Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2017-222-RC2, 2018

Geoscientific Model Development Discussions

Interactive comment on "CTDAS-Lagrange v1.0: A high-resolution data assimilation system for regional carbon dioxide observations" *by* Wei He et al.

Anonymous Referee #2

Received and published: 2 April 2018

Review of the paper entitled "CTDAS-Lagrange v1.0: A high-resolution data assimilation system for regional carbon dioxide observations" by He et al.

General comment:

The regional assimilation system presented here is the first semi-operational atmospheric inversion of carbon fluxes at the mesoscale. Compared to previous inversion systems, CTDAS-Lagrange has been significantly improved for high-resolution problems thanks to a state-of-the-art atmospheric model and a comprehensive optimization framework including the problem of boundary conditions. The inversion framework is similar to CarbonTracker but has been adapted to the optimization of a pixel-based

state vector. The sensitivity experiments performed here shows that the system is capable of reproducing the observed variability in CO2 mixing ratios across North America. The continental fluxes are consistent with global-scale inversions, and the robustness of the optimized surface fluxes to assumptions made in the inversion has been tested carefully. Therefore, we recommend this paper for publication after considering the general and specific comments listed here.

Boundary Conditions: Three products have been used here to describe the CO2 mixing ratios coming from outside the simulation domain. The results show that two of them (i.e. CT2013B and EMP) produce very similar results with nearly identical posterior fluxes. However, the third one (i.e. CTE2014) produces an offset in the late growing season which is not corrected for after inversion. The major concern here is related to the representativity of the three products. The fact that one of them remains significantly different despite the optimization process suggests that the uncertainties in the boundary conditions are not removed by the assimilation of data. Hence, the ensemble has to be representative of the actual uncertainties in boundary conditions to be properly propagated into the flux uncertainties. The authors have not clearly demonstrated that these three products represent the actual errors coming the boundary conditions. This study needs to demonstrate the value of the sensitivity tests and address more carefully the actual error propagation into the posterior fluxes and their uncertainties.

Multiplicative versus additive methods: Because the prior error variances differ significantly between the two experiments, the differences between the two methods depend on the prior error variance more than the actual method to invert the fluxes. The authors need to perform another simulation with similar prior errors to produce convincing evidences that the method used is the fundamental problem.

Correlation length scale in prior errors: In a system with a fairly small degree of freedom such as CTDAS, the spatial attribution of flux corrections may still be sensitive to the definition of the prior errors but the total flux is likely to remain unchanged. The convergence of the system using different length scales can be an artefact due to the degree of freedom but the spatial distribution of flux corrections may vary across the domain. The authors need to include maps of the flux corrections for the different length scales which will show the actual impact of prior error assumptions.

Technical comments:

P2-L10: add a reference.

P2-L11 and L12: Add references to previous studies.

P2-L14: This is true for global inversions but domain-limited inversions have opened boundaries. Refine this statement.

P2-L22: Feng et al., 2016 is not an inversion study.

P2-L25: Eulerian models are often a pre-requisite to Lagrangian models, like in CT-DAS. You may refer here to the model used in the assimilation framework, typically ensemble-based methods based on Eulerian models and analytical methods with a linearized adjoint model. This statement is unclear for column-based measurements. Refine the statement.

P2-L32: Peylin et al. (2005) discuss the importance of initial conditions in a global inversion. The problem is different for the lateral flow in a domain-limited inversion which does not decrease over time. Another citation is needed here.

Figure 1: The interpolation for land cover creates artificial zones between land cover types, with halos around temperate Crops / Agriculture for example. The native resolution of the land cover plotted for each grid cell is better suited for land cover maps. Replace contours by actual grid cell colors.

P4-L21: Indicate the time period over which the flasks have been sampled with the new protocol. How did you treat data with potential biases?

P5-L5: Have you considered the possible impact of surface fluxes when sampling at 3km and above? Some of the aircraft sampling locations could be impacted by verti-

cal transport of surface fluxes. Could you use CO to detect high-altitude surface flux influence?

P6-L3: Add a reference to previous studies.

P6-L7: The description of the WRF model configuration is critical for future users of the footprints. Provide a complete description of the model, as well as the simulation domain (map projection, . . .).

P6-L21: You assume here that the prior BC errors are dominant over 10 days. Is that consistent with model-mismatches? Synoptic systems are more likely to be the main source of errors in the inflow. Justify the 10-day optimization window for boundary conditions.

P6-L24: You assume here that the errors in the BC's are identical below and above 3000m. This assumption seems very unlikely as the modeled values near the surface will differ significantly from the modeled mixing ratios in the Free Troposphere. An evaluation of the model-data mismatches compared to the altitude would help infer the actual vertical structures in BC errors. P8-L2: Which observations are assimilated here? Daily afternoon averages?

P11-L18: Justify the removal of outliers here. The factor of 3 applied to the MDM is arbitrary. Any physical reason behind this?

P12-L25: The reduction in uncertainty is a direct consequence of the optimization process but does not mean that the actual errors are reduced. Similar to the posterior CO_2 mixing ratio mismatches, the optimization was designed to reduce it, except that fluxes have no guarantee that the reduction is real. Clarify in the paragraph.

Figure 4: The aggregation of large ecoregions over the entire continent makes the interpretation of the results very difficult. The ecoregions are too wide to make sense of the seasonal cycles that correspond to multiple regions and climates. The separation of the results into smaller regions is needed here. For example, southern and northern

regions for crops would provide a better sense of what the seasonal should look like. This figure needs to represent the ecosystems defined in climate zones.

P13-L26: - Except for CTE2014, the adjustment to the boundary conditions converges to identical posterior flux values for CT2013B and EMP. These two products are fundamentally different, but the correction to the BC's are significantly different, which suggests that the inversion is able to correct and hence reduce the initial differences between the two. However, CTE2014 produces a different result, which would suggest that the BC errors are too small, or that the lack of data limits the ability of CTDAS to reduce the BC mismatches. Provide more elements here to understand what caused the final disagreement between CTE2014 and the two other BC's. - Even for the CTE2014 case, the limited impact on the optimized surface fluxes is also surprising. Less than 0.25 PgC/yr is small compared to the large posterior uncertainties (about 1.7 PgC/yr). Is your ensemble truly representative of the BC errors? The BC's need to be illustrated here with the differences across the different products before inversion. Is the simplification into four boundaries too limiting in the optimization?

Section 3.5: The flux adjustment compares two different prior errors with significantly different values in the shoulder season. Therefore, the methods disagree mainly when the prior errors are too small to let the inversion adjusts the fluxes. The comparison is presented based on the nature of the correction, but it turns out to illustrate the importance of larger errors when the NEE value is small, and hence the scaled prior variances are small as well. Here, you illustrated the dependence to prior error variances instead of a method-dependence.

Section 3.6: This exercise requires a map of the posterior corrections. Assuming that the number of towers is sufficient to constrain most of North America, the total carbon budget will remain fairly similar. It also confirms that your degree of freedom is small, as described earlier in the paper. In this case, changing the correlation length scale will not lead to any significant changes because the inversion is over-constrained. Only biases will alter the results, such as systematic offsets in the BC. A figure illustrating

the spatial distribution of flux corrections is needed here.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-222, 2017.