Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-123-RC2, 2018

Geoscientific Model Development Discussions

Interactive comment on "A Lagrangian Approach Towards Extracting Signals of Urban CO_2 Emissions from Satellite Observations of Atmospheric Column CO_2 (XCO_2): X-Stochastic Time-Inverted Lagrangian Transport model ("X-STILT v1.1")" *by* Dien Wu et al.

Anonymous Referee #2

Received and published: 24 August 2018

This study is timely as the OCO2 satellite has begun producing data and relevant analyses are being conducted. I think the manuscript can contribute to the OCO2 community and, in general, the GHG community as well. The author did a lot of work including different sensitivity tests, and I think this work deserves publication after addressing issues I raise below.

Main comments:

The paper covers a lot of aspects of comparing modeled column simulations and observations. The main manuscript is long and sometimes deviates from the main story to tell; even boring although this paper is technical by nature and the information can be useful. I recommend that the authors remove some sections and technical details to the Supplement and consolidate the main text for a coherent story. Another issue is that the authors do not link the text with figures well; some of the figure captions are enormously long. In many places, the authors finish the sentences with "see Fig. X" without explaining the content of the figure well enough. I strongly recommend that the authors identify more important results (even move some figures to the Supplement, e.g., Figure 4 or 5, 7) and convey those main results with more care and clarity; please explain the figures! For example, Section 3.2 is useful (I am glad that the authors did this), but not essential for the main story given the length of the manuscript. The authors can spend the space (after moving some details) in explaining figures associated with the main results. Third, I am not quite satisfied with the transport error analysis. The problem is that the errors (mostly winds) for WRF and GDAS are not clearly defined, so it is hard to understand how good or bad the transport is and how the error can be related to signals (e.g., low winds to high signals or the impact of wrong wind directions – not presented clearly). The authors spend a lot of space to explain transport but it needs some improvement. Referring to the unpublished paper too much is not a good idea. Last, I would like the authors to comment on the utility of OCO2 for urban studies based on this work, because there is some skepticism about OCO2's capability for estimating urban emissions with relatively small areas.

Detailed comments:

P1, L19. Global assimilation data seems to be too coarse for the urban scale CO2 simulation. Why use GDAS?

P1, L21. "68 % in posterior scaling factor" should be "68 % in posterior signal" because here the bias in background is in the units of signal. Also, it is not clear what 68 % in posterior scaling factor means. Posterior uncertainty in 1-sigma? Or Does it mean the bias in background resulted in 68% higher or lower bias in the posterior scaling factor?

P1, L22. It seems to me that the authors are referring to signal calculation, and the impact of uncertainty and bias on the urban signals by "Based on these results". I wonder if the authors can add a couple sentences that are more significant than these. If I put it differently, are these results the most important results we take home from this study?

P3, L31. Please add references related to "minimal guidance". The authors can simply add few references on uncertainties associated with atmospheric column simulations.

P3, L33 – 34: The authors underestimate recent developments in inverse modeling. There are several atmospheric inverse studies that consider transport errors and use full error matrix (not just diagonal), in particular non-CO2 studies (e.g., regional methane studies). The references there are old and does not support the statement. The authors need to be specific. I may agree that there are not many studies to incorporate full error characterizations for column-observation inversion studies, but there are now many studies to consider errors more carefully. The authors should be careful in this statement and need corrections.

Also, I am surprised that the authors use a very simple inversion – later in the section I find they are not well formulated but rudimentary – I don't see the benefit of including the inversion result in the study. Please note that there are many sophisticated inversion methods that are much more amenable to error characterizations – please do some literature review.

P4, L8-9: I don't quite understand "Most of these studies aim at extracting relatively large CO2 changes at a fixed level within the PBL or due to large emissions such as of wildfire". Which studies are the authors referring to? The point is tower vs. column or large signal vs. small? Are the authors suggesting that the study site in this work has very little CO2 changes (exchanges?)? The study areas in this study are different from other urban areas in previous studies in terms of CO2 variations or signal-to-noise

ratio? Also, related to this, why did the authors choose this study area instead of some US large cities?

P4, L13: It is not clear why the authors introduce a new background estimation method. I guess this has to do with column simulations, but please state the reason more clearly.

P4, L28: Please define "prior profile" since many "priors" are used in this paper.

P5, L4: It seems "ratios of the pressure difference between adjacent model levels over that between adjacent retrieval levels" needs more clarification. Once PW is interpolated to model levels, then the pressure difference between model levels (as the scaling factor) should be enough? Please clarify.

P5, L12: I wonder what "When WRF fields were available" means. WRF is not used for all days/hours? For the comment on the abstract, I added that GDAS alone is not sufficient for the urban scale.

Also, more importantly, the authors must add the minimum description of the WRF model, e.g., vertical and horizonal resolutions unless stated somewhere later in the sections. It is not appropriate to toss everything to another unpublished reference.

P5, L15: Is GDAS the primary choice?

P5, L19: Remove "a certain height", but directly use an explicit one, e.g., "the maximum release height" - unnecessary vagueness. I see a few places in this paper that use such a vague expression.

P5, L20: Please state what constitutes "different setups" so that the reader has a clear sense of the setups that might differ. As written, it is not clear.

P6, L26: Define "BP".

P7, L3: Please say so, if 0.1 degree is the final resolution for signal calculation, which could be coarse for a urban region.

P7, L10: Please comment on the 1-degree bio flux relative to the size of the study area and its potential impact (due to coarse resolution) on the inversion.

P8, L22: Please add comments on the potential impact of transport over the city when using Method 3. I note that the authors discussed the potential transport error for Method 1 (i.e., endpoint method). Wind direction could be a serious problem for Method 3. Enough overpasses (both up- and down-wind) are available for Method 3.

P9, L20: I agree with the authors that STILT configurations can affect the results. But I don't understand the use of bootstrapping here. The original sample here is from the 401 levels (too many in my opinion). However, what we are interested in is the results from different set-ups, e.g., 20, 40, levels, which can be different from the original samples of the 401 levels. In practice, 401 levels are unrealistic, e.g., for annual analysis.

P11, L15: It is surprising that MAXAGL < 2.5 km did not fully capture CO2 enhancements. I would expect that there is not much surface influence above 2 km. Is it because the study region is associated with really high PBLH? As the authors stated in L30-32, the lower portion of the column should matter most. Then why would MAXAGL of $\sim$ 2.5 km not capture the full enhancement of CO2? Please add sentences that discuss the reason for this. Actually, looking at Figure 8(a), I realize that there are only two cases below 2.5 km. So, 2.5 km itself looks fine. My guess is that even 2 km should be fine. I think the authors give the reader somewhat wrong information here, cosidering the fact that using a higher altitude for MAXAGL increases the computational cost significantly. My understanding from this is: 1) use 100 – 200 m vertical resolutions between 0 – 2 km and 2) above 2 km, use 500 m. If the authors can show even MAXAGL of 2 km is comparable to 2.5 or 3 km, this will reduce the computational cost significantly. I don't understand why the authors use 100 m for up to 3 km given the result shown in Figure 8(a), which in my opinion is too much without good reasoning. I think that some other studies will easily show denser vertical resolutions between 0 – 2 km is good enough.

P11, L34: Please clarify what the fractional uncertainty means here. How did the total particle number become >12500 with 100 particle every 100 m within 3 km?

P12, L32: "incorporates both" to "both incorporates"

P12, L37: I wonder what "we added a wind error component to broaden the urban plume (Sect. 2.4.3 and Sect. 2.6) that helps reduce the inclusion of enhanced values in the background region" means. I can understand this could help reduce strong local sources under the assumption that broadening plumes with additional errors reproduces the reality more accurately. But broadened background does not necessarily solve the bias in the wind direction that is directly related to the enhancement in the background region.

P13, L3-6: How did the authors judge which one is the more accurate background that is assumed to be close to the (unknown) true background? The impact of the background bias (0.56 ppm here) on the emission estimation depends on the magnitude of the observation; it can have only a small impact when the local observations are large.

P13, Section 3.4.1 Comparisons against OCO-2 XCO2 at selected soundings: What is the small conclusion here? After all the analysis, the authors state "we suspect that mismatch in the model-data enhancement widths is primarily due to errors in wind speeds". I expected that the authors state, e.g., "model X is better or worse than model Y in terms of wind' simulations compared to observations, and we also see better or worse in model X or Y for 'signal' comparison between model simulations and observations". Any advantage of WRF due to higher resolutions?

P13, L34: It depends on which wind observations are used. The number of sites for wind obs. in this study is too small to make a statement as shown here.

P17, L30: How large was the random error (S_lambda) relative to the background-subtracted enhancements? The 5 x 5 error matrix (if this is the model-data mismatch error covariacne, i.e., the irreducible error component in the linear model) suggests that only 5 obs were used? If it is true, that seems to be too small, even for a simple linear regression. The scaling factor suggests the prior emissions are consistent with the observation. Is this the conclusion and what the authors expect from the comparison between modeled XCO2 and obs? The description for this simple inversion doesn't sound good at all.

P18, L4-6: I wonder if the background estimation for column CO2 from OCO2 can be improved. Somewhat disappointing.

I hope to see some discussions (a few sentences) on the utility of OCO2 for urban studies including the retrieval error (this urban region has relatively low enhancements, difficult for OCO2 to tell something), not only for this study area, but for future other regions, more generally.

Figure 7. The trajectories seem to be stratified, with each streak (looks like thick streak) somewhat disconnected from each other, which looks strange. Any explanation? Is it because of different levels?

Figure 8-e: Please use the same labels for the legend, e.g., M3.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-123, 2018.