Supplementary Information for: Variational multivariate data assimilation improves sea-ice analysis and forecasting

Vasily Ivanov Liubov Antiufrieva Vladimir Vanovskiy Evgeny Burnaev

Supplementary Note: Theoretical results and proofs

This note states and proves the consolidated theorem cited in the Methods section of the main manuscript.

Let $x_b \in \mathbb{R}^n$ be the background state, $y \in \mathbb{R}^m$ the observation vector, and $z \in \mathbb{R}^p$ the latent variable. Let

$$E : \mathbb{R}^n \to \mathbb{R}^p, \qquad D : \mathbb{R}^p \to \mathbb{R}^n, \qquad H : \mathbb{R}^n \to \mathbb{R}^m$$

be the encoder, decoder, and observation operator, and define $z_b = E(x_b)$. Consider the objective

$$\mathscr{L}(z) = w_y \|H(D(z)) - y\|_2^2 + w_b \|D(z) - x_b\|_2^2 + w_z \|z - z_b\|_2^2, \tag{1}$$

where $w_y, w_b, w_z > 0$; this is the general multivariate form of the objective, with the background penalty acting on all decoded fields (the specialization to the concentration-only penalty used in the reported experiments is given after the proof). Assume that (1) admits a global minimizer

$$z^\star \in \arg\min_{z \in \mathbb{R}^p} \mathscr{L}(z).$$

Set $x_a = D(z^\star)$.

For the multivariate decoder, write the decoder output as a block decomposition

$$D(z) = (D_1(z), \ldots, D_K(z)),$$

where each map $D_j : \mathbb{R}^p \to \mathbb{R}^{n_j}$ corresponds to one physical field and $n_1 + \cdots + n_K = n$. Assume in addition that, for each $j \in \{1, \ldots, K\}$, the component map D_j is Lipschitz on a neighborhood of z_b containing $z^\star$: there exists $L_j > 0$ such that

$$\|D_j(z) - D_j(z')\|_2 \le L_j \|z - z'\|_2 \qquad \text{for all } z, z' \text{ in that neighborhood.} \tag{2}$$

Denote the background-reconstruction and analysis components by

$$x_{0,j} = D_j(z_b), \qquad x_{a,j} = D_j(z^\star).$$

Here $x_{0,j}$ is the j-th block of the decoded reconstruction $D(z_b)$ of the background; it differs from the corresponding block of x_b itself by the encoder–decoder reconstruction error.

Theorem 1 (Latent regularization constrains decoded multivariate increments)**.** *Even when the observation operator H acts on only part of the state, the shared latent representation together with the reconstruction and latent regularization terms in $\mathscr{L}$ constrain the magnitude of the induced correction in every decoded field $x_{a,j} = D_j(z^\star)$, as well as in the full analysis $x_a = D(z^\star)$ and the optimized latent state $z^\star$. Quantitatively, all four magnitudes are bounded by the observation mismatch and reconstruction error evaluated at the encoded background z_b: under the assumptions above* *and the componentwise Lipschitz condition* (2)*,*

$$\mathscr{L}(z^\star) \le w_y\|H(D(z_b))-y\|_2^2 + w_b\|D(z_b)-x_b\|_2^2, \tag{3}$$

$$\|x_a - x_b\|_2 \le \sqrt{\frac{w_y\|H(D(z_b))-y\|_2^2 + w_b\|D(z_b)-x_b\|_2^2}{w_b}}, \tag{4}$$

$$\|z^\star - z_b\|_2 \le \sqrt{\frac{w_y\|H(D(z_b))-y\|_2^2 + w_b\|D(z_b)-x_b\|_2^2}{w_z}}, \tag{5}$$

$$\|x_{a,j} - x_{0,j}\|_2 \le L_j\sqrt{\frac{w_y\|H(D(z_b))-y\|_2^2 + w_b\|D(z_b)-x_b\|_2^2}{w_z}} \qquad (j=1,\dots,K). \tag{6}$$

If, in addition, $D(z_b) = x_b$ (so that $x_{0,j} = x_{b,j}$ for every j), then

$$\mathscr{L}(z^\star) \le w_y\|H(D(z_b))-y\|_2^2, \tag{7}$$

$$\|x_a - x_b\|_2 \le \sqrt{\frac{w_y}{w_b}}\,\|H(D(z_b))-y\|_2, \tag{8}$$

$$\|z^\star - z_b\|_2 \le \sqrt{\frac{w_y}{w_z}}\,\|H(D(z_b))-y\|_2, \tag{9}$$

$$\|x_{a,j} - x_{0,j}\|_2 \le L_j\sqrt{\frac{w_y}{w_z}}\,\|H(D(z_b))-y\|_2 \qquad (j=1,\dots,K). \tag{10}$$

Proof. Since $z^\star$ is a global minimizer,

$$\mathscr{L}(z^\star) \le \mathscr{L}(z) \qquad \text{for all } z \in \mathbb{R}^p.$$

Choosing $z = z_b$ and using $\|z_b - z_b\|_2^2 = 0$, we obtain

$$\mathscr{L}(z^\star) \le \mathscr{L}(z_b) = w_y\|H(D(z_b))-y\|_2^2 + w_b\|D(z_b)-x_b\|_2^2,$$

which proves (3).

Next, all terms in (1) are nonnegative. Hence

$$w_b\|D(z^\star)-x_b\|_2^2 \le \mathscr{L}(z^\star), \qquad w_z\|z^\star - z_b\|_2^2 \le \mathscr{L}(z^\star).$$

Combining these inequalities with (3), dividing by w_b and w_z, respectively, and taking square roots gives (4) and (5).

For the componentwise bound, fix any $j \in \{1,\dots,K\}$. By definition, $x_{a,j} - x_{0,j} = D_j(z^\star) - D_j(z_b)$. Applying the Lipschitz property (2) with $z = z^\star$ and $z' = z_b$,

$$\|x_{a,j} - x_{0,j}\|_2 = \|D_j(z^\star) - D_j(z_b)\|_2 \le L_j\|z^\star - z_b\|_2.$$

Substituting the latent bound (5) yields (6).

Finally, if $D(z_b) = x_b$, then the reconstruction term in (1) vanishes at z_b, so (3) reduces to $\mathscr{L}(z^\star) \le w_y\|H(D(z_b)) - y\|_2^2$. Substituting this simpler upper bound into the divide-by-w_b, divide-by-w_z, and Lipschitz steps above gives the four simplified bounds. □

Specialization to the concentration-only penalty. The experiments reported in the main manuscript restrict the background penalty to the sea-ice-concentration field: the background term in (1) is replaced by $w_b\|P_{\mathrm{sic}}(D(z)-x_b)\|_2^2$, where P_{sic} is the linear selector of the concentration field on valid ocean cells. The proof applies verbatim to this restricted objective, with the initial loss $\mathscr{L}(z_b)$ on the right-hand sides evaluated for the restricted background term: the total-loss bound (3), the latent bound (5), and the per-field bounds (6) are unchanged, while the background bound (4) then controls only the penalized concentration component $\|P_{\mathrm{sic}}(x_a - x_b)\|_2$. No full-state background bound follows from the restricted background term; the unobserved fields are constrained through the latent bound and the Lipschitz constants alone. The empirical comparison of the two variants is given in the Supplementary Note on the background-penalty ablation.

Interpretation. The theorem shows that the analysis x_a, the optimized latent state $z^\star$, and every individual decoded field inherit explicit bounds through the regularization terms and the decoder Lipschitz constants; the intuitive trust-region reading is given after the theorem statement in the main manuscript. In the concentration-only variant used in the reported experiments, the background term directly penalizes only the concentration field; the remaining decoded fields are then constrained indirectly, through the latent term and the per-field Lipschitz constants.

Supplementary Note: Background-penalty ablation

This note compares the multivariate background penalty of the assimilation loss (Methods, main manuscript) with its restriction to the sea-ice-concentration field (the *concentration-only* penalty) adopted in all reported experiments. Each of the five four-field VAE variants of Supplementary Table S2 was run twice in paired 200-day model-to-model experiments that were identical except for the background term; within each pair the background-model metric series coincide.

Skill is defined as in the Methods section of the main manuscript; we report Δskill, the multivariate-penalty skill minus the concentration-only skill, with positive values favoring the multivariate penalty. Paired 95% percentile intervals use the same moving-block bootstrap as in the Methods section (14-day blocks, 5000 replicates), resampling both runs jointly. Dense metrics are evaluated on the full model grid and track metrics at the SRAL-track locations; IIEE denotes the integrated ice-edge error, the fraction of grid cells whose 0.15-thresholded ice masks disagree. Supplementary Table S1 summarizes the results.

Supplementary Table S1. Paired 200-day model-to-model comparison of the multivariate and concentration-only background penalties over the five four-field VAE variants. Positive Δskill favors the multivariate penalty. SIC denotes sea-ice concentration (*siconc*), ice-temperature sea-ice temperature (*sitemp*), ice-thickness sea-ice thickness (*sithic*), and SST sea-surface temperature (*sosstsst*); IIEE denotes the integrated ice-edge error.

Metric	Range of Δskill	Pairs with interval excluding zero
Dense SIC RMSE	−0.021 to −0.007	4 of 5
Dense SIC MAE	−0.040 to −0.014	5 of 5
Dense SIC IIEE	−0.077 to −0.052	5 of 5
Track SIC RMSE	−0.027 to −0.001	1 of 5
Track SIC IIEE	−0.096 to −0.046	5 of 5
Ice-temperature RMSE	+0.002 to +0.043	4 of 5
Ice-thickness RMSE	−0.046 to +0.031	0 of 5
SST RMSE	+0.007 to +0.064	1 of 5

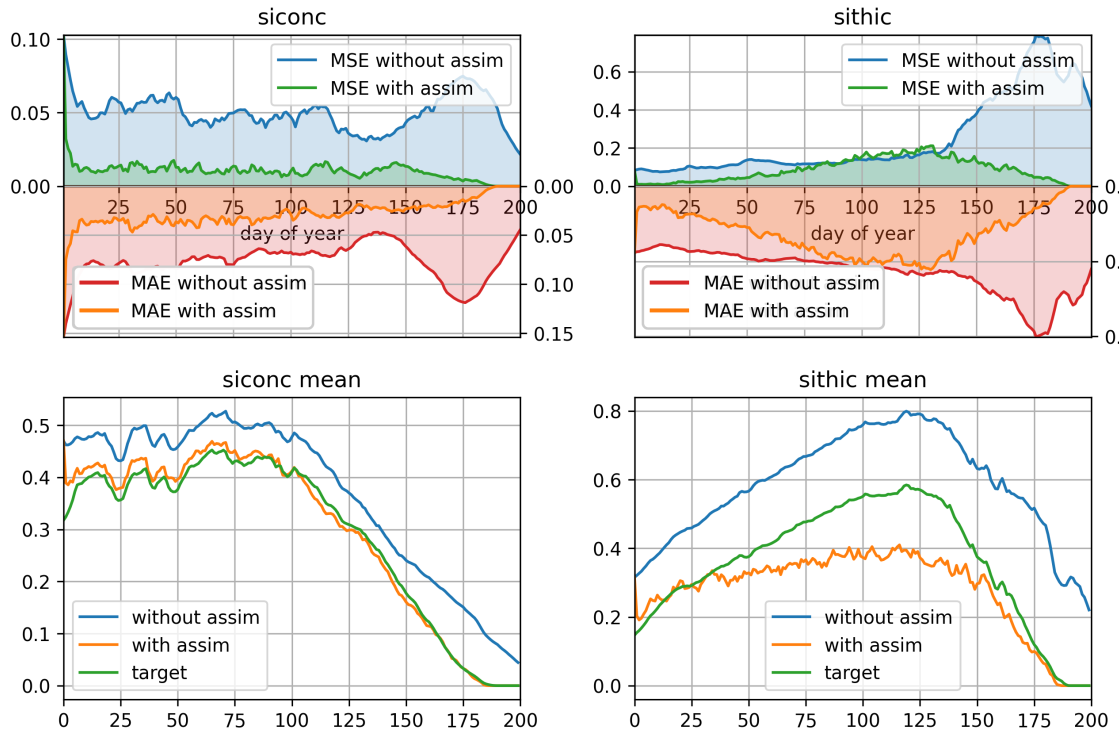
The multivariate penalty reduced sea-ice-concentration skill in all five variants, most strongly at the ice edge, while improving the unobserved ice-temperature analysis; thickness and SST differences were mostly not significant. For the selected *vae_4f_um_multiscale_sobel_binary* checkpoint, dense concentration RMSE, MAE, and IIEE skill changed by −0.021, −0.035, and −0.057, against +0.043 in ice-temperature RMSE skill. Concentration and the ice edge are what this study targets, so we use the concentration-only penalty throughout.

Two caveats. The background term carries a single configured weight, so averaging over four fields instead of one cuts the effective per-entry weight on concentration roughly fourfold; the comparison is therefore between the two implemented objectives, not between penalties at matched concentration weight. The satellite-to-model and forecast-restart experiments used the concentration-only penalty exclusively.

Supplementary Note: Twin forecast extension experiment

We performed a twin forecast experiment, a standard validation protocol in data assimilation in which the verification trajectory is taken from a separate model integration. We initialized from the 2022 state and compared two runs using 2023 atmospheric forcing and boundary conditions: a baseline run, and a second run that additionally assimilated 2023 observations into the 2022 initial state. This configuration does not constitute independent observational validation, but it tests whether sparse assimilation increments can improve a forecast-like trajectory under controlled conditions.

In this twin setting, for *siconc*, the MSE decreased from 0.050 to 0.010 and the MAE decreased from 0.077 to 0.028. For *sithic*, the MSE decreased from 0.250 to 0.085 and the MAE decreased from 0.229 to 0.123. The resulting time series is shown in Supplementary Fig. S1.



Supplementary Fig. S1. Twin forecast-extension diagnostics over 200 forecast days for sea-ice concentration (left column) and sea-ice thickness (right column). In the top row, MSE is plotted above zero (blue, baseline without assimilation; green, assimilated run), whereas MAE is mirrored below zero (red, baseline; orange, assimilated run) and should be read by magnitude from the right-hand axis. The bottom row shows the spatial mean of each field for the baseline (blue), assimilated run (orange), and verification target (green). Sea-ice concentration is a fraction and sea-ice thickness is in metres; MSE therefore has squared field units, whereas MAE and the spatial means retain the field units. Both forecasts start from the 2022 state and use 2023 atmospheric forcing and boundary conditions; the assimilated run additionally assimilates 2023 observations into the 2022 initial state before integration.

Supplementary Note: Classical baseline configurations

This note lists the parameter settings of the classical data assimilation baselines (3D-VAR, BLUE, and EnKF) used in the main manuscript. The covariance and localization choices were tuned on the same development setup as the VAE experiments.

The 3D-VAR background-error covariance used a compact-support polynomial correlation function (referenced in the Methods section of the main manuscript) with length-scale parameter 100 km and background-error variance parameter 0.001; 3D-VAR was optimized for 20 LBFGS steps. For the model-to-model experiments, the 3D-VAR observation-error variance parameter was 0.001. BLUE used the same background covariance, observation standard deviation 0.02, and Tikhonov regularization 10^{-6}. EnKF used 40 ensemble members with spatially correlated background perturbations whose correlation radius of 25 grid cells is comparable to the 100-km covariance length scale, background standard deviation 0.05, observation standard deviation 0.05, inflation 1.0, and polynomial localization with length-scale parameter 100 km. In the satellite-to-model comparison, the classical observation-error settings were scaled to match the stronger observation influence used for the VAE experiments: 3D-VAR used observation-error variance parameter $0.001/6$, BLUE used observation standard deviation $0.02/\sqrt{6}$, and EnKF used observation standard deviation $0.05/\sqrt{6}$. All classical baseline analyses were clipped to the physical sea-ice concentration range $[0, 1]$.

Supplementary Tables

Supplementary Table S2. Model-to-model assimilation results for fields other than sea-ice concentration. SIT denotes sea-ice thickness (*sithic*), SST sea-surface temperature (*sosstsst*), and ice-temp. sea-ice temperature (*sitemp*). The persistence row reports the error of the unassimilated background field against the same verification target. Entries marked "–" correspond to baselines or single-field VAE variants that do not produce a coupled multivariate analysis for that field.

model	SIT MAE	SST MAE	ice-temp. MAE	SIT RMSE	SST RMSE	ice-temp. RMSE
persistence	0.1782	0.6448	1.3872	0.2959	0.9274	2.2068
EnKF	–	–	–	–	–	–
vae_1f	–	–	–	–	–	–
vae_1f_d512	–	–	–	–	–	–
vae_3f_um_emb	0.1093	–	1.3138	0.1807	–	2.1233
vae_4f_um	0.1012	0.4489	1.2759	0.1707	0.6447	2.0893
vae_4f_um_multiscale	0.1078	0.4559	1.2976	0.1796	0.6528	2.1422
vae_4f_um_multiscale_block	0.1111	0.5449	1.3032	0.1877	0.7835	2.1323
vae_4f_um_sobel_binary	0.1078	0.5212	1.3195	0.1802	0.7587	2.1593
vae_4f_um_multiscale_sobel_binary	0.1080	0.5875	1.3571	0.1811	0.8575	2.2051
3D-VAR	–	–	–	–	–	–
BLUE	–	–	–	–	–	–

Supplementary Table S3. VAE reconstruction accuracy. SIC denotes sea-ice concentration (*siconc*), SIT sea-ice thickness (*sithic*), SST sea-surface temperature (*sosstsst*), and ice-temp. sea-ice temperature (*sitemp*).

model	SIC MAE	SIC RMSE	SIT MAE	SST MAE	ice-temp. MAE
vae_1f	0.0081	0.0255	–	–	–
vae_1f_d512	0.0227	0.0675	–	–	–
vae_3f_um_emb	0.0271	0.0757	0.0891	–	0.0849
vae_4f_um	0.0242	0.0688	0.0830	0.2643	0.0860
vae_4f_um_multiscale	0.0306	0.0699	0.0783	0.2342	0.0924
vae_4f_um_multiscale_block	0.0246	0.0701	0.0984	0.3918	0.0948
vae_4f_um_sobel_binary	0.0213	0.0665	0.0963	0.2880	0.0894
vae_4f_um_multiscale_sobel_binary	0.0214	0.0654	0.0893	0.3036	0.0998