ParticleDA.jl v.1.0: A distributed particle filtering data assimilation package

Response 3

January 13, 2024

We would like to thank the reviewers and editor for their helpful comments and feedback. Our response to their points are outlined below. An updated version of the article with highlights of the changes made has been uploaded.

1 Reviewer 3

I appreciate the authors' responses in answering my questions in the previous iteration. The revised manuscript addresses most of my concerns. In my opinion, the manuscript is nearly ready for publication, with a few minor issues below that need addressing.

1. Line 336: I am curious about why the locally optimal proposal cannot be applied for the non-linear observation operator? Even if the linear assumption is invalid, shouldn't it still be applicable? (like we can still apply the ensemble Kalman filter to assimilate non-linear observation even if it's not optimal) Please clarify if there's anything that I might be missing here.

   Author's response: This is an interesting question, and in part depends what is specifically meant by the *locally optimal proposal*.

   If we take the *locally optimal proposal* to specifically refer to the distribution with density defined by equation (8) in the paper, then assuming a non-linear observation model (but continuing to assume Gaussian observation noise with a fixed variance), that is $\boldsymbol{y}_t \mid \boldsymbol{x}_t \sim \mathcal{N}(h_t(\boldsymbol{x}_t), R)$, an immediate issue in evaluating the density in equation (8) or sampling from the corresponding distribution is that we now have a nonlinear operator $h_t : \mathbb{R}^{d_x} \to \mathbb{R}^{d_y}$ rather than a linear operator represented by a matrix $H \in \mathbb{R}^{d_x \times d_y}$. One obvious approach to evaluating the matrix products involving H in equation (8) would be to use a linearization of h_t around the input as a substitute for H, that is $H \approx \partial h_t(\boldsymbol{x}_t)$, the Jacobian of h_t evaluated at $\boldsymbol{x}_t$. If we adopted this approach then we could evaluate the density in equation (8) and sample from the corresponding distribution, however, the resulting proposal distribution would no longer be the locally optimal (in the sense of minimising the variance of the importance weights) nor would the expression for the importance weights in equation (9) be valid (as terms in the density ratio which exactly cancel in the linear case would no longer do so in the non-linear approximation). We could use the underlying definition of the importance weights as a density ratio given in the equation in step 5 of Algorithm 1 to compute valid importance weights, but importantly these would depend on both $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t-1}$ (sampled state proposal and previous state) unlike the true locally optimal proposal, for which by construction the importance weights depend only on the previous state (and so have no variance contribution from the proposal distribution). So we can use a proposal which is an approximation to the locally optimal proposal for the linear-Gaussian observation case, but the resulting proposal is not the locally optimal proposal for the model.

   If we instead consider the general definition of the density of the locally optimal proposal in equation (6) and corresponding expression for the importance weights in equation (7), then for an observation model $\boldsymbol{y}_t \mid \boldsymbol{x}_t \sim \mathcal{N}(h_t(\boldsymbol{x}_t), R)$, the integral appearing in both equations (6) and (7) is analytically intractable. We could potentially estimate this integral and use this to form an approximation to the locally optimal proposal distribution, with in comparison to the approach above we at least in this situation directly approximating the distribution of interest, rather than a proxy to that distribution. Depending on exactly on how we estimate the integral though, the resulting proposal distribution may no longer be from a known parametric family and so it may be challenging to generate independent samples from.

2. Line 344: "The time averaged RMSE results are plotted in 5." -¿ Figure 5.

   Author's response: The missing Fig. has been added.

3. Line 344-345: "The effect of the non-linear observation operator can be clearly seen, where the time averaged RMSE for the linear case is lower in all set-ups".

   This statement is inaccurate. There are a lot of factors determining the magnitude of RMSE. Assimilating a linear observation (i.e., with linear observation operator) with very large observation error standard deviation can also lead to a large RMSE. When the observation error standard deviation is fixed, using different observation operator leads to different shape of likelihood function, and therefore different magnitude of RMSE. I recommend rephrasing this sentence or just removing it.

   Author's response: Thank you for pointing this out, this statement has now been dropped.

4. Line 449: what are the units for the standard deviation (especially for ps)?

   Author's response: The units have been added.

5. Line 459: "10 hPa" noises in ps seems to be quite large

   Author's response: Please see the response below.

6. Figure 9: I am still skeptical about the RMSE values in these experiments. Even in a model run without DA (the lower left panel), the RMSE are unrealistically large, suggesting the large RMSE is not related to the observing system. I suspect that this is a result of the experiment setup that the ps standard deviation in Q is set to 100 (hPa) and the observation error is set to 10 (hPa). I think setting the standard deviation as 100 (Pa) = 1 (hPa) and 10 (Pa) = 0.1 (hPa) could lead to a more realistic representation.

   Author's response: To clarify the existing experiment was run with the ps standard deviation in Q set to 100 Pa and the standard deviation of the additive observation noise set to 1000 Pa. The authors apologise for the confusion and have added the units to the appropriate section. The authors would also like to further highlight a mistake made in the previous round. As stated by the reviewer the L_2 error is indeed too large. Unfortunately, the appropriate scaling of the pressure fields was not carried out correctly when producing the figures. The corrected figure (Fig. 1) is produced below. We would like to thank the reviewer for pointing this out and apologise for the oversight.

   To ensure that we have fully engaged with the reviewer's feedback Fig. 2 is the output of the suggested set up, where the observation error standard deviation is set to 10 Pa. Differences between Fig. 1 and Fig. 2 can be seen but when assessing the estimated ensemble size for this run (not shown here) we see occurrences of particle degeneracy. Therefore the original setup with the updated figure (Fig. 1) has been retained in the manuscript.

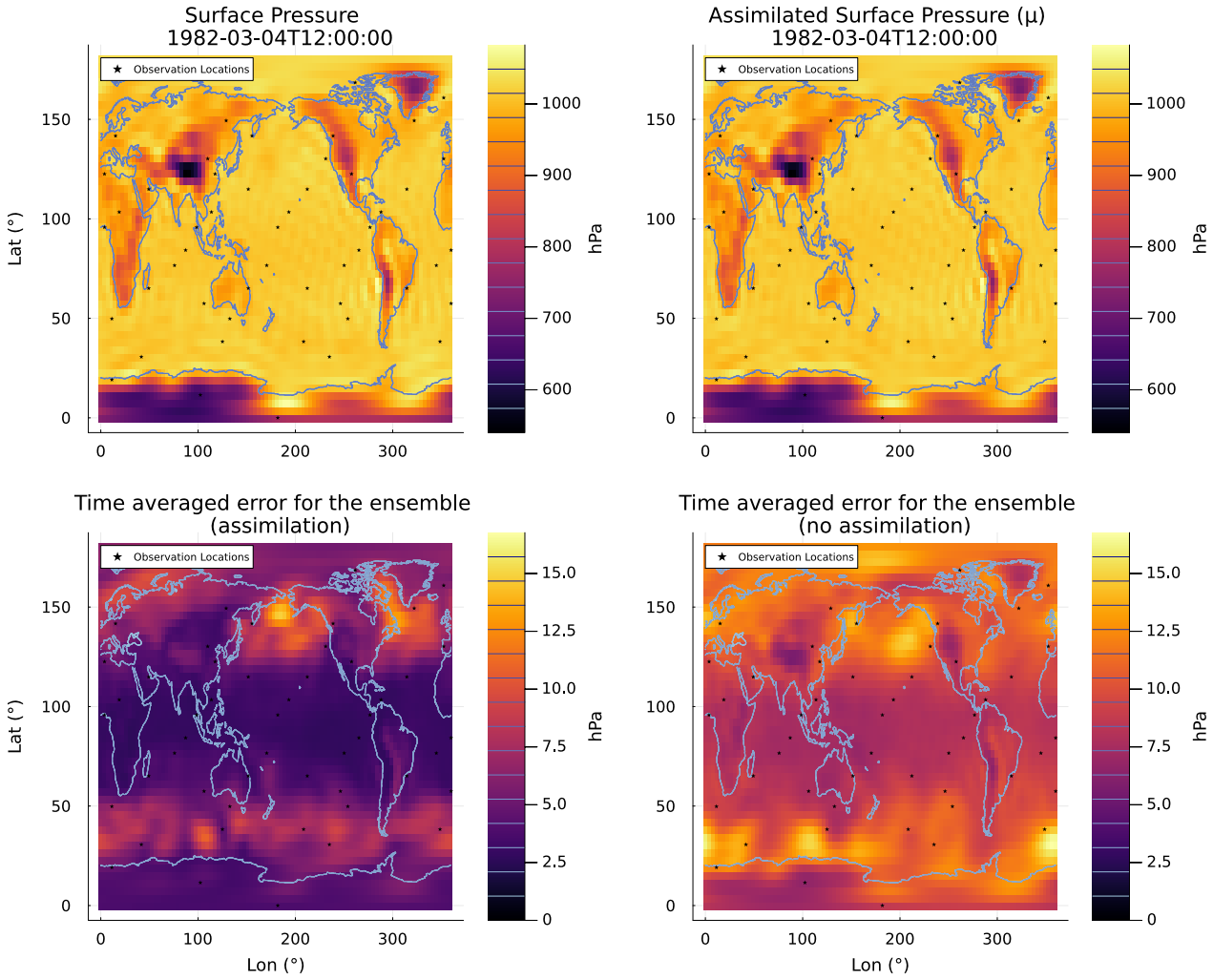

Figure 1: Corrected Fig 9. from the manuscript.

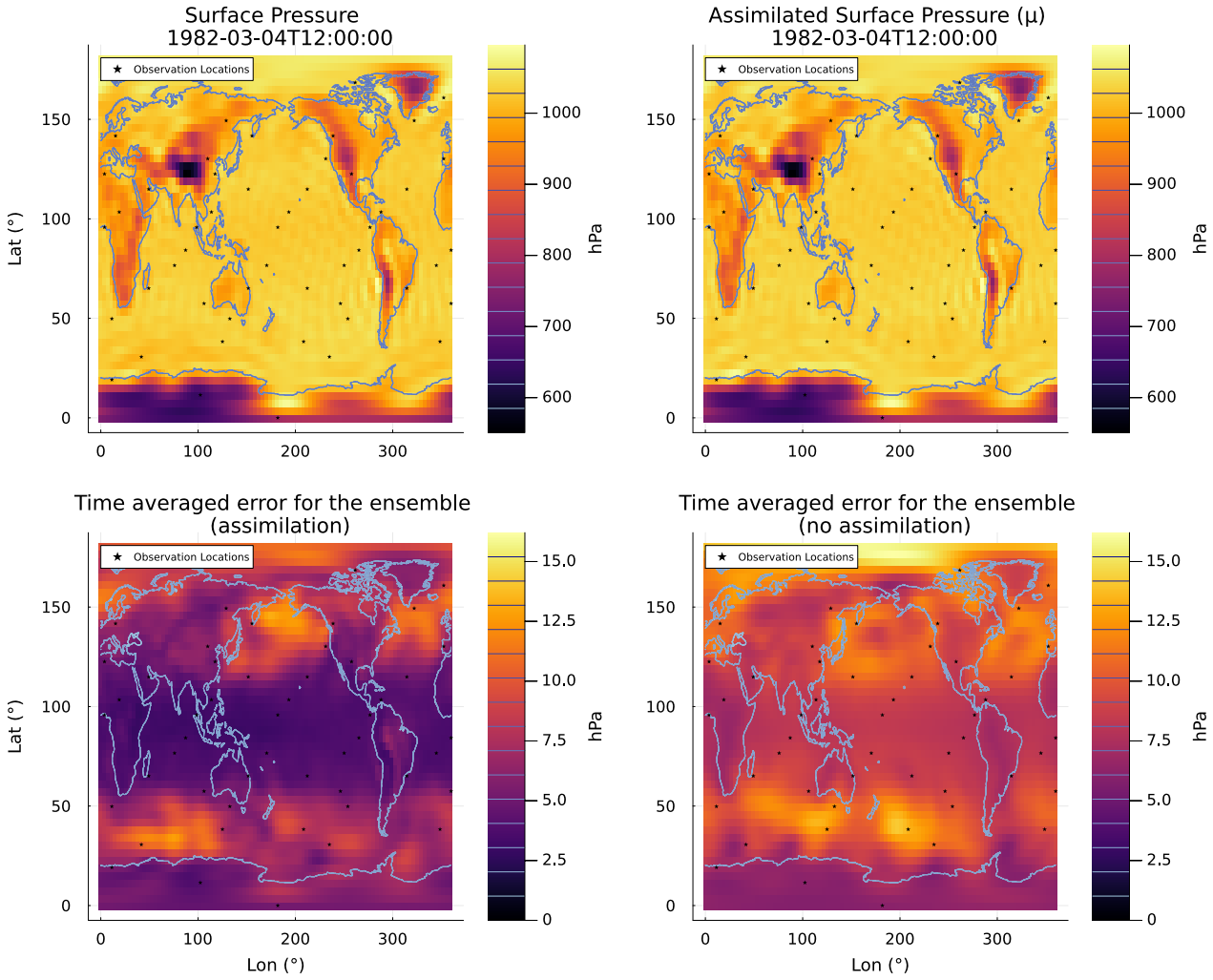

Figure 2: Reviewer's suggested setup with the standard deviation of the additive observation noise set to 10 Pa.