Review of gmd-2023-38: *ParticleDA.jl v.1.0: A real-time data assimilation software platform*

By Daniel Giles, Matthew M. Graham, Mosè Giordano, Tuomas Koskela, Alexandros Beskos, and Serge Guillas

13 Apr 2023

Verdict: **Major Revision**

In this paper the authors introduce a new data assimilation software framework, ParticleDA.jl. This uses the particle filter algorithm, which is among the class of ensemble data assimilation algorithms and in principle has the best general performance in nonlinear settings where Gaussianity cannot be assumed. They have developed this software framework in Julia in a modular way which makes it easy to integrate it to arbitrary nonlinear dynamical systems, and they demonstrate its applicability to a series of models of increasing complexity. The authors have also focused somewhat on the high-performance capability of their package, and demonstrate weak scaling performance up to 16 nodes, though I would consider this just a start and tests with much higher node counts are required. The paper was easy to read and provides sufficient detail to reproduce the experiments. I would recommend acceptance after the below major comment is addressed, which pertains to further verification that their assimilation methodology is working correctly for the highest complexity model considered, SPEEDY. I am looking forward to reading their response.

Major Comments

- Line 312: My only major comment was triggered by this statement: “It can be seen that the areas of greatest percentage error coincide with areas that lack observation stations.” I want to challenge this statement as I don’t see such a strong coincidence. Instead the error patterns seem to be dominated by patterns of midlatitude weather systems, which is what you would expect to see if you compared two snapshots of surface pressure in the same model run at different times. In fact, I don’t see sufficient evidence that the assimilation is actually working. It probably is, but it’s hard to tell unless you compare also with a run without any assimilation. A number of further plots would help to make this section more complete:
    1. A snapshot of surface pressure for the truth run some time into the run (after the assimilation error has asymptoted).
    2. A corresponding snapshot of the surface pressure averaged across the particle ensemble (or a randomly chosen particle, as you prefer — they should be similar if the ensemble has mostly converged on the truth). This should be close to plot 1. if assimilation is working.
    3. A corresponding snapshot of the surface pressure for a run without any assimilation (just a single run) to demonstrate how much divergence would be expected.

4. To complete the 4-plot square, you could also show an error plot. I would suggest to produce a plot like Fig. 7 right, but averaged over time. This should filter out the midlatitude variability and actually show an error pattern correlated with the observation locations.

Note that surface pressure is an odd variable to plot due to the dominance of mountainous areas. You might prefer to use 850 hPa temperature or something else with fewer regional nonuniformities. The above is just a guide and I leave it to the authors to demonstrate that their assimilation methodology is working correctly for the SPEEDY model.

Minor Comments

- Line 65: typo — “targeted”.
- Line 169: typo — “RandomFields.jl”.
- Line 221: Could you add one or two more sentences to elaborate on the significance of the ESS? Why is this an interesting thing to note?
- Line 232: typo — “warning centre’s”.
- Section 6.2: Considering, say, the 32 ranks per node case, the biggest test runs on only 4 nodes (128 / 32). Yet the parallel efficiency has already dropped to only around 25%. This is a much bigger drop that I would expect, intuitively. Am I missing something? Scaling tests often run into the hundreds of nodes before encountering such limits.
- Figure 6, left: Could you clarify which measure of weak scaling parallel efficiency you employ here? I am guessing it is E(N) = T(1)/T(N), where T(i) is the wall time for running on i processors.
- Line 310: “The standard deviation of the model and observation errors are set to 1 and 10 hPa respectively.” — isn’t the observation error already stated on line 309? Also what is the “model error” in this case? I thought that the SPEEDY model is integrated without a model error term?