https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Latent Space Data Assimilation of GLDAS Soil Moisture and Soil Temperature into E3SM Coupled Land-Atmosphere Simulations

Yilin Fang[1], Gautam Bisht[2], Lingcheng Li[2], Hoang V Tran[2], Shixuan Zhang[2], and L. Ruby Leung[2]

[1]Earth System Science Division, Pacific Northwest National Laboratory, Richland, WA, USA

[2]Atmospheric, Climate, and Earth Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA

Correspondence: Yilin Fang (yilin.fang@pnnl.gov)

Abstract

Soil moisture (SM) and soil temperature (ST) have strong memory and influence surface-energy, water, and carbon fluxes. However, assimilating them into Earth system models remains challenging because they are high-dimensional, climatologically inconsistent across products and models, and partially observed. Here we develop and evaluate a latent space data assimilation (DA) framework for assimilating GLDAS-informed multilayer, monthly SM and ST anomalies into Energy Exascale Earth System Model (E3SM) land model (ELM) simulations. The framework uses an autoencoder to represent ELM states in latent space, an observation-to-latent network to map noisy, masked GLDAS-derived constraints into the same space, and an element-wise latent Kalman update to combine background and GLDAS-derived constraints before decoding the analysis back to physical states. We evaluate the method using synthetic validation, offline DA experiments, and Atmospheric Model Intercomparison Project simulations with online DA coupling. GLDAS comparisons assess assimilation consistency, while independent runoff, snow water equivalent, precipitation and near-surface air temperature products diagnose downstream impacts on unassimilated variables. Results show that latent space DA efficiently adjusts ELM SM and ST anomalies toward GLDAS-informed wet/dry and warm/cold departures while allowing subsequent evolution to remain governed by ELM physics. Independent responses are regionally and seasonally heterogeneous, indicating that land-state updates can influence unassimilated hydrological processes, precipitation and near-surface temperature, but do not guarantee uniform improvement across downstream variables. These results highlight the need to evaluate both land state consistency and downstream coupled impacts. The framework provides a scalable and computationally efficient pathway for high-dimensional land state updates in Earth system models.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

1. Introduction

Soil moisture and soil temperature are important land-state variables that regulate surface energy and water fluxes, runoff generation, vegetation water stress, boundary-layer development, and land–atmosphere coupling (Seneviratne et al., 2010; Rahmati et al., 2024). Because these variables evolve more slowly than atmospheric states, particularly the weeks-to-months memory of root-zone soil moisture and deeper soil temperature, errors in soil moisture and soil temperature can persist for weeks to seasons and influence climate variability and prediction skill (Ghannam et al., 2016; Qiu et al., 2024; Rahmati et al., 2024; Zhang et al., 2021). Improving land initial states is therefore an important step toward improving subseasonal-to-seasonal prediction and land process attribution in Earth system models (Ardilouze and Boone, 2024; Duan et al., 2025).

Data assimilation (DA) provides a formal framework for combining model background states with observational constraints (Kalnay, 2003; Li et al., 2023). Traditional land DA approaches include variational methods and ensemble Kalman filter methods (Li et al., 2023; Reichle et al., 2002). Three-dimensional variational data assimilation estimates the model state at a particular analysis time by balancing consistency with the prior model state and agreement with observations, while four-dimensional variational DA extends this optimization across a time window, using the model dynamics to fit observations collected at multiple times (Courtier et al., 1998; Kalnay, 2003). Ensemble Kalman filter approaches, including land data assimilation applications implemented through DART and NASA's Land Information System (LIS), estimate flow-dependent background covariance from an ensemble and use this covariance to spread observational information across state variables and grid cells (Anderson et al., 2009; Evensen, 2009; Kumar et al., 2008; Zhang et al., 2014). These methods are powerful and physically interpretable, and scalable solvers, domain decomposition, and ensemble level parallelism can substantially reduce their wall-clock cost.

However, global multilayer land DA remains difficult because the state dimension is large, land variables are strongly coupled across space and depth, and background and observation errors can have complex spatial, vertical, and cross-variable structure. Ensemble-based methods are additionally affected by finite sample covariance error and therefore often require adequate ensemble sizes, localization, inflation, and carefully designed observation operators and cross-variable update strategies (He et al., 2020; Lei & Whitaker, 2017). Particle-filter methods provide a more general Bayesian alternative, and recent advances in localization, improved probability distributions, hybrid formulations, and evolutionary or MCMC-based approaches have reduced particle degeneracy in nonlinear and non-Gaussian hydrologic applications. Nevertheless, representing high-dimensional posterior distributions with finite particle sets remains computationally and statistically challenging, particularly for global, multilayer land models (Fearnhead & Künsch, 2018; Abbaszadeh et al., 2018; van Leeuwen et al., 2019).

Assimilating externally derived land state products introduces an additional challenge because the assimilated product and the target Earth system model may differ in land-surface physics, meteorological forcing, soil properties, vertical discretization, spatial resolution, and climatological state. Such differences can produce systematic discrepancies between modeled and externally derived soil states. Because conventional DA generally assumes approximately

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

unbiased model and observation errors, directly assimilating raw physical values may generate persistent analysis increments that correct differences between the two modeling systems rather than the land anomalies of interest (Yilmaz & Crow, 2013; Crow et al., 2024). Consequently, land data assimilation studies commonly rescale observations to the model climatology or assimilate anomalies to reduce mean state incompatibility (Yilmaz & Crow, 2013; Shi et al., 2024).

Recent machine-learning-based DA methods have explored different ways to reduce the limitations of conventional covariance modeling. Generative-prior approaches, including plug-and-play and diffusion-model DA methods, use pretrained denoising, flow-matching, or score-based models to introduce more flexible prior information than Gaussian covariance assumptions alone (Huang et al., 2024; Manshausen et al., 2024; Qu et al., 2025; Hodyss and Morzfeld, 2025). These methods can represent richer prior structure but often require iterative denoising, sampling, or optimization during the analysis step. In contrast, learned latent space DA approaches provide a complementary direction by performing the analysis in a reduced representation of the model state, where the dominant spatial and multivariate structures can be retained while covariance estimation and analysis updates become more tractable (Peyron et al., 2021; Pasmans et al., 2025; Fan et al., 2025a,b; Fan et al., 2026). For example, latent space Kalman approaches retain an interpretable Bayesian update in a learned low-dimensional space, enabling covariance estimation to be simplified relative to full physical space DA (Peyron et al., 2021; Pasmans et al., 2025). Previous studies have demonstrated the potential of learned latent representations to improve the scalability of ensemble and Kalman-based data assimilation in geophysical systems (Peyron et al., 2021; Pasmans et al., 2025; Fan et al., 2025b; Fan et al., 2026).

Building on these latent space DA developments, this study develops a framework for jointly assimilating soil-moisture and soil-temperature information into global, multilayer E3SM land states. Rather than operating directly on the full physical state, the framework represents the land state in a compact learned latent space in which dominant multivariate and spatial structures can be captured with substantially fewer dimensions. This design extends latent space assimilation to global, multilayer land states and examines whether soil moisture and soil temperature from an external land product can be used to improve the initial conditions of the Energy Exascale Earth System Model (E3SM) land model (ELM) while reducing the covariance-estimation burden of physical-space assimilation. For methodological context, Table S1 provides a conceptual comparison between the proposed latent space DA framework and representative alternatives, including bias correction, variational DA, ensemble Kalman filtering, particle filtering, and generative model-based DA.

The goal of this study is to efficiently update ELM soil moisture and soil temperature anomalies while allowing ELM to continue evolving according to its own physical parameterizations. This provides a useful capability for generating land initial conditions for understanding the role of land processes and land-atmosphere interactions in Earth system predictability through hindcast experiments. We evaluate the framework through synthetic experiments, offline DA tests, and Atmospheric Model Intercomparison Project (AMIP) coupled land-atmosphere simulations with data assimilation. Independent runoff, snow water equivalent, precipitation and near-surface air-

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

temperature products are then used to assess whether the assimilation improves physically relevant land–atmosphere behavior beyond agreement with the assimilated land state product.

2. Data and preprocessing

2.1 E3SM background simulations

The model background state and training data were obtained from previously completed three-member E3SMv3 coupled land-atmosphere simulations following the Atmospheric Model Intercomparison Project (AMIP) configuration, with prescribed sea-surface temperature and sea-ice concentrations and an interactive land component (Xie et al., 2025). Results from 1980 to 2014 of the three AMIP members are used to characterize background variability. The ELM output is represented on the native 0.5° horizontal grid used in these simulations. The DA state includes soil moisture (SM) and soil temperature (ST) from the upper 10 native ELM soil layers, giving 20 input channels in total. These layers are selected because they represent the near-surface to root-zone soil most directly connected to land-atmosphere water and energy exchanges and most comparable to the depth range constrained by GLDAS SM and ST information. The same 10-layer structure is used for both SM and ST so that water and thermal states are updated consistently across the multilayer soil column. AMIP fields are separated into samples from 1980-2008 for training, with 2009-2011 used for validation and 2012-2014 used for testing. The training period is used to compute monthly model climatology and standard deviation for anomaly standardization.

For each variable and grid cell, AMIP monthly climatology and standard deviation are computed from the training period. Soil moisture and soil temperature standard deviations are lower bounded to avoid unrealistically large z-scores in dry or frozen regions. The standardized AMIP anomaly x' at time t is defined as

$$x'(\mathrm{t}) = \frac{x(t) - \mu_m}{\tilde{\sigma}_m}, \tilde{\sigma}_m = max(\sigma_m, \sigma_{min}) \quad (1)$$

where m denotes calendar month corresponding to time t, μ_m is the AMIP training climatology, and σ_m is the AMIP training standard deviation, σ_{min} is the variable-specific lower bound for soil moisture or soil temperature. We use "z-scored" to refer to this standardized anomaly.

2.2 GLDAS and ERA5 land states

Global Land Data Assimilation System (GLDAS) soil moisture and soil temperature (Rodell et al., 2004) are used as the primary observation-like constraints. Because GLDAS and E3SM have different climatological frames, GLDAS fields are transformed into an AMIP-compatible anomaly space before assimilation. The GLDAS monthly climatology is estimated over the training period and removed from the GLDAS fields, while the resulting GLDAS monthly anomalies are normalized by the AMIP training standard deviation. This produces GLDAS-derived standardized anomalies that are comparable to the AMIP anomaly states used to train the autoencoder. ERA5 (Hersbach et al., 2020) land states are used together with GLDAS and AMIP

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

in triple-collocation (TC) analysis to estimate approximate GLDAS-derived monthly soil moisture and soil temperature anomaly uncertainty. TC provides error-variance estimates for three collocated datasets without requiring any one product to be treated as the true reference (Stoffelen, 1998; McColl et al., 2014), and it has been widely applied to soil moisture product evaluation and merging (Gruber et al., 2016; Gruber et al., 2017). GLDAS, ERA5, and AMIP monthly soil moisture and soil temperature fields were prepared on a common AMIP-compatible grid and soil layer structure before analysis. Where needed, fields were remapped to ensure consistent horizontal and vertical dimensions using bilinear interpolation horizontally, linear interpolation with depth for soil temperature, and depth-weighted averaging for soil moisture. This preprocessing enabled grid and layer anomaly calculation, triple-collocation uncertainty estimation, and construction of inputs.

2.3 Downstream coupled impact evaluation datasets

Evaluation of the coupled AMIP-DA experiments focuses on unassimilated variables and downstream land-atmosphere and hydrologic responses. Monthly precipitation is evaluated against the Global Precipitation Climatology Centre (GPCC) 0.5° V7 Full Reanalysis product, a gauge-based global land precipitation dataset (Becker et al., 2013). Near-surface air temperature is evaluated against the NOAA/NCEP Global Historical Climatology Network–Climate Anomaly Monitoring System (GHCN-CAMS) monthly temperature product, a gridded 0.5° land air temperature analysis (Fan et al., 2008). To assess hydrologic impacts more directly related to the assimilated soil moisture and soil temperature states, monthly runoff is evaluated against the Global Runoff Reconstruction (GRUN), an observation-based global gridded monthly runoff dataset covering (Ghiggi et al., 2019a,b). Snow water equivalent (SWE) is evaluated against the CanSISE observation-based ensemble of northern hemisphere terrestrial SWE, a daily gridded ensemble product constructed from multiple SWE datasets and reanalysis (Mudryk et al., 2015; Mudryk et al, 2017). Because the DA framework assimilates GLDAS soil moisture and soil temperature rather than precipitation, near-surface air temperature, runoff, or SWE, these comparisons provide independent tests of whether land state updates produce physically meaningful impacts beyond agreement with the assimilated GLDAS fields.

3. Method

Using the AMIP background states, GLDAS SM/ST constraints, and standardized anomaly variables defined in Section 2, we develop a latent space DA framework for updating multilayer ELM land states. The proposed framework belongs to a broader class of latent space DA methods, which retain the background–observation weighting concept of traditional Kalman and variational DA but perform the analysis in a learned reduced-dimensional representation rather than directly in the full physical state space (Table S1). This formulation is motivated by the high dimensionality and strong multivariate structure of global multilayer land states, where explicitly specifying or estimating full physical-space error covariances is difficult. By learning a compact representation of ELM soil moisture and soil temperature anomalies, the DA update can be applied in a space where the state dimension is reduced and where dominant spatial, vertical, and cross-variable relationships are represented implicitly by the neural networks.

3.1 Overview of latent space DA

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

We adopt the latent space DA terminology used by Fan et al. (2025a,b), in which the model background, observation-derived estimate, and analysis are represented in a learned latent space before being mapped back to the physical state space. Following this terminology, the encoded ELM background state is referred to as the latent background, the O2LNet output from GLDAS derived inputs is referred to as the observation derived latent estimate, and the updated latent variable after the Kalman-style analysis is referred to as the latent analysis.

The complete framework includes both offline preparation and online assimilation. The offline preparation states include anomaly standardization, autoencoder (AE) training, O2LNet training, and estimation of latent error variances. These steps define the standardized input variables, trained mappings, and latent-space uncertainty estimates used during DA. The online latent space DA framework is illustrated in Figure 1, in which the standardized background state is encoded to form the latent background, the GLDAS-derived anomaly constraint is mapped into the latent space to form the observation derived latent estimate, and the two are combined through a Kalman-style updated before being decoded back to physical SM/ST anomaly space.

The AMIP background anomaly state is first converted to standardized anomaly space (Eq. 1). The autoencoder (AE) encoder maps the background state to latent space:

$$z_b = E(x_b') \tag{2}$$

where x_b' is the standardized AMIP/ELM background anomaly state and $E(\cdot)$ is the AE encoder. GLDAS-derived anomaly constraints are standardized in the same anomaly framework and passed through O2LNet,

$$z_o = O2L(y', m_q) \tag{3}$$

where y' is the standardized GLDAS-derived anomaly constraints and m_q is the observation mask or quality mask, with $m_q = 1$ for grid cells where the GLDAS and ELM land masks overlap and $m_q = 0$ elsewhere in this study. O2L$(\cdot)$ denotes the O2LNet. The latent analysis z_a is computed as

$$z_a = z_b + K_z \odot (z_o - z_b) \tag{4}$$

where K_z is the element-wise Kalman gain, i.e., each value is updated independently using its own scalar gain, and $\odot$ denotes element-wise multiplication. The updated latent state is decoded back to physical model space:

$$x_a' = D(z_a) \tag{5}$$

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

where D(·) is the AE decoder. The decoded analysis x'_a is then inverse-standardized to physical soil moisture and soil temperature and used either for offline analysis or to update ELM states in memory.

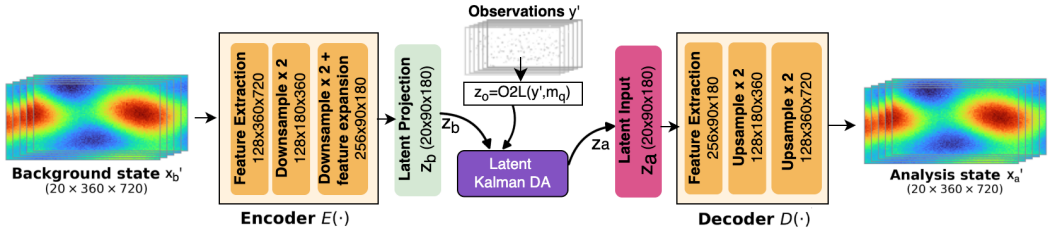


Figure 1. Schematic of the online latent space data assimilation framework for ELM soil moisture and soil temperature anomalies. The 20-channel background state x'_b, including 10 soil-moisture layers and 10 soil-temperature layers, is encoded by the convolutional autoencoder from the physical grid 20×360×720 to the latent grid 20×90×180. GLDAS-derived standardized anomaly observations y' are mapped by O2LNet into the same latent space to provide the observation-derived latent estimate, which is then combined with the latent background z_b using an element-wise latent Kalman update. The resulting latent analysis z_a is decoded back to the physical-grid analysis state x'_a.

3.2 Autoencoder training

The autoencoder (AE) is trained to reconstruct AMIP z-scored SM/ST anomaly states. The input and target are therefore the same standardized anomaly state, and the reconstruction loss measures the mismatch between the decoder output and the input. In addition to this reconstruction loss, we include a latent channel decorrelation penalty. For each training batch, the encoded latent state is used to compute a channel-by-channel correlation matrix, and the penalty encourages this matrix to be close to the identity matrix. This regularization discourages redundant latent channels, so that different latent dimensions encode more distinct information. Similar redundancy reduction or covariance regularization ideas are widely used in representation learning, where the learned feature dimensions are encouraged to be less correlated (Zbontar et al., 2021; Bardes et al., 2022). In the present application, the decorrelation term is used only as a soft regularization, i.e., it reduces channel redundancy but does not guarantee independent latent variables or remove all spatial correlations within each latent channel. Then the AE loss combines a reconstruction loss ($\mathcal{L}_{rec}$) with a latent-channel decorrelation penalty:

$$\mathcal{L}_{AE} = \mathcal{L}_{rec} + \lambda_{decor}\mathcal{L}_{decor} \tag{6}$$

where $\mathcal{L}_{rec}$ measures the reconstruction error of AMIP z-scored anomalies, $\mathcal{L}_{decor}$ encourages the latent-channel correlation matrix to be close to the identity matrix, and λ_{decor} is the latent-channel decorrelation penalty, which controls the strength of this regularization. The decorrelation penalty is included to make the latent representation less redundant and more

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

suitable for a diagonal gain approximation, but it should not be interpreted as enforcing full statistical independence.

The AE is trained only on model-space AMIP data. After training, the encoder defines the latent space used by the DA update, and the decoder maps updated latent states back to physical soil moisture and soil temperature. It is trained using standardized AMIP/ELM SM/ST anomaly states x', where each sample contains 20 channels on the 360×720 model grid. The encoder maps x' to a latent tensor ($z = E(x') \in \mathbb{R}^{20\times90\times180}$), and the decoder reconstructs the physical-grid anomaly state. This corresponds to a fourfold reduction in each horizontal dimension, or a 16-fold reduction in the number of horizontal grid cells, while retaining 20 latent channels. The compression factor was selected as a compromise between reducing the dimensionality of the DA update and preserving the large scale spatial structure needed to reconstruct multilayer SM/ST anomaly fields. The adequacy of this choice is assessed through the AE reconstruction diagnostics and downstream DA experiments. Sensitivity to alternative latent resolutions is left for future work. The AE was trained with the AdamW optimizer using a learning rate of 2×10^{-4}, a weight decay of 10^{-5}, and a cosine-annealing learning-rate scheduler. In the present configuration, the AE was trained for 150 epochs with a batch size of 4.

3.3 O2LNet training

Because GLDAS is itself a land data assimilation product rather than a direct observation, we use "GLDAS-derived constraint" to describe the SM/ST information assimilated in this study. Accordingly, "observation" refers to the GLDAS-derived constraint field. Each clean or unperturbed AMIP anomaly sample is paired with a GLDAS-derived valid data mask, and the input to O2LNet is produced by applying this mask to the AMIP anomaly field and adding synthetic perturbations to the AMIP anomaly sample. These perturbations are intended to represent effective uncertainty in the GLDAS-derived constraint rather than direct instrument observation error. The amplitude of the perturbations is scaled using triple-collocation derived uncertainty, while the mask excludes grid cells where valid GLDAS and ELM land points do not overlap. In this study, spatially correlated synthetic noise is used, where random perturbations are smoothed over land before being added to the masked AMIP anomaly fields. This provides an approximate representation of spatially coherent uncertainty in the GLDAS-derived SM/ST constraint, including potential contributions from forcing, model structure, representativeness differences, and data assimilation procedures within GLDAS. The TC-derived uncertainty is used to define the amplitude of synthetic perturbations for O2LNet training and for estimating the latent observation/O2LNet-error variance. Because TC assumes mutually independent errors and linear relationships with the unknown truth, these uncertainty estimates are interpreted as approximate diagnostics of effective constraint uncertainty rather than exact GLDAS error variances.

O2LNet provides the observation-side counterpart to the AE encoder by mapping noisy and masked GLDAS-derived SM/ST anomaly constraint into the same latent space as the AMIP background anomalies. This observation-to-latent mapping is a key capability of the framework because it allows incomplete and uncertain externally derived anomaly fields to be assimilated directly in the learned latent space. O2LNet is trained after the AE is pretrained and frozen. During training, clean AMIP standardized anomaly states x' are used to generate synthetic

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

GLDAS-derived observations following the procedure described above. The resulting noisy and masked AMIP-derived observation y' is passed to O2LNet, while the AE-encoded clean or unperturbed AMIP state $E(x')$ provides the latent training target. The O2LNet loss is the latent space prediction error:

$$\mathcal{L}_{O2L} = \left\| O2L\left(y', m_q\right) - E(x') \right\|^2 \tag{7}$$

with spatial weighting based on the observation mask m_q.

The O2LNet input is formed by concatenating the 20-channel GLDAS-derived SM/ST anomaly constraint field $y'$with the corresponding 20-channel observation mask m_q, giving a 40-channel input on the 360×720 grid. As defined in Section 2.1, the 20 SM/ST anomaly channels correspond to the 10 soil moisture and 10 soil temperature layers. O2LNet uses gated convolutional blocks and two stride-2 down sampling stages to produce the observation-derived latent estimate ($z_o = O2L\left(y', m_q\right) \in \mathbb{R}^{20\times90\times180}$). The same AdamW optimizer, learning rate, weight decay, and cosine-annealing learning-rate scheduler as AE are used for O2LNet training, with 150 training epochs in the present configuration.

3.4 Estimation of latent error variances

The latent DA update requires estimate of the background error variance, B_z, and observation-derived/O2LNet-error variance, R_z. In a conventional physical space DA, these quantities would correspond to error covariances among all SM and ST layers and grid cells, including spatial, vertical, and cross -variable dependencies. Estimating such a covariance matrix is infeasible for the global multilayer ELM state considered here, especially with only three AMIP ensemble members. In high-dimensional geophysical DA, ensemble covariance estimates from small ensembles are strongly affected by sampling error and typically require localization, inflation, or other regularization (Hamill et al., 2001; Evensen, 2003; Anderson, 2012; Lei and Whitaker, 2017).

The present framework avoids estimating a full physical space covariance by first mapping the SM/ST state into the learned AE latent space and then applying a diagonal variance approximation here. This approximation assumes that the learned latent variables are sufficiently weakly correlated for an element-wise update to provide a useful and computationally efficient analysis. Similar latent space DA approaches have used learned reduced representations to make Kalman-style updates and covariance treatments more tractable in high-dimensional geophysical systems (Peyron et al., 2021; Fan et al., 2025a; Fan et al., 2026). In this study, the diagonal treatment is therefore interpreted as a practical approximation, not as a claim that latent errors are fully independent. We will evaluate its adequacy in Section 5.

The climatological background error variance B_z is estimated from AMIP ensemble member differences in latent space. AMIP states are first encoded into a compact AE latent space, and squared latent differences among member pairs are pooled over the training record to estimate B_z for each latent channel and spatial location. This produces a stable climatological estimate of background uncertainty while avoiding reliance on a noisy flow-dependent covariance estimates

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

from only three members. Here, flow-dependent covariance refers to an error covariance that changes from one DA cycle to another according to the current ensemble spread or evolving model state. Pooling over the training period mitigates sampling noise, but it does not remove the limitation that B_z is derived from a limited number of AMIP members. As a result, B_z may underrepresent background uncertainty in some regimes and may affect the latent gain, leading to locally over or under confident analysis weights.

The latent observation error variance R_z is estimated from O2LNet prediction error in the synthetic AMIP-based training framework. For clean AMIP training samples, masked and synthetically perturbed GLDAS-derived inputs are generated and passed through O2LNet. The resulting latent prediction is compared with the AE-encoded clean AMIP target:

$$R_z = \mathrm{E}\left[\left(O2L\left(y', m_q\right) - E(x')\right)^2\right] \tag{8}$$

Thus, R_z measures the effective uncertainty of the observation to latent mapping under the synthetic training assumptions, including the imposed mask structure and TC-scaled perturbations. This estimate does not directly measure O2LNet error for actual GLDAS-derived inputs, because no independent global "true" SM/ST latent state exists for the real GLDAS case. Applying O2LNet to GLDAS therefore assumes that the standardized, masked, and perturbed AMIP-based training inputs are sufficiently representative of the GLDAS-derived fields used during DA. This assumption is evaluated indirectly through offline GLDAS consistency diagnostics and coupled downstream impact experiments, but it remains an important limitation of the present framework.

3.5 Latent Kalman update

With diagonal latent covariance assumptions, the latent Kalman gain is computed element by element:

$$K_z = \frac{B_z}{B_z + \gamma R_z} \tag{9}$$

The innovation d_z is the difference between the observation-derived latent state and the background latent state:

$$d_z = z_o - z_b \tag{10}$$

The latent analysis is then calculated using Equation 4.

K_z determines the relative weighting of the encoded AMIP background state and the O2LNet-derived latent estimate. Large B_z or small R_z increases K_z, causing the analysis to trust the GLDAS/O2LNet latent observation more strongly. Small B_z or large R_z decreases K_z, causing the analysis to remain closer to the AMIP background. Because B_z is estimated from a limited

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

AMIP ensemble, even after pooling member differences over the training period, the resulting gain should be interpreted as an approximate climatological weighting rather than an optimal flow-dependent Kalman gain. If B_z is underestimated, the update may be too conservative and vice versa. Multiplicative inflation γ can be applied to R_z to tune the update strength. Inflating R_z reduces the Kalman gain and weakens the observation influence, which can help control overshoot or wrong-direction increments. In the present study, no inflation was applied, i.e., $\gamma = 1$. The present implementation is analogous to static-covariance 3D-Var or a fixed-gain Kalman analysis in the sense that the latent gain is estimated offline and held fixed during assimilation. However, the gain is applied in a learned latent space rather than in the full physical state space, and the observation information is provided by an O2LNet mapping instead of a conventional observation operator.

3.6 Posterior uncertainty

Although uncertainty diagnostics are not evaluated in the present analysis, the latent DA formulation naturally provides a posterior latent analysis error variance. Under the diagonal latent covariance approximation, B_z, R_z, and K_z are stored as element-wise latent variance and gain fields with shape 20×90×180, rather than as full covariance matrices. Thus, each latent channel location element is updated independently as a scalar. Under this approximation, the posterior latent analysis error variance is given by

$$A_z = (1 - K_z)B_z \tag{11}$$

A_z represents the remaining latent uncertainty after combining the background and observation-derived latent state. Physical-space uncertainty can be estimated by perturbing the latent analysis by $\sigma_{A_z} = \sqrt{A_z}$, decoding the perturbed states, and computing the resulting spread in soil moisture and soil temperature. This provides a diagnostic uncertainty estimate for each DA cycle. For trajectory uncertainty, an initial-condition ensemble can also be used, in which each ELM member is updated independently during the monthly DA cycle and ensemble spread is analyzed through time.

3.7 Integration of latent space DA within ELM

DA is implemented in memory through a Cython-based Python bridge that connects ELM with the trained AE/O2LNet data assimilation workflow. Cython enables Python-based scientific workflows to be compiled into efficient extension modules and to interface with existing C and Fortran code, making it suitable for coupling Python ML components with compiled Earth system model code (Behnel et al., 2011). At each monthly DA cycle, the ELM soil moisture and soil temperature background states are constructed from grid cell accumulated monthly mean history fields, consistent with the fields used for training. The update is applied at the end of the month, before the history/restart write, using the GLDAS SM/ST anomaly for the just-completed calendar month. The ELM background and GLDAS anomaly are standardized using the AMIP training climatology and standard deviation, then mapped into the learned latent space using the AE encoder and O2LNet, respectively. After the latent Kalman update, the updated latent state is decoded back to physical soil moisture and soil temperature. ELM then applies the resulting grid

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

cell increments additively to all columns within each grid cell, preserving subgrid column differences while shifting the grid cell mean SM and ST states carried into the next month. The update modifies liquid soil water and soil temperature, so the current implementation introduces an unbalanced water and energy increment at the DA step, balance checks are therefore skipped for the DA step and the following time step. This design avoids writing and restarting full model state files at every DA cycle and allows the trained ML-based DA workflow to be embedded in the ELM simulation.

4. Experimental design and evaluation metrics

4.1 Validation of the DA framework using synthetic experiments

Synthetic validation is conducted to test whether the AE, O2LNet, and latent Kalman gain can recover a known clean AMIP states from a noisy and masked inputs. In this experiment, AMIP states from one of the AMIP members is treated as background x'_b, and the average of the other two members is treated as truth x'_{truth}. The synthetic truth is then masked using the GLDAS-derived valid data mask and perturbed with TC-scaled synthetic noise to generate observation-like inputs y'. O2LNet maps this noisy and masked input to the observation-derived latent estimate z_o, which is combined with the encoded background state z_b through the latent Kalman update. The decoded analysis x'_a is then compared with the know clean target x'_{truth}.

This experiment evaluates the three main components of the framework: AE reconstruction skill, O2LNet observation-to-latent mapping skill, and the ability of the latent Kalman update to combine the AMIP background and GLDAS-derived SM/ST anomaly constraint information. Metrics include AE reconstruction root mean square error (RMSE), O2LNet latent prediction error, analysis RMSE, anomaly correlation, and depth-resolved performance for soil moisture and soil temperature. Because the true AMIP target is known in this experiment, synthetic validation provides the most direct test of whether the latent DA update improves the background state under controlled masking and perturbation assumptions. However, the design is constrained by the availability of only three AMIP members, so it should be interpreted as a controlled consistency test rather than a comprehensive sampling of E3SM background uncertainty.

4.2 Offline DA evaluation

Offline DA is performed using saved AMIP background states and GLDAS SM/ST anomaly without feeding the analysis back into a running ELM simulation. This experiment isolates the behavior of the latent DA update. We evaluate how the analysis changes relative to the AMIP background and GLDAS, including RMSE, anomaly correlation, and increment-ratio diagnostics:

$$\mathrm{r} = \frac{x'_a - x'_b}{\mathrm{y}' - x'_b} \tag{12}$$

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Values 0<r<1 indicate updates in the correct direction without overshoot, r>1 indicates overshoot, r<0 indicates wrong-direction updates, and $r \approx 0$ indicates a weak update. This diagnostic is used to quantify whether the physical-space analysis behaves as a simple interpolation between the background and observation-derived anomaly, or whether the nonlinear latent update and decoder produce overshooting or opposite-direction increments. To avoid unstable ratios when the background constraint departure is near 0, r is computed only for samples satisfying $|y' - x'_b| > \varepsilon$, where ε is a small threshold in standardized anomaly units, which is 0.01 for this study.

Because the present implementation uses climatological latent error variances, the resulting latent gain is fixed in time, although it varies across latent channels and spatial locations. We therefore separate the increment-ratio diagnostics by the latent variance ratio (B_z/R_z) and anomaly regimes. These diagnostics are used to assess whether overshooting or wrong-direction increments are concentrated in high-gain or extreme-anomaly regimes, which would indicate regimes where a purely climatological gain may be too aggressive or insufficiently adaptive. Such patterns do not uniquely diagnose covariance error, because they may also reflect O2LNet mapping uncertainty, decoder nonlinearity, or residual product–model inconsistencies, but they provide useful evidence for whether flow-dependent or regime-dependent latent variance estimates may improve future implementations.

4.3 AMIP simulations with in-memory DA

The trained DA system is coupled with ELM to perform 5-year AMIP simulations starting from 1980. The initial condition is from one of the AMIP members used in training. Monthly DA cycles update ELM soil moisture and soil temperature states using GLDAS SM/ST anomaly. The resulting AMIP-DA simulations are compared with the original free-running AMIP background simulation to assess whether SM/ST assimilation affects unassimilated land-hydrology variables and determine whether land state updates persist and influence precipitation and near-surface temperature. Evaluation focuses on both direct land state changes and downstream coupled responses. Land state diagnostics include soil moisture and soil temperature anomaly changes, and increment persistence.

We evaluate runoff and snow water equivalent (SWE) against independent reference datasets. Runoff is compared with GRUN, and SWE is compared with the CanSISE SWE product. Because neither runoff nor SWE is assimilated, these diagnostics provide an independent assessment of downstream hydrological impacts from the land-state updates. We also compare AMIP and AMIP-DA precipitation against GPCC and compare near-surface temperature against an independent gridded station-based temperature product. This evaluation tests whether assimilating soil moisture and soil temperature produces physically meaningful effects beyond improved agreement with the assimilated GLDAS land states. Improvements in independent precipitation or near-surface temperature correlation metrics would suggest that the land state updates influence broader land–atmosphere interactions, while neutral or degraded metrics would indicate that improved land-state consistency does not necessarily translate into improved atmospheric or hydrologic responses.

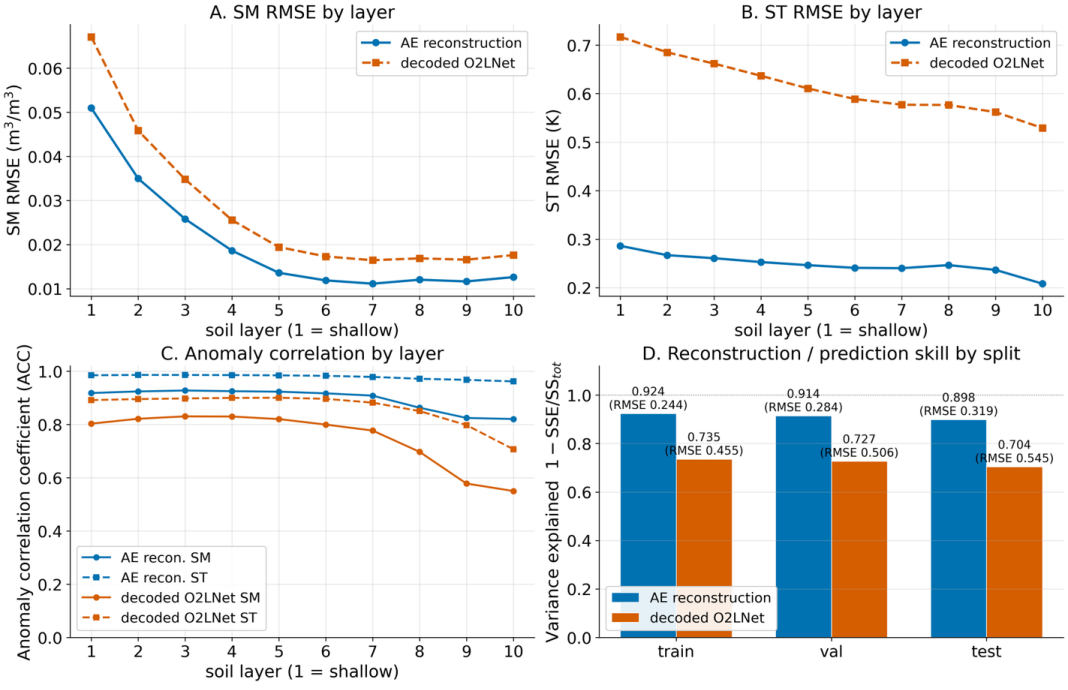
5. Results

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

5.1 AE and O2LNet performance

The AE and O2LNet were first evaluated independently of the Kalman update (Fig. 2). The AE reconstructs AMIP soil moisture (SM) and soil temperature (ST) anomaly states with low error across the 10 soil layers included in the DA state. AE reconstruction RMSE generally decreases with depth for both variables (Fig. 2A,B), a pattern consistent with smoother and more vertically coherent deeper layer anomaly structures that are easier for the AE to reconstruct in standardized anomaly space. However, because this depth dependence may also reflect the vertical structure of the standardized AMIP anomalies themselves, the RMSE profile should be interpreted together with the anomaly correlation results. Anomaly correlation remains high, exceeding about 0.8 for SM and about 0.95 for ST across layers (Fig. 2B). This indicates that the AE preserves the main multilayer anomaly structure while compressing the 20-channel land state into latent space. In the split-level diagnostic, variance explained (VE) is high: 0.92 in training, 0.91 in validation and 0.90 in testing, indicating stable reconstruction behavior on unseen periods (Fig. 2D).

O2LNet has a more difficult task than the AE because it maps masked and synthetically perturbed observation-like inputs into the AE latent space. When the O2LNet predicted latent state is decoded back to physical anomaly space, its RMSE is larger than the AE reconstruction error for both SM and ST (Fig. 2A,B), but the decoded fields retain substantial anomaly correlation with the clean AMIP target (Fig. 2C). Decoded O2LNet performance is lower than AE, but still meaningful, and VE is 0.74 in training, 0.73 in validation, and 0.70 in testing (Fig. 2D). This means decoded O2LNet captures about 70% of the clean AMIP anomaly variability in the test period, despite starting from masked and noisy inputs.



https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Figure 2. Performance of AE and O2LNet before the latent Kalman update. (A) Soil moisture (SM) RMSE by soil layer for the AE reconstruction and the decoded O2LNet prediction. (B) as in (A) but for soil temperature (ST).(C) Anomaly correlation by soil layer for the AE reconstruction and the decoded O2LNet, shown separately for SM and ST. (D) Reconstruction/prediction skill for training, validation, and test samples, measured as the fraction of variance explained (VE = 1-SSE/SS-tot, where SSE is the sum of squared errors relative to the clean AMIP target and, SS_{tot} is the total sum of squared deviations of the clean AMP target about its mean). Both estimators are evaluated in the same z-scored input space: the AE reconstruction and the decoded O2Net. Values above the bars show VE, with RMSE in parentheses. The training period is 1980-2008, with 2009-2011 used for validation and 2012-2014 used for testing.

5.2 Synthetic validation of the latent DA update

The synthetic experiment evaluates the full AE→O2LNet→latent Kalman update→decoder system against a known AMIP target. In this experiment, AMIP member 0 is used as the background, the mean of AMIP members 1 and 2 is treated as the known truth, and GLDAS-dervied masks and synthetic observation noise are applied to the truth state to generate pseudo-observations. This setup allows the O2L-only estimate, i.e., $D(z_o)$, and the DA analysis (x'_a, Eq. 5) to be compared directly with a known clean target.

Both the O2L-only estimate and the DA analysis substantially improve over the background. RMSE decreases and anomaly correlation increases for both soil moisture and soil temperature, with improvements evident across all soil layers. The O2L-only estimate performs best in this controlled setting, especially for soil temperature. This is expected because O2LNet is trained to recover the clean AMIP latent state from synthetic noisy and masked inputs. The DA analysis remains much closer to the truth than the background but is slightly less accurate than the O2L-only estimate because the latent Kalman update blends the observation-derived latent state with the less accurate background latent state. Thus, the DA update behaves conservatively rather

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

than simply replacing the background with the O2LNet estimate.

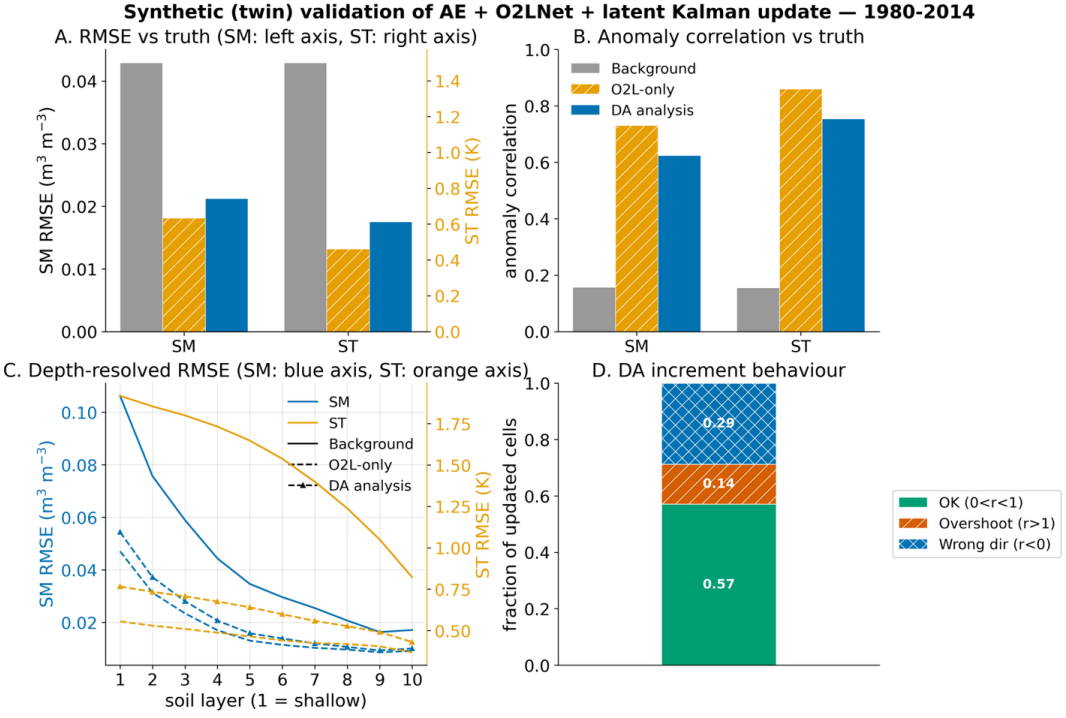


Figure 3. Synthetic validation of the AE, O2LNet, and latent Kalman update over 1980–2014. (A) RMSE against the known AMIP truth for the background, O2L-only decoded estimate, and DA analysis. (B) Anomaly correlation against the known AMIP truth. (C) Depth-resolved RMSE for SM and ST. (D) DA increment behavior fractions (*r* in Eq. 12) in standardized anomaly space, showing acceptable updates (0<r<1), overshooting (r>1), and wrong-direction increments (r<0).

Increment-behavior diagnostics further show that most synthetic DA increments move the analysis toward the truth without overshooting. Approximately 57% of updated samples fall within the nominal range (0<r<1), while about 14% overshoot and 29% move in the opposite direction after decoding. Because these ratios are evaluated after nonlinear decoding back to soil moisture and soil temperature anomaly space, the categories reflect the full AE→O2LNet→latent Kalman update→decoder system rather than the bounded latent Kalman update alone.

5.3 Offline GLDAS assimilation

The offline DA experiment isolates the trained latent update by applying it to saved AMIP background states and GLDAS-derived SM/ST anomaly constraint without feeding the analysis back into a running ELM simulation. Mean analysis increments show spatially coherent soil moisture and soil temperature adjustments rather than random grid-scale noise (Fig. 4A,B). Soil

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

moisture increments are strongest in northern high latitudes and several semi-arid regions, while soil temperature increments show broad regional warming and cooling structures.

Relative to the AMIP background, the offline analysis shows much stronger consistency with the assimilated GLDAS-derived anomaly constraint (Fig. 4C). Layer aggregated SM anomaly RMSE decreases from 0.043 to 0.02 m^3/m^3, a 54.6% reduction, while SM anomaly correlation increases from 0.04 to 0.57 (Fig. 4C). ST anomaly RMSE decreases from 1.86 to 1.04 K, while ST anomaly correlation increases from 0.08 to 0.72. Across soil layers, the analysis reduces anomaly RMSE and increases anomaly correlation for both soil moisture and soil temperature (Fig. 4D,E). The spatial win-rate (the percentage of land cells where the analysis reduced the local anomaly RMSE relative to the background) diagnostic (Fig. 4F) shows that DA improves agreement with the GLDAS-derived constraint over most valid land grid cells across soil layers, indicating that the offline consistency improvement is spatially widespread rather than dominated by a small number of regions. Because GLDAS-derived SM/ST fields anomaly are assimilated, agreement with GLDAS is interpreted as assimilation consistency rather than independent validation.

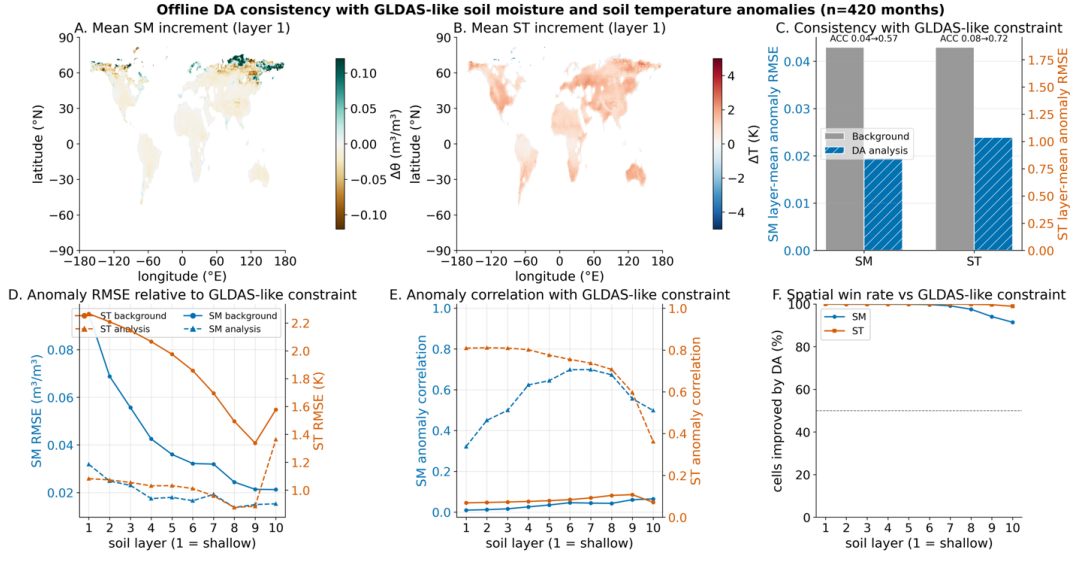


Figure 4. Offline DA consistency with assimilated GLDAS-derived soil moisture and soil temperature anomaly constraints. (A,B) Time-mean DA increments, computed as analysis minus AMIP background, for representative surface-layer SM and ST. (C) Layer-mean anomaly RMSE and anomaly correlation relative to the GLDAS-derived constraint for the AMIP background and DA analysis. (D,E) Depth-resolved anomaly RMSE and anomaly correlation relative to the GLDAS-derived anomaly constraint. (F) Spatial win rate, defined as the percentage of valid land grid cells where the DA analysis has smaller anomaly error than the AMIP background relative to the GLDAS-derived anomaly constraint.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

5.4 Increment-ratio diagnostics

To evaluate the implications of using climatological latent error variances, we diagnose the standardized anomaly-space increment ratio r shown in Equation 12. Across the offline DA period, the mean ratio is approximately 0.51, indicating that the analysis applies a moderate average correction toward the GLDAS-derived anomaly constraint. However, the decoded physical anomaly response is not equivalent to a simple scalar interpolation: approximately 44% of samples fall within 0<r<1, while about 28% overshoot and 28% move in the opposite direction (Fig. 5).

The dependence of increment behavior on the latent variance ratio (B_z/R_z) suggests that the fixed climatological gain affects overshooting (Fig. 5B). The fraction of acceptable updates increases slightly at moderate B_z/R_z. Overshoot fractions increase with larger B_z/R_z, while wrong-direction fractions generally decrease weakly. This indicates that overshooting is more likely when the climatological background variance B_z is large relative to the observation/O2LNet error variance R_z, causing the analysis to place greater weight on the GLDAS-derived latent state.

Separation by anomaly regime shows additional state dependence, with some extreme soil moisture and soil temperature regimes exhibiting larger overshoot fractions (Fig. 5C). Extreme SM and warm ST regimes show larger fractions of acceptable updates, whereas cold ST regimes show larger overshooting fractions. These patterns support the use of the climatological latent gain as an efficient and stable baseline, but they also indicate that flow-dependent (estimates that vary with the current model spread) or regime-dependent (estimates that vary across land-state conditions such as wet/dry, warm/cold, seasonal, or regional regimes) latent error variances may improve future implementations. At the same time, wrong-direction increments should not be attributed solely to covariance error because they can also arise from O2LNet uncertainty, nonlinear decoding, weak-signal cases, or residual product-model inconsistencies.

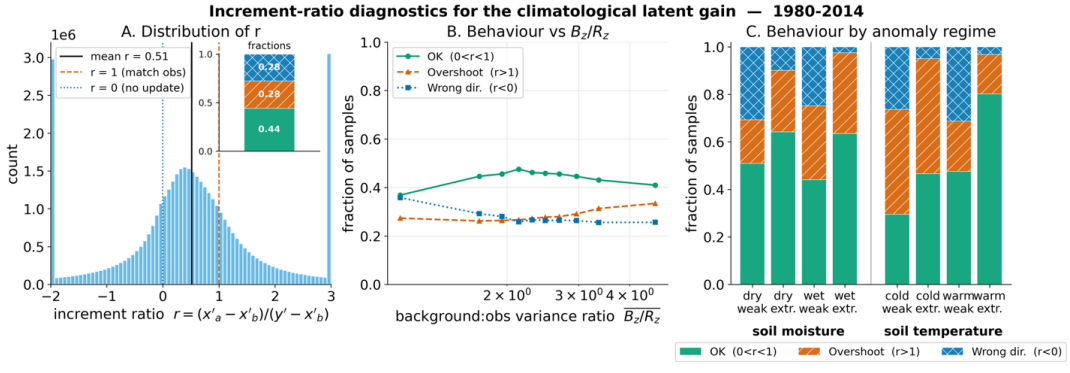


Figure 5. Increment-ratio diagnostics for the climatological latent-gain update during offline DA over 1980–2014. (A) Distribution of the standardized anomaly-space increment ratio, r (Eq. 12); values outside the plotting range are accumulated in the edge bins and therefore indicate the presence of extreme ratios rather than exact values at the plotted bounds. The inset shows the fractions of samples classified as acceptable updates (0<r<1), overshooting (r>1), and wrong-direction increments (r<0). (B) Behavior fractions as a function of the climatological latent

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

variance ratio B_z/R_z. (C) Behavior fractions separated by weak and extreme SM and ST anomaly regimes. Weak anomalies are defined as |anomaly| ≤ 2, and extreme anomalies are defined as |anomaly| > 2. Negative and positive anomalies are labeled as dry/wet for SM and cold/warm for ST, respectively.

5.5 Coupled AMIP-DA land-state impacts

The trained DA system is then applied in an in-memory coupled AMIP-DA simulation and compared with a free-running AMIP simulation over 1980-1984 (Fig. 6). The coupled DA produces time-varying domain-mean AMIP-DA minus free-running AMIP differences in surface layer SM and ST, indicating that the DA system does not impose a constant offset but responds to monthly GLDAS-derived anomaly constraints and subsequent ELM evolution. The spatial maps of AMIP-DA minus AMIP show coherent land-state differences, with soil moisture changes concentrated in high-latitude and semi-arid regions and soil temperature changes showing organized regional cooling and warming patterns.

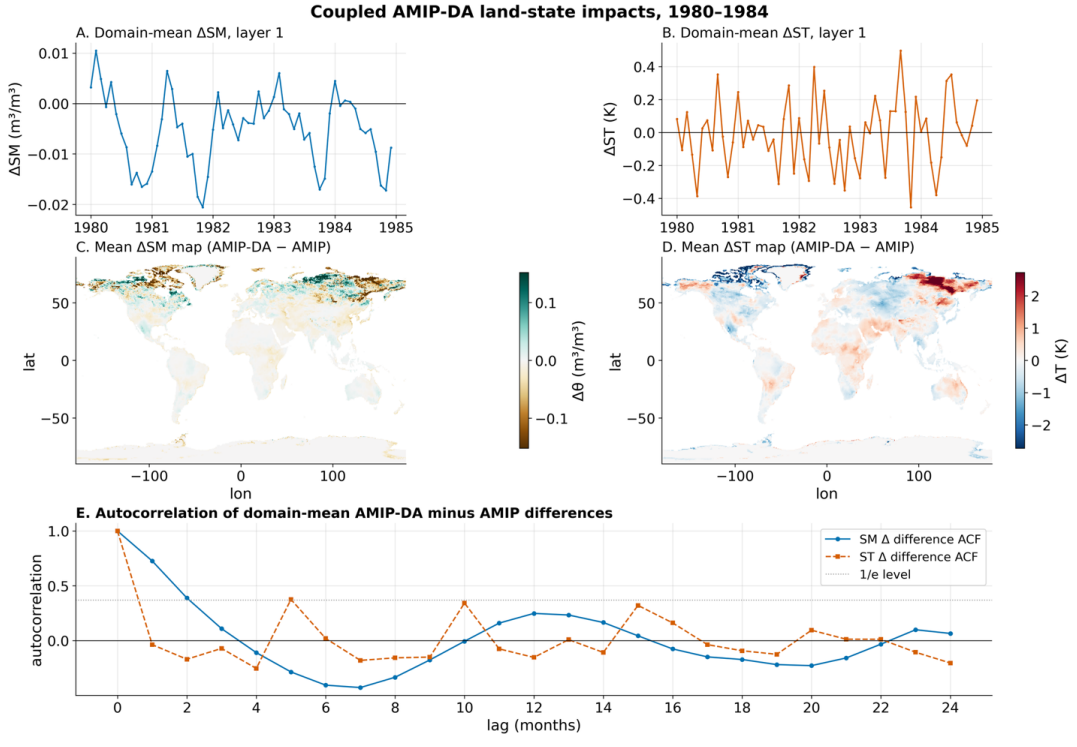


Figure 6. Coupled AMIP-DA impacts on ELM soil states during 1980-1984. (A,B) Monthly area-weighted domain-mean AMIP-DA minus free-running AMIP differences for surface-layer SM and ST. (C,D) Time-mean spatial maps of AMIP-DA minus free-running AMIP differences for the same SM and ST layers. (E) Autocorrelation of the monthly domain-mean AMIP-DA minus AMIP difference (Δ difference) series (ACF) at the surface layer, used as a diagnostic of temporal coherence in the coupled land-state response. These differences reflect the combined

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

effect of the monthly DA updates and subsequent ELM evolution, rather than the instantaneous DA increment alone. The 1/e level, approximately 0.37, is used as a reference threshold for estimating how long DA-induced land-state differences persist.

These coupled differences reflect the combined effect of the monthly DA updates and subsequent ELM land-state evolution, rather than the instantaneous DA increment alone. The autocorrelation of domain-mean AMIP-DA minus AMIP differences indicates that soil moisture differences retain temporal coherence for longer than soil temperature differences. This behavior is physically plausible because soil moisture generally has longer memory than near-surface soil temperature. Overall, the coupled experiment demonstrates that the Python-based latent DA update can be applied during a running ELM simulation and can produce persistent, spatially structured land-state impacts.

5.6 Independent evaluation of unassimilated hydrological responses

To evaluate whether SM/ST assimilation produces downstream hydrological impacts, we compared runoff and SWE from free-running AMIP and AMIP-DA against independent reference datasets over 1980-1984 (Fig. 7). Hydrological diagnostics for unassimilated runoff and SWE showed that DA produces a clear improvement in runoff relative to GRUN. The land mean runoff RMSE is reduced by 4.58 mm/month, with 66% of land grid cells showing lower RMSE than the free-running AMIP simulation (Fig. 7a). The absolute runoff mean error is also reduced, with a land mean decrease of 3.44 mm/month and improvement over 70% of the land (Fig. 7b). Improvements are particularly evident over parts of South America, central and southern Africa, and South and Southeast Asia. Some degradation occurs in parts of high latitude North America and East Asia.

SWE responses are more spatially heterogeneous. AMIP-DA reduces the land mean SWE RMSE by 1.35 mm and the absolute mean error by 1.13 mm relative to the observed SWE product, but only 47-48% of grid cells improve (Fig. 7c-d). Improvements are visible over parts of northern Eurasia, while degradation occurs in parts of western North America and other regions. These mixed SWE responses suggest that the influence of SM/ST assimilation snowpack is less direct than for runoff and likely depends on regional temperature, precipitation, and snow process biases.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

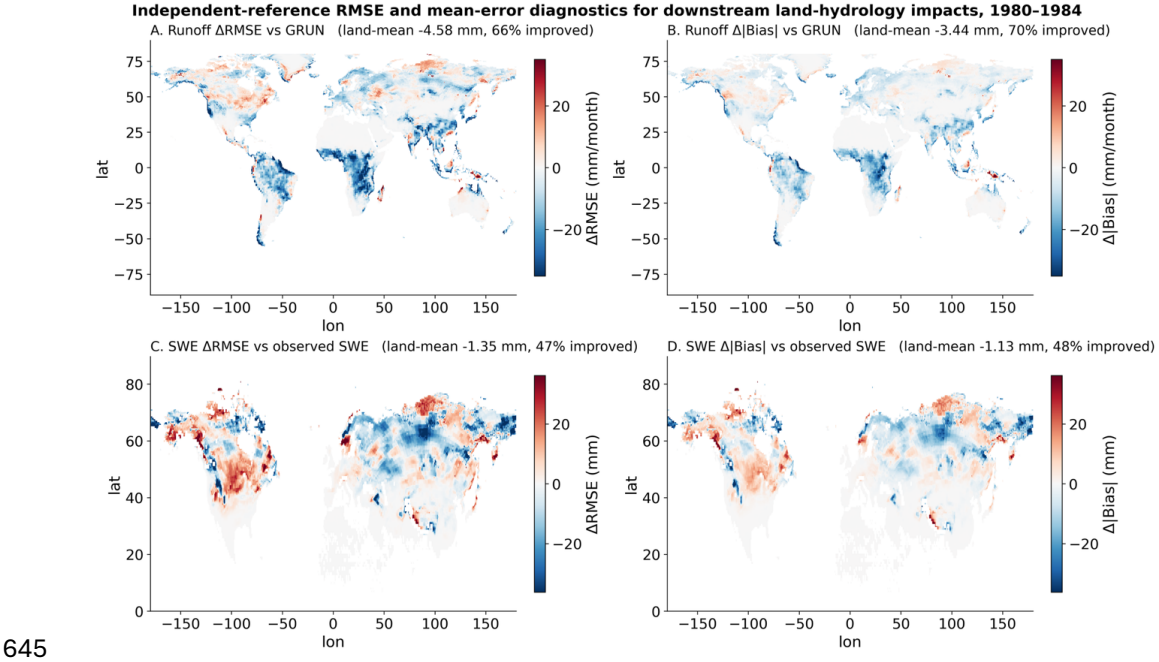


Figure 7. Independent-reference RMSE and mean-error diagnostics for unassimilated hydrological responses, 1980–1984. Maps show AMIP-DA minus AMIP changes in error relative to independent references for runoff and snow water equivalent (SWE). (A) runoff ΔRMSE and (B) runoff Δ|Bias| against GRUN, both in mm month^{-1}, and (C) SWE ΔRMSE = RMSE(AMIP-DA) − RMSE(AMIP), and (D) SWE Δ|Bias| = |Bias(AMIP-DA)| − |Bias(AMIP)|, against observed SWE, both in mm. Negative values indicate that AMIP-DA reduces RMSE or mean-bias magnitude relative to free-running AMIP, while positive values indicate degradation. Each panel title reports the land-mean change and the percentage of land cells improved.

5.7 Independent precipitation and near-surface temperature evaluation

Finally, the coupled AMIP-DA simulation is evaluated against independent precipitation and near-surface air temperature products that are not assimilated in the DA system (Fig. 8). Precipitation skill changes relative to GPCC are spatially heterogeneous. The global land-mean change in monthly precipitation anomaly correlation is small and slightly positive, approximately +0.006, with about 52% of land grid cells showing improvement (Fig. 8A). However, the robust cell mean is slightly negative, approximately -0.011, indicating that precipitation improvements are not spatially uniform and are sensitive to the free-running AMIP baseline anomaly correlation threshold. AMIP-DA slightly reduces RMSE relative to GPCC, with a land mean decrease of 0.34 mm/month and improvement over 53% of land grids (Fig. 8E). The absolute precipitation mean error is also reduced, with a land-mean decrease of 0.52 mm/month and improvement over 54% of grid cells (Fig. 8F). Improvements occur over parts of western and central North America, central and southern Africa, and portions of South and Southeast Asia, while degradation is also evident in parts of tropical South America, South Asia, and Australia.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Near-surface air temperature response relative to GHCN-CAMS is also heterogeneous, but the mean response is negative over the short coupled evaluation period (Fig. 8B). The global land-mean anomaly-correlation change is approximately -0.069, with about 39% of land grid cells showing improvement, and the robust cell mean is approximately -0.104. Seasonal and latitude-zone summaries show that both precipitation and near-surface temperature responses depend strongly on region and season (Fig. 8C,D). For example, some midlatitude and high-latitude seasons show positive skill changes, while other seasons show degraded skill. These results indicate that updating ELM soil moisture and soil temperature can influence unassimilated atmospheric variables, but improved land-state consistency does not guarantee uniform improvement in precipitation or near-surface air temperature skill. AMIP-DA slightly increases land mean RMSE relative to GHCN-CAMS by 0.097 K, with 43% of land grids improved (Fig. 8G). In contrast, the absolute mean-error magnitude is nearly unchanged at the global land mean, with a change of (-0.001) K and 50% of grid cells improved (Fig. 8H). The temperature response shows stronger regional structure than precipitation, with degradation in parts of high-latitude North America and northern Eurasia, but improvements in selected midlatitude and low-latitude regions.

As an illustrative event-based diagnostic, we compare standardized regional anomalies for free-running AMIP, AMIP-DA, and reference datasets during two selected events within the coupled evaluation period: the 1980 central/eastern U.S. drought and heat wave and the 1980–1981 U.S. low-snow winter (Fig. S1). These diagnostics evaluate whether the simulations capture the timing and sign of regional anomaly evolution rather than absolute magnitude or mean bias, because each dataset is standardized by its own 1980–1984 mean and standard deviation. AMIP-DA generally preserves the broad seasonal evolution seen in AMIP while modifying the magnitude and timing of some event-period anomalies. During the 1980 drought and heat-wave case, both AMIP and AMIP-DA capture the strong warm-season 2 m air-temperature anomaly, but precipitation and runoff responses are mixed, and AMIP-DA is not consistently closer to the reference than AMIP. The GLDAS surface soil-moisture reference shows stronger drying than either simulation, indicating that AMIP-DA does not fully reproduce the regional soil-moisture anomaly for this case. During the 1980–1981 low-snow winter, AMIP-DA broadly follows the seasonal evolution of temperature, runoff, and SWE anomalies, but it does not consistently improve agreement with the reference datasets. Given the five-year coupled AMIP-DA period, these diagnostics are interpreted as illustrative examples rather than a formal evaluation of drought, heatwave, or snow-drought skill.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

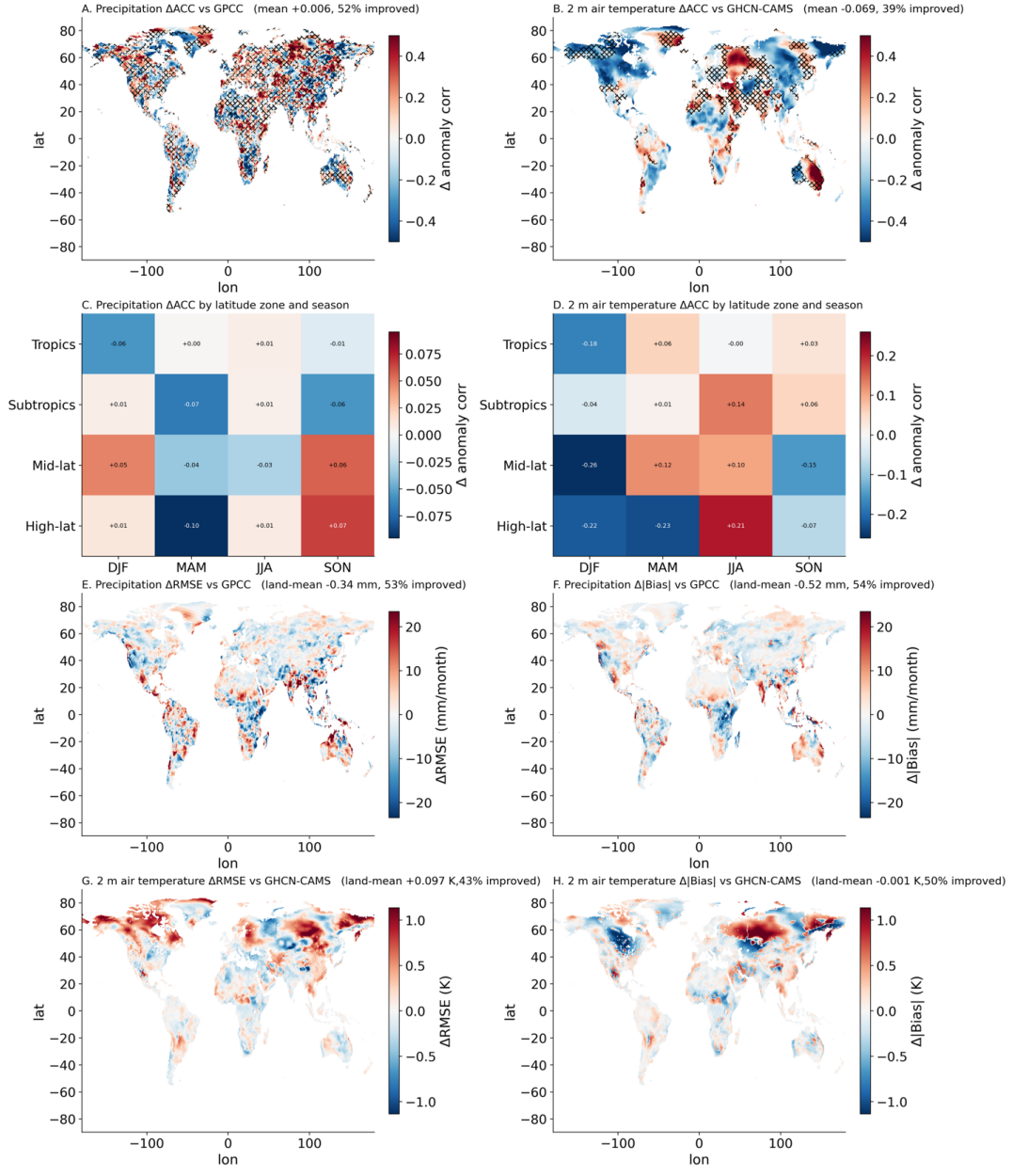


Figure 8. Independent evaluation of coupled AMIP-DA impacts on unassimilated precipitation and near-surface (2-m) air temperature during 1980-1984. (A) Change in precipitation anomaly correlation relative to GPCC, computed as AMIP-DA minus free-running AMIP. (B) Change in near-surface air temperature anomaly correlation relative to GHCN-CAMS, computed as AMIP-DA minus free-running AMIP. Positive values indicate improved agreement in AMIP-DA relative to free-running AMIP. Hatching marks grid cells where the free-running AMIP baseline

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

anomaly correlation has small absolute magnitude, $|ACC_{AMIP}|<0.1$, so skill changes in these regions should be interpreted cautiously. (C, D) Latitude-zone and seasonal averages of the same anomaly-correlation changes for precipitation and 2 m air temperature. Panels E–H show absolute error diagnostics computed from full monthly values: ΔRMSE = RMSE(AMIP-DA) − RMSE(AMIP) and Δ|Bias| = |Bias(AMIP-DA)| − |Bias(AMIP)|, where blue (negative) indicates DA reduces the error or bias magnitude. (E) Precipitation ΔRMSE and (F) Δ|Bias| vs GPCC (mm month^{-1}); (G) 2 m air temperature ΔRMSE and (H) Δ|Bias| vs GHCN-CAMS (K). Each panel title reports the land-mean change and the percentage of land cells improved.

6. Discussion

6.1. Latent space DA as a scalable framework for E3SM land-state assimilation

This study demonstrates a latent space data assimilation framework for updating high-dimensional ELM soil moisture and soil temperature states using GLDAS-derived land-state constraints. The framework combines three components: an autoencoder that maps ELM land states into a compact latent representation, an observation-to-latent network that maps masked GLDAS-derived anomaly constraint inputs into the same latent space, and an element-wise latent Kalman update that combines background and observation-derived information. This design shifts the analysis step from the full physical state space to a learned representation where simplified covariance treatment is more tractable.

This approach does not replace conventional DA principles with a purely neural correction. Rather, it preserves an interpretable Bayesian update structure while using machine learning to reduce dimensionality and handle incomplete observation-like inputs. The AE provides an ELM-consistent latent representation of multilayer soil moisture and soil temperature variability, while O2LNet allows externally derived, masked, and noisy fields to be projected into that same representation. The resulting latent update avoids full physical-space covariance estimation and matrix inversion, which are major barriers for global, multilayer land DA. This framework is well suited to high-dimensional land-state assimilation problems where ensemble sizes are limited, observations are incomplete, and model–product inconsistencies are expected.

The results show that the trained system can produce stable and spatially coherent soil moisture and soil temperature updates, and that the workflow can be coupled back to a running ELM simulation through an in-memory interface. These features make latent space DA a practical pathway for land-state initialization and coupled land-model correction in Earth system modeling applications. The framework is potentially generalizable beyond land DA because the analysis is performed on learned latent representations rather than on component-specific physical variables directly. With appropriate training data and observation-to-latent mappings, the same AE→O2LNet→latent-update structure could potentially be applied to atmospheric, ocean, sea-ice, or coupled multicomponent assimilation problems with tailored treatment of observations, uncertainty estimates, and physical constraints.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

6.2. Interpretation of assimilation consistency and coupled impacts

Because GLDAS-derived soil moisture and soil temperature anomaly provide the assimilated constraints, improved agreement with GLDAS should be interpreted as assimilation consistency rather than independent validation. AMIP–GLDAS differences quantify the pre-DA background–observation mismatch, which includes both ELM background uncertainty and product/model differences between AMIP and GLDAS. The independent GRUN runoff, observed SWE, GPCC precipitation and GHCN-CAMS near-surface temperature comparisons therefore provide a complementary diagnostic of whether land-state updates influence unassimilated atmospheric and hydrological variables in AMIP-style coupled land-atmosphere simulations. The mixed regional and seasonal responses indicate that improved land-state consistency can affect downstream variables, but the benefits depend on land–atmosphere coupling regime, season, and background-state conditions, and the specific response variable.

6.3. Computational practicality and in-memory coupling

The computational design is one of the main strengths of the method. Once the AE, O2LNet, and latent variance fields are trained or estimated, each DA cycle requires only standardization, three neural-network forward passes, an element-wise latent update, decoding, and transformation back to physical units. The standalone CPU benchmark for one global monthly DA cycle, using a 20-channel (360×720) SM/ST state and a 20×90×180 latent state, shows that the full update requires approximately 0.90 s per cycle. Most of the runtime comes from AE/O2LNet inference and decoding, while the element-wise latent update requires less than 1millisecond. Because the DA update is applied once per simulated month, the added computational cost is small for the present in-memory ELM coupling workflow.

This computational practicality is important for Earth system model applications. A DA method that requires repeated optimization, iterative sampling, or large matrix operations at every analysis cycle may become difficult to integrate into a production land or Earth system model. In contrast, the direct latent update used here is lightweight and predictable. In the coupled AMIP-DA experiment, monthly output timestamps show little discernible difference in wall-clock progression relative to the free-running AMIP simulation, suggesting that the DA overhead is small in the present monthly coupled configuration.

6.4. Limitations and future directions

Several limitations should guide interpretation and future development. First, the current system uses diagonal climatological latent variances. This makes the method stable and efficient, but it does not explicitly represent cross-latent-variable error correlations or event-specific uncertainty. Future work should test hybrid climatological/flow-dependent gains and regime-dependent uncertainty estimates.

Second, O2LNet is trained using synthetic GLDAS-derived anomaly constraint generated from AMIP states with GLDAS-derived masks and prescribed noise. This provides controlled training, but it cannot fully capture all real GLDAS–ELM inconsistencies, including representativeness differences (e.g., resolution and depth), residual product bias, and structural

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

differences between land models. Triple-collocation-derived uncertainty provides a useful first-order noise amplitude, but it should not be interpreted as a complete description of observation error.

Third, the current implementation uses diagonal climatological latent variance fields for B_z and R_z. This choice is intentionally simple and reflects the practical constraint of having only three AMIP ensemble members. With such a small ensemble, estimating a full or strongly flow-dependent high-dimensional covariance would be highly vulnerable to sampling error. Estimating stable element-wise latent variances from the training period provides a computationally efficient baseline and avoids noisy instantaneous covariance estimates.

We further evaluate this approximation using latent-correlation diagnostics (Fig. S2). After excluding ocean and inactive latent cells, off-diagonal spatial correlations are weak, while cross-channel correlations are modestly larger. These values exceed the expected mean absolute correlation from random sampling alone for N=36 validation samples, indicating that the latent variables are not fully independent. The diagonal latent-variance assumption should therefore be interpreted as a tractable approximation rather than an exact covariance representation.

The latent correlation diagnostics indicate that the diagonal approximation is reasonable but not exact, and the increment ratio diagnostics (Fig. 5) show that this approximation is useful but imperfect. In latent space, the update is bounded by the element-wise gain, but after nonlinear decoding, the physical-space analysis does not always behave like a simple interpolation between the background and the GLDAS-derived anomaly constraint. Overshooting and wrong-direction increments can occur because the decoder mixes information across channels and neighboring grid cells, and because O2LNet itself introduces mapping uncertainty. Thus, physical-space increment behavior reflects the full AE→O2LNet→latent-update→decoder system, not only the scalar value of K_z.

The dependence of increment behavior on B_z/R_z suggests that the fixed climatological gain may be too aggressive in some regimes and too conservative in others. Because the gain is fixed in time, it cannot respond to event-specific uncertainty, seasonal differences, or regime-dependent observation reliability. This motivates future extensions using flow-dependent or regime-dependent latent uncertainty estimates. A flow-dependent gain would vary from cycle to cycle according to current ensemble spread or model state uncertainty. A regime-dependent gain would use different uncertainty estimates for wet versus dry, warm versus cold, seasonal, regional, or high-departure conditions.

These limitations point to clear opportunities for methodological improvement. The present results show that a simple climatological latent-gain approximation can produce useful land-state adjustments, while the diagnostics indicate how future versions could become more adaptive. Future work should test sensitivities to B_z and R_z inflation, observation-noise assumptions, gain caps that limit the maximum latent update strength, and flow-dependent or regime-dependent latent variance estimates to better separate covariance-specification errors from O2LNet mapping uncertainty and decoder nonlinearity.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Overall, the results support latent space land DA as a promising framework for scalable Earth system model assimilation. The method does not eliminate the assumptions of Kalman-style DA and does not guarantee uniform downstream atmospheric improvement. The main contribution of this framework is that it provides an efficient, interpretable, and diagnosable way to update global multilayer land states in a learned compact representation while remaining compatible with ELM coupling. This makes it a useful foundation for future flow-dependent, regime-aware, and physically constrained land DA systems.

7. Conclusions

We developed a latent space DA framework for updating E3SM/ELM soil moisture and soil temperature using GLDAS-derived anomaly constraints. The method combines an autoencoder, an observation-to-latent network, and an element-wise latent Kalman update to assimilate masked GLDAS-derived constraint information in a compact learned space before decoding the analysis back to physical land states. The framework addresses high dimensionality and product–model climatological inconsistency by estimating simple latent error variances and assimilating standardized anomalies rather than raw GLDAS values. Synthetic, offline, and coupled AMIP-DA experiments show that the method can recover known land states, produce coherent SM/ST increments, improve consistency with assimilated GLDAS-derived anomaly constraint, and run in memory within ELM. Evaluations using independent runoff, SWE, precipitation and temperature datasets show heterogeneous downstream impacts, indicating that land-state updates can affect unassimilated variables but do not guarantee uniform downstream variable improvement. Overall, this approach provides an efficient and diagnosable anomaly-correction framework for high-dimensional E3SM land states. Future work should test adaptive latent error estimates, larger ensembles, and uncertainty diagnostics to assess impacts on land initialization, surface fluxes, and land–atmosphere interactions.

Code and data availability

The Global Land Data Assimilation System version 2.0 (GLDAS-2.0) Noah monthly soil moisture and soil temperature data used as assimilation constraints are available from NASA GES DISC (Beaudoing and Rodell, 2020; https://doi.org/10.5067/9SQ1B3ZXP2C5). ERA5 monthly reanalysis data are available from the Copernicus Climate Data Store (Hersbach et al., 2023; https://doi.org/10.24381/cds.f17050d7). Independent precipitation evaluation uses the Global Precipitation Climatology Centre (GPCC) Full Data Monthly Version 7.0 at 0.5° resolution (Schneider et al., 2015; https://doi.org/10.5676/DWD_GPCC/FD_M_V7_050). Near-surface air temperature is evaluated using the Global Historical Climatology Network–Climate Anomaly Monitoring System (GHCN-CAMS) version 3.01 (Fan and van den Dool, 2008), available from NOAA's Physical Sciences Laboratory (https://psl.noaa.gov/data/gridded/data.ghcncams.html). Runoff is evaluated against Global Runoff Reconstruction version 1 (GRUN v1) (Ghiggi et al., 2019b; https://doi.org/10.6084/m9.figshare.9228176). Snow water equivalent (SWE) is evaluated using the monthly CanSISE dataset distributed through the International Land Model Benchmarking (ILAMB) reference-data collection (https://www.ilamb.org/ILAMB-Data/DATA/swe/CanSISE/swe.nc). The underlying CanSISE observation-based SWE ensemble is archived at the National Snow and Ice Data Center (Mudryk and Derksen, 2017;

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

https://doi.org/10.5067/96ltniikJ7vd). The E3SM/AMIP simulation outputs used to train and evaluate the latent data assimilation framework, preprocessed datasets, trained model weights, and scripts needed to reproduce the latent DA experiments can be found in Fang et al. (2026a, 2026b), which are available through Zenodo at https://doi.org/10.5281/zenodo.22680946 and https://doi.org/10.5281/zenodo.22681153.

Author contributions

YF conceptualized the structure of the paper and led the writing of the manuscript. Funding acquisition: LRL. Methodology: YF, GB, LL, SZ, and LRL. Coding and implementation: YF and HT. Writing, review, and editing: All authors.

Competing interests

The authors declare there are no conflicts of interest for this manuscript.

Acknowledgments

This research was supported by the U.S. Department of Energy Office of Science, Biological and Environmental Research, as part of the Regional and Global Model Analysis program area through the Water Cycle: Modeling of Circulation, Convection, and Earth System Mechanisms (WACCEM) Scientific Focus Area. This research was also partially supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research Earth System Model Development (ESMD) program area. Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under Contract DE-AC05-76RL01830. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported under Contract DE-AC02-05CH11231. AI tools were used for language editing, including improvements to grammar and wording, and refining and formatting code.

Financial support

This research was supported by the U.S. Department of Energy Office of Science, Biological and Environmental Research, as part of the Regional and Global Model Analysis program area through the Water Cycle: Modeling of Circulation, Convection, and Earth System Mechanisms (WACCEM) Scientific Focus Area. This research was also partially supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research Earth System Model Development (ESMD) program area.

References:

Abbaszadeh, P., Moradkhani, H., and Yan, H.: Enhancing hydrologic data assimilation by evolutionary particle filter and Markov chain Monte Carlo, Adv. Water Resour., 111, 192–204, https://doi.org/10.1016/j.advwatres.2017.11.011, 2018.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Anderson, J. L.: Localization and sampling error correction in ensemble Kalman filter data assimilation, Mon. Weather Rev., 140, 2359–2371, https://doi.org/10.1175/MWR-D-11-00013.1, 2012.

Anderson, J. L., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., and Arellano, A.: The Data Assimilation Research Testbed: A community facility, B. Am. Meteorol. Soc., 90, 1283–1296, https://doi.org/10.1175/2009BAMS2618.1, 2009.

Ardilouze, C. and Boone, A. A.: Impact of initializing the soil with a thermally and hydrologically balanced state on subseasonal predictability, Clim. Dynam., 62, 2629–2644, https://doi.org/10.1007/s00382-023-07024-x, 2024.

Bardes, A., Ponce, J., and LeCun, Y.: VICReg: Variance-invariance-covariance regularization for self-supervised learning, in: International Conference on Learning Representations (ICLR 2022), online, 25–29 April 2022, https://openreview.net/forum?id=xm6YD62D1Ub, 2022.

Beaudoing, H. and Rodell, M.: GLDAS Noah Land Surface Model L4 monthly 0.25 × 0.25 degree V2.0, NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) [data set], https://doi.org/10.5067/9SQ1B3ZXP2C5, 2020.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial trend analysis from 1901–present, Earth Syst. Sci. Data, 5, 71–99, https://doi.org/10.5194/essd-5-71-2013, 2013.

Behnel, S., Bradshaw, R., Citro, C., Dalcin, L., Seljebotn, D. S., and Smith, K.: Cython: The best of both worlds, Comput. Sci. Eng., 13, 31–39, https://doi.org/10.1109/MCSE.2010.118, 2011.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljević, D., Hamrud, M., Hollingsworth, A., Rabier, F., and Fisher, M.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation, Q. J. Roy. Meteor. Soc., 124, 1783–1807, https://doi.org/10.1002/qj.49712455002, 1998.

Crow, W. T., Kim, H., and Kumar, S. V.: Systematic modeling errors undermine the application of land data assimilation systems for hydrological and weather forecasting, J. Hydrometeorol., 25, 3–26, https://doi.org/10.1175/JHM-D-23-0069.1, 2024.

Duan, Y., Kumar, S., Maruf, M., Kavoo, T. M., Rangwala, I., Richter, J. H., Glanville, A. A., King, T., Esit, M., Raczka, B., and Raeder, K.: Enhancing sub-seasonal soil moisture forecasts through land initialization, npj Clim. Atmos. Sci., 8, 100, https://doi.org/10.1038/s41612-025-00987-0, 2025.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean Dynam., 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Evensen, G.: Data assimilation: The ensemble Kalman filter, 2nd edn., Springer, https://doi.org/10.1007/978-3-642-03711-5, 2009.

Fan, H., Bai, L., Fei, B., Xiao, Y., Chen, K., Liu, Y., Qu, Y., Ling, F., and Gentine, P.: Physically consistent global atmospheric data assimilation with machine learning in latent space, arXiv [preprint], https://doi.org/10.48550/arXiv.2502.02884, 8 July 2025a.

Fan, H., Liu, Y., Huo, Z., Liu, Y., Shi, Y., and Li, Y.: A novel latent space data assimilation framework with autoencoder–observation to latent space (AE–O2L) network. Part I: The observation-only analysis method, Mon. Weather Rev., 153, 1335–1348, https://doi.org/10.1175/MWR-D-24-0057.1, 2025b.

Fan, H., Nathaniel, J., Xiao, Y., Bian, C., Ling, F., Fei, B., Bai, L., and Gentine, P.: Accurate and efficient hybrid-ensemble atmospheric data assimilation in latent space with uncertainty quantification, arXiv [preprint], https://doi.org/10.48550/arXiv.2603.04395, 4 March 2026.

Fan, Y. and van den Dool, H.: A global monthly land surface air temperature analysis for 1948–present, J. Geophys. Res.-Atmos., 113, D01103, https://doi.org/10.1029/2007JD008470, 2008.

Fang, Y., Bisht, G., Li, L., Tran, H., Zhang, S., and Leung, L. R.: Latent space data assimilation in ELM, Zenodo [code], https://doi.org/10.5281/zenodo.22680946, 2026a.

Fang, Y., Bisht, G., Li, L., Tran, H., Zhang, S., and Leung, L. R.: Dataset for latent space data assimilation in ELM, Zenodo [data set], https://doi.org/10.5281/zenodo.22681153, 2026b.

Fearnhead, P. and Künsch, H. R.: Particle filters and data assimilation, Annu. Rev. Stat. Appl., 5, 421–449, https://doi.org/10.1146/annurev-statistics-031017-100232, 2018.

Ghannam, K., Nakai, T., Paschalis, A., Oishi, A. C., Kotani, A., Igarashi, Y., Kumagai, T., and Katul, G. G.: Persistence and memory timescales in root-zone soil moisture dynamics, Water Resour. Res., 52, 1427–1445, https://doi.org/10.1002/2015WR017983, 2016.

Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: An observation-based global gridded runoff dataset from 1902 to 2014, Earth Syst. Sci. Data, 11, 1655–1674, https://doi.org/10.5194/essd-11-1655-2019, 2019a.

Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: Global Runoff Reconstruction (Version 1, GSWP3 forcing), Figshare [data set], https://doi.org/10.6084/m9.figshare.9228176, 2019b.

Gruber, A., Su, C.-H., Zwieback, S., Crow, W. T., Dorigo, W. A., and Wagner, W.: Recent advances in soil moisture triple collocation analysis, Int. J. Appl. Earth Obs. Geoinf., 45, 200–211, https://doi.org/10.1016/j.jag.2015.09.002, 2016.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

970 Gruber, A., Dorigo, W. A., Crow, W. T., and Wagner, W.: Triple collocation-based merging of
971 satellite soil moisture retrievals, IEEE T. Geosci. Remote, 55, 6780–6792,
972 https://doi.org/10.1109/TGRS.2017.2734070, 2017.

973 Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error
974 covariance estimates in an ensemble Kalman filter, Mon. Weather Rev., 129, 2776–2790,
975 https://doi.org/10.1175/1520-0493(2001)129%3C2776:DDFOBE%3E2.0.CO;2, 2001.

976 He, H., Zhang, F., Zhang, Q., and Chen, D.: An offline framework for high-dimensional
977 ensemble Kalman filters to reduce the time to solution, Geosci. Model Dev., 13, 3607–3625,
978 https://doi.org/10.5194/gmd-13-3607-2020, 2020.

979 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J.,
980 Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo,
981 G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D.,
982 Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L.,
983 Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C.,
984 Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5
985 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803,
986 2020.

987 Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz-Sabater, J., Nicolas, J.,
988 Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-
989 N.: ERA5 monthly averaged data on single levels from 1940 to present, Copernicus Climate
990 Change Service (C3S) Climate Data Store [data set], https://doi.org/10.24381/cds.f17050d7,
991 2023.

992 Hodyss, D. and Morzfeld, M.: Using diffusion models to do data assimilation, Mon. Weather
993 Rev., 154, 165–182, https://doi.org/10.1175/MWR-D-25-0125.1, 2026.

994 Huang, L., Gianinazzi, L., Yu, Y., Dueben, P. D., and Hoefler, T.: DiffDA: A diffusion model
995 for weather-scale data assimilation, in: Proceedings of the 41st International Conference on
996 Machine Learning, PMLR, 235, 19798–19815,
997 https://proceedings.mlr.press/v235/huang24h.html, 2024.

998 Kalnay, E.: Atmospheric modeling, data assimilation and predictability, Cambridge University
999 Press, https://doi.org/10.1017/CBO9780511802270, 2003.

1000 Kumar, S. V., Reichle, R. H., Peters-Lidard, C. D., Koster, R. D., Zhan, X., Crow, W. T.,
1001 Eylander, J. B., and Houser, P. R.: A land surface data assimilation framework using the Land
1002 Information System: Description and applications, Adv. Water Resour., 31, 1419–1432,
1003 https://doi.org/10.1016/j.advwatres.2008.01.013, 2008.

1004 Lei, L. and Whitaker, J. S.: Evaluating the trade-offs between ensemble size and ensemble
1005 resolution in an ensemble-variational data assimilation system, J. Adv. Model. Earth Syst., 9,
1006 781–789, https://doi.org/10.1002/2016MS000864, 2017.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Li, X., Liu, F., Ma, C., Hou, J., Zheng, D., Ma, H., Bai, Y., Han, X., Vereecken, H., Yang, K., Duan, Q., and Huang, C.: Land data assimilation: Harmonizing theory and data in land surface process studies, Rev. Geophys., 61, e2022RG000801, https://doi.org/10.1029/2022RG000801, 2023.

Manshausen, P., Cohen, Y., Pathak, J., Pritchard, M., Garg, P., Mardani, M., Kashinath, K., Byrne, S., and Brenowitz, N.: Generative data assimilation of sparse weather station observations at kilometer scales, J. Adv. Model. Earth Syst., 17, e2024MS004505, https://doi.org/10.1029/2024MS004505, 2025.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., and Stoffelen, A.: Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target, Geophys. Res. Lett., 41, 6229–6236, https://doi.org/10.1002/2014GL061322, 2014.

Mudryk, L. R. and Derksen, C.: CanSISE observation-based ensemble of Northern Hemisphere terrestrial snow water equivalent (Version 2; NSIDC-0668), NASA National Snow and Ice Data Center Distributed Active Archive Center [data set], https://doi.org/10.5067/96ltniikJ7vd, 2017.

Mudryk, L. R., Derksen, C., Kushner, P. J., and Brown, R.: Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010, J. Climate, 28, 8037–8051, https://doi.org/10.1175/JCLI-D-15-0229.1, 2015.

National Oceanic and Atmospheric Administration, Physical Sciences Laboratory: GHCN-CAMS gridded 2 m temperature (land), version 3.01, NOAA Physical Sciences Laboratory [data set], https://psl.noaa.gov/data/gridded/data.ghcncams.html (last access: 12 November 2026), n.d.

Pasmans, I., Chen, Y., Finn, T. S., Bocquet, M., and Carrassi, A.: Ensemble Kalman filter in latent space using a variational autoencoder pair, Q. J. Roy. Meteor. Soc., 152, e70070, https://doi.org/10.1002/qj.70070, 2026.

Peyron, M., Fillion, A., Gürol, S., Marchais, V., Gratton, S., Boudier, P., and Goret, G.: Latent space data assimilation by using deep learning, Q. J. Roy. Meteor. Soc., 147, 3759–3777, https://doi.org/10.1002/qj.4153, 2021.

Qiu, Y., Feng, J., Wang, J., Xue, Y., and Xu, Z.: Memory of land surface and subsurface temperature initial anomalies over the Tibetan Plateau in different land models, Clim. Dynam., 62, 2703–2718, https://doi.org/10.1007/s00382-021-05937-z, 2024.

Qu, Y., Blanke, M., Shamekh, S., and Gentine, P.: PnP-DA: Towards principled plug-and-play integration of variational data assimilation and generative models, arXiv [preprint], https://doi.org/10.48550/arXiv.2508.00325, 1 August 2025.

Rahmati, M., Amelung, W., Brogi, C., Dari, J., Flammini, A., Bogena, H., Brocca, L., Chen, H., Groh, J., Koster, R. D., McColl, K. A., Montzka, C., Moradi, S., Rahi, A., Sharghi, E., and

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.

Vereecken, H.: Soil moisture memory: State-of-the-art and the way forward, Rev. Geophys., 62, e2023RG000828, https://doi.org/10.1029/2023RG000828, 2024.

Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, Mon. Weather Rev., 130, 103–114, https://doi.org/10.1175/1520-0493(2002)130%3C0103:HDAWTE%3E2.0.CO;2, 2002.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, B. Am. Meteorol. Soc., 85, 381–394, https://doi.org/10.1175/BAMS-85-3-381, 2004.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Ziese, M.: GPCC Full Data Monthly Product Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data, Global Precipitation Climatology Centre, Deutscher Wetterdienst [data set], https://doi.org/10.5676/DWD_GPCC/FD_M_V7_050, 2015.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Sci. Rev., 99, 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004, 2010.

Shi, P., Leung, L. R., Wang, B., Zhang, K., Hagos, S. M., and Zhang, S.: The 4DEnVar-based weakly coupled land data assimilation system for E3SM version 2, Geosci. Model Dev., 17, 3025–3040, https://doi.org/10.5194/gmd-17-3025-2024, 2024.

Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, J. Geophys. Res.-Oceans, 103, 7755–7766, https://doi.org/10.1029/97JC03180, 1998.

van Leeuwen, P. J., Künsch, H. R., Nerger, L., Potthast, R., and Reich, S.: Particle filters for high-dimensional geoscience applications: A review, Q. J. Roy. Meteor. Soc., 145, 2335–2365, https://doi.org/10.1002/qj.3551, 2019.

Xie, S., Terai, C. R., Wang, H., Tang, Q., Fan, J., Burrows, S., Lin, W., Wu, M., Song, X., Zhang, Y., Taylor, M. A., Golaz, J.-C., Benedict, J. J., Chen, C.-C. J., Feng, Y., Hannah, W. M., Ke, Z., Shan, Y., Larson, V. E., Liu, X., Prather, M. J., Richter, J. H., Shrivastava, M., Wan, H., Zhang, G. J., Zhang, K., Bradley, A. M., Cameron-Smith, P., Damiano, L., Debusschere, B. J., Donahue, A. S., Easter, R. C., Eldred, M. S., Griffin, B. M., Guba, O., Guo, Z., Huang, X., Lee, J., Lee, H.-H., Lou, S., Mahfouz, N., Moncrieff, M., Mülmenstädt, J., Qian, Y., Rasool, Q. Z., Roberts, A. F., Santos, S. P., Sargsyan, K., Shpund, J., Singh, B., Tao, C., Xie, J., Yang, Y., Zeng, X., Zhang, C., Zhang, M., Zhang, S., Zhang, T., Zheng, X., Jacob, R. L., Leung, L. R., McCoy, R. B., and Bader, D. C.: The Energy Exascale Earth System Model version 3: 1. Overview of the atmospheric component, J. Adv. Model. Earth Syst., 17, e2025MS005120, https://doi.org/10.1029/2025MS005120, 2025.

https://doi.org/10.5194/egusphere-2026-5703
Preprint. Discussion started: 18 November 2026
© Author(s) 2026. CC BY 4.0 License.



1078 Yilmaz, M. T. and Crow, W. T.: The optimality of potential rescaling approaches in land data
1079 assimilation, J. Hydrometeorol., 14, 650–660, https://doi.org/10.1175/JHM-D-12-052.1, 2013.

1080 Zbontar, J., Jing, L., Misra, I., LeCun, Y., and Deny, S.: Barlow Twins: Self-supervised learning
1081 via redundancy reduction, in: Proceedings of the 38th International Conference on Machine
1082 Learning, PMLR, 139, 12310–12320, https://proceedings.mlr.press/v139/zbontar21a.html, 2021.

1083 Zhang, Y.-F., Hoar, T. J., Yang, Z.-L., Anderson, J. L., Toure, A. M., and Rodell, M.:
1084 Assimilation of MODIS snow cover through the Data Assimilation Research Testbed and the
1085 Community Land Model version 4, J. Geophys. Res.-Atmos., 119, 7091–7103,
1086 https://doi.org/10.1002/2013JD021329, 2014.