https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

Score-based Filtering for Land Data Assimilation

Zhilong Fan[1,2], Lu Li[1], Zhongwang Wei[1], Shupeng Zhang[1], Sisi Chen[1], Xingjie Lu[1], Wei Shangguan[1], Nan Wei[1], Hua Yuan[1], Pan Zhang[1,2], Jiaxin Tian[3,4], Jianhong Zhou[5], Baoxiang Pan[6], Wei Han[4], and Yongjiu Dai[1]

[1]School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, Guangdong, China
[2]School of Mathematics (Zhuhai), Sun Yat-sen University, Zhuhai, Guangdong, China
[3]State Key Laboratory of Severe Weather Meteorological Science and Technology (LaSW), Beijing, China
[4]CMA Earth System Modeling and Prediction Centre (CEMC), Beijing, China
[5]Chongqing Jinfo Mountain Karst Ecosystem National Observation and Research Station, School of Geographical Sciences, Southwest University, Chongqing, China
[6]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Correspondence: Lu Li (lilu83@mail.sysu.edu.cn) and Yongjiu Dai (daiyj6@mail.sysu.edu.cn)

Abstract. Land data assimilation (DA) improves estimates of land states by integrating observations with land surface model (LSM) simulations. Ensemble Kalman filters (EnKFs) are widely used but assume Gaussian forecast-error distributions, limiting their ability to represent non-Gaussian errors under nonlinear dynamics. Score-based filters (SFs) relax this assumption by using diffusion models to estimate the score function of non-Gaussian forecast distribution and assimilating observations through an approximate likelihood score that guide reverse diffusion process to generate analysis samples, providing a flexible non-Gaussian DA framework that has shown promise in idealized experiments. However, their application to land DA with complex LSMs remains limited by forecast-score estimation errors at computationally affordable ensemble sizes and empirically calibrated likelihood-score approximations that can be sensitive to noisy satellite retrievals. To address these limitations, we develop LandSF which uses stochastic differential equation editing for reverse-diffusion initialization with reduced sensitivity to forecast-score errors, and moment-matching posterior sampling to explicitly approximate likelihood-score without empirical calibration. We evaluate LandSF against the ensemble transform Kalman filter (ETKF) and a representative SF using a Kolmogorov-flow experiment and a realistic land DA experiment that assimilates SMAP soil moisture retrievals into the Common Land Model (CoLM). In the Kolmogorov-flow experiment, both SFs maintain low analysis errors as the forecast-error distribution evolves from Gaussian to non-Gaussian, whereas ETKF deteriorates rapidly. Compared with the benchmark SF, LandSF reduces the average RMSE by about 58% and substantially suppresses spatial error artifacts. These advantages carry over to realistic land DA, where LandSF reduces ubRMSE relative to the benchmark SF at 64% of the independent validation stations, with a mean reduction of about 2.28%. LandSF also consistently improves both ubRMSE and R over open-loop simulation across all three station networks, while the benchmark SF is less robust across networks. Diagnostic experiments further attribute these improvements to the two methodological advances, which better preserve forecast-distribution statistics and reduce sensitivity to observational noise. These results demonstrate that LandSF provides a practical framework for realistic land DA applications with complex process-based LSMs.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

1 Introduction

Land data assimilation (DA) integrates land surface model (LSM) forecasts with imperfect observations to refine estimates of land states and quantify their uncertainty, thereby improving simulations of land–atmosphere interactions (Houser et al., 1998; Rodell et al., 2004; Reichle et al., 2008b). Over the past decades, the Ensemble Kalman Filter (EnKF) and its variants (e.g., Ensemble Transform Kalman Filter, ETKF) have been widely used in global and regional land reanalysis systems (Reichle et al., 2008b, 2017; Kumar et al., 2022). A major advantage of the EnKF is that it uses a Monte Carlo ensemble of model forecasts to estimate and propagate flow-dependent forecast-error covariance (Reichle et al., 2002; Crow and Van Den Berg, 2010). However, non-Gaussian forecast error in complex LSMs violate the Gaussian assumptions underlying the linear analysis update of the EnKF, leading to suboptimal assimilation performance (Aksoy et al., 2009; Noh et al., 2011; Yan and Moradkhani, 2016). The particle Filters (PFs) relaxes the Gaussian assumption and uses sequential importance sampling to approximate non-Gaussian distributions with weighted particles (Abbaszadeh et al., 2020; Dasgupta et al., 2021; Dong et al., 2015). In high-dimensional systems, however, standard PFs struggle with the curse of dimensionality because avoiding weight degeneracy requires the number of particles to grow exponentially with the problem dimension (Farchi and Bocquet, 2018; Berg et al., 2019).

Score-based diffusion models offer an alternative way to learn and sample non-Gaussian distributions without relying on importance-weighted particles (Ho et al., 2020; Song et al., 2021). Diffusion models gradually perturb data with Gaussian noise through a forward diffusion process, generating noise-perturbed states whose distributions evolve over different noise levels. They learn the corresponding score functions—the gradients of the log probability densities—which then guide the reverse diffusion process to generate samples from the target distribution. This capability has led to a growing family of diffusion-based methods for non-Gaussian DA, with growing applications in atmospheric (Yang et al., 2025; Huang et al., 2024; Qu et al., 2024; Sun et al., 2025a, b; Andry et al., 2025; Manshausen et al., 2025), oceanic (Martin et al., 2025), and hydrologic systems (Foroumandi and Moradkhani, 2025).

A prominent class of these methods is score-based filters (SFs), which follows the recursive forecast–analysis cycle of conventional filters (Bao et al., 2024a, b; Ding et al., 2025). In each assimilation cycle, the previous analysis ensemble is propagated forward through the dynamical model, producing a forecast ensemble from which the forecast score is estimated. In the analysis step, the likelihood score is estimated from the observations and combined with the forecast score to form the analysis score. The analysis score then guides the reverse diffusion process, which transforms Gaussian noise into analysis samples. Numerical experiments have demonstrated the advantages of SFs over EnKFs and PFs in several idealized nonlinear systems, including a double-well system, a bearing-tracking problem, and the Lorenz-96 model (Bao et al., 2024a; Ding et al., 2025). Subsequent studies have advanced SFs through training-free score estimation (Bao et al., 2024b), latent-space representations (Si and Chen, 2024), iterative analysis score correction (Zhang et al., 2025) and sequential Langevin sampling (Ding et al., 2025). SFs have since been evaluated in intermediate-complexity systems, including Kolmogorov flow (Ding et al., 2025), shallow-water dynamics (Si and Chen, 2024), surface quasigeostrophic dynamics (Bao et al., 2024b; Xiong et al., 2026), and two-phase flow in porous media (Hu et al., 2025).

Despite these advances, applying existing SFs to complex process-based LSMs faces two major challenges. First, the computational cost of LSM integrations restricts the ensemble size available for forecast-score estimation, whereas idealized systems generally permit much larger ensembles (McLaughlin et al., 2006; Chen et al., 2023; Li et al., 2024). Existing SFs

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

generate analysis samples from Gaussian noise through reverse diffusion guided by the analysis score, which incorporates the estimated forecast score. Errors in the forecast-score estimate can therefore bias the analysis ensemble, degrading its representation of flow-dependent uncertainty (Finn et al., 2024) and compromising its physical consistency (Wang et al., 2025). Second, assimilating observations into SFs requires the likelihood score corresponding to the noise-perturbed state at each step of reverse diffusion, which is generally unavailable in closed form (Rozet et al., 2024; Chung et al., 2023). Existing SFs address this intractability by approximating this likelihood score with a prescribed damping function (Bao et al., 2024a, b). Such a heuristic approximation controls the contribution of observational information to the analysis samples, but its optimal form remains unknown (Bao et al., 2024a). This limitation is particularly important for practical land DA, which often assimilates satellite retrievals with imperfectly characterized errors (Gruber et al., 2020; Kumar et al., 2022), whereas idealized experiments typically generate synthetic observations using prescribed observation-error distributions (Bao et al., 2024a). The prescribed weighting does not explicitly represent the joint effects of observation errors and uncertainty associated with the noise-perturbed state, and can therefore lead the analysis to overfit noisy observations, reducing robustness in land DA applications.

To facilitate the practical application of score-based filtering to complex process-based LSMs, we develop LandSF, a score-based filtering framework for land DA that addressing these limitations. LandSF retains the forecast-analysis cycle of existing SFs but improves two components of the analysis step (Section 2.2). First, it employs stochastic differential equation editing (SDEdit; Meng et al., 2022), using partially noise-perturbed forecast members instead of Gaussian noise to initialize reverse diffusion process, thereby reducing sensitivity to forecast-score estimation errors while preserving flow-dependent information. Second, it uses moment-matching posterior sampling (MMPS; Rozet et al., 2024) to approximate the diffusion-time likelihood score at each reverse diffusion step while explicitly accounting for observation errors and noise-perturbed state uncertainty, without requiring an empirical calibrated approximation. We evaluate LandSF against ETKF, a widely used EnKF variant in land DA, and a representative SF proposed by Bao et al. (2024a; hereafter B24). We first establish the theoretical connection between LandSF and the EnKF by showing that LandSF recovers the Kalman analysis under linear–Gaussian assumptions (Section 2.2). We then use a Kolmogorov-flow experiment to evaluate its performance relative to ETKF and B24 under non-Gaussian conditions (Section 4.1). Finally, we couple LandSF with the Common Land Model (CoLM) and conduct a realistic land DA experiment assimilating SMAP surface soil moisture retrievals into CoLM (Sections 4.2). Across both experiments, we separately assess the two methodological advances and show that LandSF is less sensitive than B24 to finite-ensemble forecast-score estimation errors and observation noise and is therefore better suited for DA in complex dynamical systems. To our knowledge, this is the first study to develop and explore the benefits of SFs for practical land DA with a complex process-based LSM.

2 Method

2.1 General formulation of score-based filters (SFs)

Bayesian filtering recursively combines model forecasts with sequential observations to estimate the evolving system state. At each assimilation time k, the analysis distribution follows Bayes' rule:

$$p(\boldsymbol{x}^k|\boldsymbol{y}^{[k]}) \propto p(\boldsymbol{x}^k|\boldsymbol{y}^{[k-1]})p(\boldsymbol{y}^k|\boldsymbol{x}^k), \quad (1)$$

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

where $\boldsymbol{x}^k$ denotes the model state, $\boldsymbol{y}^k$ is the observation at time k, and $\boldsymbol{y}^{[k]}$ denotes all observations available up to time k. The terms $p(\boldsymbol{x}^k|\boldsymbol{y}^{[k-1]}), p(\boldsymbol{y}^k|\boldsymbol{x}^k), p(\boldsymbol{x}^k|\boldsymbol{y}^{[k]})$ denote the forecast distribution, observation likelihood, and analysis distribution, respectively. This Bayesian update is implemented recursively through forecast and analysis steps. In the forecast step, the previous analysis distribution ($p(\boldsymbol{x}^k|\boldsymbol{y}^{[k-1]})$) is propagated through the dynamical model according to the Chapman–Kolmogorov equation:

$$p(\boldsymbol{x}^k|\boldsymbol{y}^{[k-1]}) = \int p(\boldsymbol{x}^{k-1}|\boldsymbol{y}^{[k-1]})p(\boldsymbol{x}^k|\boldsymbol{x}^{k-1})d\boldsymbol{x}^{k-1}, \tag{2}$$

where the transition density $p(\boldsymbol{x}^k|\boldsymbol{x}^{k-1})$ is determined by the dynamical model ($\boldsymbol{f}_{k-1}$) and its representation of model uncertainty ($\boldsymbol{\varepsilon}_m^k$):

$$\boldsymbol{x}^k = \boldsymbol{f}_{k-1}(\boldsymbol{x}^{k-1}, \boldsymbol{\varepsilon}_m^k). \tag{3}$$

In the analysis step, the forecast distribution is updated using the current observation according to Bayes' rule (Equation (1)). The likelihood $p(\boldsymbol{y}^k|\boldsymbol{x}^k)$ is defined by the observation operator ($\boldsymbol{\mathcal{H}}_k$) and an additive observation-error model:

$$\boldsymbol{y}^k = \boldsymbol{\mathcal{H}}_k(\boldsymbol{x}^k) + \boldsymbol{\varepsilon}_o^k. \tag{4}$$

where $\boldsymbol{\varepsilon}_o^k$ denotes the observation error and follows a prescribed distribution, typically a Gaussian distribution.

In practice, the forecast and analysis distributions generally cannot be evaluated analytically and are therefore approximated using finite Monte Carlo ensembles. In the forecast step, N-member analysis ensemble at previous time $\{\boldsymbol{x}_a^{k-1,(i)}\}_{i=1}^N$ is used to represent the analysis distribution and each analysis member is propagated through the dynamical model to generate the forecast ensemble $\{\boldsymbol{x}_f^{k,(i)}\}_{i=1}^N$. The resulting ensemble provides a Monte Carlo approximation of the forecast distribution in Equation (2). In the analysis step, the observation is incorporated through a method-specific approximation of the Bayesian update in Equation (1), producing the analysis ensemble $\{\boldsymbol{x}_a^{k,(i)}\}_{i=1}^N$ that represents the analysis distribution and initializes the next assimilation cycle.

Most ensemble filtering methods (e.g., EnKF) approximate the forecast distribution ($p(\boldsymbol{x}^k|\boldsymbol{y}^{[k-1]})$) as Gaussian and estimate its mean and covariance from forecast ensemble. In contrast, SFs represent this distribution through its score function, which can be estimated from ensemble samples without imposing a Gaussian or other prescribed parametric form. Taking the logarithm of Equation (1) and differentiating with respect to $\boldsymbol{x}^k$ yields its score-based form:

$$\nabla_{\boldsymbol{x}^k} \log p(\boldsymbol{x}^k|\boldsymbol{y}^{[k]}) = \nabla_{\boldsymbol{x}^k} \log p(\boldsymbol{x}^k|\boldsymbol{y}^{[k-1]}) + \nabla_{\boldsymbol{x}^k} \log p(\boldsymbol{y}^k|\boldsymbol{x}^k), \tag{5}$$

where the three terms are the analysis, forecast and likelihood scores, respectively. SFs use diffusion models to estimate these scores and generate analysis samples. Diffusion models introduce a pseudo diffusion time ($t \in \{0, \cdots, T\}$), distinct from the physical assimilation time k, which parameterizes a continuous perturbation process that transform the target distribution (at $t = 0$) toward a standard Gaussian distribution (at $t = T$). Accordingly, Equation (5) is extended across diffusion time as:

$$\nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k]}) = \nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k-1]}) + \nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{y}^k|\boldsymbol{x}_t^k), \tag{6}$$

where $\boldsymbol{x}_t^k$ is the noise-perturbed model state at diffusion time t, and p_t denotes its probability density.

During the forecast stage, the forecast ensemble is progressively perturbed with Gaussian noise through the forward diffusion process, and denoising score matching is then used to estimate the forecast score across different noise levels

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

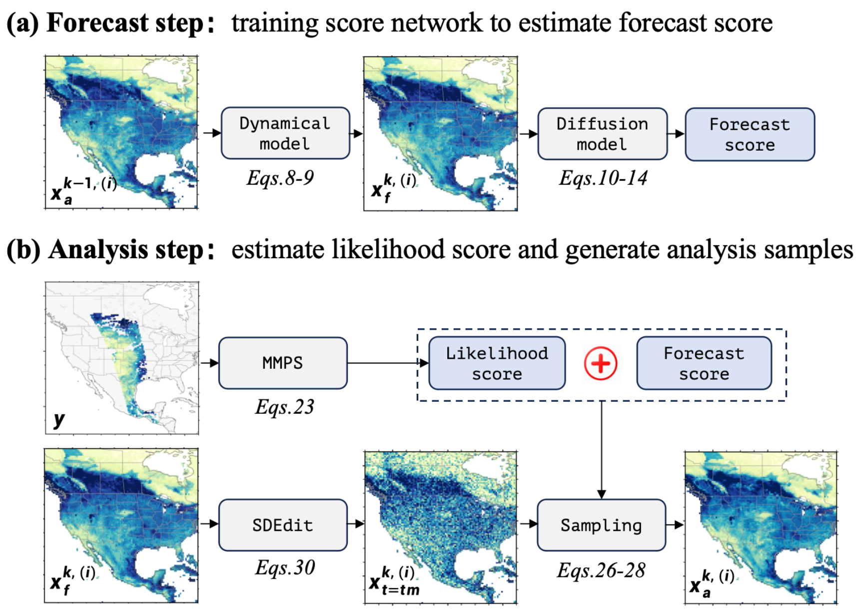

Figure 1. Schematic of the LandSF forecast–analysis cycle. (a) In the forecast step, the previous analysis ensemble is propagated through the dynamical model to generate the forecast ensemble, which is used to train the diffusion model and estimate the forecast score. (b) In the analysis step, SDEdit initializes reverse diffusion from partially perturbed forecast members, while MMPS approximates the likelihood score from the observations. The likelihood and forecast scores are combined to guide reverse diffusion and generate the analysis ensemble. Superscripts k and i denote the assimilation time and ensemble member, respectively.

$(\nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k-1]}))$ (Section 2.2.1). In the analysis step, observations are incorporated at each reverse diffusion step through the likelihood score corresponding to the current noise-perturbed state. We therefore require the likelihood $p_t(\boldsymbol{y}^k|\boldsymbol{x}_t^k)$, hereafter refereed to the noise-perturbed likelihood, which can be written by marginalizing over the possible unperturbed states:

$$p_t(\boldsymbol{y}^k|\boldsymbol{x}_t^k) = \int p(\boldsymbol{y}^k|\boldsymbol{x}^k) p(\boldsymbol{x}^k|\boldsymbol{x}_t^k) d\boldsymbol{x}^k. \tag{7}$$

Importantly, the noise-perturbed likelihood differs from the observation likelihood $p(\boldsymbol{y}^k|\boldsymbol{x}^k)$ in Equations (4), which is defined for the unperturbed state. Evaluating the noise-perturbed likelihood requires the conditional distribution $p(\boldsymbol{x}^k|\boldsymbol{x}_t^k)$, which describes the possible unperturbed states corresponding to a given noise-perturbed state. Because this conditional distribution is generally intractable, the corresponding noise-perturbed likelihood score must be approximated (Section 2.2.2). The forecast and likelihood score are then combined according to Equation (6) to form the analysis score, which guides the reverse diffusion process to generate the analysis ensemble (Section 2.2.2).

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

2.2 The proposed framework for land data assimilation

We developed a score-based filter, namely LandSF, for land DA in process-based LSMs that follows the recursive forecast-analysis cycle (Figure 1). The forecast step uses a diffusion model to estimate the forecast score from the model-propagated ensemble (Figure 1a), while the analysis step introduces two methodological modifications to improve robustness under the finite ensembles and noisy observations encountered in practical land DA (Figure 1b and S1b). First, we use SDEdit to initialize reverse diffusion from partially noise-perturbed forecast members rather than Gaussian noise. This forecast-informed initialization retains member-specific information from the model-propagated ensemble and reduces the dependence of reverse diffusion sampling on an imperfectly estimated forecast score. Second, we use MMPS to approximate the diffusion-time likelihood score by explicitly accounts for both the observation-error covariance, and the uncertainty of the noise-perturbed state, avoiding the prescribed damping function used in existing SFs.

2.2.1 Forecast step: estimate forecast score using diffusion model

At assimilation time k, the N-member forecast ensemble, $\{\boldsymbol{x}_f^{k,(i)}\}_{i=1}^N$, provides a Monte Carlo approximation of the forecast distribution $p(\boldsymbol{x}^k|\boldsymbol{y}^{[k-1]})$. To estimate its score across different noise levels, a forward diffusion process maps the unperturbed state $\boldsymbol{x}^k$ to a noise-perturbed state $\boldsymbol{x}_t^k$ through the transition kernel $q_t(\boldsymbol{x}_t^k|\boldsymbol{x}^k)$. The corresponding noise-perturbed forecast distribution is:

$$p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k-1]}) = \int q_t(\boldsymbol{x}_t^k|\boldsymbol{x}^k)p(\boldsymbol{x}^k|\boldsymbol{y}^{[k-1]})d\boldsymbol{x}^k. \tag{8}$$

Using the forecast samples, this integral is approximated by:

$$p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k-1]}) \approx \frac{1}{N}\sum_{i=1}^{N} q_t(\boldsymbol{x}_t^k|\boldsymbol{x}_f^{k,(i)}). \tag{9}$$

We adopt a variance-preserving forward diffusion process (Song et al., 2020), whose transition kernel is Gaussian:

$$q_t(\boldsymbol{x}_t^k|\boldsymbol{x}_f^{k,(i)}) = \mathcal{N}\left(\boldsymbol{x}_t^k; \sqrt{\overline{\alpha}(t)}\boldsymbol{x}_f^{k,(i)}, [1-\overline{\alpha}(t)]\boldsymbol{I}\right). \tag{10}$$

Accordingly, a noise-perturbed forecast member can be generated directly as:

$$\boldsymbol{x}_t^{k,(i)} = \sqrt{\overline{\alpha}(t)}\boldsymbol{x}_f^{k,(i)} + \sqrt{1-\overline{\alpha}(t)}\boldsymbol{\epsilon}^{(i)}, \quad \boldsymbol{\epsilon}^{(i)} \sim \mathcal{N}(0,\boldsymbol{I}). \tag{11}$$

Here, $\boldsymbol{x}_t^{k,(i)}$ denotes the perturbed version of $\boldsymbol{x}_f^{k,(i)}$ at diffusion time t, $\mathcal{N}$ denotes a Gaussian distribution and $\boldsymbol{I}$ is the identity matrix. The prescribed noise schedule $\beta_t \in (0,1)$ controls the noise introduced at each diffusion step, with $\alpha_t = 1-\beta_t, \quad \overline{\alpha}_t = \prod_{j=1}^t \alpha_j$. The diffusion step $t \in \{0,\ldots,T\}$, where T is the total number of diffusion steps. At $t=0$, $\overline{\alpha}_0 = 1$, and the noise schedule is selected such that $\overline{\alpha}_T$ approaches zero. Consequently, $\boldsymbol{x}_T^k$ approximately follows a standard Gaussian distribution (i.e., $p_T(\boldsymbol{x}_T^k|\boldsymbol{y}^{[k-1]}) \approx \mathcal{N}(\boldsymbol{x}_T^k;0,\boldsymbol{I})$).

Our objective is to estimate the score of the noise-perturbed forecast distribution ($\nabla_{\boldsymbol{x}_t^k}\log p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k-1]})$). Although this marginal score is unavailable analytically, the denoising score matching identity expresses it as the conditional expectation of the score of the known forward transition kernel (Vincent, 2011; Song et al., 2021):

$$\nabla_{\boldsymbol{x}_t^k}\log p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k-1]}) = \mathbb{E}\left[\nabla_{\boldsymbol{x}_t^k}\log q_t(\boldsymbol{x}_t^k|\boldsymbol{x}^k) \mid \boldsymbol{x}_t^k, \boldsymbol{y}^{[k-1]}\right]. \tag{12}$$

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

For the Gaussian forward transition kernel used in Equation (10), the conditional score is available analytically

$$\nabla_{\boldsymbol{x}_t^k} \log q_t(\boldsymbol{x}_t^k | \boldsymbol{x}_f^{k,(i)}) = -\frac{\boldsymbol{x}_t^k - \sqrt{\overline{\alpha}(t)}\boldsymbol{x}_f^{k,(i)}}{1-\overline{\alpha}(t)}. \tag{13}$$

We therefore parameterize the forecast score ($\nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{x}_t^k | \boldsymbol{y}^{[k-1]})$) using a neural network $s_{\theta_k}(\boldsymbol{x}_t^{k,(i)}, t)$. At each training iteration, we construct the noise-perturbed state $\boldsymbol{x}_t^{k,(i)}$ according to Equation (11), and the neural network is then optimized by minimizing the denoising score matching loss:

$$\mathcal{J}_{\text{DSM}}(\theta_k) = \mathbb{E}_{i,t,\boldsymbol{\epsilon}}\left[\left\| s_{\theta_k}(\boldsymbol{x}_t^{k,(i)}, t) + \frac{\boldsymbol{x}_t^{k,(i)} - \sqrt{\overline{\alpha}(t)}\boldsymbol{x}_f^{k,(i)}}{1-\overline{\alpha}(t)} \right\|_2^2\right]. \tag{14}$$

Training across diffusion steps yields the following neural approximation of the noise-perturbed forecast score $s_{\theta_k}(\boldsymbol{x}_t^k, t) \approx \nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{x}_t^k | \boldsymbol{y}^{[k-1]})$.

2.2.2 Analysis step: estimate likelihood score and analysis sampling

Given the forecast scores estimated in Section 2.2.1, constructing the analysis scores additionally requires the likelihood scores $\nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{y}^k | \boldsymbol{x}_t^k)$ at each reverse diffusion step, which depends on the generally intractable conditional density $p(\boldsymbol{x}^k | \boldsymbol{x}_t^k)$ (Equation (7)). Existing SFs address this difficulty by scaling the likelihood score defined for the unperturbed state using a prescribed damping function (Bao et al., 2024b, a; Zhang et al., 2025):

$$\nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{y}^k | \boldsymbol{x}_t^k) \approx \lambda_t \nabla_{\boldsymbol{x}_t^k} \log p(\boldsymbol{y}^k | \boldsymbol{x}^k), \tag{15}$$

where λ_t controls the contribution of observational information at diffusion step t. Although this approximation incorporates the prescribed observation-error covariance through $\nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{y}^k | \boldsymbol{x}_t^k)$, it does not explicitly capture the uncertainty associated with recovering the unperturbed state $\boldsymbol{x}^k$ from the noise-perturbed state $\boldsymbol{x}_t^k$.

LandSF instead employs moment-matching posterior sampling (MMPS; Rozet et al., 2024), which approximates the conditional distribution by a Gaussian distribution that matches its first two moments:

$$p(\boldsymbol{x}^k | \boldsymbol{x}_t^k) \approx \mathcal{N}\left(\boldsymbol{x}^k; \boldsymbol{\mu}_t^k, \boldsymbol{C}_t^k\right), \tag{16}$$

Under the Gaussian forward process, the two conditional moments are obtained from the forecast score and its derivative using the generalized Tweedie formulas (Efron, 2011):

$$\boldsymbol{\mu}_t^k = \mathbb{E}\left[\boldsymbol{x}^k \mid \boldsymbol{x}_t^k\right] = \frac{\boldsymbol{x}_t^k + (1-\overline{\alpha}_t)\nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{x}_t^k | \boldsymbol{y}^{[k-1]})}{\sqrt{\overline{\alpha}_t}}, \tag{17}$$

$$\boldsymbol{C}_t^k = \mathbb{V}\left[\boldsymbol{x}^k \mid \boldsymbol{x}_t^k\right] = \frac{1-\overline{\alpha}_t}{\overline{\alpha}_t}\left[\boldsymbol{I} + (1-\overline{\alpha}_t)\nabla^2_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{x}_t^k | \boldsymbol{y}^{[k-1]})\right], \tag{18}$$

We assume the additive Gaussian observation error model defined in Equation (4) substituting the moment matched Gaussian approximation in Equation (16) conditional moments into Equation (7) gives:

$$p_t(\boldsymbol{y}^k | \boldsymbol{x}_t^k) \approx \int \mathcal{N}\left(\boldsymbol{y}^k; \boldsymbol{\mathcal{H}}_k(\boldsymbol{x}^k), \boldsymbol{R}_k\right) \mathcal{N}\left(\boldsymbol{x}^k; \boldsymbol{\mu}_t^k, \boldsymbol{C}_t^k\right) d\boldsymbol{x}^k, \tag{19}$$

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

For a nonlinear observation operator, we apply a first-order linearization around the conditional mean:

$$\boldsymbol{\mathcal{H}}_k(\boldsymbol{x}^k) \approx \boldsymbol{\mathcal{H}}_k(\boldsymbol{\mu}_t^k) + \boldsymbol{H}_t^k(\boldsymbol{x}^k - \boldsymbol{\mu}_t^k), \tag{20}$$

where $\boldsymbol{H}_t^k = \nabla_{\boldsymbol{x}^k}\boldsymbol{\mathcal{H}}_k\big|_{\boldsymbol{x}^k=\boldsymbol{\mu}_t^k}$ is the local Jacobian of $\boldsymbol{\mathcal{H}}_k$. Under this linearization, Equation (19) has the Gaussian approximation:

$$p_t(\boldsymbol{y}^k|\boldsymbol{x}_t^k) \approx \mathcal{N}\left(\boldsymbol{y}^k; \boldsymbol{\mathcal{H}}_k(\boldsymbol{\mu}_t^k), \boldsymbol{S}_t^k\right), \tag{21}$$

with $\boldsymbol{S}_t^k \approx \boldsymbol{R}_k + \boldsymbol{H}_t^k \boldsymbol{C}_t^k (\boldsymbol{H}_t^k)^\mathsf{T}$. Thus, the corresponding log-likelihood is:

$$\log q_t(\boldsymbol{y}^k|\boldsymbol{x}_t^k) = -\frac{1}{2}\left[\boldsymbol{y}^k - \boldsymbol{\mathcal{H}}_k(\boldsymbol{\mu}_t^k)\right]^\mathsf{T} (\boldsymbol{S}_t^k)^{-1}\left[\boldsymbol{y}^k - \boldsymbol{\mathcal{H}}_k(\boldsymbol{\mu}_t^k)\right] - \frac{1}{2}\log|\boldsymbol{S}_t^k| + const, \tag{22}$$

where both $\boldsymbol{\mu}_t^k$ and $\boldsymbol{S}_t^k$ depend on the perturbed state $\boldsymbol{x}_t^k$. To obtain a tractable approximation, we treat $\boldsymbol{S}_t^k$ as locally constant with respect to $\boldsymbol{x}_t^k$. Applying the chain rule gives:

$$\nabla_{\boldsymbol{x}_t^k} \log q_t(\boldsymbol{y}^k|\boldsymbol{x}_t^k) \approx \left(\nabla_{\boldsymbol{x}_t^k}\boldsymbol{\mu}_t^k\right)^\mathsf{T} (\boldsymbol{H}_t^k)^\mathsf{T} (\boldsymbol{S}_t^k)^{-1}\left[\boldsymbol{y}^k - \boldsymbol{\mathcal{H}}_k(\boldsymbol{\mu}_t^k)\right], \tag{23}$$

Rather than explicitly inverting $\boldsymbol{S}_t^k$, we solve the associated symmetric positive-definite linear system using the generalized minimal residual (GMRES) method (Saad and Schultz, 1986):

$$\boldsymbol{S}_t^k \boldsymbol{u}_t^k = \boldsymbol{y}^k - \boldsymbol{\mathcal{H}}_k(\boldsymbol{\mu}_t^k), \tag{24}$$

The likelihood score is subsequently evaluated as:

$$\nabla_{\boldsymbol{x}_t^k} \log p_t(\boldsymbol{y}^k|\boldsymbol{x}_t^k) \approx \left(\nabla_{\boldsymbol{x}_t^k}\boldsymbol{\mu}_t^k\right)^\mathsf{T} (\boldsymbol{H}_t^k)^\mathsf{T} \boldsymbol{u}_t^k, \tag{25}$$

The vector-Jacobian products in Equation (25) are evaluated using automatic differentiation. Finally, the likelihood score is combined with the forecast score to obtain the analysis score at each diffusion step (Equation (6)). Pseudocode of MMPS method is shown in Table S1.

Once the analysis scores obtained across diffusion steps, the analysis ensemble can be generated by reverse diffusion process. The variance-preserving forward diffusion process (Equation (10)) can be expressed in continuous time as a stochastic differential equation (SDE):

$$d\boldsymbol{x}_t^k = -\frac{1}{2}\beta(t)\boldsymbol{x}_t^k dt + \sqrt{\beta(t)}dw_t, \tag{26}$$

where $\beta(t)$ is the continuous interpolation of the discrete noise schedule β_t, and w_t is a standard Wiener process. The corresponding reverse-time SDE based on forecast score is (Anderson, 1982):

$$d\boldsymbol{x}_t^k = \left[-\frac{1}{2}\beta(t)\boldsymbol{x}_t^k - \beta(t)\nabla_{\boldsymbol{x}_t^k}\log p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k-1]})\right]dt + \sqrt{\beta(t)}d\overline{w}_t, \tag{27}$$

where $\overline{w}_t$ is a Wiener process in reverse time. After assimilating the observation, the forecast score is updated to the analysis score and gives observation-conditioned reverse-time SDE:

$$d\boldsymbol{x}_t^k = \left[-\frac{1}{2}\beta(t)\boldsymbol{x}_t^k - \beta(t)\nabla_{\boldsymbol{x}_t^k}\log p_t(\boldsymbol{x}_t^k|\boldsymbol{y}^{[k]})\right]dt + \sqrt{\beta(t)}d\overline{w}_t, \tag{28}$$

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

Therefore, we could generate an analysis member by numerically integrating this SDE to $t = 0$.

Numerically solving the reverse diffusion process requires an initial state and a sampling scheme (Song et al., 2021). For efficient sampling, we employ the denoising diffusion implicit model (DDIM) (Song et al., 2022), which defines a non-Markovian process the same marginal distributions as the forward process. DDIM enables reverse sampling over a reduced sequence of diffusion steps, $T = t_J > t_{J-1} > \cdots > t_0 = 0$. The DDIM update is controlled by a stochasticity parameter $\eta \in [0,1]$, where $\eta = 0$ yields deterministic sampling and $\eta = 1$ recovers the variance of the corresponding diffusion process. The complete update and pseudocode are provided in Table S2. Existing SFs generally initialize the reverse process from random Gaussian noise at T, i.e. $\boldsymbol{x}_{t_J}^{k,(i)} = \boldsymbol{x}_T^{k,(i)} \sim \mathcal{N}(0, \boldsymbol{I})$. The analysis member ($\boldsymbol{x}_a^{k,(i)}$) is then generated through the DDIM sequence:

$$\boldsymbol{x}_T^{k,(i)} = \boldsymbol{x}_{t_J}^{k,(i)} \longrightarrow \boldsymbol{x}_{t_{J-1}}^{k,(i)} \longrightarrow \cdots \longrightarrow \boldsymbol{x}_0^{k,(i)} = \boldsymbol{x}_a^{k,(i)}. \tag{29}$$

However, the initial Gaussian samples are independent of the model-propagated forecast members, the forecast distribution must be recovered entirely through the estimated forecast score during reverse sampling. This strong dependence makes the generated analysis ensemble sensitive to forecast-score errors arising from a finite ensemble. LandSF reduces this sensitivity using an SDEdit-based initialization (Meng et al., 2022). We select an intermediate starting step t_m from the DDIM sampling sequence, with $1 \le m \le J$, and perturb each forecast member ($\boldsymbol{x}_f^{k,(i)}$) according to the forward diffusion process:

$$\boldsymbol{x}_{t_m}^{k,(i)} = \sqrt{\bar{\alpha}_{t_m}}\boldsymbol{x}_f^{k,(i)} + \sqrt{1 - \bar{\alpha}_{t_m}}\epsilon^{(i)}, \quad \epsilon^{(i)} \sim \mathcal{N}(0, \boldsymbol{I}), \tag{30}$$

The resulting partially noise-perturbed forecast member serves as the initial state for reverse sampling. The analysis member is then generated by applying the DDIM updates from t_m to 0 using the analysis score:

$$\boldsymbol{x}_f^{k,(i)} \longrightarrow \boldsymbol{x}_1^{k,(i)} \longrightarrow \cdots \longrightarrow \boldsymbol{x}_{t_m}^{k,(i)} \longrightarrow \boldsymbol{x}_{t_{m-1}}^{k,(i)} \longrightarrow \cdots \longrightarrow \boldsymbol{x}_0^{k,(i)} = \boldsymbol{x}_a^{k,(i)}, \tag{31}$$

This forecast-informed initialization preserves member-specific information from the model-propagated ensemble and reduces the dependence of analysis sampling on the estimated forecast score, thereby limiting the influence of forecast-score errors on the analysis ensemble.

2.2.3 Implementation details

We use a U-Net-based score network (Ronneberger et al., 2015; Ling et al., 2024) consisting of residual blocks, self-attention modules, and skip connections for multiscale feature extraction. The encoder progressively reduces the spatial resolution, and the decoder then integrates features from the corresponding encoder levels to reconstruct the score field at the original resolution (Figure S2). At each assimilation time, the network is fine-tuned on the forecast ensemble by minimizing the denoising score-matching loss in Equation (14). We use the AdamW optimizer (Loshchilov and Hutter, 2019), a batch size of 5, epoch size of 100 and a learning rate of 10^{-3} for training and 5×10^{-4} for fine-tuning. Before training, the forecast ensemble is standardized using its ensemble mean and standard deviation, and score estimation and analysis sampling are conducted in this standardized space, after which the generated analysis ensemble is transformed back to physical units. The forward diffusion process uses $T = 1{,}000$ steps and a linear noise schedule with β_t ranging from 10^{-4} to 2×10^{-2}. We use 100 DDIM sampling steps with $\eta = 1$ to introduce stochasticity during reverse sampling. The GMRES solution of Equation (24) is limited to 5 iterations, and SDEdit initializes reverse sampling at diffusion step 500.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

2.2.4 Relationship between LandSF and ensemble Kalman filtering

We show that LandSF recovers the ensemble Kalman analysis under a Gaussian forecast distribution and a linear–Gaussian observation model (Appendix A). The forward diffusion process preserves the Gaussian distribution, yielding closed-form expressions for the diffused forecast distribution and its score in Equations (A2)–(A3). Given its noise-perturbed counterpart, the unperturbed state is also conditionally Gaussian (Equation (A6)), yielding the analytical likelihood score in Equation (A5). Combining the forecast and likelihood scores and applying the Woodbury identity (Equation (A1)) yields the Gaussian analysis score in Equation (A7). The analysis mean, covariance, and Kalman gain in Equation (A8) coincide with those of the Kalman analysis. Finally, the closed-form solution of the reverse probability-flow ordinary differential equation (ODE) in Equation (A4) generates an ensemble member from the resulting Gaussian analysis distribution, as shown in Equations (A12)–(A13). This connection establishes LandSF as a non-Gaussian extension of ensemble Kalman filtering.

2.3 Benchmarking data assimilation methods

We compare LandSF with two representative filtering methods: ETKF, a Gaussian ensemble filtering method widely used in land DA systems (Bishop et al., 2001; Reichle et al., 2008a; Li et al., 2024), and B24 (Bao et al., 2024a), a widely adopted SF that serves as the primary benchmark for the formulation introduced in Section 2.1.

ETKF is a deterministic square-root EnKF that approximates the forecast distribution as Gaussian, with its mean and covariance estimated from the forecast ensemble. It applies the Kalman update to the ensemble mean and transforms the forecast anomalies to reproduce the corresponding analysis covariance. To mitigate covariance underestimation, we use the adaptive multiplicative inflation of Wang and Bishop (2003), which estimates a scalar inflation factor from observation innovations at each assimilation time.

B24 is selected as the representative SF benchmark because it provides the recursive forecast–analysis formulation on which LandSF is based (Figure S1). For a controlled comparison, B24 and LandSF use the same forecast-score estimation procedure, score-network architecture, training configuration, diffusion schedule, and sampling scheme. They differ only in the two analysis components described in Section 2.2. First, B24 initializes reverse diffusion from pure Gaussian noise (Equation (29)) rather than partially noise-perturbed forecast members (Equation (31)). Second, B24 approximates the likelihood score at each reverse diffusion step using the prescribed damping coefficient (Equation (15)). Notably, because this approximation can introduce systematic bias into the likelihood-score estimate, practical implementations combine the forecast score with the damped likelihood score and rescale the resulting analysis score when its root-mean-square magnitude exceeds a prescribed threshold γ (Bao et al., 2024a). This threshold is tuned separately for different experimental settings.

$$\boldsymbol{s}_t^k = \boldsymbol{g}_t^k \cdot \min\left(1,\ \frac{\gamma}{\left\|\boldsymbol{g}_t^k\right\|_2}\right), \tag{32}$$

where $\boldsymbol{g}_t^k$ denotes the combined analysis score. We adopt λ_t as $(T-t)/T$ following previous studies (Bao et al., 2024a, b) and optimize γ separately for each experimental setting. In contrast, LandSF does not require any tunable parameter, and the limitations associated with this case-dependent parameter are discussed in Section 5.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

3 Experimental design and materials

We evaluate LandSF using two complementary experiments designed to examine its performance under controlled and realistic conditions. First, a Kolmogorov-flow experiment with known true states enables controlled comparisons of LandSF, ETKF and B24 under non-Gaussian forecast errors (Section 3.1). Second, a representative land DA experiment assimilates SMAP Level-2 surface soil moisture retrievals into the Common Land Model (CoLM) to evaluate LandSF and B24 in a complex process-based LSM (Section 3.2).

3.1 Kolmogorov flow

Kolmogorov flow is a forced-dissipative fluid system governed by the two-dimensional incompressible Navier-Stokes equations:

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = -\nabla p + \frac{1}{Re}\nabla^2 \boldsymbol{u} + \boldsymbol{f}, \qquad \nabla \cdot \boldsymbol{u} = 0, \tag{33}$$

where $\boldsymbol{u} = (u, v)$ is the velocity field, p is the pressure, and Re is the Reynolds number. The flow is driven by a sinusoidal body forcing and dissipated by viscosity and linear drag:

$$\boldsymbol{f}(x,y) = f_0 \sin(k_f y)\boldsymbol{e}_x - \alpha \boldsymbol{u}, \tag{34}$$

where f_0 and k_f are the forcing amplitude and wavenumber, respectively, α is the linear-damping coefficient, and $\boldsymbol{e}_x$ is the unit vector in the x direction. Following established Kolmogorov-flow DA experiments (Rozet and Louppe, 2023a), the equations are integrated using JAX-CFD (Kochkov et al., 2021) on a doubly periodic domain $[0, 2\pi] \times [0, 2\pi]$, with $Re = 10^3$, $f_0 = 1$, $k_f = 4$, and $\alpha = 0.1$. The velocity field is discretized on a 64 × 64 grid, resulting in a state dimension of 64 × 64 × 2, and the model trajectory is sampled at a temporal resolution of $\Delta t = 0.2$. The reference trajectory is first spun up for 64 timesteps and then integrated for an additional 600 timesteps, with DA performed every 30 timesteps. The full velocity field is observed using an identity operator, with independent Gaussian observation errors of standard deviation $\sigma_{\mathrm{obs}} = 0.1$ (Rozet and Louppe, 2023a). The 100 initial ensemble members are generated by perturbing the true state with Gaussian noise of standard deviation $\sigma_{\mathrm{init}} = 0.2$. This initialization produces an approximately Gaussian forecast-error distribution at the beginning of the experiment. Subsequent nonlinear model propagation progressively introduces non-Gaussian structures, providing a controlled setting in which to compare ETKF, B24, and LandSF.

3.2 Land data assimilation with a complex land surface model

The Common Land Model (CoLM) (Dai et al., 2003) is a process-based LSM that integrates the strengths of the NCAR LSM (Bonan, 2005), the 1994 version of the Institute of Atmospheric Physics LSM (Yongjiu and Qingcun, 1997), and the Biosphere-Atmosphere Transfer Scheme (Dickinson et al., 1986). Previous studies have demonstrated its ability to simulate surface energy and water budgets, as well as carbon and water fluxes (Lievens et al., 2017; Li et al., 2024). We use the latest version of CoLM (CoLM202X) (Wei, 2025), driven by hourly ERA5 atmospheric forcing (Hersbach et al., 2020). The ancillary datasets used for the CoLM simulations include land cover type data (Friedl et al., 2010), soil hydraulic and thermal parameter data (Dai et al., 2019; Shangguan et al., 2014), and leaf area index data (Yuan et al., 2011).

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

Surface soil moisture retrievals are obtained from version 9 of the SMAP Level-2 passive soil moisture product (ONeill et al., 2023). The product retrieves volumetric soil moisture over approximately the upper 5 cm of the soil from SMAP L-band brightness temperatures and provides half-orbit swaths on the 36-km EASE-Grid 2.0 (ONeill et al., 2023). We use retrievals from both ascending and descending orbits and exclude those flagged for frozen soil, snow cover, excessive surface water, or dense vegetation (O'Neill et al., 2023).

CoLM is configured over 130.0°–75.0°W and 10.0°–65.0°N at a spatial resolution of 0.5° $\times$ 0.5°, resulting in 110 $\times$ 110 model grid, and is integrated at an hourly timestep (Seo et al., 2021). The complete rectangular domain is retained for score-network training, whereas the assimilation results are evaluated over the contiguous United States (CONUS). The model is spun up from 1 January 2014 to 31 March 2015, followed by the assimilation experiment from 1 April to 30 September 2015. This period is selected to focus primarily on unfrozen conditions and reduce data gaps associated with frozen-soil and snow-cover flags in the SMAP retrievals (Seo et al., 2021). To represent forecast errors, we generate a 100-member ensemble by perturbing both atmospheric forcing and soil hydraulic parameters following Ewerdwalbesloh et al. (2026). Atmospheric forcing uncertainty is represented by multiplicative perturbations to precipitation and downward shortwave radiation and additive perturbations to air temperature and downward longwave radiation. The perturbation ranges, spatial and temporal correlation scales, and cross-variable correlations follow previous studies (Reichle et al., 2019), as summarized in Table S3. Following the configuration of Ewerdwalbesloh et al. (2026), soil matric potential, porosity, the pore-size distribution parameter, and saturated hydraulic conductivity are also perturbed using texture-dependent standard deviations derived from the sand and clay fractions (Table S4). Soil moisture state uncertainty is additionally represented by additive perturbations to the liquid-water content of each soil layer, with layer-dependent standard deviations as given in Table S5 (Nearing et al., 2018). The observation operator calculate surface volumetric soil moisture by depth-weighting the corresponding CoLM soil layers and bilinearly interpolating the modeled soil moisture to the SMAP grid. To reduce systematic climatological differences between modeled and retrieved soil moisture, assimilation is performed using anomalies relative to their respective daily climatology (Rains et al., 2017; De Lannoy and Reichle, 2016). Because the SMAP record begins in 2015, the observation and model climatology are estimated from their respective available records. The SMAP climatology is derived from retrievals collected during April 2015–March 2018, whereas the model climatology is calculated from an independent CoLM simulation during April 2011–March 2015. Daily climatological values are averaged over a sliding window spanning 15 days on either side of each target day, using all years of the respective record. Both periods span complete annual cycles, allowing anomalies to be defined relative to the dataset-specific mean seasonal cycle. Observation errors are taken as Gaussian and spatially independent, characterized by a standard deviation of $0.04\ \mathrm{m}^3, \mathrm{m}^{-3}$ (ONeill et al., 2023).

LandSF and B24 are integrated with CoLM through an online coupling framework (Figure S3). CoLM propagates the ensemble in parallel using message passing interface (MPI) on CPUs, while score estimation and analysis sampling are performed using PyTorch on GPUs. Forecast and analysis ensembles are exchanged through transmission control protocol (TCP) sockets without intermediate files. Further details are provided in Text S1.

LandSF and B24 are independently evaluated using in situ soil moisture measurements from the International Soil Moisture Network (ISMN; Dorigo et al., 2013), including the Soil Climate Analysis Network (SCAN; Schaefer et al., 2007), Atmospheric Radiation Measurement network (ARM), Snowpack Telemetry network (SNOTEL), and U.S. Climate Reference Network (USCRN; Bell et al., 2013; Diamond et al., 2013). Quality control excludes missing or flagged observations, physically implausible values, measurements under frozen-soil conditions, spikes, temporal discontinuities, and prolonged constant

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

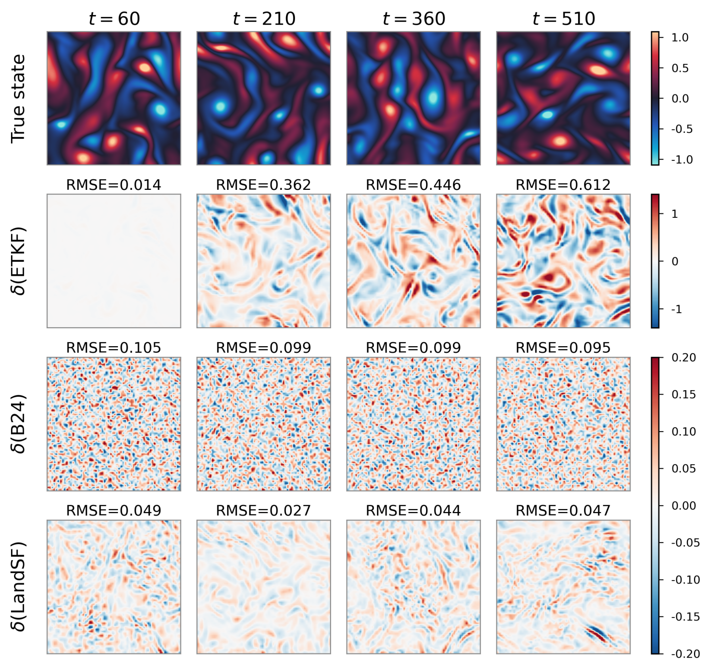

Figure 2. True vorticity fields (1st row) and spatial analysis errors for ETKF (2nd row), B24 (3rd row), and LandSF (4th row) at four assimilation times in the Kolmogorov-flow experiment. The analysis error is defined as the difference between the vorticity of the analysis ensemble mean and the true vorticity. Columns correspond to time steps 60, 210, 360, and 510. Values above the error panels indicate the corresponding spatial RMSE. All panels show the vorticity field $\omega = \nabla \times \boldsymbol{u}$.

values (Dorigo et al., 2013; Seo et al., 2021). Following Seo et al. (2021), stations are retained when valid measurements cover more than 50% of the evaluation period and are excluded when SMAP retrievals correlate substantially less well with the in-situ measurements than the open-loop simulation, defined as $R_{\mathrm{SMAP}} - R_{\mathrm{OL}} < -0.1$. The remaining hourly observations were aggregated to daily means, resulting in 110 stations from ARM (3), SCAN (34), SNOTEL (48), and USCRN (25). Modeled soil moisture is bilinearly interpolated from the 0.5° grid to each station, and performance is evaluated using the Pearson correlation coefficient (R) (Seo et al., 2021) and unbiased root-mean-square error (ubRMSE). Following Seo et al. (2021), approximate 95% confidence intervals for R were calculated at each station using the Fisher z transform and aggregated by averaging the lower and upper error widths separately and dividing by the square root of the number of stations. For ubRMSE, station-level chi-square confidence intervals were aggregated in the same way. These estimates assume independent temporal samples and stations.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

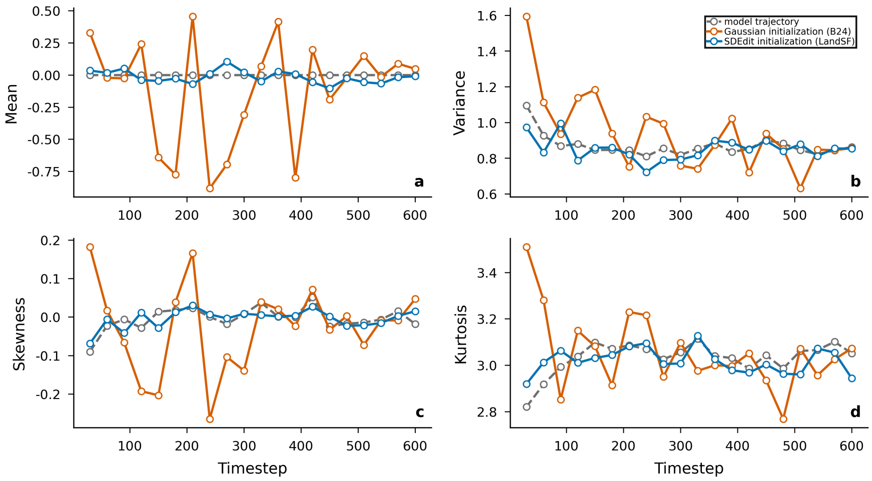

Figure 3. Evolution of the spatial averaged ensemble moments at the 30-step forecast interval, including (a) mean, (b) variance, (c) skewness, and (d) kurtosis. The model-propagated forecast ensemble is compared with the prior ensembles generated by B24 and LandSF without assimilating observations.

4 Results

4.1 Kolmogorov-flow experiments

Figure 2 compares the spatial analysis errors of ETKF, B24, and LandSF as the forecast-error distribution evolves during the Kolmogorov-flow experiment. The initial ensemble is generated by adding Gaussian perturbations to the true state (Section 3.1), resulting in an approximately Gaussian forecast-error distribution consistent with the assumption underlying the ETKF analysis. Accordingly, ETKF closely reproduces the true state and achieves the lowest RMSE (2$^{\text{nd}}$ row). As nonlinear model propagation progressively introduces non-Gaussianity, however, ETKF deteriorates rapidly and develops large, spatially coherent errors (Yang et al., 2012; Nerger, 2022). In contrast, both SFs maintain substantially smaller errors at later assimilation times, demonstrating their ability to perform DA under non-Gaussian forecast errors (Rozet and Louppe, 2023a; Bao et al., 2024b). B24 limits the overall growth of analysis errors but produces persistent small-scale, spatial artifacts across the domain (3$^{\text{rd}}$ row). LandSF further reduces the error magnitude and largely suppresses these artifacts, achieving the lowest time-mean RMSE over the assimilation period (4$^{\text{th}}$ row).

To explain the improved performance of LandSF relative to B24, we evaluate its two methodological advances separately. The forecast ensemble represents the forecast distribution and provides the training samples used to estimate its score. Because the forecast ensemble represents the target forecast distribution, an ideal reverse-sampling procedure should reproduce its statistical properties (Song et al., 2021). To evaluate the effect of the initialization strategy of reverse diffusion sampling, we use B24 and LandSF to generate ensembles without assimilating observations. Both methods use the same diffusion model to estimate the forecast score but differ only in their initialization of reverse diffusion. Figure 3 compares the moments of the generated ensembles with those of the model-propagated forecast ensemble. B24 produces substantial discrepancies, whereas LandSF consistently reproduces the forecast statistics more closely. The result indicate that SDEdit initialization reduces the influence of finite-ensemble forecast-score errors and better preserves the statistical structure of the forecast ensemble.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

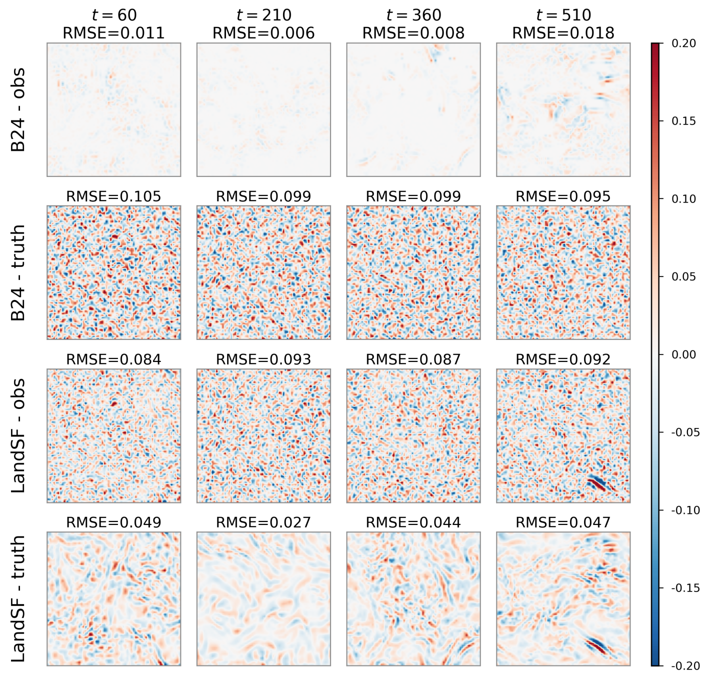

Figure 4. Spatial analysis errors of B24 and LandSF relative to the observations and true states at four assimilation times in the Kolmogorov-flow experiment. From top to bottom, the rows show B24 minus observations, B24 minus truth, LandSF minus observations, and LandSF minus truth. Columns correspond to time steps 60, 210, 360, and 510. The analysis error is measured by comparing the analysis ensemble mean with the corresponding reference field, and the values above each panel report the spatial RMSE.

We next introduce observations into the reverse diffusion process to evaluate the likelihood-score approximations of B24 and LandSF. Figure 4 shows their spatial analysis errors against the observations (1$^{\text{st}}$ and 3$^{\text{rd}}$ rows) and the true states (2$^{\text{nd}}$ and 4$^{\text{th}}$ rows). B24 fits the observations more closely but deviates substantially further from the truth, indicating overfitting to observational noise. LandSF shows the opposite pattern, remaining closer to the true states despite larger residuals against the noisy observations. Consistently, B24 deteriorates markedly as the observation noise increases, whereas LandSF remains comparatively stable (Figure S4). B24 applies a prescribed damping function to the likelihood associated with the unperturbed state but does not account for uncertainty at intermediate diffusion steps. Its guidance parameter only constrains the overall score magnitude without resolving this imbalance. LandSF instead explicitly combines the observation-error covariance with diffusion-dependent state uncertainty, thereby reducing sensitivity to noisy observations.

4.2 Land data assimilation experiments

Figure 5 evaluates the surface soil moisture against ISMN observations. The open-loop (OL) CoLM simulation shows reasonable skill, with a mean ubRMSE of $0.06\ \mathrm{m}^3\,\mathrm{m}^{-3}$ (Figure 5a). Both SFs generally improve upon the OL simulation. B24 improves 60% of the stations but produces substantial degradation at others, resulting in a mean ubRMSE reduction of only 1.78% relative to OL (Figure 5b). LandSF reduces the mean ubRMSE by a further 2.28% relative to B24 and improves 64%

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

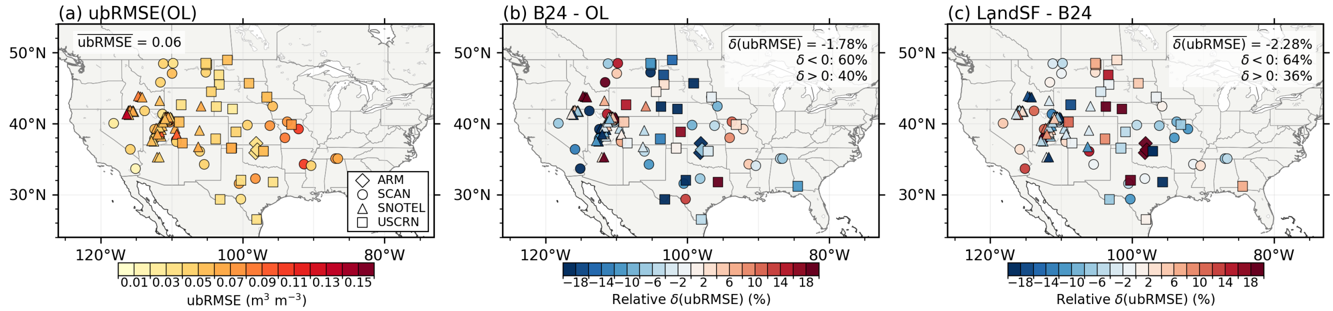

Figure 5. Station-level surface volumetric soil moisture ubRMSE evaluated against independent ISMN observations. (a) ubRMSE of the open-loop CoLM simulation (OL); (b) relative change in ubRMSE from OL to B24; and (c) relative change from B24 to LandSF. Negative changes indicate reduced ubRMSE. Marker shapes distinguish the ARM, SCAN, SNOTEL, and USCRN station networks. The annotations report the mean ubRMSE or relative change and the percentages of stations with decreased and increased ubRMSE.

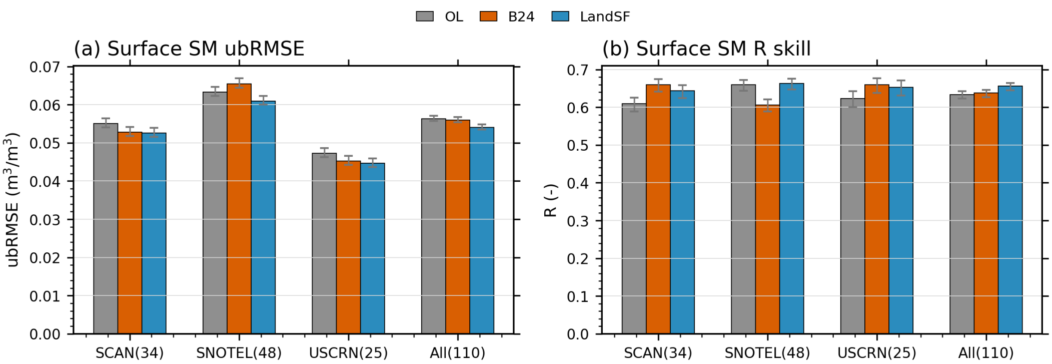

Figure 6. Surface volumetric soil moisture performance of OL, B24, and LandSF evaluated against independent ISMN observations for the SCAN, SNOTEL, and USCRN networks and all stations. Panels show (a) the mean ubRMSE and (b) the mean R. Parenthetical values denote station counts within each group. Error bars represent 95% confidence intervals.

of the stations (Figure 5c). Figure 6 compares the three experiments across station networks, with error bars indicating 95% confidence intervals. Across all stations, LandSF achieves the lowest mean ubRMSE and the highest mean R, with confidence intervals separated from those of OL and B24, indicating a statistically significant reduction. At the network level, B24 performs well for SCAN and USCRN, where it yields slightly higher R than LandSF, but deteriorates markedly for SNOTEL and even falls below OL. LandSF provides more consistent improvements across all three networks, resulting in the lowest mean ubRMSE and highest mean R across all stations. These results demonstrate the greater robustness of LandSF across heterogeneous land-surface conditions. We next evaluate the two methodological advances separately, following the diagnostic design of the Kolmogorov-flow experiment (Section 4.1). We first isolate the effect of initialization by generating ensembles from the estimated forecast score without assimilating observations. Figure 7 shows the model forecast ensemble mean and the bias of the B24- and LandSF-generated ensemble means from this forecast. B24 produces substantial spatial biases that increase over time, with positive biases over the Southwest region and widespread negative biases across the northern and eastern regions.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

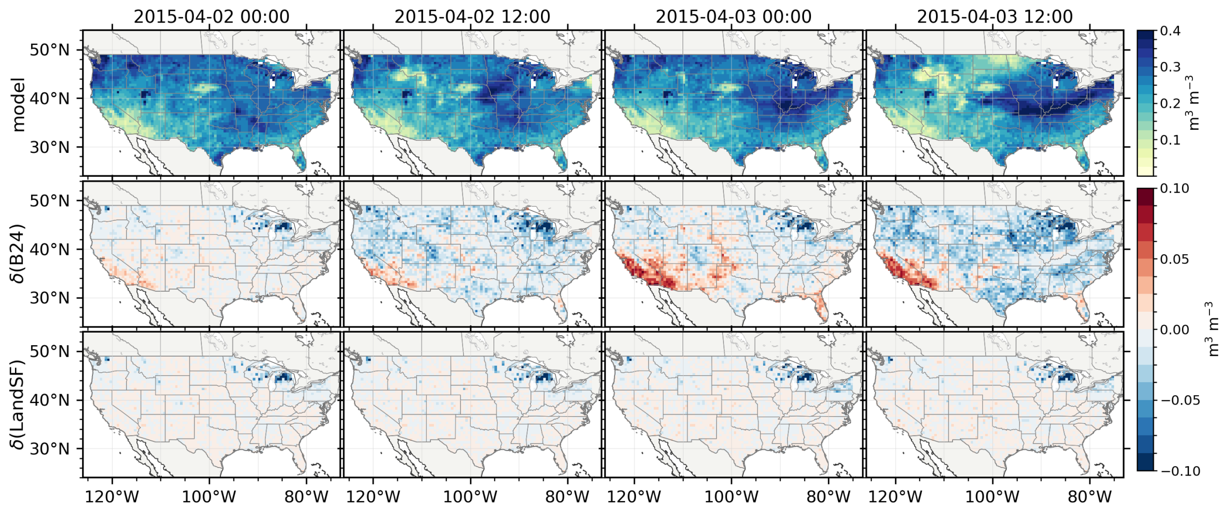

Figure 7. Spatial distributions of surface volumetric soil moisture in the model-propagated forecast ensemble mean (1st row) and the deviation of the prior ensemble means generated by B24 (2nd row) and LandSF (3rd row). The prior ensembles are generated without assimilating observations. Columns correspond to 00:00 and 12:00 UTC on 2 and 3 April 2015.

Notably, the error reaches nearly $0.10\ \mathrm{m}^3\,\mathrm{m}^{-3}$ in Southwest CONUS, leading to relative biases exceeding 100% in the driest areas. LandSF produces much smaller errors across the domain, with no systematic amplification during model integration, consistent with the Kolmogorov-flow results in Figure 3.

We then evaluate the likelihood-score approximations using surface soil moisture anomaly increments. Figure 8 compares the analysis increments of B24 and LandSF with the observation innovations at three selected SCAN, SNOTEL, and USCRN stations. B24 closely tracks individual innovations and produces abrupt fluctuations, particularly at the SNOTEL station, indicating excessive sensitivity to observational noise. This behavior is consistent with Figure 4 and is further supported by the spatial increments pattern in Figure S5. LandSF produces smoother and reasonable increments that retaining the broader temporal variations of observation innovations. This moderated response reflects the explicit treatment of observation uncertainty by MMPS and helps explain the more consistent performance of LandSF across station networks (Figures 5 and 6).

5 Discussions and conclusion

We developed LandSF, a score-based filtering framework for non-Gaussian data assimilation with complex process-based land surface models. To make score-based filtering applicable to such models, LandSF addresses two limitations of existing methods. Forecast-informed initialization reduces sensitivity to forecast-score estimation errors arising from computationally limited ensembles, and moment-matching posterior sampling explicitly accounts for observation errors when approximating

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

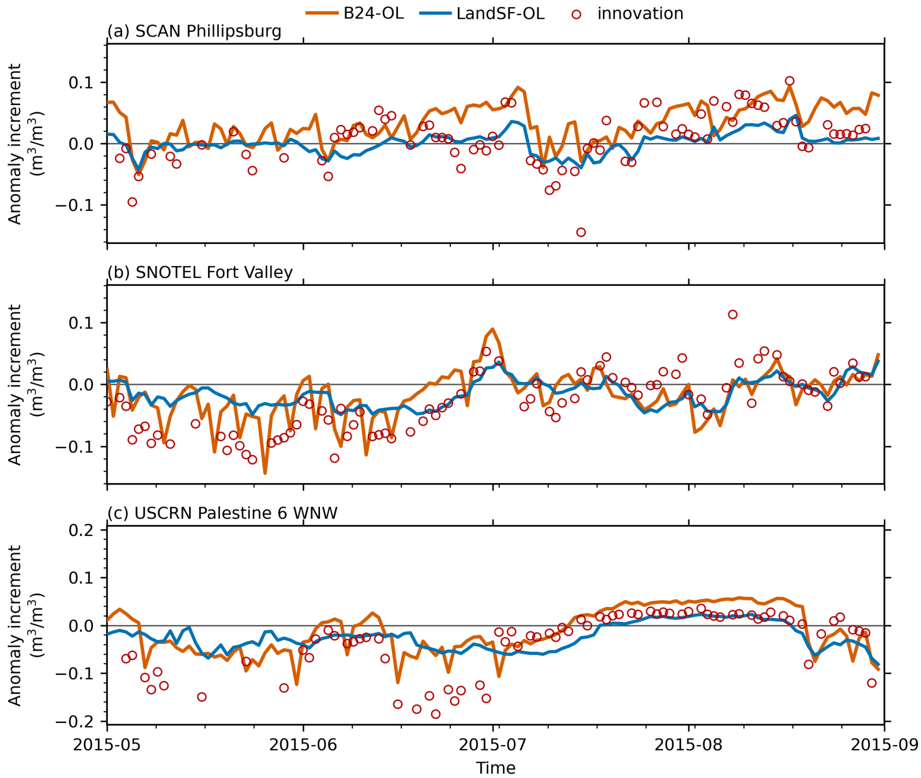

Figure 8. Time series of surface volumetric soil moisture anomaly increments produced by B24 and LandSF relative to the open-loop simulation and the corresponding observation innovations at representative stations from (a) SCAN (Phillipsburg), (b) SNOTEL (Fort Valley), and (c) USCRN (Palestine 6 WNW). Red circles denote observation innovations at assimilation times.

the intractable likelihood score at each reverse diffusion step. Its effectiveness and robustness are demonstrated through a Kolmogorov-flow experiment and a representative land DA experiment using the Common Land Model.

The Kolmogorov-flow experiment highlights the limitation of Gaussian ensemble filtering under nonlinear dynamics. ETKF performs well while the forecast-error distribution remains approximately Gaussian but deteriorates as nonlinear propagation produces non-Gaussian structures (Figure 2). Covariance inflation can compensate for underestimated ensemble spread but cannot represent distributional features beyond the first two moments. Score-based filters instead represent the forecast distribution through its score without imposing a Gaussian approximation, maintaining stable performance, even without covariance inflation (Figure 2).

The comparison with B24 demonstrates that LandSF is better suited to realistic, complex DA applications than existing score-based filters. One practical constraint is that the computational expense of process-based models restricts ensemble size, making some error in the estimated forecast score inevitable. Because B24 begins sampling from Gaussian noise, the learned score must reconstruct the forecast distribution throughout the entire generation process. Any estimation error can therefore be directly expressed as biased ensemble statistics, as observed in Figures 3 and 7. By anchoring each reverse trajectory to a

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

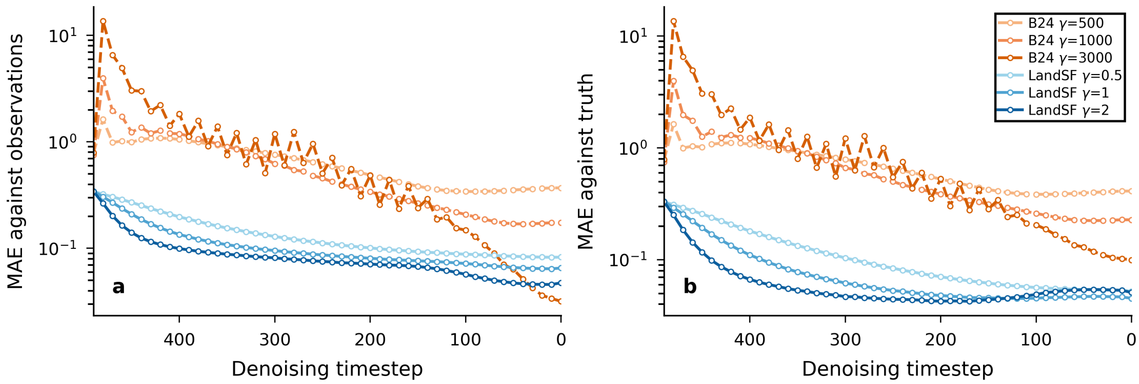

Figure 9. Evolution of the mean deviations from (a) the observations and (b) the true state during reverse diffusion at a single assimilation cycle. B24 uses the prescribed likelihood-score approximation with guidance parameters 500, 1,000, 3,000, whereas LandSF uses the MMPS likelihood-score approximation without an additional tunable guidance parameter. The denoising timestep decreases from left to right.

perturbed forecast member, LandSF retains information from the model forecast and reduces its dependence on an imperfect score estimate. A second practical challenge is the incorporation of noisy observations into score-based filters. The observation model defines the original likelihood only for unperturbed state. However, reverse diffusion process requires the likelihood for noise-perturbed states. B24 bridges this difference by scaling the original likelihood score with a prescribed damping function at each diffusion step. This treatment includes the observation-error covariance but does not account for uncertainty in the physical state represented by a noise-perturbed state. The resulting likelihood score can therefore be overconfident, causing the analysis to fit observational noise too closely. The tunable parameter γ limits the magnitude of the combined score but does not correct this missing uncertainty. Consistent with this interpretation, B24 becomes increasingly sensitive as the observation noise increases (Figures 4, 8, and S5). Figure 9 further examines the likelihood-score approximations along a single reverse-diffusion trajectory. B24 produces pronounced error oscillations during reverse diffusion and is highly sensitive to γ. Because γ bounds the magnitude of the combined forecast and likelihood scores, increasing it permits larger updates that move samples closer to the observations without reducing their error relative to the truth. Notably, tuning γ based on observational agreement may therefore encourage overfitting rather than improve analysis accuracy. Although γ can be optimized using known true states in idealized experiments, it cannot be objectively calibrated in real applications. LandSF avoids this case-dependent tuning by accounting for both observation error and conditional state uncertainty, resulting in a smoother and more stable balance between forecast and observational information.

Despite these improvements, LandSF remains dependent on the ability of the forecast ensemble to represent the underlying forecast distribution. This representation is determined by the perturbations applied to forcing, parameters, initial conditions, and model structure, together with their spatial and temporal dependencies (Ewerdwalbesloh et al., 2026). In high-dimensional process-based models, a computationally affordable ensemble may not adequately sample the relevant non-Gaussian structures. Pretrained generative priors offer a potential solution by learning a richer state distribution from long simulations or reanalysis records and reusing this information across assimilation cycles (Yang et al., 2025; Andrae et al., 2026). Moreover, video diffusion models provide another promising direction by learning the joint evolution of land states over time, thereby allowing observations from multiple time steps to jointly constrain the state trajectory rather than updating each assimilation time independently (Rozet and Louppe, 2023b). Future work should systematically compare different generative DA frameworks,

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

including score-based data assimilation (Rozet and Louppe, 2023b; Manshausen et al., 2025; Andry et al., 2025; Foroumandi and Moradkhani, 2025), score-based filters and their extensions (Bao et al., 2024b; Ding et al., 2025; Si and Chen, 2024; Zhang et al., 2025; Xiong et al., 2026; Hu et al., 2025), and data assimilation with pretrained generative priors (Yang et al., 2025; Andrae et al., 2026), to clarify their trade-offs in flow dependence, computational efficiency, and robustness to limited ensembles and noisy observations for land DA with process-based LSMs.

Author contributions.

Zhilong Fan and Lu Li conceptualized the structure of the paper, developed the LandSF methodology, and led the writing of the draft. Conceptualization: ZF and LL. Methodology: ZF and LL. Coding and implementation of the LandSF framework and its coupling with CoLM: ZF and LL. Design and execution of the Kolmogorov-flow and land data assimilation experiments: ZF and LL. Data curation: ZW, SZ, SC, XL, WS, NW, HY, and PZ. Resources: YD, LL, and the SYSU CoLM development team (ZW, SZ, SC, XL, WS, NW, HY, PZ). Funding acquisition: YD and LL. Supervision: YD and LL. Methodology discussion: all authors. Writing, review, and editing: all authors.

Code and data availability.

The Common Land Model source code used in this study is available at https://github.com/CoLM-SYSU/CoLM and archived on Zenodo (CoLM202X; Wei, 2025, https://doi.org/10.5281/zenodo.17120235).

The SMAP Level-2 passive soil moisture retrievals (Version 9) are available from the NASA NSIDC Distributed Active Archive Center (O'Neill et al., 2023, https://doi.org/10.5067/K7Y2D8QQVZ4L).

The reference implementation of the B24 score-based filter is available at https://github.com/zezhongzhang/Score-based-Filter.

The LandSF analysis code used in this study is archived on Zenodo (LandSF; Fan et al., 2026, https://zenodo.org/records/22156332?token=eyJhbGciOiJIUzUxMiIsImlhdCI6MTc4Nzk4NDQ0MSwiZXhwIjoxODMwMjExMTk5fQ.eyJpZCI6IjFkMGQ5Yzg4LTRjYWMtNDQ0MC05ODBiLWRjOGY3NGM4MTcxNiIsImRhdGEiOnt9LCJyYW5kb20iOiJiNTAzZTQxNTliODdhMDU2NWU0NzYzOTcwNGYyYjcxYyJ9.ChUj2woyrm4TWYvltlaur-k1BDOKw7g-6-iVZM5Cs7myqhX5JPz41tbqvztNmKhMwWrrIDELm3Kk3b-OS6BUbA).

Competing interests.

The contact author has declared that none of the authors has any competing interests.

Acknowledgements.

We thank the NASA National Snow and Ice Data Center Distributed Active Archive Center for providing the SMAP Level-2 soil moisture retrievals, the European Centre for Medium-Range Weather Forecasts for providing the ERA5 atmospheric forcing data, and the International Soil Moisture Network and its contributing station networks (ARM, SCAN, SNOTEL, and USCRN) for providing in situ soil moisture observations. We thank NASA for providing the MODIS remote sensing products, Wei Shangguan for providing the global soil dataset, and Hua Yuan for providing the leaf area index data used in this work. AI tools were used to assist with language editing. The authors take full responsibility for the content of the manuscript.

Financial support.

The work was partially supported by the National Key Research and Development Plan of China (Grant No. 2023YFB3002400), the Natural Science Foundation of China under Grants (42575176, 42075158, U2342227, 42375166, 42375144, 42105144, 42305164, 42575159,

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

42475172), Guangdong Basic and Applied Basic Research Foundation (2024A1515011760, 2024A1515010283), the Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (SML2023SP216), the Natural Science Foundation of Guangdong Province (2023A1515011541), the National Key Research and Development Program of China (2025YFE0217100), the CAS Project for Young Scientists in Basic Research (No. YSBR-137), Strategic Priority Research Program of the Chinese Academy of Sciences (XDA0500400).

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

References

Abbaszadeh, P., Gavahi, K., and Moradkhani, H.: Multivariate Remotely Sensed and In-Situ Data Assimilation for Enhancing Community WRF-Hydro Model Forecasting, Advances in Water Resources, 145, 103 721, https://doi.org/10.1016/j.advwatres.2020.103721, 2020.

Aksoy, A., Dowell, D. C., and Snyder, C.: A Multicase Comparative Assessment of the Ensemble Kalman Filter for Assimilation of Radar Observations. Part I: Storm-Scale Analyses, Monthly Weather Review, 137, 1805–1824, https://doi.org/10.1175/2008MWR2691.1, 2009.

Anderson, B. D.: Reverse-Time Diffusion Equation Models, Stochastic Processes and their Applications, 12, 313–326, https://doi.org/10.1016/0304-4149(82)90051-5, 1982.

Andrae, M., Larsson, E., Takao, S., Landelius, T., and Lindsten, F.: DAISI: Data Assimilation with Inverse Sampling Using Stochastic Interpolants, https://doi.org/10.48550/arXiv.2512.00252, 2026.

Andry, G., Rozet, F., Lewin, S., Rochman, O., Mangeleer, V., Pirlet, M., Faulx, E., Grégoire, M., and Louppe, G.: Appa: Bending Weather Dynamics with Latent Diffusion Models for Global Data Assimilation, https://doi.org/10.48550/arXiv.2504.18720, 2025.

Bao, F., Zhang, Z., and Zhang, G.: A Score-Based Filter for Nonlinear Data Assimilation, Journal of Computational Physics, 514, 113 207, https://doi.org/10.1016/j.jcp.2024.113207, 2024a.

Bao, F., Zhang, Z., and Zhang, G.: An Ensemble Score Filter for Tracking High-Dimensional Nonlinear Dynamical Systems, https://doi.org/10.48550/arXiv.2309.00983, 2024b.

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., Hall, M. E., Kochendorfer, J., Meyers, T. P., Wilson, T., and Diamond, H. J.: U.S. Climate Reference Network Soil Moisture and Temperature Observations, Journal of Hydrometeorology, 14, 977–988, https://doi.org/10.1175/JHM-D-12-0146.1, 2013.

Berg, D., Bauser, H. H., and Roth, K.: Covariance Resampling for Particle Filter – State and Parameter Estimation for Soil Hydrology, Hydrology and Earth System Sciences, 23, 1163–1178, https://doi.org/10.5194/hess-23-1163-2019, 2019.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects, Monthly Weather Review, 129, 420–436, https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2, 2001.

Bonan, G.: Model ArchiveLand Surface Model (LSM 1.0) for Ecological, Hydrological, Atmospheric Studies, https://doi.org/10.3334/ORNLDAAC/807, 2005.

Chen, Y., Sanz-Alonso, D., and Willett, R.: Reduced-Order Autodifferentiable Ensemble Kalman Filters, https://doi.org/10.48550/ARXIV.2301.11961, 2023.

Chung, H., Kim, J., Mccann, M. T., Klasky, M. L., and Ye, J. C.: DIFFUSION POSTERIOR SAMPLING FOR GENERAL NOISY INVERSE PROBLEMS, in: International Conference on Learning Representations (ICLR), 2023.

Crow, W. T. and Van Den Berg, M. J.: An Improved Approach for Estimating Observation and Model Error Parameters in Soil Moisture Data Assimilation, Water Resources Research, 46, 2010WR009 402, https://doi.org/10.1029/2010WR009402, 2010.

Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S., Dirmeyer, P. A., Houser, P. R., Niu, G., Oleson, K. W., Schlosser, C. A., and Yang, Z.-L.: The Common Land Model, Bulletin of the American Meteorological Society, 84, 1013–1024, https://doi.org/10.1175/BAMS-84-8-1013, 2003.

Dai, Y., Xin, Q., Wei, N., Zhang, Y., Shangguan, W., Yuan, H., Zhang, S., Liu, S., and Lu, X.: A Global High-Resolution Data Set of Soil Hydraulic and Thermal Properties for Land Surface Modeling, Journal of Advances in Modeling Earth Systems, 11, 2996–3023, https://doi.org/10.1029/2019MS001784, 2019.

Dasgupta, A., Hostache, R., Ramsankaran, R., Schumann, G. J.-P., Grimaldi, S., Pauwels, V. R. N., and Walker, J. P.: A Mutual Information-Based Likelihood Function for Particle Filter Flood Extent Assimilation, Water Resources Research, 57, e2020WR027 859, https://doi.org/10.1029/2020WR027859, 2021.

De Lannoy, G. J. M. and Reichle, R. H.: Assimilation of SMOS Brightness Temperatures or Soil Moisture Retrievals into a Land Surface Model, Hydrology and Earth System Sciences, 20, 4895–4911, https://doi.org/10.5194/hess-20-4895-2016, 2016.

Diamond, H. J., Karl, T. R., Palecki, M. A., Baker, C. B., Bell, J. E., Leeper, R. D., Easterling, D. R., Lawrimore, J. H., Meyers, T. P., Helfert, M. R., Goodge, G., and Thorne, P. W.: U.S. Climate Reference Network after One Decade of Operations: Status and Assessment, Bulletin of the American Meteorological Society, 94, 485–498, https://doi.org/10.1175/BAMS-D-12-00170.1, 2013.

Dickinson, R. E., Henderson-Sellers, A., Kennedy, P. J., and Wilson, M. F.: Biosphere-Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model, Tech. Rep. NCAR/TN-275+STR, National Center for Atmospheric Research, 1986.

Ding, Z., Duan, C., Jiao, Y., Yang, J. Z., Yuan, C., and Zhang, P.: Nonlinear Assimilation via Score-based Sequential Langevin Sampling, https://doi.org/10.48550/arXiv.2411.13443, 2025.

Dong, J., Steele-Dunne, S. C., Judge, J., and Van De Giesen, N.: A Particle Batch Smoother for Soil Moisture Estimation Using Soil Temperature Observations, Advances in Water Resources, 83, 111–122, https://doi.org/10.1016/j.advwatres.2015.05.017, 2015.

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A., Zamojski, D., Cordes, C., Wagner, W., and Drusch, M.: Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network, Vadose Zone Journal, 12, 1–21, https://doi.org/10.2136/vzj2012.0097, 2013.

Efron, B.: Tweedie's Formula and Selection Bias, Journal of the American Statistical Association, 106, 1602–1614, https://doi.org/10.1198/jasa.2011.tm11181, 2011.

Ewerdwalbesloh, Y., Montzka, C., Kusche, J., and Springer, A.: GRACE TWSA Data Assimilation in Land Surface Models-The Role of State Vector and Model Uncertainty Representation, Water Resources Research, 62, e2025WR042 324, https://doi.org/10.1029/2025WR042324, 2026.

Farchi, A. and Bocquet, M.: Review Article: Comparison of Local Particle Filters and New Implementations, Nonlinear Processes in Geophysics, 25, 765–807, https://doi.org/10.5194/npg-25-765-2018, 2018.

Finn, T. S., Disson, L., Farchi, A., Bocquet, M., and Durand, C.: Representation Learning with Unconditional Denoising Diffusion Models for Dynamical Systems, Nonlinear Processes in Geophysics, 31, 409–431, https://doi.org/10.5194/npg-31-409-2024, 2024.

Foroumandi, E. and Moradkhani, H.: Harnessing Generative Deep Learning for Enhanced Ensemble Data Assimilation, Water Resources Research, 61, https://doi.org/10.1029/2025wr040078, 2025.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X.: MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets, Remote Sensing of Environment, 114, 168–182, https://doi.org/10.1016/j.rse.2009.08.016, 2010.

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J.-C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., Draper, C., Hirschi, M., Kerr, Y., Konings, A., Lahoz, W., McColl, K., Montzka, C., Muñoz-Sabater, J., Peng, J., Reichle, R., Richaume, P., Rüdiger, C., Scanlon, T., Van Der Schalie, R., Wigneron, J.-P., and Wagner, W.: Validation Practices for Satellite Soil Moisture Retrievals: What Are (the) Errors?, Remote Sensing of Environment, 244, 111 806, https://doi.org/10.1016/j.rse.2020.111806, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 Global Reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Ho, J., Jain, A., and Abbeel, P.: Denoising diffusion probabilistic models, Advances in neural information processing systems, 33, 6840–6851, 2020.

Houser, P. R., Shuttleworth, W. J., Famiglietti, J. S., Gupta, H. V., Syed, K. H., and Goodrich, D. C.: Integration of Soil Moisture Remote Sensing and Hydrologic Modeling Using Data Assimilation, Water Resources Research, 34, 3405–3420, https://doi.org/10.1029/1998WR900001, 1998.

Hu, R., Poudel, S., Bao, F., and Lee, S.: Ensemble Score Filter for Data Assimilation of Two-Phase Flow Models in Porous Media, https://doi.org/10.48550/arXiv.2504.09245, 2025.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

Huang, L., Gianinazzi, L., Yu, Y., Dueben, P. D., and Hoefler, T.: DiffDA: A Diffusion Model for Weather-scale Data Assimilation, https://doi.org/10.48550/arXiv.2401.05932, 2024.

Kochkov, D., Smith, J. A., Alieva, A., Wang, Q., Brenner, M. P., and Hoyer, S.: Machine learning-accelerated computational fluid dynamics, Proceedings of the National Academy of Sciences, 118, e2101784 118, 2021.

Kumar, S., Kolassa, J., Reichle, R., Crow, W., De Lannoy, G., De Rosnay, P., MacBean, N., Girotto, M., Fox, A., Quaife, T., Draper, C., Forman, B., Balsamo, G., Steele-Dunne, S., Albergel, C., Bonan, B., Calvet, J.-C., Dong, J., Liddy, H., and Ruston, B.: An Agenda for Land Data Assimilation Priorities: Realizing the Promise of Terrestrial Water, Energy, and Vegetation Observations From Space, Journal of Advances in Modeling Earth Systems, 14, e2022MS003 259, https://doi.org/10.1029/2022MS003259, 2022.

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, Reviews of Geophysics, 62, e2022RG000 801, https://doi.org/10.1029/2022RG000801, 2024.

Lievens, H., Reichle, R. H., Liu, Q., De Lannoy, G. J. M., Dunbar, R. S., Kim, S. B., Das, N. N., Cosh, M., Walker, J. P., and Wagner, W.: Joint Sentinel-1 and SMAP Data Assimilation to Improve Soil Moisture Estimates, Geophysical Research Letters, 44, 6145–6153, https://doi.org/10.1002/2017GL073904, 2017.

Ling, F., Ouyang, L., Larbi, B. R., Luo, J.-J., Han, T., Zhong, X., and Bai, L.: Improving Global Weather and Ocean Wave Forecast with Large Artificial Intelligence Models, Science China Earth Sciences, 67, 3641–3654, https://doi.org/10.1007/s11430-024-1452-3, 2024.

Loshchilov, I. and Hutter, F.: Decoupled Weight Decay Regularization, https://doi.org/10.48550/arXiv.1711.05101, 2019.

Manshausen, P., Cohen, Y., Harrington, P., 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, Journal of Advances in Modeling Earth Systems, 17, e2024MS004 505, https://doi.org/10.1029/2024MS004505, 2025.

Martin, S., Manucharyan, G., and Klein, P.: Generative Data Assimilation for Surface Ocean State Estimation from Multi-Modal Satellite Observations, https://doi.org/10.31223/X5ZT6N, 2025.

McLaughlin, D., Zhou, Y., Entekhabi, D., and Chatdarong, V.: Computational Issues for Large-Scale Land Surface Data Assimilation Problems, Journal of Hydrometeorology, 7, 494–510, https://doi.org/10.1175/JHM493.1, 2006.

Meng, C., He, Y., Song, Y., Song, J., Wu, J., Zhu, J.-Y., and Ermon, S.: SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations, https://doi.org/10.48550/arXiv.2108.01073, 2022.

Nearing, G., Yatheendradas, S., Crow, W., Zhan, X., Liu, J., and Chen, F.: The Efficiency of Data Assimilation, Water Resources Research, 54, 6374–6392, https://doi.org/10.1029/2017WR020991, 2018.

Nerger, L.: Data Assimilation for Nonlinear Systems with a Hybrid Nonlinear Kalman Ensemble Transform Filter, Quarterly Journal of the Royal Meteorological Society, 148, 620–640, https://doi.org/10.1002/qj.4221, 2022.

Noh, S. J., Tachikawa, Y., Shiiba, M., and Kim, S.: Applying Sequential Monte Carlo Methods into a Distributed Hydrologic Model: Lagged Particle Filtering Approach with Regularization, Hydrology and Earth System Sciences, 15, 3237–3251, https://doi.org/10.5194/hess-15-3237-2011, 2011.

O'Neill, P., Bindlish, R., Chan, S., Chaubell, J., Njoku, E., and Jackson, T.: Algorithm Theoretical Basis Document Level 2 & 3 Soil Moisture (Passive) Data Products, Tech. rep., NASA Jet Propulsion Laboratory, California Institute of Technology, 2023.

ONeill, P., Chan, S., Njoku, E., Jackson, T., Bindlish, R., and Chaubell, J.: SMAP L2 Radiometer Half-Orbit 36 Km EASE-Grid Soil Moisture, Version 9, https://doi.org/10.5067/K7Y2D8QQVZ4L, 2023.

Qu, Y., Nathaniel, J., Li, S., and Gentine, P.: Deep Generative Data Assimilation in Multimodal Setting, in: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 449–459, IEEE, Seattle, WA, USA, ISBN 979-8-3503-6547-4, https://doi.org/10.1109/CVPRW63382.2024.00050, 2024.

Rains, D., Han, X., Lievens, H., Montzka, C., and Verhoest, N. E. C.: SMOS Brightness Temperature Assimilation into the Community Land Model, Hydrology and Earth System Sciences, 21, 5929–5951, https://doi.org/10.5194/hess-21-5929-2017, 2017.

Reichle, R. H., Walker, J. P., Koster, R. D., and Houser, P. R.: Extended versus Ensemble Kalman Filtering for Land Data Assimilation, Journal of Hydrometeorology, 3, 728–740, https://doi.org/10.1175/1525-7541(2002)003<0728:EVEKFF>2.0.CO;2, 2002.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

Reichle, R. H., Crow, W. T., and Keppenne, C. L.: An Adaptive Ensemble Kalman Filter for Soil Moisture Data Assimilation, Water Resources Research, 44, 2007WR006 357, https://doi.org/10.1029/2007WR006357, 2008a.

Reichle, R. H., Crow, W. T., Koster, R. D., Sharif, H. O., and Mahanama, S. P. P.: Contribution of Soil Moisture Retrievals to Land Data Assimilation Products, Geophysical Research Letters, 35, 2007GL031 986, https://doi.org/10.1029/2007GL031986, 2008b.

Reichle, R. H., De Lannoy, G. J. M., Liu, Q., Koster, R. D., Kimball, J. S., Crow, W. T., Ardizzone, J. V., Chakraborty, P., Collins, D. W., Conaty, A. L., Girotto, M., Jones, L. A., Kolassa, J., Lievens, H., Lucchesi, R. A., and Smith, E. B.: Global Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using Assimilation Diagnostics, Journal of Hydrometeorology, 18, 3217–3237, https://doi.org/10.1175/JHM-D-17-0130.1, 2017.

Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J. M., Kimball, J. S., Ardizzone, J. V., Bosch, D., Colliander, A., Cosh, M., Kolassa, J., Mahanama, S. P., Prueger, J., Starks, P., and Walker, J. P.: Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product, Journal of Advances in Modeling Earth Systems, 11, 3106–3130, https://doi.org/10.1029/2019MS001729, 2019.

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, Bulletin of the American Meteorological Society, 85, 381–394, https://doi.org/10.1175/BAMS-85-3-381, 2004.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, https://doi.org/10.48550/arXiv.1505.04597, 2015.

Rozet, F. and Louppe, G.: Score-Based Data Assimilation, in: Advances in Neural Information Processing Systems 36 (NeurIPS 2023), 2023a.

Rozet, F. and Louppe, G.: Score-Based Data Assimilation for a Two-Layer Quasi-Geostrophic Model, https://doi.org/10.48550/arXiv.2310.01853, 2023b.

Rozet, F., Andry, G., Lanusse, F., and Louppe, G.: Learning diffusion priors from observations by expectation maximization, Advances in Neural Information Processing Systems, 37, 87 647–87 682, 2024.

Saad, Y. and Schultz, M. H.: GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems, SIAM Journal on Scientific and Statistical Computing, 7, 856–869, https://doi.org/10.1137/0907058, 1986.

Schaefer, G. L., Cosh, M. H., and Jackson, T. J.: The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN), Journal of Atmospheric and Oceanic Technology, 24, 2073–2077, https://doi.org/10.1175/2007JTECHA930.1, 2007.

Seo, E., Lee, M.-I., and Reichle, R. H.: Assimilation of SMAP and ASCAT Soil Moisture Retrievals into the JULES Land Surface Model Using the Local Ensemble Transform Kalman Filter, Remote Sensing of Environment, 253, 112 222, https://doi.org/10.1016/j.rse.2020.112222, 2021.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., and Yuan, H.: A Global Soil Data Set for Earth System Modeling, Journal of Advances in Modeling Earth Systems, 6, 249–263, https://doi.org/10.1002/2013MS000293, 2014.

Si, P. and Chen, P.: Latent-EnSF: A Latent Ensemble Score Filter for High-Dimensional Data Assimilation with Sparse Observation Data, https://doi.org/10.48550/arXiv.2409.00127, 2024.

Song, J., Meng, C., and Ermon, S.: Denoising Diffusion Implicit Models, https://doi.org/10.48550/arXiv.2010.02502, 2022.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B.: Score-Based Generative Modeling through Stochastic Differential Equations, arXiv preprint arXiv:2011.13456, 2020.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B.: Score-Based Generative Modeling through Stochastic Differential Equations, https://doi.org/10.48550/arXiv.2011.13456, 2021.

Sun, J.-A., Fan, H., Gong, J., Fei, B., Chen, K., Ling, F., Zhang, W., Xu, W., Yan, L., Gentine, P., and Bai, L.: LO-SDA: Latent Optimization for Score-based Atmospheric Data Assimilation, https://doi.org/10.48550/arXiv.2510.22562, 2025a.

Sun, J.-A., Fan, H., Gong, J., Fei, B., Chen, K., Ling, F., Zhang, W., Xu, W., Yan, L., Gentine, P., and Bai, L.: Align-DA: Align Score-based Atmospheric Data Assimilation with Multiple Preferences, https://doi.org/10.48550/arXiv.2505.22008, 2025b.

Vincent, P.: A Connection Between Score Matching and Denoising Autoencoders, Neural Computation, 23, 1661–1674, https://doi.org/10.1162/neco_a_00142, 2011.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

Wang, H., Han, J., Fan, W., Zhang, W., and Liu, H.: PhyDA: Physics-Guided Diffusion Models for Data Assimilation in Atmospheric Systems, https://doi.org/10.48550/arXiv.2505.12882, 2025.

Wang, X. and Bishop, C. H.: A Comparison of Breeding and Ensemble Transform Kalman Filter Ensemble Forecast Schemes, Journal of the Atmospheric Sciences, 60, 1140–1158, https://doi.org/10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2, 2003.

Wei, Z.: The Common Land Model Version 202X (CoLM202X), https://doi.org/10.5281/ZENODO.17120235, 2025.

Xiong, Z., Liang, S., Bao, F., Zhang, G., and Chipilski, H. G.: Robustness of the Ensemble Score Filter to the Type of Assimilated Observation Networks, Atmospheric Science Letters, 27, e70 004, https://doi.org/10.1002/asl.70004, 2026.

Yan, H. and Moradkhani, H.: Combined Assimilation of Streamflow and Satellite Soil Moisture with the Particle Filter and Geostatistical Modeling, Advances in Water Resources, 94, 364–378, https://doi.org/10.1016/j.advwatres.2016.06.002, 2016.

Yang, S., Nai, C., Liu, X., Li, W., Chao, J., Wang, J., Wang, L., Li, X., Chen, X., Lu, B., Xiao, Z., Boers, N., Yuan, H., and Pan, B.: Generative Assimilation and Prediction for Weather and Climate, https://doi.org/10.48550/arXiv.2503.03038, 2025.

Yang, S.-C., Kalnay, E., and Hunt, B.: Handling Nonlinearity in an Ensemble Kalman Filter: Experiments with the Three-Variable Lorenz Model, Monthly Weather Review, 140, 2628–2646, https://doi.org/10.1175/MWR-D-11-00313.1, 2012.

Yongjiu, D. and Qingcun, Z.: A Land Surface Model (IAP94) for Climate Studies Part I: Formulation and Validation in off-Line Experiments, Advances in Atmospheric Sciences, 14, 433–460, https://doi.org/10.1007/s00376-997-0063-4, 1997.

Yuan, H., Dai, Y., Xiao, Z., Ji, D., and Shangguan, W.: Reprocessing the MODIS Leaf Area Index Products for Land Surface and Climate Modelling, Remote Sensing of Environment, 115, 1171–1187, https://doi.org/10.1016/j.rse.2011.01.001, 2011.

Zhang, Z., Bao, F., and Zhang, G.: IEnSF: Iterative Ensemble Score Filter for Reducing Error in Posterior Score Estimation in Nonlinear Data Assimilation, https://doi.org/10.48550/arXiv.2510.20159, 2025.

Appendix A: Deriving EnKF from the Score Perspective

Lemma A.1 (Woodbury Matrix Identity). *For any invertible matrices $\boldsymbol{A}$, $\boldsymbol{C}$ and full-rank matrices $\boldsymbol{U}$, $\boldsymbol{V}$, we have*

$$(\boldsymbol{A}+\boldsymbol{U}\boldsymbol{C}\boldsymbol{V})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{U}(\boldsymbol{C}^{-1}+\boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U})^{-1}\boldsymbol{V}\boldsymbol{A}^{-1}. \tag{A1}$$

The goal of this appendix is to derive the EnKF update formula from the score perspective under the assumptions of a Gaussian prior and linear Gaussian observations. To this end, we first present the analytical properties of a Gaussian distribution under the VP-SDE noising process.

Lemma A.2 (Properties of Gaussian Distributions under VP-SDE). *Let the target distribution be a multivariate Gaussian $p(\boldsymbol{x}_0)=\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$. Under the VP-SDE forward diffusion process, the following properties hold:*

(i) The marginal distribution is

$$p(\boldsymbol{x}_t)=\mathcal{N}\big(\sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu},\,\bar{\alpha}(t)\boldsymbol{\Sigma}+(1-\bar{\alpha}(t))\boldsymbol{I}\big); \tag{A2}$$

(ii) The score function is

$$\nabla_{\boldsymbol{x}_t}\log p(\boldsymbol{x}_t)=-\big(\bar{\alpha}(t)\boldsymbol{\Sigma}+(1-\bar{\alpha}(t))\boldsymbol{I}\big)^{-1}\big(\boldsymbol{x}_t-\sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}\big); \tag{A3}$$

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

(iii) After temporal reparameterization, the reverse PF-ODE admits an analytic solution, yielding a deterministic mapping from standard Gaussian noise to the target distribution:

$$\boxed{\tilde{\boldsymbol{x}}(1) = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{z}, \qquad \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})}. \tag{A4}$$

Proof. From (10), the forward process is a linear transformation with additive independent noise:

$$\boldsymbol{x}_t = \sqrt{\bar{\alpha}(t)}\,\boldsymbol{x}_0 + \sqrt{1-\bar{\alpha}(t)}\,\boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}),$$

where $\boldsymbol{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{\epsilon}$ is independent of $\boldsymbol{x}_0$. Since the Gaussian family is closed under linear transformation and additive independent noise, the marginal distribution (A2) follows directly. The score function (A3) is obtained by taking the log-gradient of (A2).

From Equation (27), the reverse-time SDE is

$$d\boldsymbol{x}_t = \left[-\frac{1}{2}\beta(t)\boldsymbol{x}_t - \beta(t)\nabla_{\boldsymbol{x}_t}\log p(\boldsymbol{x}_t)\right]dt + \sqrt{\beta(t)}\,d\bar{\boldsymbol{w}},$$

where $\bar{\boldsymbol{w}}$ is the standard Brownian motion in reverse time. Analogous to the forward-process PF-ODE, the corresponding PF-ODE is

$$\frac{d\boldsymbol{x}_t}{dt} = -\frac{1}{2}\beta(t)\left[\boldsymbol{x}_t + \nabla_{\boldsymbol{x}_t}\log p(\boldsymbol{x}_t)\right].$$

To eliminate the dependence on the specific noise schedule $\beta(t)$, we introduce the temporal reparameterization $\tau = 1 - \bar{\alpha}(t) \in [0,1]$. Under this transformation, $\frac{d\tau}{dt} = \beta(t)\bar{\alpha}(t)$, and we define $\tilde{\beta}(\tau) = \frac{1}{\bar{\alpha}(t)} = \frac{1}{1-\tau}$. The marginal distribution becomes

$$p(\boldsymbol{x}_\tau) = \mathcal{N}\left(\sqrt{1-\tau}\,\boldsymbol{\mu},\ (1-\tau)\boldsymbol{\Sigma} + \tau\boldsymbol{I}\right),$$

and the score function is

$$\nabla_{\boldsymbol{x}_\tau}\log p(\boldsymbol{x}_\tau) = -\left[(1-\tau)\boldsymbol{\Sigma} + \tau\boldsymbol{I}\right]^{-1}\left(\boldsymbol{x}_\tau - \sqrt{1-\tau}\,\boldsymbol{\mu}\right).$$

The PF-ODE becomes

$$\frac{d\boldsymbol{x}_\tau}{d\tau} = -\frac{1}{2}\tilde{\beta}(\tau)\left[\boldsymbol{x}_\tau + \nabla_{\boldsymbol{x}_\tau}\log p(\boldsymbol{x}_\tau)\right].$$

To generate samples, one evolves from pure noise at $\tau = 1$ backward to $\tau = 0$. Let $s = 1 - \tau \in [0,1]$ and define $\tilde{\boldsymbol{x}}(s) = \boldsymbol{x}(1-s)$. Using $\frac{d\tilde{\boldsymbol{x}}}{ds} = -\frac{d\boldsymbol{x}}{d\tau}\big|_{\tau=1-s}$ and $\tilde{\beta}(1-s) = \frac{1}{s}$, substituting the score yields the reverse ODE

$$\frac{d\tilde{\boldsymbol{x}}}{ds} = \frac{1}{2s}\left[\tilde{\boldsymbol{x}} - \boldsymbol{M}_s^{-1}\left(\tilde{\boldsymbol{x}} - \sqrt{s}\,\boldsymbol{\mu}\right)\right], \quad \boldsymbol{M}_s = s\boldsymbol{\Sigma} + (1-s)\boldsymbol{I}.$$

This linear ODE admits an analytic solution. The general solution is

$$\tilde{\boldsymbol{x}}(s) = \sqrt{s}\,\boldsymbol{\mu} + \boldsymbol{M}_s^{1/2}\boldsymbol{z}, \qquad \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}).$$

At $s = 1$ (corresponding to $t = 0$), $\boldsymbol{M}_1 = \boldsymbol{\Sigma}$, yielding (A4). □

Using Lemma A.2, the posterior score is the sum of the prior score and the likelihood score. The prior score follows directly from Lemma A.2(ii); we now compute the likelihood score.

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

Proposition A.1 (Likelihood Score under Linear Gaussian Observations)**.** *Let the prior distribution be $q(\boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{\mu}_f, \boldsymbol{P}_f)$ and the observation model be $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}_0 + \boldsymbol{\eta}$, where $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R})$ is independent of $\boldsymbol{x}_0$. Under VP-SDE noising, the likelihood score is*

$$\nabla_{\boldsymbol{x}_t} \log q(\boldsymbol{y} \mid \boldsymbol{x}_t) = \sqrt{\bar{\alpha}(t)}\,\boldsymbol{C}(t)^{-1}\boldsymbol{P}_f\boldsymbol{H}^\top\big(\boldsymbol{H}\boldsymbol{\Sigma}_{0|t}\boldsymbol{H}^\top + \boldsymbol{R}\big)^{-1}\big(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_{0|t}\big), \tag{A5}$$

where $\boldsymbol{C}(t) = \bar{\alpha}(t)\boldsymbol{P}_f + (1-\bar{\alpha}(t))\boldsymbol{I}$, and $\boldsymbol{\mu}_{0|t}$, $\boldsymbol{\Sigma}_{0|t}$ are given by (A6).

Proof. We first derive the reverse conditional distribution $q(\boldsymbol{x}_0 \mid \boldsymbol{x}_t)$. Since $(\boldsymbol{x}_0, \boldsymbol{x}_t)$ are jointly Gaussian, their covariance structure is

$$\mathrm{Cov}(\boldsymbol{x}_0, \boldsymbol{x}_t) = \sqrt{\bar{\alpha}(t)}\,\boldsymbol{P}_f, \qquad \mathrm{Var}(\boldsymbol{x}_t) = \boldsymbol{C}(t).$$

Using the Gaussian conditioning formula, we obtain

$$\begin{aligned}\boldsymbol{\mu}_{0|t} &:= \mathbb{E}[\boldsymbol{x}_0 \mid \boldsymbol{x}_t] = \boldsymbol{\mu}_f + \sqrt{\bar{\alpha}(t)}\,\boldsymbol{P}_f\boldsymbol{C}(t)^{-1}\big(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}_f\big),\\ \boldsymbol{\Sigma}_{0|t} &:= \mathbb{V}[\boldsymbol{x}_0 \mid \boldsymbol{x}_t] = \boldsymbol{P}_f - \bar{\alpha}(t)\,\boldsymbol{P}_f\boldsymbol{C}(t)^{-1}\boldsymbol{P}_f.\end{aligned} \tag{A6}$$

Since $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}_0 + \boldsymbol{\eta}$ is a linear function of $\boldsymbol{x}_0$, conditional on $\boldsymbol{x}_t$, $\boldsymbol{y}$ remains Gaussian:

$$\mathbb{E}[\boldsymbol{y} \mid \boldsymbol{x}_t] = \boldsymbol{H}\boldsymbol{\mu}_{0|t}, \quad \mathbb{V}[\boldsymbol{y} \mid \boldsymbol{x}_t] = \boldsymbol{H}\boldsymbol{\Sigma}_{0|t}\boldsymbol{H}^\top + \boldsymbol{R},$$

$$q(\boldsymbol{y} \mid \boldsymbol{x}_t) = \mathcal{N}\big(\boldsymbol{y}; \boldsymbol{H}\boldsymbol{\mu}_{0|t},\ \boldsymbol{H}\boldsymbol{\Sigma}_{0|t}\boldsymbol{H}^\top + \boldsymbol{R}\big).$$

Taking the gradient of $\log q(\boldsymbol{y} \mid \boldsymbol{x}_t)$ with respect to $\boldsymbol{x}_t$ (noting that $\boldsymbol{\mu}_{0|t}$ depends linearly on $\boldsymbol{x}_t$ with gradient $\sqrt{\bar{\alpha}(t)}\,\boldsymbol{P}_f\boldsymbol{C}(t)^{-1}$) yields (A5).

Note: Equation (A5) can also be obtained directly from the MMPS Equation (23); the two are equivalent. □

Summing the prior score and the likelihood score yields the posterior score. Simplifying via the Woodbury identity allows one to identify its Gaussian structure and the corresponding moment parameters.

Theorem A.1 (Score-Based Derivation of EnKF)**.** *Let the prior be $q(\boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{\mu}_f, \boldsymbol{P}_f)$ and the likelihood be $q(\boldsymbol{y} \mid \boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{H}\boldsymbol{x}_0, \boldsymbol{R})$, where $\boldsymbol{H}$ is a linear observation operator. Under VP-SDE noising, the posterior score takes the following Gaussian form:*

$$\nabla_{\boldsymbol{x}_t} \log q(\boldsymbol{x}_t \mid \boldsymbol{y}) = -\big(\bar{\alpha}(t)\boldsymbol{\Sigma}_a + (1-\bar{\alpha}(t))\boldsymbol{I}\big)^{-1}\big(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}_a\big), \tag{A7}$$

where

$$\boldsymbol{\mu}_a = \boldsymbol{\mu}_f + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_f), \qquad \boldsymbol{\Sigma}_a = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{P}_f, \qquad \boldsymbol{K} = \boldsymbol{P}_f\boldsymbol{H}^\top(\boldsymbol{H}\boldsymbol{P}_f\boldsymbol{H}^\top + \boldsymbol{R})^{-1}. \tag{A8}$$

That is, $\boldsymbol{\mu}_a$ and $\boldsymbol{\Sigma}_a$ are precisely the analysis mean and covariance of the EnKF.

Proof. By Lemma A.2(ii), substituting $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $(\boldsymbol{\mu}_f, \boldsymbol{P}_f)$ gives the prior score:

$$\nabla_{\boldsymbol{x}_t} \log q(\boldsymbol{x}_t) = -\boldsymbol{C}(t)^{-1}\big(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}_f\big). \tag{A9}$$

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

By Proposition A.1, the likelihood score is given by (A5).

To simplify the posterior score, we introduce the following notation:

$$\begin{aligned}
\boldsymbol{C}_a &= \bar{\alpha}(t)\boldsymbol{\Sigma}_a + (1-\bar{\alpha}(t))\boldsymbol{I},\\
\hat{\boldsymbol{T}} &= \boldsymbol{P}_f\boldsymbol{H}^\top,\\
\boldsymbol{B} &= \boldsymbol{H}\boldsymbol{P}_f\boldsymbol{H}^\top + \boldsymbol{R},\\
\boldsymbol{S} &= \boldsymbol{H}\boldsymbol{\Sigma}_{0|t}\boldsymbol{H}^\top + \boldsymbol{R}.
\end{aligned}$$

From $\boldsymbol{\Sigma}_a = \boldsymbol{P}_f - \hat{\boldsymbol{T}}\boldsymbol{B}^{-1}\hat{\boldsymbol{T}}^\top$, we have

$$\boldsymbol{C}_a = \boldsymbol{C}(t) - \bar{\alpha}(t)\hat{\boldsymbol{T}}\boldsymbol{B}^{-1}\hat{\boldsymbol{T}}^\top, \qquad \boldsymbol{S} = \boldsymbol{B} - \bar{\alpha}(t)\hat{\boldsymbol{T}}^\top\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}.$$

By Lemma A.1,

$$\boldsymbol{C}_a^{-1} = \boldsymbol{C}(t)^{-1} + \bar{\alpha}(t)\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}\boldsymbol{S}^{-1}\hat{\boldsymbol{T}}^\top\boldsymbol{C}(t)^{-1}. \tag{A10}$$

Notice that

$$\begin{aligned}
\boldsymbol{C}_a\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}} &= \left(\boldsymbol{C}(t) - \bar{\alpha}(t)\hat{\boldsymbol{T}}\boldsymbol{B}^{-1}\hat{\boldsymbol{T}}^\top\right)\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}\\
&= \hat{\boldsymbol{T}} - \bar{\alpha}(t)\hat{\boldsymbol{T}}\boldsymbol{B}^{-1}\hat{\boldsymbol{T}}^\top\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}\\
&= \hat{\boldsymbol{T}}\boldsymbol{B}^{-1}(\boldsymbol{B} - \bar{\alpha}(t)\hat{\boldsymbol{T}}^\top\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}) = \hat{\boldsymbol{T}}\boldsymbol{B}^{-1}\boldsymbol{S},
\end{aligned}$$

hence $\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}\boldsymbol{S}^{-1} = \boldsymbol{C}_a^{-1}\hat{\boldsymbol{T}}\boldsymbol{B}^{-1}$.

Substituting the expression for $\boldsymbol{\mu}_{0|t}$ into (A5) and using the above relation, we simplify the likelihood score:

$$\begin{aligned}
\nabla_{\boldsymbol{x}_t}\log q(\boldsymbol{y}\mid\boldsymbol{x}_t) &= \sqrt{\bar{\alpha}(t)}\,\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}\boldsymbol{S}^{-1}\left(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_f - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{H}\boldsymbol{P}_f\boldsymbol{C}(t)^{-1}(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}_f)\right)\\
&= -\bar{\alpha}(t)\,\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}\boldsymbol{S}^{-1}\hat{\boldsymbol{T}}^\top\boldsymbol{C}(t)^{-1}\left(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}_f\right)\\
&\quad + \sqrt{\bar{\alpha}(t)}\,\boldsymbol{C}_a^{-1}\hat{\boldsymbol{T}}\boldsymbol{B}^{-1}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_f).
\end{aligned} \tag{A11}$$

Summing the prior score (A9) and (A11):

$$\begin{aligned}
\nabla_{\boldsymbol{x}_t}\log q(\boldsymbol{x}_t\mid\boldsymbol{y}) &= \nabla_{\boldsymbol{x}_t}\log q(\boldsymbol{x}_t) + \nabla_{\boldsymbol{x}_t}\log q(\boldsymbol{y}\mid\boldsymbol{x}_t)\\
&= -\left[\boldsymbol{C}(t)^{-1} + \bar{\alpha}(t)\boldsymbol{C}(t)^{-1}\hat{\boldsymbol{T}}\boldsymbol{S}^{-1}\hat{\boldsymbol{T}}^\top\boldsymbol{C}(t)^{-1}\right]\left(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}_f\right)\\
&\quad + \sqrt{\bar{\alpha}(t)}\,\boldsymbol{C}_a^{-1}\hat{\boldsymbol{T}}\boldsymbol{B}^{-1}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_f)\\
&= -\boldsymbol{C}_a^{-1}\left(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}_f\right) + \sqrt{\bar{\alpha}(t)}\,\boldsymbol{C}_a^{-1}\hat{\boldsymbol{T}}\boldsymbol{B}^{-1}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_f) \qquad \text{(by (A10))}\\
&= -\boldsymbol{C}_a^{-1}\left(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}(\boldsymbol{\mu}_f + \hat{\boldsymbol{T}}\boldsymbol{B}^{-1}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_f))\right)\\
&= -\left(\bar{\alpha}(t)\boldsymbol{\Sigma}_a + (1-\bar{\alpha}(t))\boldsymbol{I}\right)^{-1}\left(\boldsymbol{x}_t - \sqrt{\bar{\alpha}(t)}\,\boldsymbol{\mu}_a\right).
\end{aligned}$$

The last step uses $\boldsymbol{\mu}_a = \boldsymbol{\mu}_f + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_f)$ and $\boldsymbol{K} = \hat{\boldsymbol{T}}\boldsymbol{B}^{-1}$. □

https://doi.org/10.5194/egusphere-2026-5358
Preprint. Discussion started: 10 September 2026

Theorem A.1 shows that the posterior score has a Gaussian form. Applying the reverse PF-ODE solution from Lemma A.2(iii) directly yields the posterior sampling formula, establishing an explicit correspondence between the score-based method and the EnKF analysis ensemble update.

Corollary A.1 (Analysis Ensemble Update)**.** *From the score perspective, the posterior distribution is $\mathcal{N}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)$, where $\boldsymbol{\mu}_a$ and $\boldsymbol{\Sigma}_a$ are given by* (A8)*. Hence the posterior covariance of the score-based method is exactly the EnKF analysis covariance $\boldsymbol{\Sigma}_a$, and the analysis ensemble members are updated as*

$$\boldsymbol{x}_a^{(i)} = \boldsymbol{\mu}_a + \boldsymbol{\Sigma}_a^{1/2} \boldsymbol{P}_f^{-1/2} (\boldsymbol{x}_f^{(i)} - \boldsymbol{\mu}_f), \qquad i = 1, 2, \ldots, N_e. \tag{A12}$$

Proof. By Theorem A.1, the posterior score (A7) coincides with the score function of $\mathcal{N}(\sqrt{\bar{\alpha}(t)}\boldsymbol{\mu}_a, \bar{\alpha}(t)\boldsymbol{\Sigma}_a + (1 - \bar{\alpha}(t))\boldsymbol{I})$. By Lemma A.2(iii), the reverse PF-ODE solution of this Gaussian distribution at $\tau = 0$ (i.e., $t = 0$) is

$$\boldsymbol{x}_0 = \boldsymbol{\mu}_a + \boldsymbol{\Sigma}_a^{1/2} \boldsymbol{z}, \qquad \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}). \tag{A13}$$

Since $\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$, we immediately obtain $\boldsymbol{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)$, i.e., the posterior mean is $\boldsymbol{\mu}_a$ and the posterior covariance is $\boldsymbol{\Sigma}_a$.

For the prior ensemble samples $\{\boldsymbol{x}_f^{(i)}\}_{i=1}^{N_e}$, standardizing them as $\boldsymbol{z}^{(i)} = \boldsymbol{P}_f^{-1/2}(\boldsymbol{x}_f^{(i)} - \boldsymbol{\mu}_f)$ and substituting into (A13) yields (A12).

In practical ensemble Kalman filter implementations, one often uses the equivalent perturbation matrix transformation. Given $\boldsymbol{X}_f = (\boldsymbol{x}_f^{(1)} - \boldsymbol{\mu}_f, \ldots, \boldsymbol{x}_f^{(N_e)} - \boldsymbol{\mu}_f)$, the analysis ensemble mean and perturbations are updated separately as

$$\boldsymbol{x}_a^{(i)} = \boldsymbol{\mu}_a + \boldsymbol{X}_a(:, i), \qquad \boldsymbol{X}_a = \boldsymbol{X}_f \boldsymbol{T},$$

where $\boldsymbol{T}\boldsymbol{1} = \boldsymbol{0}$, $\boldsymbol{T}\boldsymbol{T}^\top = (N_e - 1)\tilde{\boldsymbol{P}}_a$, $\tilde{\boldsymbol{P}}_a = [(N_e - 1)\boldsymbol{I} + \boldsymbol{Y}_f^\top \boldsymbol{R}^{-1} \boldsymbol{Y}_f]^{-1}$, and $\boldsymbol{Y}_f = \boldsymbol{H}\boldsymbol{X}_f$. The construction of $\boldsymbol{X}_a$ ensures $\frac{1}{N_e - 1}\boldsymbol{X}_a \boldsymbol{X}_a^\top = \boldsymbol{\Sigma}_a$. Both formulations produce an ensemble representation of the same posterior distribution $\mathcal{N}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)$, i.e., the analysis ensemble members follow the same posterior distribution. □