Referee comments on manuscript gmd-2023-15
Anonymous: 28 April 2023

In this manuscript, an alternative *a priori* flux constraint is presented in the context of a global CO_2 flux inversion performed using an ensemble Kalman filter (EnKF) with a short assimilation window. Observing system simulation studies (OSSEs) are preformed to give an idea of how this alternative constraint might function when used with real data in a real inversion. The flavor of EnKF used is the local ensemble transform Kalman filter (LETKF), as implemented in the Carbon in Ocean–Land–Atmosphere (COLA) data assimilation system, a global CO_2 flux inversion based on the GEOS-Chem transport model.

The alternative flux constraint is formulated in terms of the spatial gradient of the fluxes: finite differences of flux using adjacent grid boxes in the model. These spatial gradients are then added as new measurements in the measurement vector, as opposed to additional constraints in the traditional *a priori* state vector. Gradients used in this manner could capture the bulk of the flux constraint (its spatial and temporal patterns), while at the same time cutting the tie to the absolute value of the flux -- i.e. its overall constant offset or long-term mean. This in turn could be useful when using priors for which the variability is more robust than the long-term mean -- for example, the terrestrial biosphere models used as priors for CO_2 fluxes over land in global flux inversions, which do a good job getting the seasonality of the fluxes right (e.g., using satellite measurements vegetation greenness, plus assumptions on the timing of respiration) but a less-good job of estimating the integrated flux across a full year. By getting rid of the constraint to the long-term mean of the prior, the flux estimate might be freer to move to the long-term mean given by the data and not suffer from being biased in the direction of the incorrect or inaccurate prior. This of course would be at the cost of losing any benefit that that long-term prior mean might provide.

In general, a flux constraint of this nature should be able to be implemented as a measurement in the measurement vector, as is done here, assuming that the measurement uncertainty used gives the constraint the same weight as it would have had if it had been implemented more traditionally in the *a priori* state vector. One would have to avoid double counting by not also having the traditional flux prior in force at the same time.

In their OSSE experiments, the authors compare the effectiveness of this flux spatial gradient constraint against the usual prior flux constraint (i.e. in terms of the actual flux value itself, not the spatial gradient) implemented either in the measurement vector or, more traditionally, as part of the *a priori* state vector; in the latter case, a couple different forms for the first guess of the flux at the new measurement time are used: either 1) a combination of the prior flux at the given time plus the flux estimate from the EnKF at the two immediately-earlier times, or 2) just the prior flux at the new time. This is done using one land biospheric model (VEGAS) to generate the 'true' measurements, and a second model (CASA) to be used as the prior flux. The authors find that, in general, when the flux gradient prior is used, the EnKF does a better job estimating the true fluxes than when three other approaches based on the absolute fluxes themselves (i.e., not gradients) are used.

While these results look promising, there are some inconsistencies in the results that I would like explained. Also, I suggest modified OSSEs in which the ocean fluxes are allowed to be corrected along with the land fluxes, in order to give a more realistic test of the new constraint. Finally, there is a lack of detail in the description of the methods used that makes it difficult for me as a reviewer to assess the full meaning of the results. I suspect that the general reader will have similar questions. I suggest that the authors add these needed details to the manuscript, address the points that I raise below, and resubmit, at which point I will re-review it and decide on final publication.

Comments:

First, the authors should describe in detail [with equations] the meaning of the terms 'assimilation window' and observation window', since how these terms are used in the context of the LETKF is not generally known. The reader should not have to go back to the previous LETKF papers to find this. Does the 1-day assimilation window mean that the filter is stepped forward in time a day at a time, each day allowing the new measurements to update the fluxes across the 7-day measurement window (i.e. the current day plus six previous days)? If so, the weight given to the flux constraint (or flux prior constraint) for each of those 7 days ought to be reduced, so that the integrated effect of the seven days of measurement updates affecting the fluxes on a given day is equivalent to the weight given to a single days' flux prior in some other estimation method (e.g. a variational method or a matrix-inversion-based Bayesian synthesis method).

Second, the weights given to the spatial gradient constraint in the inversion relative to the straight flux constraint cases ought to be given. Perhaps the spatial gradient case does a better job because it has a looser (or tighter) weighting than the other cases. A tighter flux prior usually results in a worse fit to the measurement data; or, vice versa, the inversion can over-fit the measurement data at the cost of too great a change from the flux prior. Knowing the weights assumed in the inversion for the gradient case *vis a vis* the straight flux case could help assess this. Similarly, some information on how good the fit to the measurement data is for the four cases could help.

Third, if the flux constraint can be implemented equally as well in the measurement vector as in the *a priori* state vector, then the two cases in which the straight flux prior are implemented these two ways should give the same flux results. That is, the EXP-NP case, in which the flux prior is applied normally, as the *a priori* constraint on the fluxes in the state vector, and the EXP-AP case, in which the flux prior is assimilated as a measurement in the measurement vector, should give the same flux estimates. But they don't -- they give quite different answers, as seen by the turquois and orange lines in Figures 3 through 5. What is it about the different implementation of the prior that causes these differences? Different weights used in each case? A different number of times that the constraint is applied (if fluxes at multiple times are updated by measurements at a single time)? Similarly in Figures 6 and 7, the EXP-NP case gives much worse RMSEs for flux and flux spatial gradient than does EXP-AP. Why is this, if the two ways of implementing the prior are equivalent? I can understand why, with a short-window inversion, the EXP-NP case might have higher values for these metrics (i.e. a flux error frozen in at a given assimilation step would need to be corrected by a balancing error at the next step of opposite sign, resulting in a lot of noise in time), but what is it about the EXP-AP implementation that prevents this?

Fourth, because the OSSE experiments use the same ocean fluxes in the truth and assimilation runs, there is effectively no error coming from the oceans and no need to allocate any flux corrections there in the inversions. This is effectively the same thing as holding the oceans fixed and only allowing flux changes over the land areas. This significantly simplifies the inversion and gives an overly-optimistic view of how well the inversions can retrieve the land fluxes. However, even worse, it may favor the spatial gradient prior constraint more than the straight flux prior constraint, since, with the ocean corrections fixed to zero, the fluxes bordering the oceans are then strongly constrained by the spatial gradient constraint, and the fluxes in the interior similarly prevented from moving as much as they otherwise would. With the straight flux constraint, however, the fluxes are still allowed to trade off corrections between continents. It would be interesting to see whether these same favorable results with the EXP-ASG case are achieved if more realistic errors are allowed over the oceans (i.e., if separate ocean flux models were used in generating the truth and prior, as has been done with the land biospheric fluxes here).

Fifth, it would be useful for the authors to discuss how specific their results are to the flux inversion method they use (a short-window EnKF). Would they anticipate that the alternative flux spatial gradient constraint would give similar improvements in methods that allow the transport model to link measurements and flux corrections across a longer span? Similarly, since this reliance on the transport model is less important when there is more data coverage, would the results obtained here still hold were a less-dense observing network (the *in situ* CO_2 network instead of a CO_2-measuring satellite, say) to be used?

More-detailed comments:

14: “dynamic constraints” I do not believe that the reason the inversion problem is ill-posed is because of the lack of explicit dynamical constraints in the setup. Really it is due to the sparse data.

16-17: "Ensemble Kalman filter-based inversion algorithms usually weigh a priori flux to the background or directly replace the background with the a priori flux." It is not very clear what this means. Please reword. What do you mean by ‘background’?

21: spell out "AAPO"? It is not clear why you use this combination of letters for what you are describing.

38: I wouldn't say the problem is ‘ill-posed’ because of transport errors or retrieval biases -- those just bias the result. Ill-posedness is more due to lack of a sufficient data constraint, for example, trying to solve for more unknowns than can be constrained by a given number of data points.

49: "the LETKF with a short assimilation window and long observation window setting"
I do not see this described later in the text. Please describe what these ‘window’ terms refer to, for example in terms of the filter time stepping, what span of data is assimilated at each time step, and what span of fluxes is allowed to change per time step; preferably with equations.

54-56: "On the other hand, even though *a priori* information includes biases, it could be used to further improve the SCF estimation in COLA because it includes important dynamic information generated by terrestrial models, which is missing in the top-down inversion system."
It is not clear why you think that dynamic information generated by the terrestrial models is not represented in the top-down inversion systems. Insofar as it is used to generate the *a priori* SCFs, it is in there. Do you mean to say that the dynamical constraint of the *a priori* fluxes is not represented explicitly as a dynamic model in the Kalman filter, i.e. as a formal constraint?

75: add "at" after "including"

77-81 ”Similar to the other EnKF, the LETKF prefers a short assimilation window to produce accurate model state analysis, which reduces noise within the background for parameter estimation. On the other hand, parameter estimation requires a long training period to enhance the model response to the estimated parameter (the signal). Therefore, COLA implements a new version of LETKF with a unique feature of a short assimilation window (1 day) and a long observation window (7 days) to enhance the SCF estimation (Liu et al., 2019).”

It is not clear how these various ‘windows’ relate to the fluxes being solved for. You should write out with equations what is being solved for, how the time stepping is done, what observations are assimilated in which time step with which weights, etc. And point out which spans are the ‘observation window’ versus the ‘assimilation window’. This may be detailed in previous LETKF papers, but the reader shouldn’t have to go back to them to understand what is being used here.

119: “In COLA, the main purpose of applying *a priori* regularization is to introduce the dynamic constraint for SCF estimation.”

It is not at all clear that you have now introduced a better dynamic constraint by changing from using the prior flux value to using spatial gradients instead. Nothing involving dynamics has been changed by this. All you have succeeded in doing is removing the link to the overall absolute value of the prior flux (the long-term mean). That may indeed have value, but don’t confuse it with dynamics. Any dynamics that were or were not in the original flux prior are still there with this new constraint. Please reword to reflect this, here and elsewhere in the document where ‘dynamics’ are discussed.

138-147: You are free to add dynamical noise to your propagation of information forward in time in your model. You should discuss why you choose not to add dynamical noise that reflects errors in your transport model and/or variability in the land fluxes not captured by a forward propagation based on persistence. Why do you instead add an inflation term that is based more on the technical needs of your EnKF rather than a physically-based dynamical error?

149-150: “COLA assimilates the *a priori* SCF spatial gradients into the system, which needs to define the *a priori* uncertainty. In this study, we simply set the *a priori* uncertainty proportional to the uncertainty of the analysis ensemble uncertainty.” Please describe what this analysis ensemble uncertainty looks like. Does it differentiate between forested areas that are likely to have larger fluxes and flux uncertainty and desert areas that are likely to have smaller ones? (Or similarly for flux gradients?) A sensitivity study done using uncertainties proportional to the magnitude of the fluxes in either the VEGAS or CASA models, or based on the difference between VEGAS and CASA (and preferably other models), would be welcome to test the dependence of your results on this assumption.

165-166: “We set the CO2 observation localization radius to 4000 kilometers.” Since the general reader probably will not understand what this means, please say what this means, practically, in your inversion setup. Does it mean literally that each observation has zero impact on any flux farther away than 4000 kilometers at a given time? What about at previous times?

168-174: By using the same fossil fuel, ocean, and wildfire fluxes in both the truth and
prior, the simulation is artificially rosy: terrestrial fluxes are solved for
using only differences there by permitting flux corrections only over the land and not over the ocean. By not considering the impact of ocean flux errors, this will give you
lower error estimates for the land fluxes than you'd get otherwise. It would be a useful sensitivity study to look at the impact of considering ocean flux errors, as well.
Figures 6 & 7: The difference between the EXP-NP and EXP-AP cases still needs to be explained. Yes, the short window of the COLA setup results in over-fitting of the data and noisy fluxes (and spatial gradients) in the EXP-NP case. But how does applying the prior flux constraint via the measurement vector prevent this?

290: What does 'dynamically' in 'dynamically assimilated' indicate? Is this some special sort of assimilation method? Also, define what the acronym 'AAPO' refers to.

297-304: "However, the advantage of error transport is partly sacrificed or abandoned by introducing the *a priori* flux information to the background in most of the EnKF-based CO_2 inversion methods (Peters et al., 2007; Feng et al., 2009). This is because of the loss of a dynamic model to provide the background and the background covariance estimations. Different from most EnKF-based systems, COLA maintains the mean and error transport advantages of the EnKF by including the dynamic information constraints of the a priori flux spatial gradient and using an additive covariance inflation method (Liu et al., 2022)."
I agree that the loss of the dynamical model for the fluxes in most of our flux inversion methodologies is unfortunate. I do not believe, however, that you are remedying that with your spatial gradient constraint here. Nothing has changed regarding the dynamics in using this constraint. Your only change is to cut the tie to the long-term mean, allowing your estimate to be shifted up or down as a whole more easily.

310: 'unique strategy'? Maybe referring to it as a 'new strategy' would be better.