

The advantages of data assimilation in parametric space rather than classic grid space

Solène Dealbera^{1,2}, Stéphane Raynaud³, Carlos Granero Belinchon^{1,2}, Brahim Boussidi⁴, Clément Le Goff⁴, and Pierre Tandeo^{1,2}

¹IMT Atlantique, Lab-STICC, UMR 6285, 29238, CNRS, Brest, France

²ODYSSEY Team-Project, INRIA Ifremer IMT-Atl., 35042, CNRS, Brest, France

³Shom, 29200 Brest, France

⁴LEGOS Observatoire Midi-Pyrénées, 14 avenue Édouard Belin, 31400 Toulouse, France

Correspondence: Solène Dealbera (solene.dealbera@imt-atlantique.fr)

Abstract. Data assimilation (DA) combines observations and model background to estimate the state of geophysical systems. When coherent structures such as ocean eddies are present, linear Gaussian DA performed in gridded space may struggle to accurately represent structural properties, including position and amplitude, particularly if displacement uncertainty dominates. In this work, we explore a DA framework in a reduced parametric space, where the state is defined by a set of parameters describing the structure. Using an idealized one-dimensional ocean eddy, we compare single-step analyses performed both in gridded and parametric spaces. Three tutorial experiments with prescribed parametric uncertainties illustrate how linear Gaussian updates in gridded space can distort coherent structures. A fourth experiment, starting from noisy gridded observations, provides a more balanced end-to-end evaluation by accounting for the inverse nonlinear mapping from grid space to parameter space. The parametric approach allows a direct estimation of the structure’s parameters, preserves the structure consistency, and may reduce computational cost in this low-dimensional idealized setting. However, it requires reliable parameter extraction, and its benefits depend on the number of structures relative to the grid resolution.

1 Introduction

Data assimilation (DA) is widely used in geosciences to combine observations with information from a dynamical model (the background) in a statistically optimal way (Carrassi et al., 2018). In its standard formulation, DA aims to estimate the full system state on a spatial grid.

In this study, we do not focus on the full gridded state, but rather on the estimation of specific geophysical structures (GSs), such as ocean eddies or cyclones, viewed as coherent objects characterized by a small set of parameters. When sequential DA is applied in gridded space, linear Gaussian updates can struggle to correctly represent the position or shape of coherent structures, potentially altering their physical properties (Chen and Snyder, 2007). However, these limitations arise specifically when displacement or shape uncertainty dominates, and thus do not constitute a general failure of gridded data assimilation. To address these challenges, the literature includes non-Gaussian methods such as particle filters (van Leeuwen, 2010; Poterjoy, 2016) that can, in principle, handle the nonlinearity associated with displacement and shape uncertainty, as well as displacement-aware

transformations of the state, including alignment or morphing (Ravela et al., 2007; Beezley and Mandel, 2008; Zhen et al., 2025). Feature-informed and Lagrangian-based DA approaches have also been proposed, where the characteristics of coherent structures explicitly guide the assimilation (Srivastava et al., 2023; Maclean et al., 2017).

The goal of this study is to investigate, in a controlled and idealized framework, how the choice of representation space influences the behavior of a standard Gaussian linear DA update when applied to coherent structures with uncertain position and shape. In particular, we compare assimilation performed in the classic gridded space with assimilation performed directly in a reduced parametric space describing the structure.

The idea of working in a low-dimensional parameter space rather than on a full-state vector has a long history in the control and stochastic filtering literature (Jazwinski, 1970). However, in these frameworks, the parameters to be estimated are typically internal coefficients of a dynamical model, rather than geometric descriptors of coherent structures. The present work applies a related principle to an explicitly geometric and physically interpretable representation of a coherent structure, defined here by amplitude, position and radius.

This parametric perspective is also related to reduced-order and subspace approaches, where the system state is represented in a lower-dimensional space, such as empirical orthogonal functions or learned latent variables (Hannachi et al., 2007; Alvera-Azcárate et al., 2025). In contrast to these approaches, which define reduced coordinates in an abstract basis, the present work considers an explicit physically interpretable parameterization of the structure. This allows a direct interpretation of assimilation updates in terms of physically meaningful geometric properties.

We emphasize that this framework does not aim to replace existing Lagrangian, feature-based, or non-Gaussian DA methods, but rather to provide a simplified setting to analyze how standard Gaussian updates behave when applied either in gridded space or in a low-dimensional parameter space representing coherent structures. To the best of our knowledge, this specific comparison of the assimilation behavior between gridded space and an explicitly defined geometric parameter space of a coherent structure has not been analyzed in the literature. In this sense, the study is primarily methodological and diagnostic, focusing on the mechanism of error representation rather than on operational forecasting performance.

In the proposed framework, the state of the system is fully described by the parameters of a single idealized eddy. This allows a direct analysis of how uncertainties in amplitude, position, and radius propagate through the analysis step of the assimilation in both representations. While the parametric formulation may offer computational advantages due to its reduced dimensionality in this idealized context, practical applications would require additional steps such as eddy detection, parameter fitting, and feature association, which may significantly affect the overall computational cost.

We present a set of idealized numerical experiments designed to compare a single analysis step performed in the classic gridded space and in this reduced parametric space. The eddy representation and assimilation methodology are detailed in Sect. 2, the experimental setup in Sect. 3, and the results are presented in Sect. 4. Conclusions and perspectives are discussed in Sect. 5.

The objective of this paper is to compare two methods of blending two sources of information, a background and an observation, using a basic DA algorithm. We focus exclusively on the analysis step and do not perform repeated assimilation cycles (forecast-analysis-reanalysis). The stability of the system over multiple cycles is beyond the scope of this study and will be investigated in future work.

60 We consider a simplified one-dimensional cross-shore section of an ocean eddy, described by three parameters: peak amplitude a , peak position p , and maximal velocity radius r , such that $\mathbf{x}^{\text{param}} = (a, p, r)$. In practice, these parameters can be estimated using standard eddy detection and fitting algorithms, such as local maxima identification combined with Gaussian profile fitting (e.g., Chelton et al., 2011).

For our idealized experiments, however, the parameters are prescribed manually rather than obtained from a detection algorithm, in order to control uncertainties and systematically study the assimilation behavior. The physical profile along the cross-section is given by the mapping

$$\mathbf{x}^{\text{grid}} = h(\mathbf{lon}; \mathbf{x}^{\text{param}}) = a \exp \left(-\frac{1}{2} \left(\frac{\mathbf{lon} - p}{r} \right)^2 \right), \quad (1)$$

where $\mathbf{lon}$ is the vector of longitudinal positions along the cross-section. Here, h serves as a mapping function from the parametric space $\mathbf{x}^{\text{param}}$ to the physical gridded space $\mathbf{x}^{\text{grid}}$. It is neither an observation operator nor a numerical dynamical model, but simply defines the physical profile corresponding to a given set of parameters.

Two spaces are considered for the eddy representation: the *parametric space* (`param`), of dimension $n = 3$ corresponding to (a, p, r) , and the *gridded space* (`grid`), of dimension m equal to the length of the longitude vector. The mapping function h allows conversion from the parametric to the gridded space. The notation used throughout this manuscript is summarized in Table 1.

75 The background state in each space is modeled as a perturbation around a reference value:

$$\mathbf{x} = \mathbf{x}^b + \boldsymbol{\eta}^b, \quad (2)$$

where $\mathbf{x}^b$ is the reference background state and $\boldsymbol{\eta}^b \sim \mathcal{N}(\mathbf{0}, \mathbf{B})$ is the background error, assumed Gaussian with covariance matrix $\mathbf{B}$. The observation is related to the true state by:

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}, \quad (3)$$

80 where $\mathcal{H}$ is the observation operator and $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ is the observation error with covariance $\mathbf{R}$. Equations (2) and (3) are written generically and apply separately in each space with the appropriate superscript (`param` or `grid`) on all variables (see Table 1).

In our case, since the state vector is fully observed in both spaces, the operator $\mathcal{H}$ is the identity matrix, leading to a simple linear observation equation.

Table 1. Symbols and associated spaces.

Symbol	Meaning	Space
$\mathbf{x}^{t,\text{param}}$	True state vector	param
$\mathbf{x}^{t,\text{grid}}$	True state vector	grid
$\mathbf{x}^{b,\text{param}}$	Background state vector	param
$\mathbf{x}^{b,\text{grid}}$	Background state vector	grid
$\mathbf{x}^{a,\text{param}}$	Analysis state vector	param
$\mathbf{x}^{a,\text{grid}}$	Analysis state vector	grid
$\mathbf{y}^{\text{param}}$	Observation vector	param
$\mathbf{y}^{\text{grid}}$	Observation vector	grid
$\mathbf{B}^{\text{param}}$	Background error covariance	param
$\mathbf{B}^{\text{grid}}$	Background error covariance	grid
$\mathbf{R}^{\text{param}}$	Observation error covariance	param
$\mathbf{R}^{\text{grid}}$	Observation error covariance	grid
$h(\cdot; \cdot)$	Mapping from param to grid	n/a

85 The data assimilation is performed in each space with the appropriate superscript (param or grid) on all variables using the following cost function J :

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{x}). \quad (4)$$

Under the assumptions of linearity and Gaussian errors, the minimizer of J is given analytically by the Best Linear Unbiased Estimator (BLUE). No iterative minimization algorithm is therefore required; the solution is computed directly for each configuration. Let $\mathbf{X}^b = [\mathbf{x}_1^b, \dots, \mathbf{x}_N^b]$ and $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N]$ denote the background and observation ensemble matrices, each of size $n \times N$ (or $m \times N$ in grid space), where N is the ensemble size. The covariance matrices are estimated from these ensembles as:

$$\mathbf{B} = \text{Cov}(\mathbf{X}^b, \mathbf{X}^b), \text{ and } \mathbf{R} = \text{Cov}(\mathbf{Y}, \mathbf{Y}), \quad (5)$$

following the Ensemble Optimal Interpolation (EnOI) framework. The BLUE update of each ensemble member i reads:

$$95 \quad \mathbf{x}_i^a = \mathbf{x}_i^b + \mathbf{K}(\mathbf{y}_i - \mathbf{x}_i^b), \quad i = 1, \dots, N, \quad (6)$$

from which the analysis ensemble mean $\bar{\mathbf{x}}^a = \frac{1}{N} \sum_i \mathbf{x}_i^a$ is derived, and the analysis error covariance $\mathbf{P}^a$ is given by:

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K})\mathbf{B}, \quad (7)$$

where $\mathbf{K} = \mathbf{B}(\mathbf{B} + \mathbf{R})^{-1}$ is the Kalman gain. Equations (5)-(7) apply separately in each space with the appropriate superscript (param or grid) on all variables.

Two experimental frameworks are considered, differing in how the uncertainty is introduced and how the covariance matrices are specified, as described in the following subsections. In the tutorial framework (Sect. 2.1), the uncertainty originates in parameter space and is propagated to grid space via h , which naturally favors the parametric approach. In the realistic framework (Sect. 2.2), the uncertainty originates in grid space and must be inverted back to parameter space through a fitting procedure, providing a more balanced end-to-end comparison.

2.1 Tutorial framework

This framework is designed as a controlled tutorial example to isolate the mechanism by which linear Gaussian updates can distort coherent structures in gridded space. It is not intended as a balanced end-to-end comparison of practical DA workflows, since the uncertainty is generated in parameter space and then mapped to grid space, which inherently favors the parametric approach.

In this framework, the parameters are assumed to be fully independent and to have equal uncertainties for both the background and the observation. The error covariance matrices in parameter space are therefore diagonal:

$$\mathbf{B}^{\text{param}} = \mathbf{R}^{\text{param}} = \text{diag}(\sigma_a^2, \sigma_p^2, \sigma_r^2) = \boldsymbol{\sigma}^2 \mathbf{I}_n, \quad (8)$$

where $\boldsymbol{\sigma} = (\sigma_a, \sigma_p, \sigma_r)$ contains the prescribed standard deviations for amplitude, position and radius, respectively.

Ensemble members in `param` space are generated by independently sampling each parameter from a Gaussian distribution around its reference value with the prescribed standard deviation. The corresponding ensemble members in `grid` space are then obtained by applying the mapping h (Eq. 1) to each member, so that the gridded covariance matrices $\mathbf{B}^{\text{grid}}$ and $\mathbf{R}^{\text{grid}}$ are fully determined by the nonlinear push-forward of the parametric uncertainties. As a result, these gridded covariances are in general non-diagonal and reflect the nonlinear geometry of the eddy profile, even though the parametric uncertainties are uncorrelated.

Figure 1 illustrates the overall workflow comparing assimilation in the parametric and gridded spaces.

2.2 Realistic framework

In realistic applications, observations are available as gridded fields rather than direct eddy parameters. In this second framework, both the background and the observation start as noisy gridded profiles, and the parametric representation must be inferred from them through a fitting procedure. This constitutes a more balanced end-to-end test, since the uncertainty now originates in grid space and the inverse nonlinear mapping from grid to parameter space is explicitly accounted for.

The observed and background gridded fields are defined as:

$$\mathbf{y}^{\text{grid}} = h(\mathbf{lon}; \mathbf{y}^{\text{param},*}) + \boldsymbol{\varepsilon}^{\text{grid}}, \quad (9)$$

$$\mathbf{x}^{b,\text{grid}} = h(\mathbf{lon}; \mathbf{x}^{b,\text{param},*}) + \boldsymbol{\eta}^{b,\text{grid}}, \quad (10)$$

where $\mathbf{y}^{\text{param},*}$ and $\mathbf{x}^{b,\text{param},*}$ are the reference observation and background parameter vectors, and $\boldsymbol{\varepsilon}^{\text{grid}} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^{\text{grid}})$, $\boldsymbol{\eta}^{b,\text{grid}} \sim \mathcal{N}(\mathbf{0}, \mathbf{B}^{\text{grid}})$ are the respective gridded noise terms. The gridded covariance matrices are prescribed using an ex-

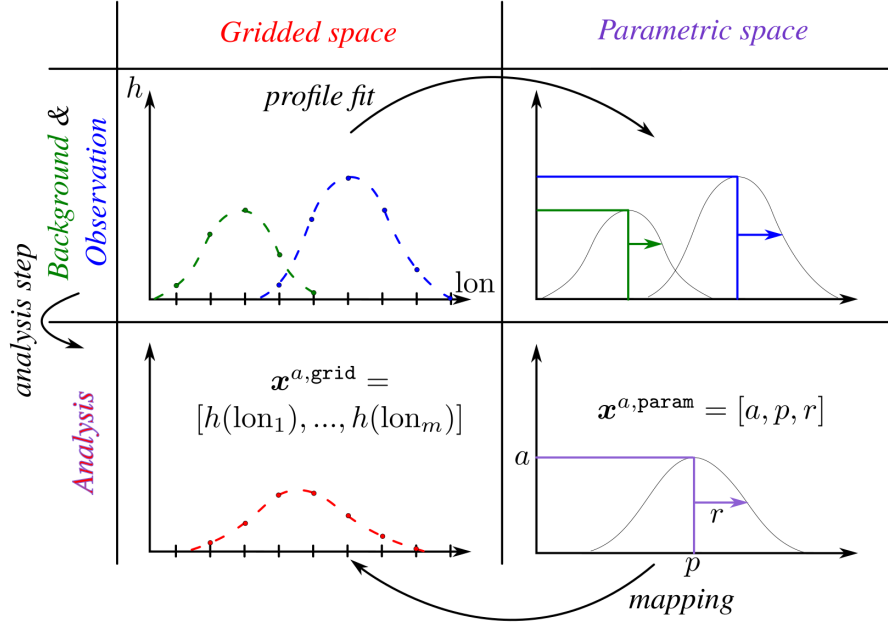


Figure 1. Schematic of the data assimilation framework comparing gridded space (left column, red) and parametric space (right column, purple). Top row: background (green) and observation (blue) eddies represented as gridded profiles (top left) and as parameter vectors (top right). The *profile fit* arrow indicates the nonlinear least-squares mapping from gridded observations to parameter space, as used in Experiment #4. Bottom row: analysis state from the analysis step in gridded space $\mathbf{x}^{a,\text{grid}} = [h(\text{lon}_1), \dots, h(\text{lon}_m)]$ (bottom left) and in parametric space $\mathbf{x}^{a,\text{param}} = [a, p, r]$ (bottom right). The *mapping* arrow indicates the projection of the parametric analysis back to gridded space via h (Eq. 1).

ponentially decaying correlation model:

$$\mathbf{B}_{ij}^{\text{grid}} = \mathbf{R}_{ij}^{\text{grid}} = \sigma_h^2 \exp\left(-\frac{|\text{lon}_i - \text{lon}_j|}{\lambda}\right), \quad (11)$$

where σ_h^2 is the surface-height variance and λ is the correlation length scale. In the gridded DA approach, $\mathbf{y}^{\text{grid}}$ is assimilated directly using Eqs. (6)-(7) with these prescribed covariances.

135 For the parametric DA approach, the parameter vectors are first estimated from the noisy gridded fields through a nonlinear least-squares fit of the Gaussian profile h . Specifically, for a given noisy gridded profile $\mathbf{u}$, the estimated parameter vector $\hat{\boldsymbol{\theta}}$ minimizes:

$$\mathcal{L}(\boldsymbol{\theta}) = \|\mathbf{u} - h(\text{lon}; \boldsymbol{\theta})\|_2^2. \quad (12)$$

The minimization is performed using the `scipy.optimize.curve_fit` algorithm with bounds $a \in [0, 10]$ m, $p \in [-50, 50]$ km, and $r \in [0.1, 50]$ km, ensuring physical solutions (positive amplitude and positive radius). The initial values for the fit are set as following: $[\max \mathbf{u}, \text{lon}_{\arg \max \mathbf{u}}, \frac{1}{2} |\text{lon}_{\arg \max \mathbf{u}} - \text{lon}_{\arg \min \mathbf{u}}|]$ where the amplitude is initialized from the maximum value

of the profile, the center from the longitude at which this maximum occurs, and the width from half the distance between the longitudes of the maximum and minimum values. This simplified initialization is appropriate given the idealized framework of the present experiment. In practice, initial values could be provided by a prior detection algorithm.

145 The fitting procedure is repeated over an ensemble of N noisy realizations drawn from Eqs. (9)-(10). The parametric covariance matrices are then estimated as the sample covariance of the fitted parameter ensembles, determined by the nonlinear push-forward of the gridded height uncertainties. Unlike the tutorial framework, these covariance matrices are in general non-diagonal, reflecting the nonlinear coupling between parameters introduced by the fitting procedure. The parametric DA is then performed using Eqs. (6)-(7) in `param` space with these estimated covariances.

150 3 Experiments

Four experiments are presented, organized into two sets corresponding to the two frameworks described in Sect. 2. Experiments #1, #2 and #3 follow the tutorial framework (Sect. 2.1) and are designed to isolate the effect of a single uncertain parameter on the assimilation outcome. Experiment #4 follows the realistic framework (Sect. 2.2) and provides a more balanced end-to-end evaluation by starting from noisy gridded observations. The parameter values for all four experiments are given in Table 2.

155 3.1 Tutorial experiments (#1, #2, #3)

In each of the three tutorial experiments, a single parameter carries the uncertainty, while the other two are set to their true values and are identical in the background and the observation. The uncertain parameter is, respectively: amplitude a (Experiment #1), position p (Experiment #2), and radius r (Experiment #3). This design isolates the effect of each parameter's nonlinearity on the assimilation update (see Table 2).

160 As noted in Sect. 2.1, this setup inherently favors the parametric approach, since the uncertainty originates in parameter space and is propagated to grid space via h . The purpose of these experiments is therefore not to provide a balanced comparison of practical DA workflows, but to illustrate the mechanism by which linear Gaussian updates in gridded space can smear or distort a coherent structure when the uncertain parameter enters h nonlinearly.

165 The background and observation are constructed symmetrically around a true state $\mathbf{x}^{t,\text{param}}$ by applying shifts of equal magnitude and opposite sign:

$$\mathbf{x}^{b,\text{param}} = \mathbf{x}^{t,\text{param}} - \boldsymbol{\delta}, \text{ and } \mathbf{y}^{\text{param}} = \mathbf{x}^{t,\text{param}} + \boldsymbol{\delta}, \quad (13)$$

where $\boldsymbol{\delta}$ is the prescribed shift vector, with only the uncertain parameter having a non-zero component. The associated true gridded profile is $\mathbf{x}^{t,\text{grid}} = h(\mathbf{lon}; \mathbf{x}^{t,\text{param}})$. This symmetric construction ensures that the background and observation errors are equal in magnitude, so that the expected BLUE analysis in parameter space recovers the truth exactly: $\bar{\mathbf{x}}^{a,\text{param}} = \mathbf{x}^{t,\text{param}}$.

170 An ensemble of $N = 10000$ members is used for both the background and the observation. In `param` space, ensemble members are generated by independently sampling each parameter from a Gaussian distribution around its reference value with the prescribed standard deviation σ_a , σ_p , or σ_r (the non-perturbed parameters having zero spread). The corresponding

Table 2. True parameter values $\mathbf{x}^{t,\text{param}}$ and reference parameter values with prescribed standard deviations ($\pm\sigma$) for the background and observation in each experiment. The background and observation are constructed symmetrically around the truth via Eq. (13). For Experiment #4, no standard deviation is prescribed in parameter space; the uncertainty is instead introduced as additive gridded noise with amplitude $\sigma_h = 0.2$ m and correlation length $\lambda = r/2 = 3.5$ km (see Sect. 2.2).

Exp.	Eddy	a [m]	p [km]	r [km]
#1	Truth	1.5	0	7
	Background	1.0 ± 0.7	0	7
	Observation	2.0 ± 0.7	0	7
#2	Truth	1.0	0	7
	Background	1.0	-5 ± 7	7
	Observation	1.0	5 ± 7	7
#3	Truth	1.0	0	11
	Background	1.0	0	8 ± 4
	Observation	1.0	0	14 ± 4
#4	Truth	1.0	0	7
	Background	1.0	-5	7
	Observation	1.0	5	7

grid ensemble members are obtained by applying h (Eq. 1) to each parametric member, as described in Sect. 2.1. The radius is clipped to $r \geq 10^{-6}$ km to prevent non-physical values.

175 3.2 Realistic experiment (#4)

Experiment #4 uses the same reference parameter values as Experiment #2 (position perturbation), but the uncertainty is now introduced directly in grid space following the realistic framework of Sect. 2.2. The background and observation gridded fields are generated as:

$$\mathbf{x}^{b,\text{grid}} = h(\mathbf{lon}; (-5, 1, 7)) + \boldsymbol{\eta}^{b,\text{grid}}, \text{ and } \mathbf{y}^{\text{grid}} = h(\mathbf{lon}; (5, 1, 7)) + \boldsymbol{\varepsilon}^{\text{grid}}, \quad (14)$$

180 where the noise terms are drawn from the exponentially correlated distribution of Eq. (11) with $\sigma_h = 0.2$ m and $\lambda = r/2 = 3.5$ km.

For the parametric DA, an ensemble of $N = 10000$ noisy realizations is generated for both the background and the observation. Each realization is fitted independently using the procedure described in Sect. 2.2, yielding an ensemble of parameter vectors from which $\mathbf{B}^{\text{param}}$ and $\mathbf{R}^{\text{param}}$ are estimated. To assess the robustness of the results with respect to the noise realization, the full experiment is repeated $n_{\text{iter}} = 10$ times with independent noise draws; quantitative diagnostics are reported as means and standard deviations over these 10 noise realizations.

3.3 Evaluation

For all four experiments, two comparisons are performed:

- *Surface height reconstruction*: The analysis mean is projected into `grid` space and compared to the true profile $\mathbf{x}^{t,\text{grid}}$.
 190 For the parametric analysis, the projection is done by applying h (Eq. 1) to the analyzed parameter vector $\bar{\mathbf{x}}^{a,\text{param}}$. For the gridded analysis, the profile $\bar{\mathbf{x}}^{a,\text{grid}}$ is used directly. Two metrics are reported: the Root Mean Square Error (RMSE) of the analysis mean relative to $\mathbf{x}^{t,\text{grid}}$, and the Continuous Ranked Probability Score (CRPS) of the analysis ensemble relative to $\mathbf{x}^{t,\text{grid}}$, which jointly measures accuracy and sharpness of the probabilistic forecast.
- *Parameter estimation*: The estimated parameters of each analysis are compared to the true values $\mathbf{x}^{t,\text{param}}$. For the
 195 parametric analysis, the estimated parameters are directly $\bar{\mathbf{x}}^{a,\text{param}}$. For the gridded analysis, the analyzed profile is fitted with the Gaussian function h using the least-squares procedure of Eq. (12), yielding estimated parameters $\hat{\theta}^{\text{grid}}$. The absolute error of each analysis relative to the truth is reported for each parameter separately, defined as $|\hat{\theta}_k^{\text{grid}} - x_k^{t,\text{param}}|$ and $|\bar{x}_k^{a,\text{param}} - x_k^{t,\text{param}}|$ for $k \in \{a, p, r\}$.

4 Results and discussion

200 4.1 Tutorial experiments (#1, #2, #3)

The surface height profiles are shown in Figs. 2, 3 and 4 for the amplitude, position and radius perturbation experiments, respectively. The quantitative diagnostics for all experiments are reported in Tables 3 and 4.

4.1.1 Experiment #1 — Amplitude perturbation.

Figure 2a shows that the analyzed profiles are identical in both `param` and `grid` spaces: the analyzed amplitude is 1.5 m, the
 205 exact mean of the background (1.0 m) and observation (2.0 m), with position and radius unchanged. Figure 2b shows identical a posteriori standard deviation profiles in both spaces. This agreement reflects the fact that amplitude is a linear parameter of h (Eq. 1): the nonlinear push-forward that distinguishes the two approaches has no effect in this case.

This is confirmed by the metrics in Table 3: the RMSE is 0.00 in both spaces, and the CRPS values are 0.03. The absolute error is 0.00 for all parameters (see Table 4).

210 4.1.2 Experiment #2 — Position perturbation.

By contrast, Fig. 3a shows a clear difference between the two analyses. The `grid` analyzed profile is widened and flattened relative to both the background and the observation. A Gaussian fit of this profile yields $(a, p, r) \approx (0.77, 0.02, 9.05)$ m/km/km: the amplitude is underestimated by 23% and the radius overestimated by 29% relative to the true values, even though the position is correctly recovered. This deformation is a direct consequence of the nonlinear dependence of h on position: the

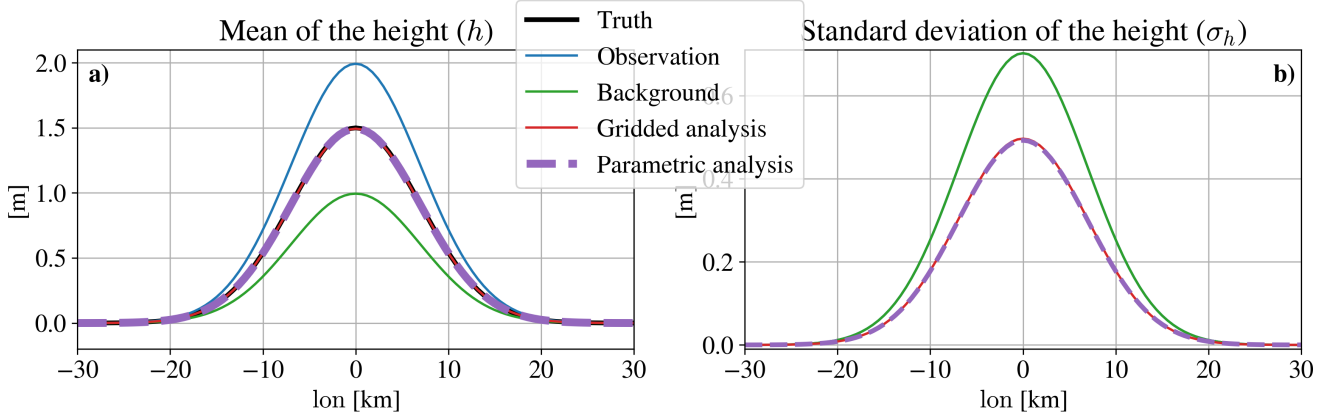


Figure 2. Results for Experiment #1 (amplitude perturbation). Left: true surface height (black), mean surface height of the observed (blue), background (green), gridded analysis (red), and parametric analysis (dashed purple) eddies along the longitude. Right: corresponding a posteriori standard deviation $\sigma_h = \sqrt{\text{diag}(\mathbf{P}^a)}$.

215 linear Gaussian update in grid space averages the two displaced profiles, producing a broader, lower peak, rather than a correctly shaped one.

In contrast, the `param` analysis recovers $(a, p, r) = (1.0, 0.0, 7.0)$ m/km/km, the exact midpoint in parameter space, with all three parameters correctly estimated. The a posteriori standard deviation in `grid` space (Fig. 3b) is also spatially underestimated relative to the `param` counterpart, indicating that the gridded analysis is overconfident about the reconstructed profile.

220 These differences are captured by the metrics: the RMSE increases to 0.10 m in the `grid` space, while it remains at 0.00 m in the `param` space. The CRPS is slightly better in the `param` space. The absolute error shows higher values for the estimation of the amplitude (0.24 m) and the radius (2.15 km) in the `grid` space, while the position is well-estimated (Table 4).

4.1.3 Experiment #3 — Radius perturbation.

Figure 4a shows that the `grid` analyzed profile has a narrower peak with a wider tail, compared to the `param` profile. A Gaussian fit yields $(a, p, r) \approx (0.96, 0.00, 11.09)$ m/km/km: the radius is correctly recovered (11.09 km, close to the expected mean of 11 km), but the amplitude is slightly underestimated (0.96 m vs. 1.0 m). More importantly, the `grid` analyzed profile is no longer strictly Gaussian: the profile shape is sub-Gaussian, with a narrower core and heavier tails than a Gaussian of the same standard deviation, a consequence of the nonlinear averaging of two Gaussian profiles of different widths in grid space. In the `param` space, the estimated radius is 11 km (the mean of 8 and 14 km) and the profile remains Gaussian. The `grid` RMSE is slightly higher (0.03 m), while the CRPS value is close (0.02 for the `grid` space and 0.03 for the `param` space).

230

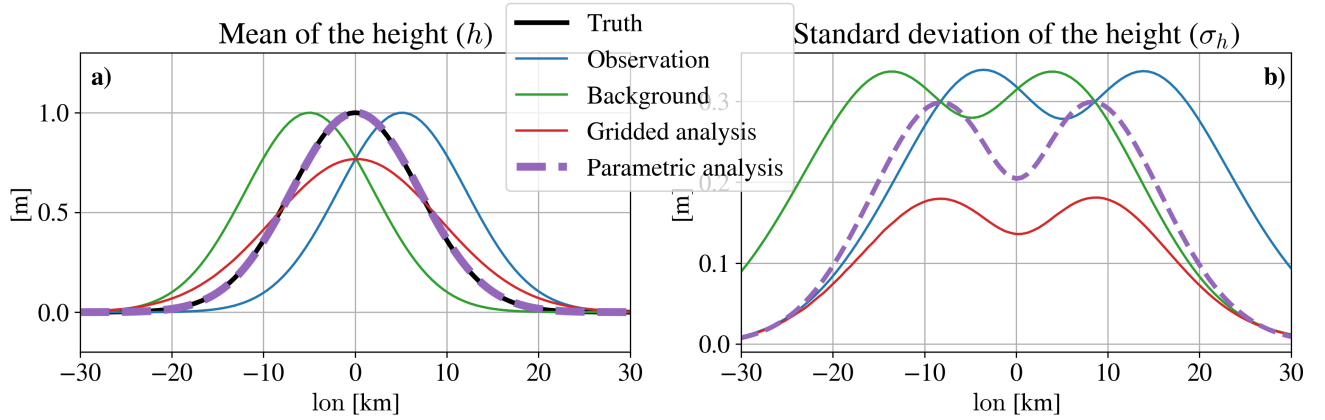


Figure 3. Results for Experiment #2 (position perturbation). Same color and panel convention as Fig. 2.

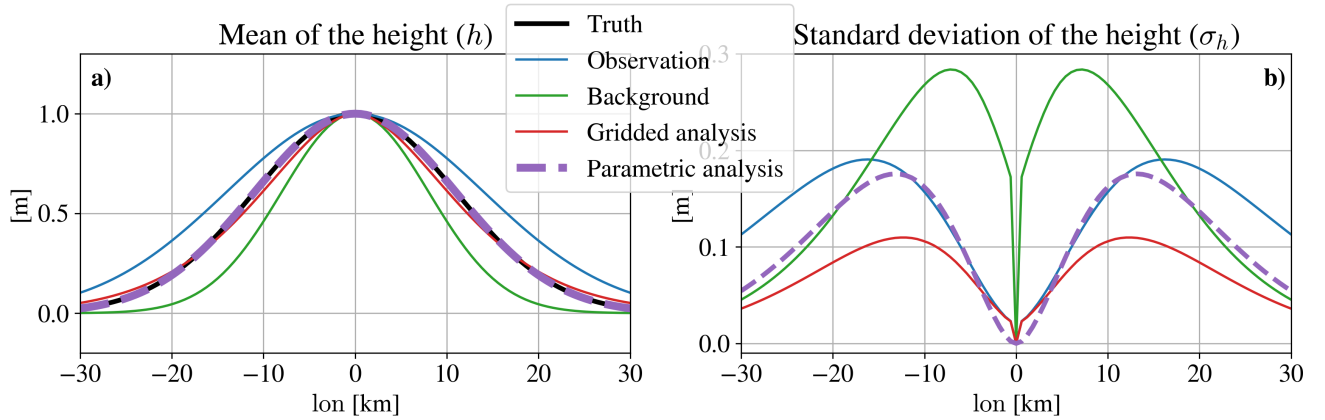


Figure 4. Results for Experiment #3 (radius perturbation). Same color and panel convention as Fig. 2.

4.2 Realistic experiment (#4)

Figure 5 shows the results for Experiment #4. The qualitative behavior is consistent with Experiment #2: the `grid` analysis produces a widened and underestimated profile (Gaussian fit: $a \approx 0.79$ m, $p \approx 0.04$ km, $r \approx 8.96$ km), while the `param` analysis recovers parameters close to the expected mean. These values are stable across the 10 noise realizations.

235 The a posteriori standard deviations differ from Experiment #2 (Fig. 5b). In the `grid` space, the standard deviation is spatially uniform at approximately $0.2/\sqrt{2}$ m, reflecting the prescribed exponentially correlated noise structure. In the `param` space, the standard deviation is spatially varying, with larger values in the region spanned by the background and observa-

Table 3. Gridded field diagnostics for all experiments. RMSE: root-mean squared error of the analysis mean relative to the true profile $\mathbf{x}^{t,\text{grid}}$ (lower is better); CRPS: continuous ranked probability score of the analysis ensemble relative to $\mathbf{x}^{t,\text{grid}}$ (lower is better). For Experiment #4, values are means over 10 noise realizations.

Experiment	Space	RMSE [m]	CRPS
#1	grid	0.00	0.03
	param	0.00	0.03
#2	grid	0.10	0.05
	param	0.00	0.04
#3	grid	0.03	0.02
	param	0.00	0.03
#4	grid	0.09	0.05
	param	0.02	0.02

Table 4. Parameter estimation diagnostics for all experiments. Values are the signed errors of the analyzed parameters relative to the truth $\mathbf{x}^{t,\text{param}}$: $|\hat{\theta}_k^{\text{grid}} - x_k^{t,\text{param}}|$ for the gridded analysis (after Gaussian fitting) and $|\bar{x}_k^{a,\text{param}} - x_k^{t,\text{param}}|$ for the parametric analysis, for $k \in \{a, p, r\}$. For Experiment #4, values are means over 10 noise realizations.

Experiment	Space	Amplitude [m]	Position [km]	Radius [km]
#1	grid	0.00	0.00	0.00
	param	0.00	0.00	0.00
#2	grid	0.24	0.01	2.15
	param	0.00	0.04	0.00
#3	grid	0.00	0.00	0.01
	param	0.00	0.00	0.09
#4	grid	0.21	0.00	1.94
	param	0.02	0.00	0.24

tion positions: this spatial structure directly reflects the positional uncertainty of the eddy, and is a distinctive feature of the parametric representation.

240 The mean correlation matrix of the fitted parameters, estimated over the 10 noise realizations, is reported in Table 5. The dominant off-diagonal correlation is between amplitude a and radius r (-0.47 ± 0.01): wider fits tend to produce lower peak amplitudes, reflecting a geometric coupling introduced by the nonlinear fitting procedure. The correlation between position p and the other parameters is negligible ($< |0.01|$), consistent with the symmetry of the Gaussian profile. This non-diagonal covariance structure means that $\mathbf{B}^{\text{param}}$ and $\mathbf{R}^{\text{param}}$ are no longer diagonal, unlike in the tutorial experiments.

245 After assimilation, the posterior covariance $\mathbf{P}^{a,\text{param}}$ shows variances approximately halved relative to the prior, consistent with the expected uncertainty reduction for equal background and observation errors.

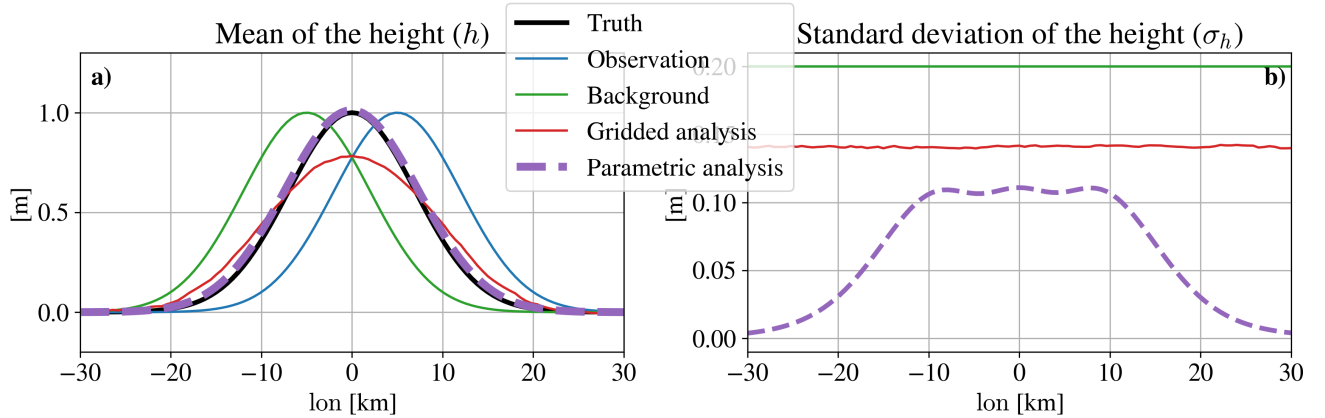


Figure 5. Results for Experiment #4 (realistic framework, position perturbation). Same color and panel convention as Fig. 2. Results are shown for a representative single noise realization; quantitative diagnostics in Tables 3 and 4 are means over 10 noise realizations.

Table 5. Mean parameter correlation matrix $\overline{\mathbf{B}}^{\text{param}}$ estimated from the fitting ensemble, averaged over 10 noise realizations (Experiment #4). Values are mean and standard deviations across realizations.

	a	p	r
a	1.00 ± 0.00	0.00 ± 0.01	-0.47 ± 0.01
p	0.00 ± 0.01	1.00 ± 0.00	-0.01 ± 0.02
r	-0.47 ± 0.01	-0.01 ± 0.02	1.00 ± 0.00

The metrics for Experiment #4 (Tables 3 and 4) confirm the qualitative observations. The RMSE increases to 0.09 m in `grid` (compared to 0.02 m in the `param` space). The CRPS in both spaces remains low, with still a higher difference between the `grid` space (0.05) and the `param` space (0.02). The parameter estimation in the `grid` space shows similar values compared to Experiment #2, with slightly improved performance (0.21 m instead of 0.24 m for the amplitude and 1.94 km instead of 2.15 km for the radius). Compared with the analytically prescribed covariance, estimating the parameter covariance from a noisy fitting procedure results in larger errors for the amplitude (0.02 m) and radius (0.24 km).

Overall, Experiment #4 shows that the parametric approach retains its qualitative advantage in recovering the coherent structure properties when starting from noisy gridded observations.

255 4.3 Physical illustration: Eddy Kinetic Energy

The profile deformations observed in Experiments #2, #3 and #4 have consequences for the physical properties of the reconstructed eddy. As an illustration of this effect, we compute the geostrophic velocity and associated Eddy Kinetic Energy (EKE)

Table 6. EKE [$kJ.m^{-2}$]: total integrated eddy kinetic energy computed from the analysis mean profile, reported as a physical illustration. For Experiment #4, values are means over 10 noise realizations.

Experiment	Truth	grid	param
#1	137	137	137
#2	607	275	607
#3	386	356	387
#4	607	304	600

from the analyzed height profiles. The geostrophic velocity is given by:

$$v_g(x) = -\frac{g}{f} \frac{\partial h}{\partial x}, \quad (15)$$

where $g = 9.81 \text{ m s}^{-2}$ is gravitational acceleration and $f = 10^{-4} \text{ s}^{-1}$ is the Coriolis parameter. The EKE per unit volume and the total integrated EKE are then:

$$\text{EKE}(x) = \frac{\rho}{2} v_g(x)^2, \quad \text{EKE}_{\text{tot}} = \int_{\text{lon}} \text{EKE}(x) dx, \quad (16)$$

with $\rho = 1000 \text{ kg m}^{-3}$.

This diagnostic is not used here as a validation metric, but as a concrete illustration of how structural distortions in the gridded analysis propagate into physically relevant quantities. Because EKE scales as the square of the height gradient, even moderate profile deformations can have a disproportionate impact on this quantity. Values are given in Table 6.

In Experiment #2, the widened and flattened `grid` analyzed profile produces smaller height gradients and therefore weaker geostrophic currents. The resulting total EKE is reduced by 55% relative to the value computed from the true profile. In Experiment #3, the sub-Gaussian shape similarly reduces the peak gradient, leading to an 8% reduction in total EKE relative to the true value. For both experiments, EKE estimations from the `param` projected profile have the same values as the truth. Similarly to Experiment #2, the flattened `grid` analyzed profile induces an EKE reduction of 50% while the slightly lower peak of the `param` projected profile reduces the EKE by only 1%. These reductions would directly affect estimates of lateral transport, air-sea interactions, and the mesoscale energy budget, if such an analysis were used operationally.

5 Conclusions

This study investigates data assimilation (DA) performed in a reduced parametric space and compares it with classic DA in gridded space, focusing on a simplified one-dimensional representation of an ocean eddy and on a single analysis step.

The results obtained in this idealized framework indicate that DA in gridded space provides satisfactory reconstructions when uncertainty mainly affects linear parameters such as amplitude, as illustrated by Experiment #1, where both approaches yield identical results (RMSE = 0.00 m, zero parameter error). However, when uncertainty involves nonlinear geometric parameters

280 such as position and radius, the gridded-space update leads to a deformation of the reconstructed structure. In Experiment #2, the gridded analysis underestimates the amplitude by 23% (error of 0.24 m) and overestimates the radius by 29% (error of 2.15 km), while the RMSE relative to the true profile reaches 0.10 m, compared with 0.00 m in parametric space. In Experiment #3, the analyzed profile is no longer Gaussian, and the RMSE in gridded space (0.03 m) exceeds that in parametric space (0.00 m). These structural distortions affect derived physical quantities: the total EKE is reduced by 67% in Experiment #2 and 285 by 11% in Experiment #3 relative to the parametric analysis. These EKE figures are reported as an illustration of the physical consequences of the profile deformation, rather than as validation metrics.

The tutorial experiments (#1-3) are useful to isolate and illustrate this mechanism, but they do not constitute a balanced end-to-end comparison, since the uncertainty originates in parameter space and is propagated to grid space. Experiment #4, which starts from noisy gridded observations, provides a more symmetric test. The parametric approach retains its qualitative 290 advantage: the RMSE of the gridded analysis (0.09 m) remains higher than that of the parametric analysis (0.02 m), and the gridded analysis underestimates the amplitude by 0.21 m and overestimates the radius by 1.94 km relative to the truth, compared to errors of 0.02 m and 0.24 km for the parametric analysis. However, the calibration of the parametric error covariances is reduced compared to the tutorial experiments, due to the additional uncertainty introduced by the nonlinear fitting step. Conclusions drawn from this study should therefore be understood as pertaining to these specific idealized configurations and 295 not as general statements about the relative performance of gridded and parametric DA approaches.

In contrast, the parametric formulation provides a direct update of the structure-defining parameters in this controlled setting. This leads to a more consistent representation of amplitude, position and radius within the assumptions of the experiments. In this idealized setting, where the three-parameter representation is exact and parameter extraction is assumed costless, the parametric formulation also operates in a much lower-dimensional space than the gridded approach. However, this dimensional 300 advantage is conditional on the availability of a reliable and efficient parameter extraction procedure; in practical applications, detection, fitting, covariance estimation and feature association would add to computational cost in ways not quantified here. In this study, no ensemble-size sensitivity analysis was performed, and therefore no general conclusion can be drawn regarding minimum ensemble requirements in either space.

An important limitation of the present work is the specification of error statistics in parameter space. The non-diagonal 305 covariance structure found in Experiment #4, with a significant a - r correlation of -0.47 , illustrates that the coupling between parameters introduced by the fitting procedure is non-negligible. The sensitivity of the results to this correlation structure, and more generally to the state dependence of $\mathbf{R}^{\text{param}}$, remains an open question that would need to be addressed in future work.

This work is intentionally limited to a controlled and idealized setting and does not include sequential assimilation cycles or prognostic model dynamics. Future work will investigate the extension of this framework to realistic, time-evolving situations, 310 where eddy parameters must be inferred from observations through detection and fitting procedures, and updated sequentially. In such a context, more general dynamical models for parameter evolution will be required, which may include stochastic parameter models, reduced-order models, tracking-based models, or hybrid approaches, rather than being restricted to any specific method. In particular, the dynamics of eddy parameters in a sequential context are expected to be strongly nonlinear,

which may necessitate ensemble-based or nonlinear assimilation methods. The stability of the system over multiple cycles will
315 also be investigated in future work, when this parametric approach is applied in an operational, sequential assimilation context.

Finally, the parametric and gridded formulations provide two complementary perspectives: the former emphasizes the evolution of physically interpretable structure parameters, while the latter operates directly on spatial fields. Depending on the application, analyses can be projected back into gridded space for diagnostic purposes or retained in parameter space for object-oriented applications that directly exploit coherent structure properties.

320 *Code availability.* The code used for the experiments is available here: <https://gitlab.imt-atlantique.fr/s24dealb/param-grid-1d-eddy>

Author contributions. SD developed the code, performed the experiments and prepared the manuscript, with contributions from co-authors; SR proposed the research topic, provided knowledge and background, performed the seminal experiments, and offered mentorship throughout this research; PT and CGB provided knowledge and assistance during code development; CL and BB helped with the writing of the paper.

Acknowledgements. This study was carried out as part of the PTD PROTEVS2, project financed by the French Ministry of Defense / DGA
325 and led by Shom.

References

- Alvera-Azcárate, A., Van Der Zande, D., Barth, A., Dille, A., Massant, J., and Beckers, J.-M.: Generation of super-resolution gap-free ocean colour satellite products using data-interpolating empirical orthogonal functions (DINEOF), *Ocean Science*, 21, 787–805, <https://doi.org/10.5194/os-21-787-2025>, 2025.
- 330 Beezley, J. D. and Mandel, J.: Morphing ensemble Kalman filters, *Tellus*, 60, 2008.
- Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, *WIREs Climate Change*, 9, e535, <https://doi.org/10.1002/wcc.535>, 2018.
- Chelton, D. B., Schlax, M. G., and Samelson, R. M.: Global observations of nonlinear mesoscale eddies, *Progress in Oceanography*, 91, 167–216, <https://doi.org/10.1016/j.pocean.2011.01.002>, 2011.
- 335 Chen, Y. and Snyder, C.: Assimilating Vortex Position with an Ensemble Kalman Filter, *Monthly Weather Review*, 135, 1828–1845, <https://doi.org/10.1175/MWR3351.1>, 2007.
- Hannachi, A., Jolliffe, I. T., and Stephenson, D. B.: Empirical orthogonal functions and related techniques in atmospheric science: A review, *International Journal of Climatology*, 27, 1119–1152, <https://doi.org/10.1002/joc.1499>, 2007.
- Jazwinski, A. H.: *Stochastic Processes and Filtering Theory*, 1970.
- 340 Maclean, J., Santitissadeekorn, N., and Jones, C. K.: A coherent structure approach for parameter estimation in Lagrangian Data Assimilation, *Physica D: Nonlinear Phenomena*, 360, 36–45, <https://doi.org/10.1016/j.physd.2017.08.007>, 2017.
- Poterjoy, J.: A Localized Particle Filter for High-Dimensional Nonlinear Systems, *Monthly Weather Review*, 144, 59–76, <https://doi.org/10.1175/MWR-D-15-0163.1>, 2016.
- Ravela, S., Emanuel, K., and McLaughlin, D.: Data assimilation by field alignment, *Physica D: Nonlinear Phenomena*, 230, 127–145, <https://doi.org/10.1016/j.physd.2006.09.035>, 2007.
- 345 Srivastava, A., Kang, W., and Tartakovsky, D. M.: Feature-informed data assimilation, *Journal of Computational Physics*, 494, 112 499, <https://doi.org/10.1016/j.jcp.2023.112499>, 2023.
- van Leeuwen, P. J.: Nonlinear data assimilation in geosciences: an extremely efficient particle filter, *Quarterly Journal of the Royal Meteorological Society*, 136, 1991–1999, <https://doi.org/10.1002/qj.699>, 2010.
- 350 Zhen, Y., Resseguier, V., and Chapron, B.: Alignment of geophysical fields: A differential geometry perspective, *Physica D: Nonlinear Phenomena*, 483, 134 997, <https://doi.org/10.1016/j.physd.2025.134997>, 2025.