Environmental Data Science (2026), 5: e36, 1–10
doi:10.1017/eds.2026.10058

CAMBRIDGE UNIVERSITY PRESS

METHODS PAPER

Spatiotemporal coherent displacements for ensemble-based neural data assimilation

Shashank Kumar Roy and Ronan Fablet

Lab-STICC, IMT Atlantique, Brest, France
Corresponding author: Shashank Kumar Roy; Email: shashank.roy@imt-atlantique.fr

Received: 26 June 2026 UTC; **Revised:** 26 June 2026 UTC; **Accepted:** 04 July 2026 UTC

Keywords: 4DVarNet; data assimilation; deep learning; neural networks; uncertainty quantification

Abstract

While deep learning (DL) methods have expanded the approach to solving state estimation problems in data assimilation (DA), a robust approach requires extending them for uncertainty quantification. In this work, we propose a method to represent location-based uncertainty through random, spatiotemporally coherent displacements in physical space, which are used to sample states in the vicinity of a given reference state. We integrate this strategy into an end-to-end variational DA framework, termed 4DVarNet-LU, and compare its performance with a UNet baseline, demonstrating improved reconstruction skill through the explicit use of prior state information. Trained deterministically in output and stochastically in input, we demonstrate that at inference, we see improvement in skill when using information from initial conditions. Using observation system simulation experiments based on a single-layer quasi-geostrophic model with sparse pseudo-nadir altimetry observations, we evaluate reconstruction mean and ensemble skill. Results show that 4DVarNet-LU consistently outperforms the UNet baseline in terms of root mean square error (RMSE) and continuous ranked probability score (CRPS), with increased robustness to initial condition bias and improved temporal stability. While the method does not propose to account for all different uncertainty sources in DA, it provides a structured way for modeling and training DL models for dealing with location-based uncertainty. Future work may extend this approach to account for additional sources of uncertainty and end-to-end ensemble generation.

Impact Statement

This research paper discusses a method for uncertainty quantification via ensemble generation for training an end-to-end data assimilation (DA) model under the 4DVarNet framework and UNet. To this end, this paper describes a novel method for ensemble generation for spatiotemporal data in the context of geosciences. It demonstrates the approach to two neural network models, 4DVarNet and UNet, that are trained for reconstruction tasks in training and shows the utility in producing ensemble outputs for uncertainty quantification. Its primary goal is to show how such structural choices that are inspired by the problem domain at hand are not only better at hidden state estimation problems but also pave the way for uncertainty quantification. The broader community may find this desirable for application and research in DA and geosciences.

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

1. Introduction

In geosciences, deep learning (DL) has emerged as a new way of approaching state reconstruction and forecasting problems. Through a training phase, deep neural architectures capture complex nonlinear relationships within a vast amount of spatiotemporal data (LeCun et al., 2015). By contrast, traditional physically based methods rely on data assimilation (DA) frameworks (Trémolet, 2007; Carrassi et al., 2018). They explicitly leverage physical knowledge, but they are often computationally expensive for multiple runs. They also generally rely on simplified models and assumptions to remain tractable (Blum et al., 2009; Asch et al., 2016; Bannister, 2017; Carrassi et al., 2018).

The DA community has also witnessed a continuously evolving landscape of model architectures for developing applications with a focus on improving different components and algorithms in the DA pipeline (Cheng et al., 2023). End-to-end neural architectures tailored to the structural knowledge of the underlying DA problem have seen new architectural discoveries and refinement of those suited best to the task (Fablet et al., 2021; Beauchamp et al., 2023a; Farchi et al., 2023; Martin et al., 2025).

However, the application of DL in DA is not without challenges: a critical challenge lies in uncertainty quantification and the model's generalization ability to unseen data. It is thus important to obtain both the best estimate and to quantify the associated uncertainty in DA (van Leeuwen et al., 2025). Ensemble methods in particular have been the cornerstone of physics-based operational DA systems (Palmer and Zanna, 2013; Carrassi et al., 2018). They rely on stochastic perturbation of physical parameters and perturbed initial conditions that are used to initialize the model (Palmer and Zanna, 2013; Bannister, 2017). In DL, modern methods have gained importance to generate new states from noise, such as probabilistic diffusion models (Ho et al., 2020). Their ability to sample physically consistent states in inverse problem setting remain a scientific challenge (Cheng et al., 2023; Martin et al., 2025).

Here, we investigate how to embed uncertainty quantification in end-to-end neural DA schemes, namely 4DVarNets introduced by Fablet et al. (2021). Leveraging physically consistent random perturbation schemes, we introduce a novel ensemble neural DA scheme, 4DVarNet-LU, extending 4DVarNets (Fablet et al., 2021) to uncertainty quantification issues. The considered random perturbations model location-based uncertainties through spatiotemporal coherent displacements in physical space using Gaussian random fields. We demonstrate that training with an initial state that reflects both prior knowledge and positional error can be used to obtain an ensemble of reconstructions from a given set of partial observations at inference. We exploit the approach to explore how such schemes could help in reducing reconstruction biases and uncertainties of DA systems. We report for one-layer quasi-geostrophic dynamics (Frezat et al., 2022) with sparse observations, seen as an intermediate-complexity testbed representation of sea surface dynamics (Le Guillou et al., 2023). Our results support the relevance of the end-to-end neural DA to enhance ensemble reconstruction when dealing with partially observed systems.

2. Methodology

We briefly introduce the problem of variational DA that we focus on in this work, called the weak-4DVar. This is the most general form of variational DA since the underlying physical model is rarely perfect (Trémolet, 2007; Fletcher, 2017). The objective here is to find the state sequence over a time window, denoted by $X^i = \{x_1, x_2, .., x_n\}$ that fits the sparse and noisy observations $Y^i = \{y_1, y_2, \ldots y_n\}$ over a time window $t_1 : t_n$ observed on some domain $\Omega = \{\Omega_1, \Omega_2, .., \Omega_n\}$. This translates to an inverse problem solved by the minimization of the weak-constraint 4DVar cost $\mathcal{J}$, given by

$$\mathcal{J}\left(X^i, Y^i\right) = \lambda_{ob} \| X^i - Y^i \|_\Omega + \lambda \| X^i - M\left(X^i\right) \|. \tag{2.1}$$

While Equation (2.1) is minimized, it is computationally expensive for large-scale optimization using adjoint computations. The computational complexity obviously increases to ensemble DA with such model-based variational schemes.

End-to-end trainable DL methods present an alternative approach for such inverse problems in space–time dynamical systems. It reformulates the inverse problem as learning a neural operator Ψ_θ to predict the

target state X^t directly from observations and partial state information. This results in the optimization problem (2.2):

$$\min_{\theta} \mathbb{E}_{X^i,X^t,Y}\left[\|X^t - \widehat{X}\|^2\right] \text{ s.t.} \widehat{X} = \Psi_\theta\left(X^i, Y^i\right). \tag{2.2}$$

While generic architectures such as UNets can be used for this purpose, recent work has shown that embedding variational DA principles within neural architectures improves physical consistency and generalization (Fablet et al., 2021; Beauchamp et al., 2023a). Building on this perspective, we now introduce the two end-to-end DL models that we use in this work.

2.1. 4DVarNet

4DVarNet is an end-to-end, data-driven neural network approach that learns both the 4DVar variational cost and its solver together (Fablet et al., 2021). It has a neural-network-parameterized 4DVar cost as part of the architecture given by

$$U_\Phi(X,Y) = \|H(X) - Y\|^2 + \lambda\|X - \Phi(X)\|^2, \tag{2.3}$$

where the first term contains the observation operator H and measures the misfit between the estimated state X and the observations Y. The second term is a regularization term $\lambda\|X - \Phi(X)\|$, which is an autoencoder that can be interpreted as a learned representation for the state sequence in a low-dimensional manifold. A trainable, long short-term memory (LSTM)-based gradient-based solver is used to minimize this cost iteratively as

$$X^n = X^{n-1} - ConvLSTM_\Gamma\left(\nabla_X U_\Phi\left(X^{n-1}, Y\right)\right). \tag{2.4}$$

The end-to-end training of 4DVarNet leads to a bi-level optimization problem (Da Li et al., 2018; Schwartz et al., 2019). The objective is to solve the inner minimization problem, which is achieved by minimizing the reconstruction performance criterion defined in Equation (2.2) during training.

2.2. 4DVarNet-LU: position-based uncertainty via *coherent random displacements*

In DA literature, it is known that the quality of initial prior state information is reflected in the quality of reconstruction for the unknown state (Carrassi et al., 2018). Hence, we propose the modified 4DVarNet architecture to output the reconstructed state by taking both the initial guess and the observations $\left(X^0, Y\right)$ as input to predict the true state X, which we refer to as 4DVarNet-LU, which we describe in Section 2.2. It differs from the standard 4DVarNet (Fablet et al., 2021; Beauchamp et al., 2023a) where the state X^0 is initialized directly from the sparse observations. A conceptual overview of the modified architecture is illustrated in Figure 1, where the initial state estimate X^0 is obtained by randomly sampling states around the true state via coherent displacements in position.

Position-based errors are well known in operational DA, where dynamically consistent structures (e.g., eddies or vortices) may be correctly represented but displaced in space and/or time. Such misalignment errors lead to the well-known double-penalty issue in DA (Lawson and Hansen, 2005; Nehrkorn et al., 2015; Bocquet et al., 2024).

Motivated by this, we propose a method to generate an ensemble of initial states by coherently perturbing a reference state, tailored to reconstruction problems in DA. To preserve physical consistency over the assimilation window, the induced errors are required to be both spatially and temporally correlated. This motivates the use of a Gaussian random field to sample a coherent displacement field, in contrast to diffusion-based probabilistic approaches (Ho et al., 2020), which rely on unstructured random noise that is subsequently mapped to physical space.

To generate a new perturbed state X^{ptb} from the true state X, we sample a coherent random displacement field $S^i\left(\vec{x}, t\right)$ from a GRF, denoted by $\mathcal{GRF}\left(\sigma_{disp}, L, \mathcal{T}\right)$ and use it to displace the field X in space and time, given by

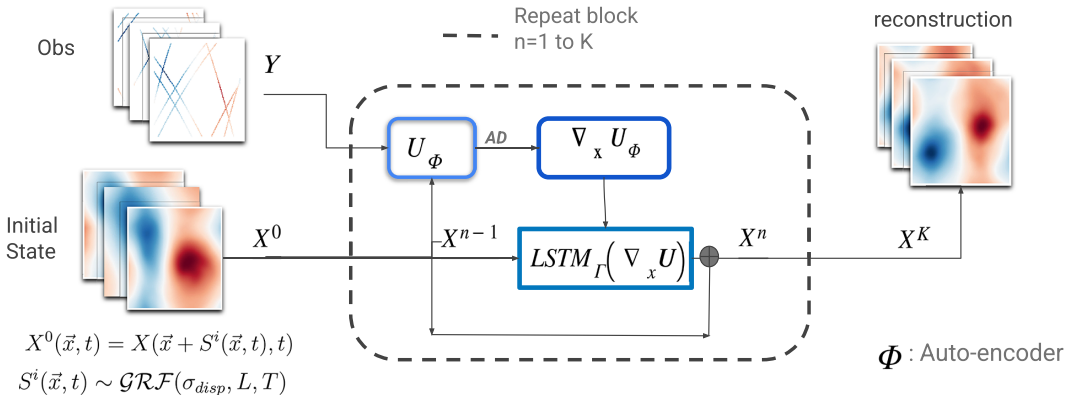


Figure 1. *Schematic of 4DVarNet-LU showing the adaptation of original architecture* (Fablet et al., 2021) *to include an initialization method based on a random coherent displacement model.*

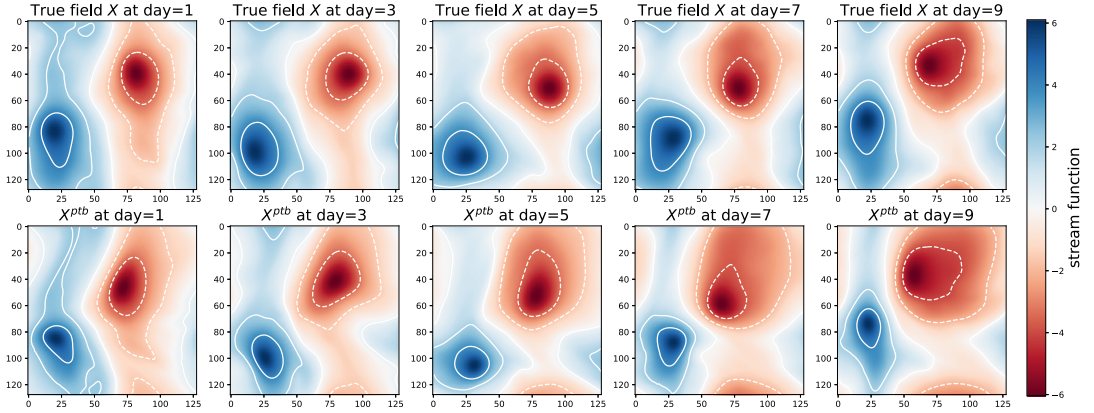


Figure 2. *The top row illustrates the true stream function field. The second row shows the perturbed stream function field obtained using the coherent random perturbations with* $L = 5.0, \mathcal{T} = 20.0$, *and* $\sigma_{disp} = 7.0$. *The contour lines show iso-curves at five levels.*

$$X^{ptb}\left(\vec{x},t\right) = X\left(\vec{x} + S^i\left(\vec{x},t\right),t\right)$$
$$S^i\left(\vec{x},t\right) = \left(S_x^i\left(\vec{x},t\right), S_y^i\left(\vec{x},t\right)\right) \sim \mathcal{GRF}\left(\sigma_{disp}, L, \mathcal{T}\right). \tag{2.5}$$

The parameter σ_{disp} controls the standard deviation of the field. L and $\mathcal{T}$ are the spatial and temporal correlation lengths, and together control the spatiotemporal consistency of $S^i\left(\vec{x},t\right)$ across the DA window. An illustrative figure of the stream function field is shown in Figure 2. As the displacement is spatially nonuniform, the mean displacement magnitude at any position $\vec{x}$ is given by

$$\mathrm{E}_{\vec{x},t}\left[\left\|S\left(\vec{x},t\right)\right\|\right] = \mathrm{E}_{\vec{x},t}\left[\sqrt{S_x\left(\vec{x},t\right)^2 + S_y\left(\vec{x},t\right)^2}\right] = \sqrt{2}\sigma_{disp}. \tag{2.6}$$

We refer to the obtained state $X^{ptb}\left(\vec{x},t\right)$ as a perturbed state corresponding to the true state $X\left(\vec{x},t\right)$. For a given observation sequence Y, Equation (2.5) enables sampling different initial state realizations X^0 from the true target state X to train 4DVarNet-LU. In training, we sample σ_{disp} from a uniform distribution $\mathcal{U}(0,\sigma_{max})$ so that the model learns to adapt to the quality of the initial state. As in traditional ensemble DA,

this approach generates the ensemble by running the models with multiple initial states independently with the same set of observations, rather than from stochasticity in the model parameters or training procedure. More details about the training and inference configurations can be found in the Supplementary Material.

2.3. UNet

We now briefly describe the UNet architecture that is used in this work. We use a standard convolutional UNet (Ronneberger et al., 2015) as a feed-forward baseline to provide a direct mapping from observations Y to the target stream function X, each of shape $[T,H,W]$. Next, we modify the input for UNet where we concatenate (Y,X^{ptb}) along the channel dimension. The architecture follows an encoder–decoder structure with skip connections and produces a single output tensor of shape $[T,H,W]$. This model is trained end-to-end in the same training configuration, training data, and loss function as 4DVarNet-LU. This UNet performs a single-pass feed-forward prediction and does not incorporate an explicit variational formulation or iterative solver structure.

3. Result

Performance is evaluated in an observation system simulation experiment (OSSE) framework based on a single-layer quasi-geostrophic model with periodic boundary conditions on a 128×128 grid. Sparse observations are generated by applying a pseudo-nadir altimetry track mask to the stream function fields. An example of the inputs, reconstructions, and reference states is shown in Figure 3. We present results for 4DVarNet-LU and UNet using identical training and testing configurations, where the training is restricted to sample $\sigma_{disp} \sim \mathcal{U}(0,\sigma_{max})$, $\sigma_{max} = 15.0$, and tested systematically for different levels of perturbations σ_{disp}. This value of σ_{max} is chosen empirically to be commensurate with the radius of the eddy features, thereby constraining perturbations to a physically plausible range of spatial variability in this system.

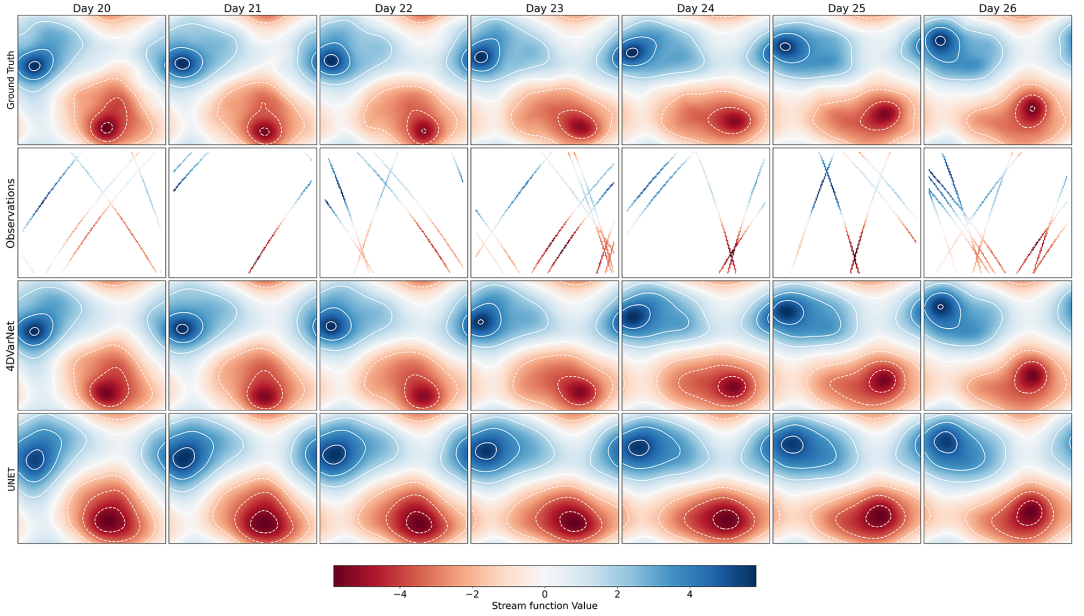


Figure 3. *The first row shows the true stream function X for 7 consecutive days in the test period of 81 days, starting from day 20. The second row displays the sparse pseudo-nadir altimetry tracks as observations Y. The third and fourth are the 4DVarNet-LU and UNet mean obtained from ensemble of size 25, respectively.*

We first consider an unbiased ensemble, obtained by sampling the first-guess state at inference around X^t, followed by a biased initial ensemble, where a biased state X^{bias} is constructed from the true state and ensemble members are sampled around X^{bias} for different values of σ_{disp}. Reconstruction quality is assessed using *RMSEs* score of ensemble mean reconstruction $\hat{X}(t)$ over the test period of 81 days. We compute its mean μ and standard deviation $\sigma[RMSEs(t)]$, where

$$RMSEs(t) = 1 - \frac{RMSE\left(\hat{X}(t)\right)}{RMS(X)}, \quad \mu = \overline{RMSEs(t)}, \quad \sigma[RMSEs(t)]. \tag{3.1}$$

To evaluate the ensembles, we use the spatially averaged continuous ranked probability score (CRPS), which penalizes both bias and overconfidence in the ensemble (Hersbach, 2000). For an ensemble of size N with members $\{x_j\}_{j=1}^{N}$ and reference truth (or analysis) x^t, the CRPS is defined as

$$\mathrm{CRPS}\left(\{x_j\}, x^t\right) = \frac{1}{N}\sum_{j=1}^{N}|x_j - x^t| - \frac{1}{2N(N-1)}\sum_{j=1}^{N}\sum_{k=1}^{N}|x_j - x_k|. \tag{3.2}$$

Importantly, we evaluate the model beyond σ_{max} at test time to assess its ability to generalize to more extreme perturbations. At inference, each ensemble member is conditioned on an independent observation noise realization, analogous to the perturbed-observation approach in the ensemble Kalman filter (EnKF) (Burgers et al., 1998). Notably, we also consider the cases where both 4DVarNet and UNet models are trained directly on observations to the ground truth.

3.1. Reducing reconstruction uncertainties from a first guess

To illustrate the mean state from 4DVarNet-LU & UNet, we plot the mean reconstruction of 25 ensemble members in Figure 3. From the left panel of Figure 4, we see that even if individual initial conditions are away from the true state, the ensemble mean obtained from these models learn to correctly reduce the errors beyond those seen in training, notably $\sigma_{ptb} = 20, 25$. Here, the spread represents the standard deviation $\sigma[RMSEs(t)]$ of the ensemble mean over the test period. We also notice a threshold for both models as σ_{disp} approaches 0, where location errors smaller than this do not gain improvement. For UNet, this threshold is between $\sigma_{disp} = 5$ and $\sigma_{disp} = 10.0$, whereas for 4DVarNet-LU it is observed below $\sigma_{disp} = 5.0$.

While 4DVarNet-LU consistently performs better than UNet at all perturbation levels, both models show similar trend: we see that their reconstruction scores are fairly stable with a slight but gradual

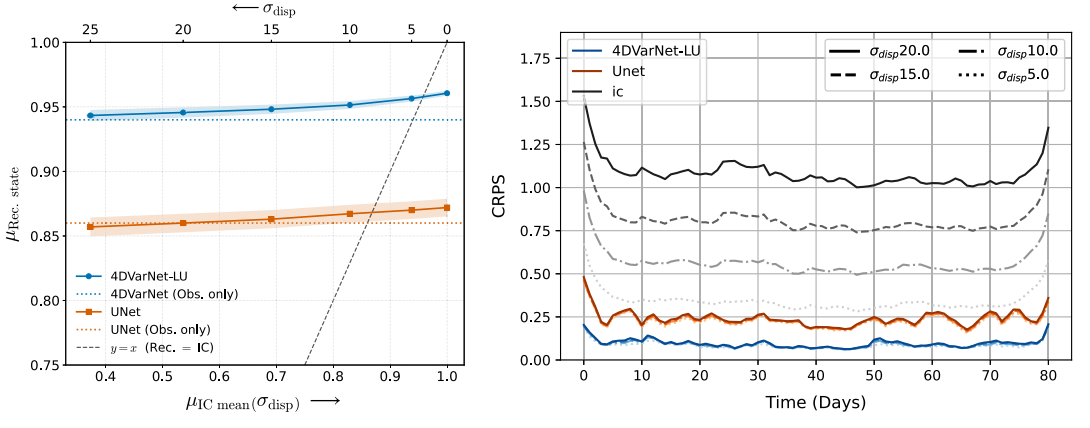


Figure 4. *Ensemble mean performance μ for 4DVarNet-LU and UNet as a function of test-time perturbation amplitude* σ_{disp}. *The plot highlights the robustness of 4DVarNet and UNet to increasing perturbations in the initial state.*

decrease in the accuracy for higher displacement levels σ_{disp}, indicating the impact of error in the initial guess. This behavior is expected since, given a sequence of observations, a better initial condition should need less correction.

The second panel for CRPS in Figure 4 shows the temporal structure of CRPS, which also complements the stability of the ensemble for both models. The CRPS value is not affected by the level of perturbation, although it degrades by a very small amount, which is not visible in the plot. For the exact value of these details, we include a separate tables in the Supplementary Material for both metrics. Figure 11 in the Supplementary Material illustrates the visual difference across the ensemble for reconstruction.

3.2. Biased DA ensemble as first guess: out-of-distribution generalization

To assess the robustness of this approach outside the training distribution, we look at the performance of both 4DVarNet-LU and UNet when the initial ensemble mean is biased. We introduce it by imposing a constant spatially coherent displacement on the true state consistent across all assimilation windows, yielding the average positional shift value b. Figure 5 shows the ensemble mean performance over the $(\sigma_{\text{disp}}, b)$ in parameter space. We see a gradual decrease in performance for both models, highlighting the joint sensitivity to initial ensemble bias and spread. We observe that the performance is highest for zero bias in the initial ensemble. Whereas if the initial ensemble is biased, there is a gradual decline in the reconstruction performance. This expected behavior points to the fact that training was entirely in the unbiased initial state scenarios. Further increases in bias magnitude result in limited additional degradation, indicating the impact of observations compensating for the error in the initial condition.

Figure 6 shows the spatially averaged CRPS for 4DVarNet-LU and UNet as a function of time for different levels of biased initial ensemble. While 4DVarNet-LU is better than UNet, for any fixed bias, both show a gradual increase in CRPS with increasing σ_{disp}. Here, we observe that the displacement bias in the initial ensemble mean has more pronounced effect, while the ensemble spread around this biased mean does not necessarily compensate for the bias (Figures 7 and 8). While smaller σ_{disp} ensembles achieve slightly lower CRPS, it has limited impact on the overall score, suggesting that the temporal trend is largely dominated by the bias component of CRPS (Figure 6).

Although CRPS is known to favor increased uncertainty for biased Gaussian forecasts, the ensemble in this case (Figure 7) represents coherent, non-Gaussian uncertainties associated with structural and position errors (Lawson and Hansen, 2005; Nehrkorn et al., 2015). As a result, increasing the positional displacement σ_{disp} does not necessarily improve reliability and instead leads to higher CRPS (Figure 6).

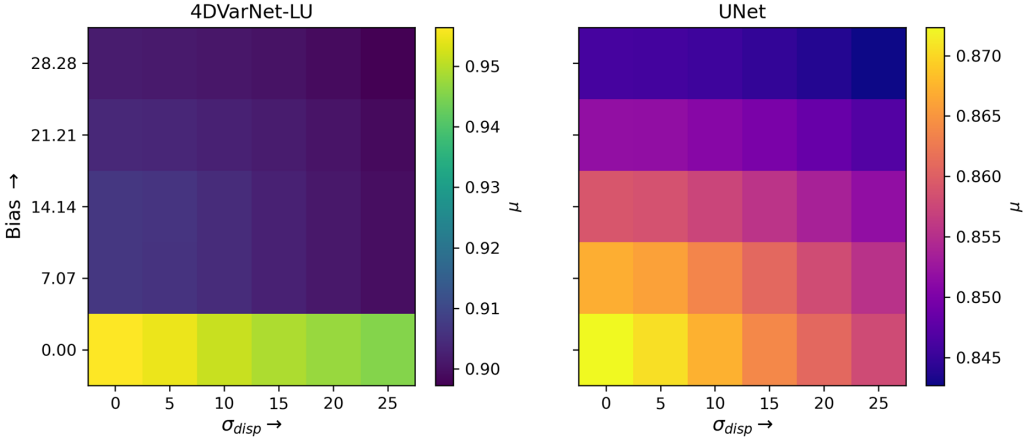


Figure 5. *Mean performance μ over the ensemble bias and ensemble spread parameter space for 4DVarNet-LU (left) and UNet (right).*

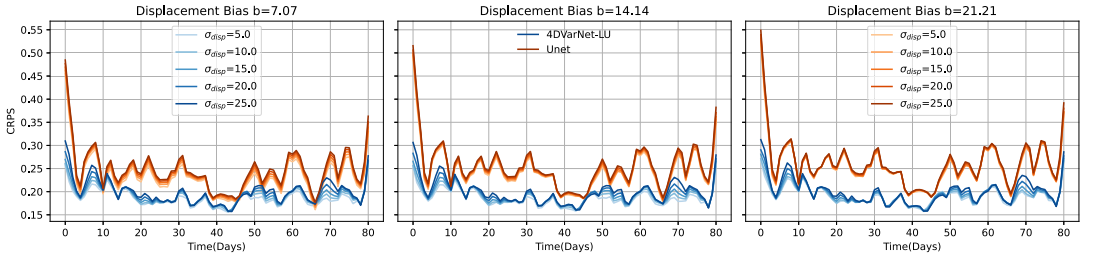


Figure 6. *Continuous ranked probability score (CRPS) as a function of time, averaged over spatial dimensions. Lower values indicate improved probabilistic performance.*

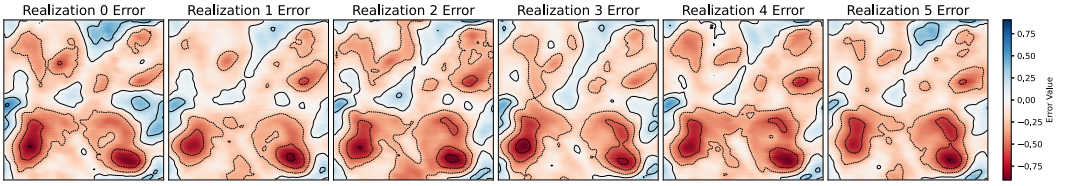


Figure 7. *Error for different ensemble members for bias = 7.07, day = 50,* $\sigma_{disp} = 15.0$.

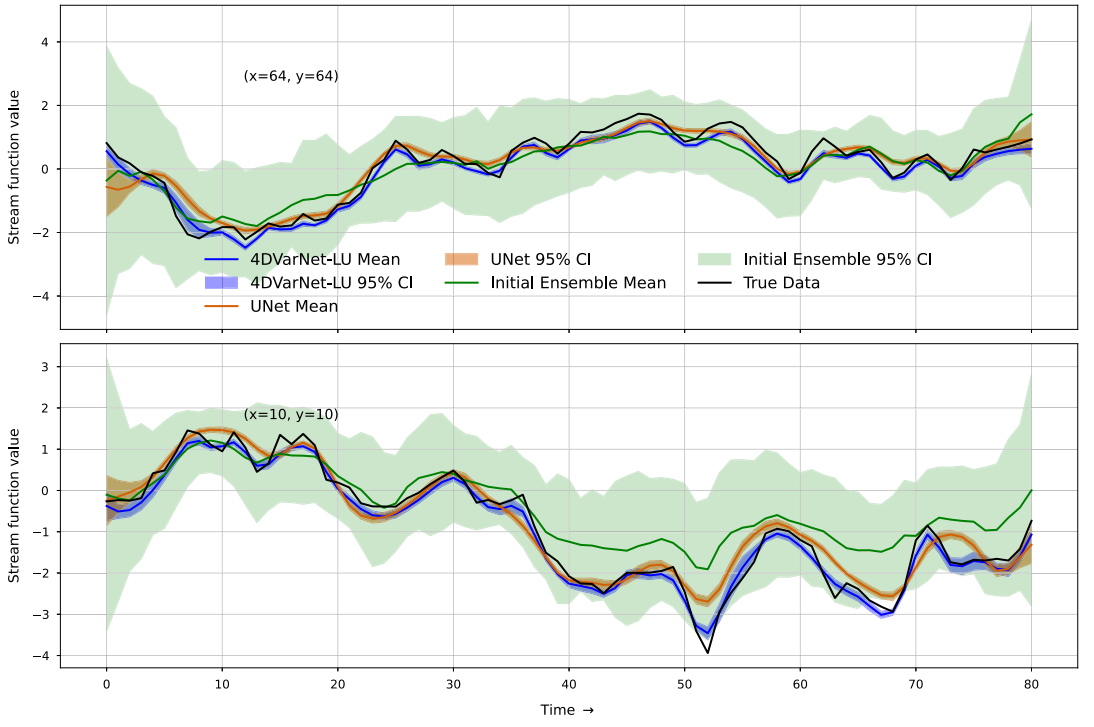


Figure 8. *Temporal behavior of the ensemble over the test period chosen at two different spatial locations for 4DVarNet-LU and UNet. The displacement bias is given by* $b = 21.21$ *and* $\sigma_{\text{disp}} = 10.0$. *The solid lines denote the ensemble mean, and the shaded regions indicate 95% confidence intervals computed using the Student-t distribution for an ensemble of size* $n = 25$.

4. Conclusions

DA methods combine observations with some prior estimate or background knowledge to solve inverse problems. We explore how initial conditions information could be used in neural DA methods, by primarily focusing on two DL models that can be trained to solve the DA task, 4DVarNet-LU and UNet. Using the ideas of position-based uncertainty in DA, we propose a method to systematically sample states around a given true state in a controllable manner using coherent random displacements in physical space. The proposed strategy provides a physically plausible mechanism to generate structured variability using a Gaussian random field model that preserves continuity and enforces spatiotemporal coherence, a key characteristic of physically realistic fields. The proposed scheme becomes a part of the new learning strategy where it preserves the large-scale coherent structures while generating randomness for ensemble generation. During inference, the trained model can be used to obtain an ensemble of reconstructions when used with an ensemble of initial conditions for the same set of observations. Our experiments on a quasi-geostrophic model show that 4DVarNet-LU consistently outperforms UNet as quantified by RMSE and CRPS metrics. The performance gap between 4DVarNet-LU and UNet can be attributed to the former's DA-based variational-formulation-aware architecture (Fablet et al., 2021; Beauchamp et al., 2023a). Importantly, even for biased initial ensembles not seen during training, 4DVarNet-LU maintains superior reconstruction skill over UNet, demonstrating good skill in correcting the bias in reconstruction to some level and generalization beyond the training distribution. While global quantities such as mass or mean field value remain preserved under our proposed scheme, local gradients and fine-scale structures may be modified due to interpolation, and further study of the effects on conservation of higher-order derived quantities such as energy and vorticity is required. This work is a step toward robust methods and strategies for understanding and further developing uncertainty-aware DL-based neural DA models.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/eds.2026.10058.

Acknowledgments. The authors acknowledge the computing resources provided by IMT Atlantique for access to the Odyssey server, which were used for model training and inference in this work.

Author contribution. Conceptualization: S.K.R., R.F.; Data curation: S.K.R.; Formal analysis: S.K.R.; Funding acquisition: R.F.; Investigation: S.K.R.; Methodology: S.K.R., R.F.; Software: S.K.R.; Supervision: R.F.; Visualization: S.K.R.; Writing—original draft: S.K.R.; Writing—review and editing: S.K.R., R.F.

Competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement. The simulation data used in this study are publicly available on Zenodo, which contains quasi-geostrophic stream function and vorticity ground truth fields, as well as satellite-track observations of the stream function (doi:10.5281/zenodo.16890868). All the model code, training configuration, inference scripts, and figure reproduction scripts are available at https://github.com/Mr-Markovian/4DVarNet-LU.

Ethics statement. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Funding statement. This work was supported by the European Union's Horizon Europe research and innovation programme under grant agreement No. 101137682 (AI4PEX—*Artificial Intelligence and Machine Learning for Enhanced Representation of Processes and Extremes in Earth System Models*; doi:/10.3030/101137682).

Provenance statement. This article is part of the Climate Informatics 2026 proceedings and was accepted in Environmental Data Science on the basis of the Climate Informatics peer review process.

References

Asch M, **Bocquet M and Nodet M** (2016) *Data Assimilation: Methods, Algorithms, and Applications*. Fundamentals of Algorithms. Philadelphia, PA: Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611974546.

Bannister RN (2017) A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society 143*(703), 607–633. https://doi.org/10.1002/qj.2982.

Beauchamp M, **Febvre Q**, **Georgenthum H and Fablet R** (2023a) 4DVarNet-SSH: End-to-end learning of variational interpolation schemes for nadir and wide-swath satellite altimetry. *Geoscientific Model Development 16*(8), 2119–2147. https://doi.org/10.5194/gmd-16-2119-2023.

Blum J, **Le Dimet F-X and Navon IM** (2009) Data assimilation for geophysical fluids. *Handbook of Numerical Analysis 14*, 385–441. https://doi.org/10.1016/S1570-8659(08)00209-3.

Bocquet M, **Vanderbecken PJ**, **Farchi A**, **Le Brazidec JD and Roustan Y** (2024) Bridging classical data assimilation and optimal transport: the 3D-Var case. *Nonlinear Processes in Geophysics 31*(3), 335–357. https://doi.org/10.5194/npg-31-335-2024.

Burgers G, **van Leeuwen PJ and Evensen G** (1998) Analysis scheme in the ensemble kalman filter. *Monthly Weather Review 126*(6), 1719–1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Carrassi A, **Bocquet M**, **Bertino L and Evensen G** (2018) Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change 9*(5), e535. https://doi.org/10.1002/wcc.535.

Cheng S, **Quilodrán-Casas C**, **Ouala S**, **Farchi A**, **Liu C**, **Tandeo P**, **Fablet R**, **Lucor D**, **Iooss B**, **Brajard J**, **Xiao D**, **Janjic T**, **Ding W**, **Guo Y**, **Carrassi A**, **Bocquet M and Arcucci R** (2023) Machine learning with data assimilation and uncertainty quantification for dynamical systems: A review. *IEEE/CAA Journal of Automatica Sinica 10*(6), 1361–1387.

Da Li YY, **Song Y-Z and Hospedales TM** (2018) Learning to generalize: Meta-learning for domain generalization. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, New Orleans, LA, February 2–7, pp. 3490–3497. AAAI Press.

Fablet R, **Chapron B**, **Drumetz L**, **Mémin E**, **Pannekoucke O and Rousseau F** (2021) Learning variational data assimilation models and solvers. *Journal of Advances in Modeling Earth Systems 13*(10), e2021MS002572. https://doi.org/10.1029/2021MS002572.

Farchi A, **Chrust M**, **Bocquet M**, **Laloyaux P and Bonavita M** (2023) Online model error correction with neural networks in the incremental 4D-Var framework. *Journal of Advances in Modeling Earth Systems 15*(9), e2022MS003474. https://doi.org/10.1029/2022MS003474.

Fletcher SJ (2017) Chapter 16 – variational data assimilation, data assimilation for the geosciences. In Fletcher SJ (ed.), *Data Assimilation for the Geosciences: From Theory to Application*. Amsterdam: Elsevier, pp. 673–703.

Frezat H, **Le Sommer J**, **Fablet R**, **Balarac G and Lguensat R** (2022) A posteriori learning for quasi-geostrophic turbulence parametrization. *Journal of Advances in Modeling Earth Systems 14*(11), e2022MS003124. https://doi.org/10.1029/2022MS003124.

Hersbach H (2000) Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting 15*(5), 559–570. https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>003E;2.0.CO;2.

Ho J, **Jain A and Abbeel P** (2020) Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020*, NeurIPS 2020, December 6–12, virtual. Red Hook, NY: Curran Associates, Inc., pp. 6840–6851.

Lawson WG and Hansen JA (2005) Alignment error models and ensemble-based data assimilation. *Monthly Weather Review 133*(6), 1687–1709. https://doi.org/10.1175/MWR2945.1.

Le Guillou F, **Gaultier L**, **Ballarotta M**, **Metref S**, **Ubelmann C**, **Cosme E and Rio M-H** (2023) Regional mapping of energetic short mesoscale ocean dynamics from altimetry: Performances from real observations. *Ocean Science 19*(5), 1517–1527. https://doi.org/10.5194/os-19-1517-2023.

LeCun Y, **Bengio Y and Hinton G** (2015) Deep learning. *Nature 521*(7553), 436–444.

van Leeuwen PJ, **Chiu J-YC and Yang C-KK** (2025) Uncertainty quantification for deep learning. *Environmental Data Science 4*, e53. https://doi.org/10.1017/eds.2025.10027.

Martin SA, **Manucharyan GE and Klein P** (2025) Generative data assimilation for surface ocean state estimation from multimodal satellite observations. *Journal of Advances in Modeling Earth Systems 17*(8), e2025MS005063. https://doi.org/10.1029/2025MS005063.

Nehrkorn T, **Woods BK**, **Hoffman RN and Auligné T** (2015) Correcting for position errors in variational data assimilation. *Monthly Weather Review 143*(4), 1368–1381. https://doi.org/10.1175/MWR-D-14-00127.1.

Palmer TN and Zanna L (2013) Singular vectors, predictability and ensemble forecasting for weather and climate. *Journal of Physics A: Mathematical and Theoretical 46*(25), 254018. https://doi.org/10.1088/1751-8113/46/25/254018.

Ronneberger O, **Fischer P and Brox T** (2015) *U-net: Convolutional networks for biomedical image segmentation*. In Navab N, Hornegger J, Wells WM and Frangi AF (eds), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (Cham)*. Cham: Springer International Publishing, pp. 234–241.

Schwartz E, **Giryes R and Bronstein AM** (2019) Deepisp: Toward learning an end-to-end image processing pipeline. *IEEE Transactions on Image Processing 28*(2), 912–923. https://doi.org/10.1109/TIP.2018.2872858.

Trémolet Y (2007) Model-error estimation in 4D-Var. *Quarterly Journal of the Royal Meteorological Society 133*(626), 1267–1280. https://doi.org/10.1002/qj.94.

Cite this article: Roy SK and Fablet R (2026). Spatiotemporal coherent displacements for ensemble-based neural data assimilation. *Environmental Data Science*, 5: e36. doi:10.1017/eds.2026.10058