Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2020-864-AC2, 2021

Atmospheric Chemistry and Physics Discussions

Interactive comment on "On the relationship between tropospheric CO and CO_2 during KORUS-AQ and its role in constraining anthropogenic CO_2" *by* Wenfu Tang et al.

Wenfu Tang et al.

afarellano@email.arizona.edu

Received and published: 17 February 2021

We thank Peter Rayner for his very helpful comments. We certainly appreciate your time in reviewing this manuscript. Most importantly, we value your comments in improving this manuscript.

We do recognize that the approach we have taken on this paper (i.e., revisiting Palmer et al. 2006) could open up discussion on the methodology for joint inverse analysis. Our intention is to bring to light some of the issues that need to be considered for future joint inversions, especially on the use of satellite retrievals for inversions, which lacks the vertical information that is critical in identifying transport-related errors that may

confound any inverse analysis. This is not only true in single-species inversions but in fact is an added complication with multi-species inversions that should be carefully considered.

———- Please see our responses to your comments below.

1) this paper examines the role of carbon monoxide (CO) measurements in a joint inversion of CO and CO2 fluxes over East Asia. The paper builds on methodologies and results from some previous papers but delves further into the specific role of CO in improving the inversion.

Response: Yes, this is a paper about the relationship between CO and CO2. We also emphasize that while our inverse analysis is the final key point of this work, the two other important key points that we are highlighting in this work are the following:

a) evaluation of CAM-Chem with CO2 and CO tags using prior fluxes from a posteriori fluxes from state-of-the-art global CO2 inversions (i.e., CAMSv17r1, CT2017, CTE2018). To our knowledge, this is the first development of a joint CO2 and CO modeling in community Earth system model, CESM/CAM-Chem. This is also one of the few studies in recent years that takes advantage of flux/source inversion products as priors in the model. We note that we have also compared our simulations to Carbon Tracker CO2 fields to ensure that our simulations are reasonable.

b) detailed description on CO and CO2 error statistics. Our intention is to highlight the need to evaluate the statistics of the CO and CO2 relationship using aircraft measurements, prior to any inverse analysis. We highlight, in particular, the need to provide spatiotemporal context on these relationships, which in previous inversion studies heavily rely on sensitivity tests. In this work, we highlight the importance of having knowledge of 'spatial error covariance localization length scales' prior to inversion.

———- 2) I found the paper difficult to assess mainly because I have never felt comfortable that I understand the methodology of Palmer et al. (2006). It's quite possible

therefore that what I'm about to say is wrong. I also want to add that I am taking advantage of the discussion period to comment on the paper, i.e. this is not a review.

Response: We recognize the potential difficulty. As mentioned, we intend to open up the discussion on such topic given the relevance of adding CO observations in reducing uncertainties on top-down estimates of ffCO2; especially with the availability of atmospheric composition (reactive gases) retrievals (CO, NO2) nowadays (i.e., TROPOMI, GOSAT-2). CO is also used mention that CO is used for BBCO2, and usually as a separate step (Liu et al., Science).We emphasize the need to better understand the joint inversion, as this has yet to be investigated rigorously and that the species relationships (and associated errors) should not be taken for granted. As mentioned in the manuscript, our approach is based on widely used Bayesian synthesis inversion introduced by Enting (2002) and used by Baker et al. (2006), and more recently applied for multi-species inversion by Boschetti et al. (2018). While Wang et al. (2018) used a similar method (for single-species CO2 inversion), they focused on investigating the impact of reduced carbon component of CO2 based on CO on their estimates.

———- 3) As I understand it the Palmer et al. (2006) methodology defines independent Jacobians for CO and CO2 i.e., there is no physical coupling between the two sources or between each source type and the other observation type. These relationships are introduced by correlations in the observation and prior covariance matrices. It's obvious this works for prior correlations and less obvious (though still true) that it works for observations.

Response: Yes, Palmer et al., (2006), as well as Boschetti et al. (2018), relied on tagging CO2 and CO components using a CTM (linearizing both CO2 and CO) to conduct a suite of Bayesian synthesis inversions. Boschetti et al. (2018), however, utilized STILT to obtain these Jacobians.

Within the CTM (where both CO2 and CO are simulated), there is, however, a physically-based coupling between these runs through the use of emission ratios in

prior emissions for ffCO2 and CO. That is, they are both emitted to the atmosphere in the model by sector-specific assumptions of emission factors per species, and associated emission ratio between species. As we showed in Tang et al. (2019b), for a particular combustion process, we can estimate the combustion products by stoichiometry:

–please see figure 1 for equations and figures–

And so, if we are to look at the equivalent atmosphere abundance of these emission tags (as they are emitted to the atmosphere), they are mostly correlated near the sources, with correlations diminishing downwind of the source. Obviously, observing the relationship downwind does not imply that the correlations (or the lack thereof) are created by this co-emission alone, but typically also through mixing, transport, and associated sinks. These are further quantified in the inverse analysis through the a priori error covariance matrix (i.e., errors in emission ratios) and observation error covariance matrix (i.e., errors in transport, mixing, among others).

We note here that a more accurate representation of both the a priori error and observation error covariance matrices is in fact the crux of the analysis, as with any other inversion methods that account for these uncertainties. We do emphasize the need to investigate the covariances and error covariances (and sensitivities) as these are the terms in the analysis that 'transfer information' to unobserved parameters that we are interested in estimating.

——— 4). It is worth comparing the approach with what happens in nature. Sources are clearly related since some CO2 fluxes are the result of combustion which may also produce CO with an uncertain yield. Writing the problem as FCO = αF, FCO2 = (1$-\alpha$)F where F is the combustion flux and α the CO yield is how an inventory model would write this. If we linearise this relationship we can, I think, rewrite the problem with separate CO and CO2 fluxes. With judicious choices of correlations we can probably ensure that the eigen-vectors of the prior covariance matrix (the space in

which the inversion really takes place) reflect the underlying physical relationships. It's not obvious to me how to do this but the generated slopes between CO and CO2 (as used in the paper) does seem a reasonable way to test it. So far so good.

Response: We agree that it is worth comparing with what happens in nature. In fact, this is one of our key points – i.e., can we evaluate this 'modeled' relationship? Are these relationships that can be derived from observations of the species abundance in nature reasonably captured by CTMs? We agree with Peter regarding the expression on sources. Please see our response to Comment 3). In essence, we are looking at α (effective emission factor in our notation), which we argue should be consistently represented in the a priori inventory for both CO and CO2. At present, with few exceptions, ffCO2 is derived from energy statistics separate from air pollution emission inventories.

We note, however, that as we previously mentioned the yield is quantified as effective emission factor, and the emission ratio is the quantity we can evaluate with observations by looking at the derived enhancement ratio (ΔCO/(ΔãĂŰCOãĂŮ_2)) from observations of CO and CO2 abundance. Note that this can only be 'appropriately' evaluated with observations near the emission sources, just like studies using field campaign measurements to estimate emission ratio. Otherwise, the enhancement ratio is confounded by mixing and transport and hence cannot be taken and interpreted as the abundance equivalent of emission ratio. See for example Mauzerall et al. (1998) Figure 4 (shown below):

–please see figure 2 figures–

where, the regression line of [X] and [Y] (i.e., (Δ[Y])/(Δ[X])) downwind is steeper than over a fresh plume. This complication arises from various sources and sinks, other than combustion which in reality are ubiquitous. And so, here we take the enhancement ratio only as a key diagnostic on the bulk relationship between CO and CO2 (resulting from a combination of emission ratio, mixing and sinks).

Palmer et al. (2006) investigated the sensitivities of different assumptions of activity,

and emission factors and concluded that errors in emission factors largely contribute to the uncertainty estimate. Rather than 'reinventing the wheel', we took a more conservative value of 0.5 (Boschetti et al. 2018 used 0.7, while Palmer et al. 2006 suggested >0.7). As we noted in the manuscript, there is little information from flux measurements to verify this error correlation. Also, note that these correlations, when used in the inversion, is not correlation between CO and CO2 abundance near the source but the correlation of the errors between modeled CO and CO2 abundance.

——— 5) The case for the observational correlations seems harder. What we shorthand as observational errors describe the differences between the modelled value with true inputs and the observed value (Rayner et al., 2019, Section 5). It encapsulates both model and instrumental/retrieval error. It is true that correlations in the observational covariance R do change the posterior uncertainty and mean for fluxes even if the fluxes are not coupled through the Jacobian.

Following notation of Rayner et al. (2019) consider the simplest case of two unknowns and two observations with an identity Jacobian H = I and an identity prior covariance matrix P = I. Assume an observational covariance matrix R = 1 α α1.

Applying the Sherman–Morrison–Woodbury formula (Cressie and Johannesson, 2008) to generate the posterior covariance A we see A_1,1=1-2/(4-αˆ2). The key point here is that the posterior uncertainty, and consequently the posterior estimates for the fluxes are quite dependent on the choices of these correlation parameters. If this is true then the paper needs to spend some more effort on either justifying or testing the sensitivity to these parameters.

Response: We agree with Peter on the proper interpretation of observation error covariances, R. We also agree on the influence of R on the Bayesian synthesis inversion mean estimate and a posteriori error covariance estimate. In fact, both P (or B) and R influence the first and second moment estimates of the conditional pdf. This is the main reason why prior to our inverse analysis (section 5), we expanded our discussion

on error statistics to make sense of R and to come up with reasonable localization of R through the use of a subset of measurements closer to the sources (section 4.1 and 4.2 leading to section 4.3).

To show a similar derivation by Peter for two observations, if we let:

–please see figure 3 for equations–

This is clearly more complex but it shows as well that the a posteriori error covariance is a function of both the a priori error and observation error covariances. While the Jacobians (tags) itself should exhibit some correlation as the species are co-emitted, we need to quantify the uncertainties, not only on error variances but more importantly on the error correlations.

We note as well that our inverse analysis is a product of this investigation on R. By way of expressing the sensitivity (enhancement ratio) with correlation (see related response in Comment 3), we get:

–please see figure 4 for equations–

The same relationship holds for the error sensitivity. In this manuscript, we focus on discussing the spatial variability of these terms, as well as the corresponding magnitudes, so as to guide us in defining the cross-species error correlation components of R (see Eq. 4 of our manuscript). In fact, we used a more conservative error correlation of 0.33 (Boschetti et al. 2018 and Palmer et al. 2006 used 0.7). We opted not to show our tests where we vary our assumptions of these values as this was already done by Palmer et al. (2006). Rather, we choose to emphasize the need to have an estimate of these values based on data, especially aircraft measurements given known issues with vertical transport/mixing in CTMs. We intend to show the role of this relationship on the inverse analysis (section 5) but only to briefly make an example. We emphasize that this is not a manuscript of an inverse analysis of ffCO2 but more of elucidating the need to investigate the relationships prior to any inversions.

As an aside, if we are to recast the first moment of p(x|y):

–please see figure 5 for equations–

——— Specific Comments:

1) L447 Is an r value of 0.51 really moderate here, 25% of variance?

Response: Thank you. We should change it to ‘fair’.

———

2) L451 You should not be quoting p values here, the p value measures the chance that there is a relationship at all which is not interesting in this case.

Response: Thank you. We should not quote p values for this.

——— References (not cited in the manuscript):

Mauzerall, D.L., Logan, J.A., Jacob, D.J., Anderson, B.E., Blake, D.R., Bradshaw, J.D., Heikes, B., Sachse, G.W., Singh, H. and Talbot, B., 1998. Photochemistry in biomass burning plumes and implications for tropospheric ozone over the tropical South Atlantic. Journal of Geophysical Research: Atmospheres, 103(D7), pp.8401-8423.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-864, 2020.

$$C_{x_1}H_{x_2}O_{x_3}N_{x_4}S_{x_5} + n_1(1+e)(O_2 + 3.76N_2) \rightarrow$$
$$n_2CO_2 + n_3H_2O + n_4O_2 + n_5N_2 + n_6CO + n_7NO + n_8NO_2 + n_9SO_2 + n_{10}C + \cdots \quad \text{Eq. (A1)}$$

where emissions of these intermediate products are typically expressed as:

$$E_x = \sum_s [A_s \cdot EF_{x,s} \cdot (1 - CE_{x,s})]$$
$$= \sum_s [A_s \cdot EEF_{x,s}] \quad \text{Eq. (A2)}$$

Here, E_x is the total mass of emissions for species x, $EF_{x,s}$ is its associated emission factor for a specific source/sector s, and A_s is the activity level of the source. $CE_{x,s}$ corresponds to effectiveness of control measure and $EEF_{x,s} = EF_{x,s} \cdot (1 - CE_{x,s})$ is the effective emission factor. When we take the ratio of emissions (Eq. A2) of co-emitted species x and y, we obtain

$$\frac{E_y}{E_x} = \frac{\sum_s [A_s \cdot EEF_{y,s}]}{\sum_s [A_s \cdot EEF_{x,s}]}$$
$$= \sum_s \left(\frac{EEF_{y,s}}{EEF_{x,s}}\right)\left(\frac{E_{x,s}}{E_{x,total}}\right) \quad \text{Eq. (A3)}$$

This ratio can be expressed as the sum of the products of the ratio of effective emission factors ($R^{EEF}_{x,y,s}$) and the fractional contribution of emission sector f for species ($f_{x,s}$).

Fig. 1. Math behind Emission Ratios

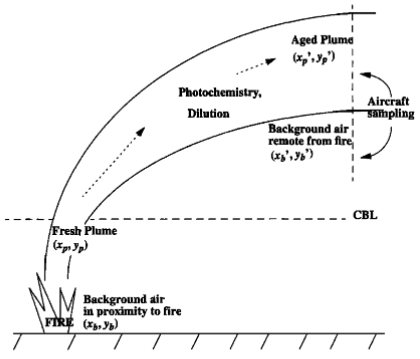

Figure 4a. Schematic of plume formation and transport. The (x_p, y_p) and (x_b, y_b) indicate mixing ratios within the fresh plume at the time of sampling and in background air in proximity to the fire, respectively. The (x_p', y_p'), and (x_b', y_b') indicate mixing ratios within the plume sampled by the aircraft and in background air in the vicinity of the sampled plume, respectively.

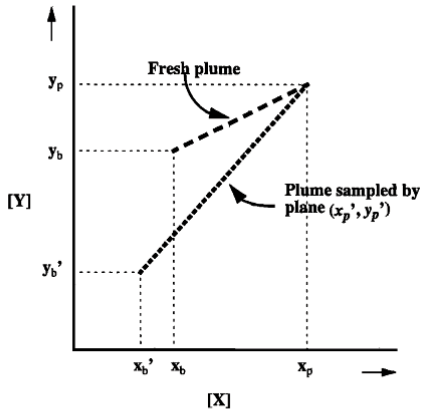

Figure 4b. Schematic of plume sampling. Long-dashed line indicates a fresh plume with enhancements (x_p-x_b, y_p-y_b) relative to background mixing ratios near the fire (x_b,y_b). Dotted line indicates a plume remote from the fire which the aircraft measures relative to local background mixing ratios (x_b', y_b'). Enhancement ratios are obtained from the slopes of these lines.

Fig. 2. Mauzerall et al 1998 Figure 4 on Enhancement Ratio

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \boldsymbol{Hx} + \boldsymbol{e} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}, \quad \boldsymbol{e} \sim N(\boldsymbol{0}, \boldsymbol{R}) \text{ where } \boldsymbol{R} = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}$$

and $\boldsymbol{x} \sim N(\boldsymbol{x^b}, \boldsymbol{B})$, where $\boldsymbol{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

then, the first ($\boldsymbol{x^a}$) and second ($\boldsymbol{A}$) moment of the conditional distribution, $p(\boldsymbol{x}|\boldsymbol{y})$ would be:

$$\boldsymbol{x^a} = (\boldsymbol{H^T R^{-1} H} + \boldsymbol{B^{-1}})^{-1}(\boldsymbol{H^T R^{-1} H y} + \boldsymbol{B^{-1} x^b}), \qquad \boldsymbol{A} = (\boldsymbol{H^T R^{-1} H} + \boldsymbol{B^{-1}})^{-1}$$

respectively. Substituting, we obtain:

$$\boldsymbol{A} = \left(\frac{1}{\alpha^2-4}\right)\begin{pmatrix} \alpha^2-2 & -\alpha \\ \alpha- & \alpha^2-2 \end{pmatrix} = \begin{pmatrix} \frac{\alpha^2-2}{\alpha^2-4} & \frac{\alpha}{4-\alpha^2} \\ \frac{\alpha}{4-\alpha^2} & \frac{\alpha^2-2}{\alpha^2-4} \end{pmatrix}.$$

If we use the other expression for $\boldsymbol{A}$:

$$\boldsymbol{A} = \boldsymbol{B} - \boldsymbol{BH^T}(\boldsymbol{HBH^T} + \boldsymbol{R})^{-1}\boldsymbol{HB} = (\boldsymbol{I} - \boldsymbol{GH})\boldsymbol{B}$$

this produces the same elements of $\boldsymbol{A}$, i.e.,

$$\boldsymbol{A} = \begin{pmatrix} 1 - \frac{2}{4-\alpha^2} & \frac{\alpha}{4-\alpha^2} \\ \frac{\alpha}{4-\alpha^2} & 1 - \frac{2}{4-\alpha^2} \end{pmatrix}.$$

Hence, $A_{1,1} = \frac{\alpha^2-2}{\alpha^2-4} = 1 - \frac{2}{4-\alpha^2}$ as pointed out by Peter.

Now if we let $\boldsymbol{B} = \begin{bmatrix} 1 & \beta \\ \beta & 1 \end{bmatrix}$ we get:

$$\boldsymbol{A} = \left(\frac{1}{\alpha^2+\beta^2+\alpha\beta-4}\right)\begin{pmatrix} \alpha^2+\beta^2-2 & -\alpha+\alpha\beta^2+\alpha^2\beta-b \\ -\alpha+\alpha\beta^2+\alpha^2\beta-b & \alpha^2+\beta^2-2 \end{pmatrix}$$

Fig. 3. Math behind Error Correlations

C11

$$\frac{\Delta CO}{\Delta CO_2} = \frac{cov(CO, CO_2)}{var(CO_2)} = \rho_{CO,CO_2}\frac{\sigma_{CO}}{\sigma_{CO_2}}$$

Fig. 4. Relationship between Enhancement Ratio and Correlation

$$\boldsymbol{x^a = (H^T R^{-1} H + B^{-1})^{-1}(H^T R^{-1} H y + B^{-1} x^b)}$$

or

$$\boldsymbol{x^a = x^b + BH^T(HBH^T + R)^{-1}(y - Hx^b)}$$

to:

$$\boldsymbol{x^a - x^b = BH^T(HBH^T)^{-1}(HBH^T)(HBH^T + R)^{-1}(y - Hx^b)}$$

we can easily recognize that the increment $\boldsymbol{x^a - x^b}$ is a function of three terms:

1) Innovation: $\boldsymbol{(y - Hx^b)}$, which is the mismatch of observed and modeled abundance
2) Least squares fit $\boldsymbol{(HBH^T)(HBH^T + R)^{-1}}$ which is the ratio of the error variances
3) Linear regression: $\boldsymbol{BH^T(HBH^T)^{-1} = cov(He_x, e_y)/var(e_y)}$ where $\boldsymbol{B = cov(e_y, e_y{}^T)}$ which is the slope of a line.

In this manuscript, we highlight the need to investigate 3).

Fig. 5. Posterior Mean Estimate Recasted