2003), Version 1.0*. application/xml. NERC EDS Centre for Environmental Data Analysis, 2023. DOI: 10.5285/0AEBA0C203C2447B9553A78F99D3A276.