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

February 23, 2026

Overall assessment

The manuscript provides a clear and pedagogically effective demonstration of a known difficulty in geophysical data assimilation (DA): when uncertainty is dominated by displacement and/or shape variability of coherent structures (e.g., eddies, cyclones), Gaussian/linear analysis updates performed in gridded state space can smear modes and shape.

However, in its current form I find the comparison between grid-space DA and parameter-space DA to be asymmetric in the assumed observation model and error structure. As a result, the experiments convincingly illustrate a representation/coordinate effect (Gaussian updates behave better in coordinates aligned with the dominant uncertainty modes), but do not yet support several broad claims made in the abstract and conclusions ("new DA scheme", "classic gridded DA fails" in general, and "considerably reduces computational cost").

Recommendation: Major revisions

The work could become publishable in NPG either by (i) adding an end-to-end/symmetric observation experiment and stronger quantitative support, or (ii) reframing the manuscript explicitly as a conceptual/didactic note with more moderate claims.

Brief summary of the manuscript

This manuscript considers a 1D Gaussian eddy profile parameterized by amplitude a, position p, and radius r, and compare DA performed in:

1. Parametric space: assimilating the 3D parameter vector $\theta = (a, p, r)$;
2. Grid space: assimilating the discretized profile $h(x; \theta)$ on a 1D grid.

Through numerical experiments where individual parameters are perturbed one at a time, namely linear case: a; nonlinear cases: p and r, the manuscript shows that grid-space Gaussian updates can yield distorted eddy shapes, whereas parameter-space updates preserve eddy properties such as shape and position much better.

Major comments

1. The central methodological comparison appears asymmetric in what is treated as "observed".

   In Parameter-space DA the experiments effectively assume access to observations in parameter space, i.e. a noisy (a, p, r) with simple Gaussian errors (diagonal, equal-variance structure). Under these assumptions, the parameter update is close to a textbook low-dimensional Gaussian analysis. In that regime the analysis can indeed be interpreted as the Bayesian posterior mean (under linear-Gaussian assumptions).

   Ath the same time in Grid-space DA the gridded state distribution is induced by pushing parameter uncertainty through the nonlinear map $\theta \mapsto h(\cdot; \theta)$. For displacement (p) or scale (r) uncertainty, this induced distribution over h is nonlinear and thus inherently non-Gaussian (mixture-like / multi-modal). A Kalman/EnOI-style update uses only mean/covariance information, i.e. implicitly "Gaussianizes" the non-Gaussian problem, which is known to produce smearing for displaced coherent structures.

   It reads to me like the presented experiments primarily demonstrate that the Gaussian analysis behaves better in coordinates where uncertainty is closer to Gaussian, rather than providing an end-to-end fair comparison between DA schemes. In a practical setting there appears to be no reason to assume the parameters, predicted and observed, would be Gaussian, while the grid prediction and observations are not.

   The assumption that (a, p, r) are so neatly Gaussian, needs justification, if even possible. Please clarify the intended practical workflow:

   (a) Are parameter observations (a, p, r) assumed to be provided externally by a detection/fitting pipeline with known uncertainty?

   (b) If yes, how is this uncertainty quantified in practice (correlations, state dependence, missed detections, multi-feature association)?

   (c) If no, how are parameters inferred from raw observations within the DA system?

2. To strengthen the claims and make the comparison more symmetric, I recommend adding at least one experiment in which both methods start from the same noisy gridded observations:

   $$y_h(x) = h(x; \theta^\star) + \epsilon(x),$$

   with specified sampling and noise model (potentially sparse/irregular).

   For Grid DA assimilate y_h directly (as currently). For Param DA derive $\hat{\theta} = g(y_h)$ via the stated fitting/detection method and estimate an observation error covariance R_θ for $\hat{\theta}$ (likely non-diagonal and state dependent; e.g. a–r coupling). Then assimilate $\hat{\theta}$ in parameter space.

   This would test whether the parametric approach remains advantageous once it "pays" for the inverse nonlinearity that is currently idealized away. This likely also mirrors a real practical setting more closely, and reflects the real computational cost more realistically then the current setting.

3. The abstract states that classic DA in gridded space fails to estimate position and intensity in the presence of coherent structures. This is plausible for Gaussian/linear updates without displacement correction, but reads too broadly. There exists a literature on displacement-aware approaches (alignment/morphing/registration, feature-based increments, mixture methods, particle filters, etc.). I recommend either narrowing the claim to the class of Gaussian/linear

updates without alignment corrections, or including at least one baseline that attempts a minimal displacement-aware correction in grid space (e.g. shift registration pre-processing) to avoid a strawman comparison.

4. It is further claimed that parameter-space DA considerably reduces computational cost. This seems obvious in the discussed synthetic setting given the dimension of the parameter space is lower than that of the grid space, but ignores that in practical settings obtaining the parameters, predicted and observed, along with their covariance comes at a cost as well. The current setup, also uses large ensembles in either case, and no timing/complexity evidence is provided. In realistic applications, parameter DA also presupposes a detection/fitting pipeline and (potentially) multi-feature association, which carry costs.

   I suggest either: adding explicit wall-clock timing / complexity scaling (dimension m vs. $n = 3$, ensemble size dependence, cost of fitting/detection), or softening the statement to a conditional claim ("potentially reduces cost when reliable parameter extraction is available and the number of features is small relative to grid dimension").

Minor comments

1. The claim that DA in reduced paramter/feature space has not been studied seems overreaching and needs softening. Rather elaborate how this fits in with existing feature- based DA and displacement aware DA.

2. State explicitly that the experiments focus on a single analysis step (no forecast–analysis cycling), and clarify what would be required to integrate the method into a full DA cycle.

3. If possible, include discussion (or sensitivity tests) showing the impact of correlated parameter errors (notably a–r coupling) and state dependence of R_θ. Quantifiable metrics to compare quality and parameter dependence of each method would be a major improvement.

Conclusion

I appreciate the clarity and pedagogical value of the manuscript. With major revisions that either (i) add a symmetric observation experiment / forward-operator parameter inference, or (ii) reframe the manuscript as a conceptual demonstration with moderated claims, this work could become a useful contribution to NPG.