Reply to reviewer #1

Scot M. Miller, et. al.

We would like to thank the reviewer for suggestions and comments on the manuscript. The reviewer's detailed suggestions have been very helpful in improving the manuscript. Below, we have included the reviewers comments (in bold) along with our reply and the associated changes/updates to the manuscript.

• My main criticism if focused on case study #1 and the framing of this experiment as a way to quantify the bias detection limit of carbotracker (CT). But CT does not work on a month by month basis lie your lambda scaling factoor, it does not consider signals site-by-site like here but instead a whole network, and it does not scale flux signals locally at each site like in your FSSR but over a large spatial area that is also seen by other sites. [...] Framing this experiment as a way to determine the balance between large-scale flux influences and transport errors is in that sense more appropriate, and I think describes better what was actually done.

The reviewer brings up a great suggestion here, and we have re-framed case study #1 accordingly. As the reviewer points out, the goal of this investigation is not to re-estimate the uncertainty bounds on CarbonTracker. Rather, our goal is to understand the magnitude of these transport uncertainties relative to the fluxes. To that end, this case study provides ones means to relate these two entities (the transport uncertainties and fluxes) in the absence of an explicit model adjoint. We no longer frame the case study as a means to quantify the bias detection limit of CarbonTracker. Rather, as the reviewer suggests (below), we have re-framed the case study to examine the following question: how does the magnitude of the transport uncertainties compare against the afternoon, atmospheric CO_2 signal from regional surface fluxes? We have modified the manuscript text and figures accordingly.

• Specifically, your comparison of transport noise (SSR) and flux biases (FSSR) is done in squared residual space which only measures the magnitude of a signal, but does not account for its sign. A bias in fluxes would typically manifest itself as a consistent over- or underestimate of the true concentrations observed and even if these are small (say 0.5 ppm) compared to the more random transport uncertainties (say 3 ppm), their consistency in sign over longer periods of time would make them detectable. In fact, in a Bayesian inversion the system would try to overcome this small bias as by design it strives for zero mean residuals even in the presence of large observation error covariances.

The reviewer raises a very good point here, and we have clarified this point in the revised manuscript. As the reviewer explains, a Bayesian inversion will optimize the fluxes to minimize or remove any biases between the model and the observations. If a transport model is completely unbiased relative to the actual atmosphere, then the CO_2 budget

estimated by an inversion should also be unbiased. (This statement assumes that other components of the inversion, including the observations, are unbiased.)

In contrast, the inversion may estimate an erroneous or biased budget if the atmospheric transport model is biased. For example, imagine a hypothetical transport model that consistently overestimates vertical mixing. One could construct a Bayesian inversion to optimize CO_2 fluxes using that transport estimate. The inversion will optimize the fluxes to minimize any model-measurement bias. However, the resulting flux estimate is unlikely to be correct; the inversion would erroneously increase the magnitude of the fluxes to compensate for errors in vertical mixing. The model would appear to match the CO_2 measurements, but the estimated fluxes would nonetheless be biased relative to the true fluxes. In this case, the bias in the fluxes would be undetectable with respect to the atmospheric observations. Stephens et al. (2007) adeptly discuss this topic in the context of atmospheric inverse modeling.

We have re-designed case study one in the manuscript to make this comparison more direct. Among other changes, the revised case study no longer uses squared residuals. We hope the revision makes this point about biases more transparent.

• To overcome this criticism, I would suggest one of two approaches:

(1) is to try and change the metric so that it includes more sites at once and includes also spatial covariances between residuals. The new metric then also needs to account in some way for the sign of the residuals.

(2) Is to write the question of this case study differently and to say that you'll try to estimate to what variation in flux magnitude the meteorological uncertainty corresponds for each site given a realistic surface flux from CT. This also means that most of the use of the word "bias" gets replaced by "flux signal".

The reviewer makes two good suggestions for revising the manuscript. We have re-framed the case study according to the reviewer's second suggestion. In addition, we have also included multi-site comparisons in the revised manuscript, as per the reviewer's first suggestion. To this end, we have revised sections 2.4, 3.3, and Fig. 4.

• I find the discussion section a bit too short, and would like to see some more connections to other studies in this field. For example, some reflection could be added on the LETKF methods used by these authors in the past, and about the possible gain of co-simulating CO2 and transport errors. Also, there is room for some reflection on the covariations of CO2 surface fluxes, and those that shape the weather conditions (water and energy and momentum fluxes). What would the next step with this type of system look like when surface fluxes also become a function of the weather variables?

We have lengthened the discussion section to include these points, as suggested by the reviewer. For example, in section 3.4 of the revised manuscript, we discuss the possible gain of co-simulating CO_2 surface fluxes and transport errors. That approach could provide a more complete picture of how meteorological uncertainties affect CO_2 fluxes from the origin of the fluxes to the locations where we actually measure atmospheric CO_2 . For example, Lin et al. (2011) explored how uncertainties in flux model drivers affected fluxes estimated for Canadian boreal forests. They found that uncertainties in downward shortwave radiation contributed to the largest uncertainties in the simulated fluxes. Similarly, Law et al. (2002) and Gourdji et al. (2012), among many others, have

shown that both air temperature and specific humidity are drivers of CO_2 fluxes. These meteorological variables (e.g., downward shortwave radiation, temperature, and specific humidity) correlate with the persistent atmospheric transport uncertainties discussed in section 3.4. A future study could connect these uncertainties (in transport and flux estimation) to gain an even broader picture of how meteorological uncertainties affect CO_2 flux modeling and ultimately top-down CO_2 flux estimates.

• Furthermore, these findings can nicely be connected to the error budgets presented in Pino et al., (2011) and in Williams et al (2011). Both take a look at the driving forces behind variations in CO2 in the PBL, one from a local and one from a larger perspective.

We have added references to both papers in the revised manuscript. Pino et al. (2012) argue that estimated morning PBL heights play a critical role to modeled CO_2 concentrations during midday. They examined transport errors at diurnal scales but point out that the role of different boundary layer processes could change when examined over longer time scales. Our analysis examines transport errors at both the diurnal and monthly time scales and can extend the arguments presented by Pino et al. (2012) to these longer time scales.

Williams et al. (2011) argue that previous meteorological model discrepancies are usually due to overestimated vertical mixing. According to the authors, "However, the simple inverse proportionality between errors in vertical gradients and mixing only works when there are no systematic errors in the surface flux, horizontal advective transport, or nonlinear vertical advective transport (i.e., synoptic-scale eddies)." In our analysis, we place these individual error sources, like those invested by Williams et al. (2011), in the context of other transport processes or uncertainties at sub-daily to monthly time scales.

• Note that I remain a bit puzzled on the implementation of the SSR vs FSSR metric in equations 4 and 5 + the explanation in the supplement and would like to see some clarification.

We have simplified the approach in this section of the manuscript to make the methods easier to follow and more transparent.

• p.23684: I could not find where the range is actually applied instead of the SDV

We have removed the phrase "or alternately the range" from the manuscript. In the revised manuscript, we primarily refer to the 95% confidence interval throughout the manuscript.

• p.23684: What is the temporal resolution of these fluxes?

We reformatted CarbonTracker fluxes to a 6-hourly resolution. This resolution is identical to the CAM model time step. We use this 6-hourly resolution for all model simulations presented in the manuscript. Figure 2, by contrast displays monthly-averaged CT fluxes. The primary objective of this figure is to illustrate the spatial and seasonal distribution of the fluxes. We do not use these monthly-averages in the actual model runs or analysis. We have clarified this point throughout the manuscript.

• p.23684: So this means that the feedbacks of meteorological errors on carbon exchange are not accounted for? In other words, different weather does lead to different water exchange, but not other carbon fluxes. Okay, I got it.

The reviewer is correct here. We have added a sentence to section 2.3 clarifying this point.

• p.23685: Larger than most means more than 32 if k=64 members?

We agree with the reviewer that this text is ambiguous as written. We have re-framed this section of the methods accordingly.

• So this suggests that for p to get to 0.05, there must be 64*0.05 = 3.2 elements in A (eq 5). And when there are four or more SSRs in the set that are larger than FSSR then you have proven the null-hypothesis that bias in fluxes is indistinguishable above transport uncertainties. This seems quite strict to me.

Oh wait, I think there might simply be a typo here and you actually meant 0.5 instead of 0.05? Sorry, I spotted this kind of late because 0.05 is such a typical p-value in statistics...

We have simplified the approach in this section of the methods and no longer use a hypothesis test or associated p-values. In the revised manuscript, we estimate confidence intervals in modeled atmospheric CO_2 and compare those uncertainties against the surface flux signal. We no longer test an explicit hypothesis.

• Can you elaborate in the main text how this temporal covariance is accounted for. I am sure the Supplement gives info but I'd rather like to understand it here.

The reviewer makes a great suggestion here, and we have elaborated on this point in section 2.2 of the manuscript.

Both spatial and temporal covariance are built into the transport errors estimated by CAM-LETKF. The CAM-LETKF system includes 64 different ensemble members. At the first time step, we launch 64 weather forecasts simultaneously, one for each ensemble member. At the end of the first 6-hour time step, we optimize these ensemble members collectively to match meteorological observations, and the spread of these ensemble members represents our posterior uncertainty in the meteorology. We then use these optimized ensemble members as initial conditions for the next time step and re-launch 64 simultaneous weather forecasts.

Transport uncertainties within one ensemble member can easily persist over many time steps. For example, if the PBL height in one ensemble member is lower than the ensemble average at one time step, it will probably be lower than average at the next time step. In this way, transport uncertainties or errors can persist over many time steps.

• p.23687: This suggests you indeed used fluxes including a diurnal cycle.

That statement is correct. We have updated the methods section to make this point clearer to the reader.

• p.23688: I think this is an absolutely wonderful conclusion to draw, and hope it will get a prominent place in the abstract and conclusions

Thank you for the encouraging suggestion! We have modified case study #1 to focus more specifically on these conclusions. Furthermore, we have made these points more prominent in both the abstract and conclusion.

• p.23688: I do not think this case study uses an appropriate question, as your test is not a correct metric to determine the minimum size of flux biases that are detectable through atmospheric CO2.

We agree with the reviewer here. We have re-framed case study #1 based upon the reviewer's suggestions above.

• p.23688: This effect of measurement bias was explored by Masarie et al., (2011), please reference.

This is a good suggestion, and we have included this reference in the revised manuscript accordingly.

• p.23689: What does the number 0.3 represent?

A correlation coefficient of $R^2 = 0.3$ does not represent any specific threshold. Rather, we simply wanted to show the meteorological variables that correlate best with the transport uncertainties (instead of including 60 different scatterplots). We have modified this section of the revised manuscript. Instead, we now show the two variables that correlate most closely over land regions and over the ocean (four total variables).

- p.23689: Since this point is now mentioned a second time, a reference to Pino et al., (2012) is in place as he already showed such PBL-CO2 error relations. We have included this reference in the revised manuscript.
- p.23689: Again, your analysis is very nice but this conclusions is not correct. Since one of the authors is associated with the CT group at NOAA, perhaps a synthetic inversion could be done to prove this statement beyond my doubt?

We agree and have re-framed case study #1 accordingly.

• p.23689: This second part is very nice. Can you speculate how this conclusion might change if the interactions between the meteorological variables and the CO2 fluxes themselves were included in a follow-up study?

The reviewer poses an interesting question: what would be the effect of including these meteorological uncertainties in the bottom-up or biogeochemical model that generates the CO_2 fluxes? The uncertainties in estimate CO_2 fluxes would likely increase. We have added a discussion on this point to section 3.4. Refer to the discussion earlier in this reply for more detail on this point.

• p.23690: I find the discussion section a bit too short, and would like to see some more connections to other studies in this field.

We have expanded the discussion accordingly (see the discussion earlier in this reply for more detail).

• p.23693: You could compare these to the posterior flux uncertainty in CT and show that they are at least as large indeed.

As per the reviewer's suggestion, we have re-framed case study #1 to de-emphasize any direct comparison against the posterior uncertainties in CarbonTracker. As such, we would hesitate to make that comparison explicit here.

Furthermore, it might be difficult to make a direct comparison in this instance. CAM– LETKF estimates the variances and covariances due to transport errors. This information is often incorporated into one of the covariance matrices in a Bayesian inversion. This matrix is often termed the 'model-data mismatch matrix' or 'observational error covariance matrix'. This covariance matrix is then combined with the prior covariance matrix to compute the posterior uncertainty. Hence, this suggestion would require comparing somewhat different quantities. In other words, if we compared transport uncertainties against the posterior flux uncertainty we would be comparing two very different covariances matrices to one another.

• p.23694: What do the letters below the x-axis indicate?

We have removed these letters from any analogous plots in the revised manuscript. The letters below the x-axis that figure indicated whether the CO_2 measurement sites were marine ("M"), short towers ("S"), or tall towers ("T").

• p.23694: Why do we only see the land CV? Was the constant flux also only applied over land? This was not clear to me from the description yet.

In the revised manuscript, we discuss these results over both land and ocean regions (in section 3.4 and Figs. 6-7).

• p.23695: The variables 1,2 and 4 look very similar as one would expect from meteorological principles. In the same way, 5 and 6 are closely related. What is perhaps more interesting is that (1) the PBL height which in the end is most directly related to the CO2 mixing ratios is not shaped the same as these primary drivers. This stresses the need for a meteorological model to calculate the (co)variances of transport errors rather than to just use some simple proxy. And (b) is that the CV of temperature and CO2 are very similar which is because they are shaped by the same large scale synoptic systems. This is also discussed in the Williams et al., (2011) paper, and the driving power behind the LETK methods shown by Kang, Kalnay, Liu, and Fung (co-authors here). Perhaps this is worth to mention in the discussion.

The reviewer makes a great point here. We have also added an analysis over ocean regions, and the errors here correlate most closely with zonal winds. This added analysis further supports the reviewer's comment above on the role of synoptic scale systems. Also, these variables cannot explain all of the uncertainties, and this result stresses the need for a meteorological model to calculate transport errors over the use of a single proxy for transport errors (like PBLH). We have added a discussion of these points to section 3.4 of the revised manuscript.

• p.19: This 5% I guess corresponds the p=0.05 probability stated in the main text. That suggests this was not just a typo, and I remain confused on equations 5 and the use of this test.

We have removed the hypothesis test from case study #1 to make the analysis simpler and more straightforward. Concomitantly, we have removed most equations to streamline and simplify the revised text.

• p.19: This is a nice illustration of the properties of the SSR, which I think correctly assumes transport errors to be normally distributed around a zero mean. But the problem I have is in the comparison to FSSR, which for a biased flux would not just be a residual around some mean, but an actual signal with a sign and a spatial pattern. See for instance the figures S9, S11, and S14 that both represent winter conditions. A shift of the fluxes by 10% upwards would lift both lines for the ensemble mean upwards by 0.5-2.0 ppm and reveal a systematic offset (if the model mean was a bit more unbiased which it is not without data assimilation of the fluxes) at three locations.

If the atmospheric transport errors were completely uncorrelated from one model time step to another, then it might be relatively easy to distinguish a bias in modeled concentrations caused by an erroneous flux estimate. However, the atmospheric transport errors estimated in this study are often correlated in both space and time. In other words, these errors are modeled as a multivariate normal distribution, and the covariances in this distribution can be large. As a result, transport errors could bias the model relative to the measurements over many time steps. In that case, it could be very difficult to distinguish the difference between sustained model-data differences due to the fluxes or due to transport errors. We have revised and re-framed case study #1 to better explain and more prominently feature the role of spatially and temporally correlated transport errors.

References

- Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO₂ exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, doi:10.5194/bg-9-457-2012, 2012.
- Law, B., Falge, E., Gu, L., Baldocchi, D., Bakwin, P., Berbigier, P., Davis, K., Dolman, A., Falk, M., Fuentes, J., Goldstein, A., Granier, A., Grelle, A., Hollinger, D., Janssens, I., Jarvis, P., Jensen, N., Katul, G., Mahli, Y., Matteucci, G., Meyers, T., Monson, R., Munger, W., Oechel, W., Olson, R., Pilegaard, K., U, K. P., Thorgeirsson, H., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation, Agricultural and Forest Meteorology, 113, 97 120, doi:10.1016/S0168-1923(02)00104-1, 2002.
- Lin, J. C., Pejam, M. R., Chan, E., Wofsy, S. C., Gottlieb, E. W., Margolis, H. A., and McCaughey, J. H.: Attributing uncertainties in simulated biospheric carbon fluxes to different error sources, Global Biogeochemical Cycles, 25, doi:10.1029/2010GB003884, 2011.
- Pino, D., Vilà-Guerau de Arellano, J., Peters, W., Schröter, J., van Heerwaarden, C. C., and Krol, M. C.: A conceptual framework to quantify the influence of convective boundary layer development on carbon dioxide mixing ratios, Atmospheric Chemistry and Physics, 12, 2969– 2985, doi:10.5194/acp-12-2969-2012, 2012.
- Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.: Weak Northern and Strong Tropical Land Carbon Uptake from Vertical Profiles of Atmospheric CO₂, Science, 316, 1732–1735, doi:10.1126/science.1137004, 2007.
- Williams, I. N., Riley, W. J., Torn, M. S., Berry, J. A., and Biraud, S. C.: Using boundary layer equilibrium to reduce uncertainties in transport models and CO₂ flux inversions, Atmospheric Chemistry and Physics, 11, 9631–9641, doi:10.5194/acp-11-9631-2011, 2011.

Reply to reviewer #2

Scot M. Miller, et. al.

We would like to thank the reviewer for suggestions and comments on the manuscript. The reviewer's detailed suggestions have been very helpful in improving the manuscript. Below, we have included the reviewers comments (in bold) along with our reply and the associated changes/updates to the manuscript.

• General comment: Not all sources of transport model uncertainty are captured by an ensemble of forecasts with a single transport model. Somewhere in the paper it should clearly be listed which uncertainties are not included (e.g. spatial representation error/model resolution, uncertainty arising from imperfect parameterizations of turbulent processes and cloud transport, other structural model errors such as numerical diffusion).

The reviewer makes a good point here. We have added text accordingly in sections 2.2 to clarify this point.

• P23684 L12: "correlated errors can bias" I suggest to replace this by "spatially correlated errors can bias"

We have changed this statement to "temporally and/or spatially correlated errors can bias".

• P23694 L4-6: The case that the ensemble does not encapsulate the CO2 measurements might also be related to differences in the transport models used here and for CT (TM5). This should be mentioned.

We have added comments to this effect in sections 2.2 and 3.1 of the revised manuscript.

• P23695 L17-18: "most existing top-down studies will underestimate the uncertainties in estimated CO2 fluxes" here references should be given as this is quite a strong statement. Some inverse modelling systems e.g. use error inflation to allow for covariance on timescales shorter than a week (e.g. Rödenbeck et al., 2003).

We have re-written this statement in the revised manuscript. In that statement, we wanted to communicate the importance of accounting for spatial and/or temporal correlations in the transport errors. For example, an inversion that includes these covariances would estimate larger uncertainties in the fluxes relative to one that uses a diagonal covariance matrix. We have revised that statement to clarify our intended meaning. Furthermore, we have expanded case study #1 in the revised manuscript to better indicate how these spatial and/or temporal error covariances can affect the estimated fluxes.

• P23695 L25: I suggest dropping the comma after "top-down"

We have updated the manuscript accordingly.

• P23696 L5-8: It might not a property of the tall towers to be more or less sensitive, but a property of the transport model. It should be mentioned that there is not really a difference expected, given the vertical resolution of the transport model. In that context, it would be appropriate to mention the number of vertical levels in the lowest km as this information seems hard to find for the reader.

We have clarified this point in the manuscript. In most regions there are 3 vertical levels within the lowest kilometer of CAM. (These three levels are centered at 929.6, 970.6, and 992.6 hPa over regions where the land/water surface is at sea level.). Hence, some CO_2 observation sites are associated with the lowest vertical level of the model while others are associated with the next vertical level. We have removed any comments from the manuscript on differences between short versus tall towers.

• P23696 L8-9: I have difficulties averaging the bar plots for marine sites to 76%. There are three bars that are of scale, and the others average to something around 35% in February and 45% in July.

We have updated the analysis with more observation sites and have modified the associated figure.

• P23698 L3-5: Figs. S16 and S17 do not really provide any information regarding the uncertainty represented in the meteorological ensemble, as they only show monthly mean values for each of the variables. A parameter that might be interesting in this regard is the coefficient of variation for the boundary layer height (PBLH), as a small uncertainty in PBLH will lead to a large uncertainty in tracer in regions with low average PBLH.

We have added several additional plots to the supplement that visualize additional meteorological variables and their uncertainties (including the coefficient of variation for the boundary layer height).

• Supplement S1, P1, first line of 3rd paragraph: suggest replacing "for each for the" by "for each of the"

We have changed this text in the supplement.

• Supplement S2, P4, 4th paragraph: I don't quite understand why there is a need for manually setting inflation factors to 0.4 (the lowest values globally); in the text "unphysical temperature estimates near the tropopause" are mentioned. Are there no satellite data in this region available that are assimilated? Kalnay et al., (1996) mentions that TOVS sounder data are assimilated; also there should be a few radiosonde data in that region.

This issue is due to an enigmatic temperature instability in the meteorological model. In the forecast stage of the CAM model, the ensemble's temperature spread in this region can increase rapidly if the initial conditions (i.e., the posterior estimate from the previous time step) have a sufficiently large spread.

Normally, one might expect the adaptive inflation to correct for this issue; the adaptive inflation adjusts the variance of the meteorology model ensemble to match the actual model-data residuals. In theory, this procedure should prevent the ensemble spread from exploding (given sufficient data). However, the inflation factor by design cannot change suddenly from one time step to another. The adaptive inflation procedure uses the previous time step as the prior inflation estimate, and that prior estimate has a finite uncertainty (in this case, a prior standard deviation of 0.03 – similar to the values used by Miyoshi (2011)). Because of this prior uncertainty, the adaptive inflation factor must evolve slowly over many days (if it changes at all). In most cases, this property is desirable because it prevents a single (or small number of) observation(s) from making dramatic changes to the evolution of the model-data system. However, in the case of this temperature instability, the instability in the model develops over 4-5 model time steps, much faster than the response time of the adaptive inflation factor.

The adaptive inflation procedure requires an initial inflation estimate for the first time step of the model run (i.e., an initial condition). The adaptive procedure then updates this estimate at the each model time step (e.g., Eq. S10). For this initial estimate or initial condition, we set a small value (0.4) for the equatorial western Pacific. During the one-month model spin-up period, the estimated inflation value evolves substantially from the initial estimate in most regions of the globe (e.g., Fig. S2). Over this region of the Pacific, however, the estimated inflation factor does not evolve or change very much; either this initial estimate is consistent with the actual model-data residuals or the meteorological data (and the adaptive inflation procedure) are not very informative over the region. In either case, this small initial condition prevents the ensemble spread from becoming unstable over the region.

We have added more explanation on these points within section S2 of the article supplement.

• Supplement S3, P8, figures S7 and S8: The colour scale labelling seems to be wrong; I would expect a significantly smaller range for monthly averaged concentrations than for 6 hourly concentrations

Thank you for point out this mistake! The legend on these figures should be identical to Fig. 2c and 2d. We have corrected these figures in the revised manuscript.

• Supplement S5, table S1: it should be mentioned (in the legend or in the text on page 16) that the locations for each of the sites can be seen in Figure 4, panel a).

This is a good suggestion. We have added many more observation sites to the analysis. As a result, we removed Fig. 4a and instead now list all of the site locations in table S1.

• Supplement S6, Figure S15: SSR should have units (ppm²), those should be added

We have revised case study #1 and no longer use sum of squared residuals (SSR). We have updated the supplemental figures accordingly.

References

Miyoshi, T.: The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter, Monthly Weather Review, 139, 1519–1535, doi:10.1175/2010MWR3570.1, 2011.

Reply to reviewer #3

Scot M. Miller, et. al.

We would like to thank the reviewer for suggestions and comments on the manuscript. The reviewer's detailed suggestions have been very helpful in improving the manuscript. Below, we have included the reviewers comments (in bold) along with our reply and the associated changes/updates to the manuscript.

• These uncertainties are calculated based on a coarse resolution meteorological model, which has a spatial resolution of 2.50 longitude x 1.90 latitude. In the reality, there are other additional error terms introduced due to fine-scale variations that cannot be captured by the coarse model. These additional terms will be more significant depending on the regions and/or periods you sample.

Uncertainties in the posterior meteorology estimate include uncertainties in the model, uncertainties due to measurement errors, and uncertainties due to meteorological patterns that are smaller in scale than the model resolution. The latter two uncertainties are incorporated into the posterior estimate via the **R** covariance matrix (e.g., Hunt et al. 2007). This matrix, often referred to as the nugget covariance matrix, is used as an input into the meteorology model-data assimilation and the posterior uncertainty calculation. We estimate the elements of the **R** matrix directly from the meteorological data using an adaptive approach outlined by Li et. al. (2009). This adaptive approach estimates the collective variance due to measurement error and uncertainties due to meteorological processes that occur at scales smaller than the model resolution. Hence, errors due to small-scale processes are a component of the posterior meteorology and CO_2 estimates. However, we cannot resolve the spatial distribution of these fine-scale errors at sub-grid scale.

In addition, one goal of this study is to run simulations that are analogous to commonlyused, top-down global CO_2 flux estimates like CarbonTracker. The grid used in this study is comparable, if not smaller, than many existing global CO_2 inversion studies. For example, CarbonTracker has a 2° latitude by 3° longitude global resolution (Peters et al., 2007, http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/). Other global inversion studies, like Mueller et al. (2008) and Gourdji et al. (2008) used a resolution of 3.75° by 5°, and Basu et al. (2013) used a 4° by 6° resolution. One could argue that there are advantages to estimating global CO_2 fluxes using a model with finer spatial resolution. With that said, the resolution used here is analogous to that used by common top-down CO_2 flux products like CarbonTracker and would be able to speak more directly to the types of transport errors that would be encountered in those efforts.

• The mentioned model ensemble method cannot account for these fine-scale spatial variations, given that the weights (to match the meteorological observations) are estimated for each grid box using observations within a radius about 1500 km.

We do not compare the model estimate in one grid box against wind or temperature observations taken 1500km away. As the reviewer points out, that approach would be ill-advised. We have added text to the supplement (section S1) to clarify and further explain this point.

In the LETKF, we estimate a set of weighting factors for the 64 ensemble members such that the weighted ensemble best matches the meteorological observations. To achieve this, we first interpolate the gridded model output to the observation locations and times. We then estimate a unique weighting factor for each individual grid box. If we estimated the weights using only model-measurement pairs in the grid box of interest, several problems could arise. First, there may not be many relevant observations that are sensitive to that specific grid box, particularly over the open ocean or near the poles. In those circumstances, the estimated weights could be inaccurate. Second, that approach could produce vastly different weights in adjacent grid boxes, a result that is unlikely to be physically realistic. For example, the estimated weights for one model grid box over eastern North Dakota should look somewhat similar to the weights for a grid box over western North Dakota. If the two sets of weights were completely unrelated, one could argue that the optimization would be an over-fit.

Instead, we use model-measurement pairs within a certain geographic radius to compute each set of weights. This approach ensures coherence among adjacent grid boxes and ensures that the optimization is not an over-fit to the data. We further taper the influence of model-observation pairs on the optimization depending on their distance from the grid box in question (using a Blackman window function as described by Oppenheim and Schafer (1989) and Liu et al. (2012)). Hence, model-measurement pairs located within the model grid box of interest will influence the optimization much more strongly than model-observation pairs located 1000km away. A radius of 1500km for the Blackman window function is comparable to values used throughout the meteorological literature. For example, Liu et al. (2011) and Liu et al. (2012) also used a 1500km radius. Furthermore, Miyoshi (2011) set a 1825 km radius of influence, Miyoshi and Kunii (2012) used a 1460km radius, and Szunyogh et al. (2008) used an 800km radius.

• I am not sure how nugget variance (R) is constructed and whether it necessarily represents all errors due to these fine-scale variations.

Many existing meteorology studies that implement an ensemble Kalman filter have used the published measurement error for \mathbf{R} (e.g., Szunyogh et al., 2008; Liu et al., 2012). In reality, \mathbf{R} also includes a number of other errors, including errors due to meteorological features that are smaller than the model resolution (as discussed above). To capture this entire spectrum of errors, we estimate these errors directly from the meteorological data, an advance over previous efforts that used only the published measurement error. These calculations for \mathbf{R} , by definition, will capture any variability in the measurements that cannot be incorporated into the model ensemble. This variability includes both measurement errors and errors due to fine-scale meteorological processes. This approach is detailed in Eq. S11 and in Li et al. (2009).

• Moreover, I am not much convinced how a single inflation factor for each model grid box works fine for all model parameters.

The use of a single inflation factor per grid box has been a common practice in ensemble Kalman filters applied to weather models (e.g., Szunyogh et al., 2008; Liu et al., 2011, 2012; Miyoshi and Kunii, 2012; Kang et al., 2012). In our study, we use a relatively new technique known as adaptive inflation to estimate the inflation factors. This approach

estimates inflation factors based upon actual model-data residuals (Miyoshi, 2011). The traditional approach has been to choose inflation factors subjectively based upon 'expert knowledge.' In fact, previous studies used zonally-constant inflation factors (e.g., Szunyogh et al., 2008; Saito et al., 2011; Liu et al., 2011, 2012; Yang et al., 2012). Miyoshi (2011), in contrast, argues that this zonally-constant approach is not ideal because it cannot differentiate between ocean and terrestrial regions. The statistical approach implemented here is therefore an advancement over previous efforts because we estimate spatially- and temporally-variable inflation factors directly from the data.

In practice, adaptive inflation can be very challenging to implement; the inflation factors that best match the model-data residuals can, in some cases, cause instabilities in meteorological model that result in incompatible combinations of meteorological parameters. These instabilities often crash one or more of the ensemble members. Furthermore, the approach performs poorly when observations are sparse (e.g., Miyoshi, 2011). When we estimate a single inflation factor per box, we can leverage more observations to make a more stable inflation estimate. Hence, we felt that this framework would require more development before we could reliably estimate unique, grid-scale inflation factors for many different meteorological parameters.

The meteorological data-assimilation community is moving toward adaptive inflation techniques that can accomplish this task (e.g., Zheng et al., 2013). However, this kind of in-depth methodological development is beyond the scope of our study.

• Hence I fear that the values reported for CO2 transport uncertainty (globally) can be far away from reality. This could be one of the reasons why Fig. 2 does not generally show high transport related uncertainties in the coastal sides (sea/land breeze effects?).

We do see larger uncertainties in zonal winds along many coastal regions, presumably related to sea breezes. We have added a new plot to the supplement that illustrates these features (Fig. S17). These uncertainties are particularly prominent across the west coast of North America where sea breezes are an important component of coastal weather. In our simulations, uncertainties in zonal winds at the coastline do not always translate into large uncertainties in modeled CO_2 concentrations. For example, uncertainties in both zonal and meridional winds are high along the coast of British Columbia and Alaska in February (Fig. S17). Since those regions have small CO_2 fluxes in winter, large uncertainties in the winds do not translate into large uncertainties in 6-hourly modeled atmospheric CO_2 (Fig. 2a).

• The authors may wish to provide more detailed discussion regarding this aspect and it is worthwhile to mention explicitly the significant limitations of this approach.

We have added text to the methods section 2.2 that describes both the advantages and limitations of the meteorology model-data assimilation (e.g., the model cannot resolve the spatial patterns of meteorological features at sub-grid scale).

• In the given design and set up, I would certainly consider that the flux bias estimations in the case study 1 are overestimated values, because of unrealistically "too strict" constraints.

We have reformulated case study #1 in a way that no longer uses a hypothesis test, and we no longer make definitive statements on whether the observations would be able to 'see' biases in a CO_2 flux estimate. Instead, we visually display the 95% confidence intervals in modeled atmospheric CO_2 and compare those uncertainties against the afternoon boundary layer enhancement in CO_2 at various observation sites.

• The current inversion approaches followed by many modeling groups take into account the transport uncertainties to some extent and the method is not as simplified as the approach given here.

We have clarified this point in the revised manuscript. Most current inversion approaches do account for transport uncertainties. However, the majority of existing inversion studies assume that the transport uncertainties are uncorrelated in space and time. In other words, existing studies typically use a diagonal covariance matrix to describe errors due to atmospheric transport, measurements, and model resolution, etc. A central question in our paper is to understand how transport errors are correlated in both space and in time, and we find that these correlations or covariances are substantial. An inversion study that ignores these covariances could either underestimate uncertainties in the CO_2 fluxes or propagate transport errors into the estimated fluxes. We have revised the setup for case study #1 to make this point clearer within the manuscript.

• I am a bit surprised to see totally different patterns between these two mean values. I could not find very direct and convincing reasons for these differences from the manuscript. Perhaps I missed some details. In that case, the authors may wish to bring this point clearly in the discussion part.

Monthly-scale error patterns depend upon error covariances in the 6-hourly model output. Different regions will have greater temporal error covariances than others. These differences in the covariances will result in different error patterns at the 6-hourly versus monthly scale. The underlying question is why the error covariances are so much higher over the oceans and Arctic than over regions with large fluxes (Fig. 2).

Uncertainties in the month-long mean concentrations (Fig. 2) are most influenced by transport errors that occur over sustained time periods. When CO_2 is transported from source/sink regions to remote regions, that transport is likely to be associated with synoptic time scales, and any transport errors would likely be sustained over multi-day time periods. At these longer time scales, the surface fluxes are transported away from the surface grid box where they occurred and can manifest as transport errors in regions that are remote from large fluxes.

In regions with large fluxes, surface concentrations will additionally be influenced by grid-scale winds or boundary layer mixing. Transport errors at this grid-scale may have a shorter decorrelation time compared to errors in large-scale flow. In addition, sustained transport errors over regions of large biosphere flux would be more likely to cancel out at longer time scales – due to the diurnal cycle of biosphere CO_2 uptake and release (i.e., transport errors times of CO_2 uptake and release will have opposite sign.). Hence, transport errors in regions with large fluxes would likely average out or cancel to a greater degree than those in remote areas.

We have added additional explanation on this point to section 3.2 in the revised manuscript.

• p.23692, line 13: ".. from surface sources is strong" - ".. from surface sources and sinks is strong"

We have updated the manuscript accordingly.

• p. 23696, line 9: "At marine sites, in contrast, the minimum detectable bias is far larger". Why? transport uncertainties are comparatively shown lower over coastal areas!?

Marine sites are often located relatively far from regions with large CO_2 fluxes. At these marine sites, the signal-to-noise ratio is therefore smaller. We have added a similar explanation to this section of the revised manuscript.

• p. 23696, line 11: ".. large sources are better .." - ".. large sources and sinks are better .."

We have changed this text accordingly.

References

- Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO₂ fluxes estimated from GOSAT retrievals of total column CO₂, Atmospheric Chemistry and Physics, 13, 8695–8717, doi:10.5194/acp-13-8695-2013, 2013.
- Gourdji, S. M., Mueller, K. L., Schaefer, K., and Michalak, A. M.: Global monthly averaged CO₂ fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, Journal of Geophysical Research: Atmospheres, 113, doi:10. 1029/2007JD009733, 2008.
- Kang, J.-S., Kalnay, E., Miyoshi, T., Liu, J., and Fung, I.: Estimation of surface carbon fluxes with an advanced data assimilation methodology, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2012JD018259, 2012.
- Li, H., Kalnay, E., and Miyoshi, T.: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter, Quarterly Journal of the Royal Meteorological Society, 135, 523–533, doi:10.1002/qj.371, 2009.
- Liu, J., Fung, I., Kalnay, E., and Kang, J.-S.: CO₂ transport uncertainties from the uncertainties in meteorological fields, Geophysical Research Letters, 38, doi:10.1029/2011GL047213, 2011.
- Liu, J., Fung, I., Kalnay, E., Kang, J.-S., Olsen, E. T., and Chen, L.: Simultaneous assimilation of AIRS XCO₂ and meteorological observations in a carbon climate model with an ensemble Kalman filter, Journal of Geophysical Research: Atmospheres, 117, doi: 10.1029/2011JD016642, 2012.
- Miyoshi, T.: The Gaussian approach to adaptive covariance inflation and its implementation with the Local Ensemble Transform Kalman Filter, Monthly Weather Review, 139, 1519–1535, doi:10.1175/2010MWR3570.1, 2011.
- Miyoshi, T. and Kunii, M.: The Local Ensemble Transform Kalman Filter with the Weather Research and Forecasting Model: Experiments with real observations, Pure and Applied Geophysics, 169, 321–333, doi:10.1007/s00024-011-0373-4, 2012.
- Mueller, K. L., Gourdji, S. M., and Michalak, A. M.: Global monthly averaged CO₂ fluxes recovered using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements, Journal of Geophysical Research: Atmospheres, 113, doi:10.1029/2007JD009734, 2008.

- Oppenheim, A. and Schafer, R.: Discrete-time Signal Processing, Prentice-Hall signal processing series, Prentice Hall, 1989.
- Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., Werf, G. R. v. d., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proceedings of the National Academy of Sciences, with updates documented at http://carbontracker.noaa.gov, 104, 18 925–18 930, doi:10.1073/pnas.0708986104, 2007.
- Saito, K., Hara, M., Kunii, M., Seko, H., and Yamaguchi, M.: Comparison of initial perturbation methods for the mesoscale ensemble prediction system of the Meteorological Research Institute for the WWRP Beijing 2008 Olympics Research and Development Project (B08RDP), Tellus A, 63, doi:10.1111/j.1600-0870.2010.00509.x, 2011.
- Szunyogh, I., Kostelich, E., Gyarmati, G., Kalnay, E., Hunt, B., Ott, E., Satterfield, E., and Yorke, J.: A local ensemble transform Kalman filter data assimilation system for the NCEP global model, Tellus A, 60, doi:10.1111/j.1600-0870.2007.00274.x, 2008.
- Yang, S.-C., Kalnay, E., and Miyoshi, T.: Accelerating the EnKF Spinup for Typhoon Assimilation and Prediction, Weather and Forecasting, 27, 878–897, 2012.
- Zheng, X., Wu, G., Zhang, S., Liang, X., Dai, Y., and Li, Y.: Using analysis state to construct a forecast error covariance matrix in ensemble Kalman filter assimilation, Advances in Atmospheric Sciences, 30, 1303–1312, doi:10.1007/s00376-012-2133-5, 2013.

Discussion Paper

Manuscript prepared for Atmos. Chem. Phys. Discuss. with version 2014/05/30 6.91 Copernicus papers of the LATEX class copernicus.cls. Date: 8 February 2015

The potential for regional-scale bias in top-down CO₂ flux estimates due to effects of persistent atmospheric transport errors on uncertainties in modeled atmospheric CO₂

S. M. Miller¹, M. N. Hayek¹, A. E. Andrews², I. Fung³, and J. Liu⁴

¹Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA
 ²Global Monitoring Division, Earth Systems Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA
 ³Department of Earth and Planetary Sciences, University of California–Berkeley, Berkeley, CA, USA
 ⁴Earth Science Division, Jet Propulsion Laboratory, NASA, Pasadena, CA, USA

Correspondence to: S. M. Miller (scot.m.miller@gmail.com)

Discussion Paper

Abstract

5

10

Estimates of CO2-

Estimates of CO_2 fluxes that are based on atmospheric data-measurements rely upon a meteorological model to simulate atmospheric CO_2 - CO_2 transport. These models provide a quantitative link between surface fluxes of CO_2 and atmospheric the fluxes and CO_2 measurements taken downwind. Therefore, any errors in the meteorological model can Errors in the meteorology can therefore propagate into atmospheric CO_2 - CO_2 transport and ultimately bias the estimated CO_2 fluxes. These errors, however, have traditionally been CO_2 fluxes. Errors that covary in space and/or time are particularly worrisome because they are more easily confused with the actual signal from surface CO_2 fluxes and are difficult to characterize. To examine the effects of CO_2 transport errors on estimated CO_2 fluxes, we use In this paper, we leverage

- a global meteorological model-data assimilation system known as 'CAM-LETKF' to quantify two aspects of the transport errors: error variances (standard deviations) and temporal error correlations. Furthermore, we develop two case studies. In the first-model combined with a
- ¹⁵ data assimilation system to estimate atmospheric CO_2 transport errors. In one case study, we examine the extent to which CO_2 transport uncertainties can bias CO_2 flux estimates. In particular, we use a common flux estimate known as Carbon Tracker to discover the minimum hypothetical bias that can be detected above the CO_2 transport uncertainties. In the estimate the magnitude of monthly-averaged CO_2 transport errors relative to CO_2 boundary layer
- ²⁰ enhancements and quantify how that answer changes if we either include or remove error covariances. In a second case study, we then investigate which meteorological conditions may contribute to are associated with covarying errors at this month-long biases in modeled atmospheric CO₂ transporttime scale.

We estimate 6-hourly CO₂ transport uncertainties in the model surface layer that range from 0.15 to 9.6 ppm (standard deviation), depending on location , and we estimate an average error decorrelation time of ~2.3 days at existing CO₂ In the first case study, we estimate uncertainties of 0.5 to 7 ppm in monthly averaged CO₂ concentrations, depending upon location (95% confidence interval). These uncertainties correspond to 13-150% of the afternoon CO_2 boundary layer enhancement at individual observation sites. As a consequence of these uncertainties, we find that Carbon Tracker CO_2 fluxes would need to be biased by at least 29, on average, before that bias were detectable at existing non-marine atmospheric CO_2 observation sites. Furthermore When we remove error covariances, however, the error range drops to 2-22%.

⁵ Top-down studies that ignore these covariances could therefore underestimate the uncertainties and/or propagate transport errors into the flux estimate.

In the second case study, we find that persistent, bias-type errors in atmospheric CO₂ transport are associated with consistent low net radiation, low energy boundary layer conditions. The meteorological model is not necessarily more uncertain in these conditions. Rather, the extent to which meteorological uncertainties manifest as persistent atmospheric CO₂ transportbiases appears to depend, at least in part, on the energy and stability of the boundary layer. Existing CO₂ flux studiesmay be more likely to estimate inaccurate regional

fluxesunder those conditions. these persistent, month-long atmospheric transport errors are anti-correlated with temperature and planetary boundary layer (PBL) height over terrestrial

regions. In marine environments, by contrast, these persistent transport errors are more strongly associated with weak zonal winds. Many errors, however, are not correlated with a single meteorological parameter, suggesting that a single meteorological proxy is not sufficient to characterize uncertainties in atmospheric CO₂ transport. Together, these two case studies provide information to improve the setup of future top-down inverse modeling studies,
 preventing unforeseen biases in estimated CO₂ fluxes.

1 Introduction

10

25

Scientists increasingly use atmospheric CO_2 -CO₂ observations to estimate CO_2 -CO₂ fluxes at the Earth's surface (e.g., Gurney et al., 2002; Michalak et al., 2004; Peters et al., 2007; Gourdji et al., 2012). This $\frac{24}{2}$ top-down $\frac{22}{2}$ approach contrasts with $\frac{24}{2}$ bottom-up $\frac{22}{2}$ studies that rely primarily on expert knowledge of biological processes (e.g., Huntzinger et al., 2012; Raczka et al., 2013). In order to estimate the fluxes, top-down studies typically require a meteorology model to link fluxes at the surface with measurements taken downwind. Using this link, one can

estimate the fluxes even if the atmospheric measurements do not themselves directly measure the fluxes.

However, both the accuracy and effective resolution of the flux estimate hinge upon the accuracy of the meteorological model. Errors in the meteorological model may (or may not) translate into errors in $CO_2 CO_2$ transport from the location(s) of surface fluxes to the atmospheric mea-5 surement site(s). Subsequently, errors in $\frac{CO_2}{CO_2}$ cross transport may (or may not) bias estimated CO_2 -CO₂ fluxes depending upon the error characteristics and the space/time scales of interest. This caseading chain of cause and effect defines the three types of errors or uncertainties that are of primary interest in this paper: (1) errors in modeled meteorological variables, (2) errors in atmospheric CO₂ transport, as they manifest in modeled atmospheric CO₂ concentrations, and 10

(3) errors in the fluxes that result from problems in estimated transport. This study is particularly concerned with how CO₂ transport errors may propagate into the estimated fluxes.

More specifically, the effect of $\frac{CO_2}{CO_2}$ consport errors on the estimated fluxes depends upon two important factors. First, the flux estimate becomes more uncertain as the $\frac{CO_2}{CO_2}$ transport error variance (or standard deviation) increases. Top-down studies that use Bayesian 15 statistics will explicitly account for these variances when estimating fluxes (e.g., Enting, 2002; Tarantola, 2005); before estimating the fluxes, the modeler first estimates the total variance due to an array of model or data errors - due to imperfect atmospheric transport or imperfect measurements, among many other sources of error (e.g. Gerbig et al., 2003; Michalak et al., 2005; Ciais et al., 2011).

20

Second, the flux estimate becomes more uncertain as the temporal and/or spatial covariance in the errors increases error covariances increase. As the covariances increase, each CO_2 CO₂ measurement effectively provides less and less independent information to constrain the surface fluxes. Error correlations, however, are often difficult to characterize

(e.g. Lin and Gerbig, 2005; Lauvaux et al., 2009) and are omitted from most existing top-down 25 studies. These difficulties aside, correlated transport errors can have a number of impacts on the estimated greenhouse gas fluxes. First, an top-down study that does not account for these errors will typically underestimate the uncertainties in the flux estimate. Second, correlated errors can

Discussion Paper

Furthermore, these temporally- and/or spatially-correlated errors can bias the flux estimate over a region or over the entire geographic area of interest (e.g., Stephens et al., 2007).

Quantification of this complex cause-and-effect between meteorological errors and errors in estimated CO_2 -CO₂ fluxes represents an ongoing research challenge, and a number of existing studies have partly characterized characterized different aspects of these uncertain-

- ties. For example, a series of studies known as "TRANSCOM" "TRANSCOM" represents 5 one of the first coordinated projects on CO_2 CO₂ transport uncertainties (Gurney et al., 2002; Baker et al., 2006). These early studies used 13 different global atmospheric models
- 2002, Baker et al., 2000). These early studies used 13 different global atmospheric models and compared differences in top-down CO_2 – CO_2 budgets due to atmospheric model dif-ferences. These models gave an uncertainty in the northern hemisphere CO_2 budget of \pm 1.1 PgCyr⁻¹ (standard deviation; mean budget of 2.4 PgCyr⁻¹) (Stephens et al., 2007). Subsequent to the TRANSCOM project, a number of studies have focused on the effects of changing vertical mixing and/or planetary boundary layer height (PBLH) (Gerbig et al., 2008; Kretschmer et al., 2012; Parazoo et al., 2012; Pino et al., 2012; Kretschmer et al. 10 In general, these those papers found that uncertainties in PBLH can lead to errors of up to
- \sim 3ppm in modeled CO₂ \sim 3ppm in modeled CO₂. Another paper examined the effect of 15 uncertain horizontal winds (Lin and Gerbig, 2005). The authors applied a particle-trajectory model at a measurement site in Wisconsin and found that uncertainties in the horizontal winds contributed $\sim 6 \text{ ppm} \sim 6 \text{ ppm}$ (standard deviation) to the overall CO_2 -CO₂ transport uncertainty. In summary, a number of previous studies have either perturbed individual meteorological parameters or, in the case of TRANSCOM, sampled a subset of transport 20 uncertainties using 13 pre-selected atmospheric models.

Numerous questions still remain, however. For example, if one could carefully utilize all available meteorological observations, what meteorological and This study is particularly concerned with temporal and/or spatial error covariances in atmospheric CO_2 transport. To what extent do CO₂ transport uncertainties would remain? Furthermore, what is the combined effect of meteorological errors from multiple parameters (e.g., wind, boundary layer, etc.) on CO₂ transport and subsequently on CO₂ fluxes? In addition, which meteorological errors are most likely to bias regional-scale CO₂ flux estimates on month-long time scales? errors covary

25

in space and time? How large are these covariances relative to the magnitude of the surface CO_2 fluxes, and which meteorological factors drive large error covariances? These covariances are often difficult to characterize (e.g. Lin and Gerbig, 2005; Lauvaux et al., 2009) and are omitted from most existing top-down efforts.

In the present study, we We explore several facets of these questions using a global meteorology model ensemble and a meteorology data assimilation system – the Community Atmosphere Model (CAM) and an assimilation framework known as an Local Ensemble Transform Kalman Filter (LETKF) (Hunt et al., 2007; Liu et al., 2011) . CAM–LETKF (Hunt et al., 2007) . Efforts by Liu et al. (2011) and Liu et al. (2012) extended this meteorological framework to model uncertainties in atmospheric CO₂.

This framework systematically estimates meteorology and CO₂ CO₂ transport uncertainties to an extent not previously possible; this framework CAM-LEKTF explicitly represents the CO₂ CO₂ transport uncertainties that remain after assimilating several hundred thousand meteorology observations at each 6-hour 6 h model time step. To accomplish this task, CAM-LETKF uses an ensemble of weather forecasts and optimizes the ensemble to match available meteorological observations. Furthermore, CAM-LETKF adjusts the variance of the weather ensemble at each time step to match the modeling uncertainties implied by the meteorological

observations.

20

5

Using this toolkit, we construct two several case studies to understand both the possible magnitude and potential drivers of bias in top-down CO_2 flux budgets. Previous studies by Liu et al. (2011) and Liu et al. (2012) used CAM-LETKF to estimate CO_2 transport uncertainties, and this study investigates connections with top-down CO_2 flux estimation. First, we construct a case study with a commonly-used estimate of CO_2 fluxes known as Carbon Tracker (CT): how biased would regional CO_2 fluxes need to be before that bias were detectable above the meteorological uncertainties estimated by CAM-LETKF? We test this hypothesis at a number of atmospheric CO_2 monitoring sites in the US, Canada, Europe, andEast Asia. Second, we construct a case study using a synthetic atmospheric tracer. This synthetic experiment serves as a lens to explore the possible meteorological factors associated with persistent, month-long deviations in atmospheric transportdrivers of CO_2 transport error covariances – errors that persist over many time steps and/or across large regions. The next section describes CAM-LETKF and these case studies in greater detail.

2 Methods

2.1 The meteorology and CO_2 -CO₂ model

The first component of CAM-LETKF is the meteorological model. We simulate global meteorology using the Community Atmosphere Model (CAM) and Community Land Model (CLM, version 3.5), run in weather forecast mode (not climate mode) (Collins et al., 2006; Oleson et al., 2008; Chen et al., 2010). Model simulations in this study have a spatial resolution of 2.5^o olongitude by 1.9^o alatitude with 26 vertical model levels. In most regions, there are three vertical
 model levels within the lowest kilometer of the atmosphere. These model levels are centered at 929.6, 970.6, and 992.6 hPa over regions where the land/water surface is at sea level.

We save the global model output at 6-hour 6 h time increments. Furthermore, we run the model for two time periods: January – February January–February 2009 and May – July May–July 2009. The first month of each run serves as an initial spin-up for the model-data assimilation system. The next section describes this assimilation in greater detail.

20

25

2.2 The meteorological model-data assimilation framework

The second component of CAM–LETKF is the data assimilation and model optimization framework. This framework serves two purposes. First, the LETKF optimizes modeled meteorology (CAM–CLM) to match available observations. Second, the LETKF uses an ensemble of model forecasts to represent model uncertainties that remain after data assimilation (Hunt et al., 2004, 2007). We define each ensemble member and the mean of the entire ensemble as follows:

 $oldsymbol{x}_i \ = \ oldsymbol{ar{x}} + oldsymbol{X}_i \quad ext{where} \ i = 1 k$

(1)

 $\boldsymbol{x}_i = \boldsymbol{\bar{x}} + \mathbf{X}_i$ where i = 1...k

5

10

where x_i ($m_1 \times 1m \times 1$) is a single model ensemble member, \bar{x} ($m_1 \times 1m \times 1$) is the mean of the model ensemble, and X_i ($m_1 \times km \times k$) refers to the *i*th *i*th column of the matrix that defines the ensemble spread. In this paper, the variable $m_1 m$ refers to the total number of model parameters – the model estimate for a variety of meteorological variables, concatenated across the globe and across all 6-hourly 6 hourly time steps in a given model run. Furthermore, we use k = 64 total ensemble members in this setup, as was done in Liu et al. (2011) and Liu et al. (2012).

Using this ensemble, CAM-LETKF steps through time in sequential 6-hour 6 h intervals. First, the model ensemble at time t is optimized to match meteorological data (Hunt et al., 2007). To this end, we assimilate the same meteorological observations used in the 15 National Centers for Environmental Prediction-Department of Energy reanalysis 2 (Kanamitsu et al., 2002): temperature (in situ and satellite), zonal wind (in situ and satellite), meridional wind (in situ and satellite), surface pressure (in situ), and specific humidity (in situ). At each 6-hour 6 h model time step, we assimilate between $\sim 180,000 \sim 180,000$ to 330,000 observations globally. At that juncture, the ensemble mean associated with time t, $\bar{x}(t)$, represents the 20 model best guess and the ensemble members, $\bar{x}(t) + X(t)$, collectively represent the posterior variances and covariances uncertainties in the modeled meteorology (i.e., posterior variances and covariances). For the remainder of this paper, we define the 6-hourly meteorological uncertainties as the standard deviation (or alternately, the range) of each row in X. Second, we run 6-hour 6 h CAM-CLM forecasts using these realizations as initial conditions – a total of 64 25 model forecasts. This ensemble of global forecasts then becomes the prior (and prior variances and covariances) for the next LETKF assimilation cycle (Hunt et al., 2007). The 6-hour The 6 h cycle of data assimilation and model forecast then begins again.

This model ensemble, by design, is guaranteed to reflect actual uncertainties in modeled meteorology; at each 6-hour 6 h model time step, we adjust the ensemble variance such that this variance matches against the model–data residuals (Li et al., 2009; Miyoshi, 2011). The supplement Supplement describes this procedure, known as adaptive covariance inflation. For-

The model ensemble also accounts for both spatial and temporal covariances in modeled

- meteorological uncertainties; meteorological errors within one ensemble member can easily persist over many time steps. This continuity occurs because the optimized ensemble members from the one time step become the initial conditions for the weather forecast at the next time step. For example, if the PBL height in one ensemble member is lower than the ensemble average at a given time step, it will likely be lower than average at the next time step. Similarly,
 if the PBL height in one ensemble member is lower one grid box, it will likely
- ¹⁰ If the PBL height in one ensemble member is lower than average over one grid box, it will likely also be lower than average over the adjacent grid box.

Certain meteorological uncertainties, however, may not always be captured by the assimilation system, particularly uncertainties that do not manifest in the model-data residuals. For example, CAM-LETKF will not fully characterize uncertainties due to different PBL schemes (e.g., Yonsei versus Mellor-Yamada-Janjic) or due to other structural model

¹⁵ schemes (e.g., Yonsei versus Mellor-Yamada-Janjic) or due to other structural model differences. Furthermore, LETKF cannot spatially resolve uncertainties that occur at sub-grid scale (e.g., turbulent eddies or numerical diffusion). For further technical detail on the LETKF and adaptive covariance inflation, refer to the supplementSupplement, Hunt et al. (2004), Hunt et al. (2007), Li et al. (2009), Liu et al. (2011), or Miyoshi (2011).

20 2.3 CO_2 -CO₂ transport error variances and covariances

25

The CAM-LETKF system described above estimates not only meteorological uncertainties but also uncertainties in CO_2 - CO_2 transport. In this study, CO_2 - CO_2 is a passive tracer and that is not part of the data assimilation. Instead, we use biospheric, oceanic, biomass burning, and fossil fuel CO_2 fluxes from CT (version 'CT2011oi', Fig. 1)(Peters et al., 2007, http://carbontracker.noaa.gov). Furthermore, we use CT as the initial condition for global atmospheric CO_2 mixing ratios on January 1 and May 1, 2009. Each CAM ensemble member uses the same initial condition for atmospheric CO_2 , so any subsequent differences in CO_2 among the model realizations are due entirely to meteorological uncertainties, so any uncertainties in CO_2 concentrations are solely due to uncertainties in atmospheric transport.

We drive all model simulations with a published CO_2 flux estimate from CarbonTracker (CT), version "CT2011

- SUBSCRIPTNBoi" (Fig. 1, Peters et al., 2007, http://carbontracker.noaa.gov). CT is 5 a commonly-used global CO₂ flux estimate created by the US National Oceanic and Atmospheric Administration (NOAA). NOAA scientists optimize CT fluxes to match atmospheric CO_2 data, so the flux estimate is consistent with actual observations (Peters et al., 2007).
- 10
- We estimate 6-hourly CO_2 subsequently estimate 6 hourly CO_2 transport uncertainties using this setup. These uncertainties are defined as the difference between the top and bottom of the standard deviation of CO₂ concentrations across the 95% confidence interval, computed from the 64 model realizations (e.g., Fig. 2). To make this estimate, we calculate the standard deviation 2.5th and 97.5th percentiles of each row in $X_{[CO_2]}X_{[CO_2]}$, where the subscript
- " $[CO_2]$ [CO_2]" refers to the atmospheric CO_2 CO₂ concentrations estimated by the ensemble. In 15 addition, we characterize temporal covariance or correlation in transport errors (ie, in $X_{[CO_2]}$). To estimate an error decorrelation time, we use a variogram analysis. In specific, we fit an exponential variogram model to afternoon-only model output (1pm - 7pm local time) associated with a number of existing, global atmospheric CO₂ observation sites. Both Kitanidis (1997) and
- the supplement describe variograms in greater detail. The remainder of the methods section ap-20 plies this CO_2 -CO₂ and meteorology modeling framework to two case studies.
 - 2.4 Case study 1: How biased would Carbon Tracker fluxes need to be before that bias were detectable above CO_2 . The magnitude of temporally- and spatially-covarying atmospheric transport uncertainties? errors relative to a CO₂ flux estimate
- In this case study, we construct a hypothesis test to determine whether biases in CT 25 This case study explores the importance of persistent, covarying transport errors and the magnitude of those errors relative to the CO₂ fluxeswould be detectable above atmospheric transport uncertainties. CT is a commonly-used global CO₂ flux estimate created by the US National Oceanic and Atmospheric Administration (NOAA). To create CT, NOAA scientists

use atmospheric CO_2 data from observations towers and surface sites around the world and estimate regional scaling factors that optimize an initial CO_2 flux model (Peters et al., 2007).

We test whether a hypothetical bias in regional scaling factors, like those estimated by CT, would be detectable at atmospheric CO₂ observation sitesacross the globe. We build this test using the CO₂ sum of squared residuals (SSR) from the CAM-LETKF model ensemble. A number of previous statistical. In particular, we estimate uncertainties in monthly mean, afternoon, modeled CO₂ concentrations at a number of in situ atmospheric observation sites. In one case, we include temporal and/or greenhouse gas studies construct hypothesis tests

¹⁰ using squared residuals (e.g., Sheskin, 2003; Huntzinger et al., 2011). spatial covariances in the atmospheric transport errors, and in another case, we remove these covariances. We then compare these uncertainties against the modeled afternoon CO₂ boundary layer enhancement to understand the magnitude of these errors relative to the surface fluxes.

In this setup, we construct the test as follows. First, compute the SSR associated with the transport uncertainties :

 $\boldsymbol{SSR} = \sum_{i=1}^{n_2} \left(\mathbf{H}_{[CO_2]} \mathbf{X}_{[CO_2]} \right)^2$

This equation interpolates the model residuals $(\mathbf{X}_{[CO_2]})$ to the observation sites, squares these residuals, and sums them over the entire timeperiod of interest. More specifically, the variable n_2 refers to the number of hourly CO₂ observations at an observation site over the time span of the hypothesis test. In addition, $\mathbf{H}_{[CO_2]}$ $(n_2 \times m_2)$ is the matrix that interpolates or maps the ensemble deviations $(\mathbf{X}_{[CO_2]}, \text{ dimensions } m_2 \times k)$ to the CO₂ observations. Lastly, *SSR* $(1 \times k)$ are the sum of squared residuals from each of the k CAM-LETKF model ensemble members. Note that some of the ensemble members will be closer than others to the ensemble mean or The uncertainty in monthly-averaged CO₂ concentrations serves as a measure of how transport errors persist over time, a measure of error covariance. Uncorrelated transport errors

will average out, to a large degree, over many model time steps, but temporal error covariances prevent the errors from averaging down over time. Furthermore, CO_2 budgets are often reported

in month-long increments (e.g., Gourdji et al., 2012, and CT), so this time window is a relevant benchmark with respect to inverse modeling studies.

We calculate uncertainties in the monthly-averaged model output (including error covariances) via several steps. First, we select out the rows of $X_{[CO_2]}$ that correspond to

- ⁵ afternoon observations (1–7 p.m. LT) for a given month at an in situ CO_2 observation site. Second, we calculate the mean of each column in $\mathbf{X}_{[CO_2]}$. Each column corresponds to a different ensemble member. The resulting vector of length 64 is the difference between each ensemble member and the best estimate ($\bar{x}_{[CO_2]}$). Therefore, \bar{x}), averaged at the monthly scale. Lastly, we use this vector to compute a confidence interval in monthly-averaged, modeled
- ¹⁰ CO₂ (the *k* SSRs from the *k* ensemble members will not be identical and will instead form a distribution. 97.5th minus 2.5th percentiles).

Second, we compute the SSR associated with a hypothetical bias (λ)in the fluxes:-

 $\underline{FSSR} = \underline{\sum^{n_2} (\Delta CO_2)^2}$

 $\Delta CO_2 = \lambda \mathbf{H}_{[CO_2]} \left(\bar{\boldsymbol{x}}_{[CO_2, \text{surface}]} - \bar{\boldsymbol{x}}_{[CO_2, 600 \text{ hPa}]} \right)$

¹⁵ We subsequently remove covariances in the CO₂ transport errors and re-calculate uncertainties in the monthly-averaged CO₂ concentrations. As described in section 2.2, errors in one ensemble member can persist over many steps and can persist across a large geographic region. However, we can remove these error covariances by randomly re-shuffling the elements of each individual row in X_[CO2]. The variance in modeled concentrations in any row or at any given time step will remain the same. However, each column will no longer represent a single ensemble member. Rather, each column will represent a random assortment of different ensemble members, and the errors in each column will no longer covary from one time step to another or one geographic location to another.

The output of this equation, FSSR, is a scalar that estimates the squared residuals due to a biased flux estimate, summed over all observations at a given CO₂ measurement site. The variable λ represents a hypothetical bias in CT fluxes. In this study, we conduct the hypothesis test at each measurement site individually, so the variable λ is

Discussion Paper

specific to the site being consideredWe conduct this analysis at a representative selection of observation sites in North America, Asia, and Europe. This setup indicates how errors covary with time at the monthly scale. In addition, the variables in parentheses $(\bar{x}_{[CO_2,surface]} - \bar{x}_{[CO_2,600 hPa]})$ quantify the contribution of regional-scale fluxes to CO₂ at the atmospheric observation site. Many top-down studies pre-subtract free troposphere or marine 'clean air' concentrations from the CO₂ measurements or model output at the observation sites (e.g., Gerbig et al., 2003; Gourdji et al., 2012). These top-down studies then optimize regional fluxes to match the pre-subtracted CO₂ observations. In this study, we similarly subtract modeled we also conduct the analysis using multiple observation sites; we estimate monthly-averaged uncertainties at the eco-region scale and include all observation sites that lie within the given eco-region. This latter approach indicates how errors covary spatially across

multiple sites at the regional scale.

5

10

These monthly-averaged uncertainties can then be compared against the afternoon, modeled CO_2 increment from regional surface fluxes. To estimate this increment, we subtract modeled

- ¹⁵ free troposphere, "clean air" concentrations at 600hPa in the free troposphere ($\bar{x}_{[CO_2,600 \text{ hPa}]}$) from those at the CO₂ observation sites ($\bar{x}_{[CO_2,\text{surface}]}$) hPa from concentrations modeled at the surface using CT fluxes. The concentrations at 600hPa hPa are not necessarily a perfect measure of 'elean air ' clean air concentrations. Rather, this approach is an approximation similar to that used in the existing literature . This difference in concentrations is then used to estimate
- ²⁰ how a regional bias in CT fluxes would manifest at a given observation site (ΔCO_2 , in ppm). by inverse modeling studies in the literature (e.g., Gerbig et al., 2003; Gourdji et al., 2012).

Finally, we test the hypothesis. If the FSSR is larger than most of the k SSR associated with the meteorological uncertainties, then we can distinguish the flux bias (λ) above the meteorological noise. This statement can be formalized into a hypothesis test as follows:

25 $\underline{A} = \{ \mathbf{SSR} \mid \mathbf{SSR} > FSSR \}$

 $\underline{p} = \frac{|A|}{k}$

where A is the set of **SSR** that are greater than FSSR, and the expression |A| indicates the number of elements in A. If the p-value is less than 0.05, we have disproven the null hypothesis – that the hypothetical bias (λ) in CO₂ fluxes is indistinguishable above the transport uncertainties.

Note that this hypothesis test accounts for both variance and temporal covariance in the CO₂ transport uncertainties, a concept discussed in detail in the supplement. In addition, note 5 that FSSR will almost never be zero due to diurnal or daily changes in NEP, even if the monthly-averaged NEP at a given site is zero.

We conduct the hypothesis test above for both February and July 2009 at a variety of different observation sites in North America, Asia, and Europe. We report the results of this

- hypothesis test for a representative selection of global CO₂ observation sites different types 10 of observation towers located on different continents and in different biomes. Furthermore, we test this hypothesis using month-long modeled time series corresponding to afternoon data only (1pm-7pm local time). We use this month-long window because CO₂ budgets are often reported in month-long increments.
- 15
 - In summary, case study 1 quantifies the extent to which atmospheric CO_2 transport errors ean obscure any regional biases in estimated CO_2 one explores the magnitude of sustained atmospheric CO_2 transport errors or error covariances relative to the afternoon CO_2 signal from surface fluxes. The next case study, in contrast, explores the meteorological conditions under which sustained CO_2 these sustained CO_2 transport errors may be more likely to occur.

2.5 Case study 2: Which meteorological factors may be associated with sustained, 20 month-long CO₂-transport biases?

We create arelatively simple, synthetic experiment to explore the meteorological conditions under which month-long model biases in atmospheric transport may occur. The spatial patterns in the CO_2 CO₂ transport uncertainties are heavily influenced by spatial patterns in the CO_2

CO₂ fluxes (Fig. 2). In other words, regions with large fluxes or large diurnal flux variability also 25 show higher CO_2 CO₂ transport uncertainties. As a result, it is difficult to disentangle the effect of different meteorological parameters on $\frac{CO_2}{CO_2}$ CO₂ transport uncertainties. Instead, we create a synthetic tracer with constant global emissions in both space and time. This experiment serves as a lens to explore the possible effects of different meteorological parameters independent of the spatial variability in CO_2 spatiotemporal variability in CO_2 fluxes.

To this end, we initialize CAM-LETKF runs with zero atmospheric concentration of this synthetic tracer and then run CAM-LETKF forward for one month using constant global emissions 5 (e.g., for both February and July 2009). Any uncertainties in the atmospheric distribution of this tracer are solely due to meteorological parameters, not due to the spatial distribution of the underlying fluxes.

Next, we calculate the coefficient of variation (CV) associated with the monthly-averaged surface concentrations. The CV is an inverted signal-to-noise ratio; it measures the uncertainty 10 in modeled surface concentrations relative to the average surface concentration $(\frac{\sigma}{2})$. For example, an uncertainty of 1ppm in modeled concentrations is most problematic if the signal from surface fluxes is weak, and a 1ppm 1 ppm uncertainty is less problematic if the signal from surface sources and/or sinks is strong.

- For this setup, the CV equals the standard deviation in the monthly-averaged surface concen-15 trations divided by the monthly surface concentration averaged across all 64-realizations. We then plot the tracer CV against monthly-averaged meteorological parameters and their associated uncertainties from CAM-LETKF. These relationships give insight into the meteorological conditions or meteorological uncertainties that are associated with month-long biases in the modeled synthetic tracer. 20

25

Results and discussion 3

3.1 Uncertainties in the 6-hourly 6 hourly modeled CO₂-CO₂ concentrations

Before examining the two case studies in detail, we first provide context on the $\frac{CO_2}{CO_2}$ transport uncertainties estimated with CT fluxes and CAM-LETKF. Figure^{2a} and ² 2a and b visually summarize the average 6-hourly CO₂ transport uncertainties (standard deviations) 6 hourly CO₂ transport uncertainties in the model surface layer ; these – the difference between

the top and bottom of the 95% confidence intervals. These figures show how CO_2 - CO_2 transport uncertainties vary across the globe – from 0.15 to 9.6 ppm (standard deviation)0.6 to 26 ppm, depending on location. Furthermore, the transport uncertainties in Figs.2a and 2Fig. 2a and b show several distinctive features. The largest uncertainties are localized to regions where either the magnitude or the diurnal cycle of the CT fluxes is largest (e.g., the US Eastern Seaboard and southern Siberia during summertime, the Amazon, the Congo, and eastern China). CO_2 CO_2 transport uncertainties in the Eastern US and East Asia bleed, to a smaller degree, over the adjacent ocean where surface fluxes are small.

5

These transport uncertainties are in the range of the uncertainties estimated in a number of previous studies. For example, the spatial patterns in the 6-hourly 10 uncertainties are similar to those modeled by Liu et al. (2011) using CAM-LETKF and temperature-scaled CO₂ fluxes from TRANSCOM 3. In addition, a number of previous studies focused on the effects of perturbing individual meteorological parameters at specific observation sites or for individual aircraft campaigns (e.g., Gerbig et al., 2003; Lin and Gerbig, 2005; Gerbig et al., 2008; Kretschmer et al., 2012). 15 Our 6-hourly transport uncertainties, though very different in both scope and scale, are comparable in magnitude to the individual parameter uncertainties estimated by Gerbig et al. (2003), Gerbig et al. (2008), and Kretschmer et al. (2012) but are less than the uncertainties in (Lin and Gerbig, 2005). Furthermore, our estimated 6-hourly transport uncertainties also appear similar to or slightly smaller than the model-data 20 mismatch errors estimated at individual observation sites in several inversion studies (e.g., Peters et al., 2007; Schuh et al., 2010; Gourdji et al., 2012). Model-data mismatch includes not only transport errors but also any model or data errors unrelated to an imperfect initial flux estimate. This result may reflect the fact that atmospheric transport often dominates model-data mismatch errors. 25

Figure Figure 3 places these transport uncertainties in context of CO_2 - CO_2 data measured at two observation sites in the United States. These time series plots validate the model's capacity to simulate daily variations in CO_2 - CO_2 concentrations. Furthermore, the comparison illustrates the magnitude of the CO_2 - CO_2 transport uncertainties relative to the diurnal cycle in CO_2

CO₂ concentrations. For example, the uncertainties at AMT in July are ~30~30% of the diurnal range in the CO₂-CO₂ measurements. Overall, the model ensemble depicted in these plots usually encapsulates the hourly-averaged measurements. CT fluxes are estimated using these CO₂-CO₂ observations and the TM5 transport model (Tracer Model, version 5) (Peters et al., 2007), so one might expect the CAM model to fit the CO₂-CO₂ observations relatively well. In the instances when the model ensemble does not encapsulate the hourly-averaged CO₂ CO₂ measurements, one of the many other non-transport error types could be to blame; the ensemble spread only encompasses transport error errors and does not include measurement Furthermore, these instances could be due to structural differences between CAM and TM5, including differences in model resolution. The Supplement provides more example CO₂ CO₂ model-data comparisons, meteorology model validation, and data assimilation diagnostics.

3.2 CO₂ CO₂ transport uncertainties at longer time scales

25

- The uncertainty in monthly-averaged CO₂-CO₂ concentrations provides one measure of how transport errors persist over time. In other words, these uncertainties provide a metric of error correlations in CO₂ transport. Uncorrelated transport errors will average out, to a large degree, over many model time steps, but temporal correlations prevent the errors from averaging down over time. As a result, large uncertainties in monthly-averaged concentrations indicate
 the potential for persistent bias in CO₂ fluxes estimated using atmospheric observations. Such
- bias could lead to under- or over-estimation of regional-scale CO₂ budgets.

To this end, Figs. (2e and 2d), as discussed in section 2.4. Fig. 2c-d displays uncertainties in the month-long average surface concentrations for February and July -2009. In contrast to the 6-hourly 6 hourly uncertainties, these uncertainties are far more spatially-distributed; the largest uncertainties are not just associated with regions that have large fluxes. This result implies that CO_2 transport errors are correlated CO_2 transport errors covary over longer periods of time in remote regions, compared to regions with large anthropogenic or biospheric fluxes. Furthermore, surface fluxes. Observation sites that are far from large fluxes are therefore more

likely to produce a biased CO₂ budget than sites near to large surface fluxes. These "remote"

sites see a lower CO_2 signal from surface fluxes, and the transport errors at these locations generally covary over longer periods of time.

A number of factors may explain these relatively large error covariances in remote regions. ⁵ CO₂ transport over remote or oceanic regions is likely dominated by synoptic-scale weather patterns that evolve over multi-day time periods. When CO₂ is transported across the oceans or remote areas from source/sink regions, atmospheric CO₂ transport errors would likely covary at time-scales characteristic of this synoptic-scale air flow. Over large CO₂ source/sink regions, by contrast, atmospheric concentrations are likely influenced more strongly by processes that occur ¹⁰ over smaller time periods – grid-scale winds or boundary layer mixing. In addition, sustained transport errors over regions of large biosphere flux would be more likely to cancel out at longer time scales – due to the diurnal cycle of biosphere CO₂ uptake and release.

In addition to remote and ocean regions, month-long transport uncertainties are also large across the entire northern hemisphere during Februaryeven though biospheric fluxes are weak

- during that time period. Asubsequent section (3.4.) Northern Hemisphere during February. A subsequent Sect. 3.4 explores possible reasons why these month-long biases occur. In particular, the case study discussed in that section provides insight into why the month-long uncertainties may be large across the northern hemisphere during winter.
- A variogram analysis provides an additional measure of the error correlations in CO₂ transport (see section 2.3 and the supplement). Based upon this analysis, we estimate CO₂ transport error decorrelation times of 2.2 and 2.3 days at global atmospheric CO₂ observation sitesduring February and July, respectively (see Table S1). The error decorrelation times are generally longer-

3.3 Case study 1: The magnitude of temporally- and spatially-covarying atmospheric transport errors relative to a CO₂ flux estimate

25

We construct a case study to understand the importance of temporal and spatial error covariances relative to the magnitude of CO_2 surface fluxes. Figure 4 displays the results of this analysis for a selection of representative global CO_2 observation sites from Asia, Europe, and North America. The y-axis of each bar plot indicates the difference between the top and bottom of the

95% confidence interval in monthly mean modeled concentrations. We first consider the results when covariances in atmospheric CO_2 transport errors are included in the analysis (dark blue

 bars) and then compare those results to a setup in which we remove these error covariances (light blue bars).

At this selection of sites, uncertainties in the monthly mean afternoon concentrations range from 1.6 to 2.8 ppm (dark blue bars). These uncertainties are lower at marine sites (average of 2.9 and 2.7 days in February and July, respectively) or at like RYO and TTA and are higher at

- continental sites located near large biospheric fluxes, sites that are far from large CO₂ fluxes.
 For example, the longest error decorrelation times occur at coastal sites in Japan, Korea and the Canary Islands. In contrast, decorrelation times are usually shorter than average for observation sites on the European mainlandlike FSD and WBI. Note that this analysis only considers estimated uncertainties due to meteorology. The capabilities of the atmospheric observations
 would deteriorate if other errors were included, such as those due to imperfect measurements
- ¹⁵ would deteriorate if other errors were included, such as those due to imperfect measurements or due to finite model resolution (e.g., Gerbig et al., 2003; Masarie et al., 2011).

This level of temporal correlation in the We subsequently remove temporal covariances in the errors to identify the role that these covariances play in CO_2 transport uncertainties at the monthly scale. These results are displayed as light blue bars in Fig. 4. When we remove

- ²⁰ the covariances, the monthly-scale uncertainties are much smaller by a factor of 5-20 at the individual observation sites. If the CO₂ transport errors implies several large-scale conclusions for estimating CO₂ fluxes. First, observation sites that are far from large fluxes are more likely to produce a biased CO₂ budget than sitesnear to large surface fluxes. These 'remote' sites see a lower CO₂ signal from surface fluxes, and transport errors were temporally independent,
- then errors of opposite sign and different magnitude would cancel out to a degree when averaged over one month (light blue bars). Instead, the transport errors at these locations are generally correlated over longer periods of time. Second, most existing top-down studies will underestimate the uncertainties in estimated CO₂ fluxes. Existing inversions rarely account for error correlations in CO₂ transport and most likely underestimate the posterior uncertainties as a direct result. The next section (3.3) quantifies the impact of the transport uncertainties discussed

Discussion Paper

above on surface flux estimation covary in time, and this covariance prevents the errors from averaging down (dark blue bars).

A multi-site comparison in Fig. 4 additionally indicates the role of spatial covariances in the 5 transport errors; the figure shows the uncertainties in CO₂ concentrations when averaged across multiple observation sites within an eco-region. We compute the monthly average afternoon concentration across multiple sites for a given ensemble member. We then estimate a confidence interval based upon the distribution of the 64 ensemble members.

3.4 Case study 1: How biased would CT fluxes need to be before that bias were detectable 10 above the CO₂ transport uncertainties?

We use a case study from CT to understand how transport errors translate into uncertainty in a top-down. The results indicate a large degree of spatial covariance in the atmospheric CO_2 flux estimate. In specific, if the flux scaling factors estimated by CT were incorrect, how wrong would those scaling factors need to be before the problem were detectable above atmospheric transport errors?

15

20

25

Figure ?? shows the results of this case study (described in section 2.4) at a selection of global CO₂ observation sites. The y-axis of each bar plot shows the minimum bias that would be detectable using hourly-averaged CO₂ observations collected over an entire month. The mean minimum detectable bias across all non-marine sites is 29(at a month-long time scale).The results are not substantially different at short versus tall non-marine tower sites: 27and 31, respectively. In other words, the tall towers examined in this analysisare neither more or less sensitive to biased CO₂ fluxes in comparison to the set of short towers in Fig. ??. At marine sites, in contrast, errors. If the errors had no spatial covariance, those errors would average down as more and more observation sites were added to the minimum detectable bias is far larger: 76on average. analysis. However, the dark blue bars in Fig. 4 have a similar magnitude

irrespective of whether the analysis was conducted on an individual site or on a collection of many sites from an ecoregion; the errors must therefore covary in space. In contrast, the light blue bars (i.e., error covariances removed) do decrease in magnitude at the eco-region scale

relative to individual observation sites. In that case, the errors do average out when more and more sites are included in the analysis.

These results show Figure 5 places the results of case study one in the context of the surface fluxes. This figure displays the uncertainties in atmospheric CO_2 transport (the dark blue bars in

- ⁵ Fig. 4) as a fraction of the mean afternoon CO_2 boundary layer enhancement. As discussed in Sect. 2.4, this enhancement approximates the CO_2 increment due to regional surface fluxes, and a similar CO_2 increment is used by a number of additional trends across the different observation sites. In general, towers that are near large sources are better able to detect bias in the modeled fluxes. These include observation sites in the central and eastern US or in Germany
- ¹⁰ and Eastern Europe sites that are strongly influence by terrestrial (versus marine) airflow relative to other locations. Most of these towers see large signals from biospheric fluxes during summer (Figs.S9–S14). Other towers, in contrast, are less sensitive to detecting bias during the summertime (e. g., the marine towers and towers in top-down studies to estimate the surface fluxes. At the individual observation sites, the uncertainty in atmospheric CO₂ constitutes
- 13-150% of the average boundary layer CO₂ enhancement. This percentage is highest at marine sites like RYO and TTA that see a relatively small boundary layer enhancement, and the relative magnitude of the uncertainties is smallest at sites that see a very large enhancement due to large summertime vegetation fluxes (e.g., at the WBI site). The uncertainties due to atmospheric transport are substantial relative to the western US). The western US towersfluxes, but only when we include covariances in transport error. When we remove these covariances, the
- uncertainty in monthly average afternoon concentrations drops to only 2-22% of the boundary layer enhancement.

The results of this analysis hold several implications for future atmospheric inversions and/or top-down studies that optimize CO_2 fluxes. Most existing inversions account for atmospheric

²⁵ CO₂ transport errors in their statistical setup. In a Bayesian synthesis or geostatistical inversion, for example, are surrounded by weak biosphere uptake that is diluted into a larger mixed layer during summer. But during summer, transport uncertainties increase due to large seasonal fluxes in adjacent regions. The sensitivity of the marine Japanese and Korean sites also declines in the summer. At these sites, this information is incorporated into a covariance matrix, and that covariance matrix is used as an input to the equation that optimizes the CO_2 fluxes (e.g., Enting, 2002; Michalak et al., 2004; Ciais et al., 2011). However, the majority of existing studies assume that this covariance matrix is diagonal (i.e., no error covariances), in part, because these temporal and spatial covariances are challenging to estimate (e.g., Lin and Gerbig, 2005; Lauvaux et al., 2009). The present study, in contrast, indicates that both temporal and spatial error covariances play an important role in monthly-scale errors in atmospheric transport.

5

25

Ignoring these error covariances could lead to numerous challenges. When we add more data at an observation site or add more sites the analysis, the actual errors do not average down to the extent that uncorrelated errors would. Rather, adding more data or more observation sites provides more limited gains in accuracy. As a result, an inversion that overlooks the error covariances will estimate uncertainties in the signal from surface fluxes is largest in winter. During summer, biosphere uptake somewhat cancels the signal from large anthropogenic

¹⁵ emissions in China. CO₂ fluxes that are too small, and/or the inversion may erroneously map atmospheric transport errors onto the surface fluxes (e.g., Stephens et al., 2007). Future inversion studies could better account for these uncertainties by including off-diagonal terms in one of the covariances matrices used by the inversion.

The next case study (section 3.4) explores the meteorological factors that may be associated with these persistent atmospheric transport errors.

Note that this analysis only considers uncertainties due to meteorology. The capabilities of the atmospheric observations would deteriorate if other errorswere included (e. g., measurement errors or errors due to model resolution).

3.4 Case study 2: Which meteorological factors are associated with sustained, month-long atmospheric transport biases?

We now examine the results of the In this case study, we use a synthetic tracer experiment (section Sect. 2.5) to uncover possible drivers of atmospheric transport biases - at month-long time scales. The previous section (3.3) explored the importance of covariances in atmospheric

 $\rm CO_2$ transport errors, and this section investigates the meteorological conditions associated with these persistent errors.

Figure 6 displays the coefficient of variation (CV) for monthly-averaged surface concentra tions of the synthetic tracer. The CV, a unitless quantity, does not just indicate where the uncertainties are largest. Rather, the CV indicates the magnitude of these uncertainties relative to the mean modeled tracer concentration. Arguably, this noise-to-signal ratio measures the influence of transport uncertainties more effectively than a simple standard deviation. The remainder of this section focuses only on land regions because most existing top-down studies focus on land fluxes.

This coefficient shows a number of distinctive seasonal and spatial patterns. Like the uncertainties in monthly-averaged $\frac{CO_2}{CO_2}$ (Fig. 2c and 2d), the CV in Fig. 6 is highest in terrestrial boreal and arctic regions of the northern hemisphere Northern Hemisphere during winter. The CV is lowest over Europe, Australia, and the Amazon during all seasons.

The CV in Fig. 6 exhibits different spatial patterns over land and ocean regions, and these respective patterns correlate with different sets of meteorological variables. Over the oceans, for example, high CV values in Fig. 6a are clustered in zonal bands – along the equator and along 40°S. In contrast, high CV values do not cluster into zonal bands to the same degree over terrestrial regions. Rather, CV values are often high when temperatures are low (e.g., over
 20 Canada or Russian in February).

We plot the synthetic tracer CV against numerous modeled meteorological parameters to further understand the possible drivers behind the transport uncertainties. Of the atmospheric transport uncertainties averaged over these monthly time scales. To this end, we examine correlations between the tracer CV and 60 variables tested (Table S2), seven of the variables

showed correlations (R²) different meteorological parameters, including the uncertainties in the meteorological variables. Figure 7 displays the two variables that correlate most strongly with the tracer CV that are greater or equal to 0.3 (Fig. 7). Meteorological over land regions and over ocean regions, respectively.

Over land regions, meteorological conditions that lead to high atmospheric stability and low energy are most closely associated with persistent tracer uncertainties (relative to mean surface

concentrations). For example, a high tracer CV is associated with low temperatures , low net radiative flux, $(R^2 = 0.45)$ and low specific humidity $(R^2 = 0.40)$. Similarly, a high tracer CV is correlated with low net solar flux $(R^2 = 0.35)$, low planetary boundary layer height $(R^2 = 0.33)$,

⁵ correlated with low net solar flux ($\mathbb{R}^2 = 0.35$), low planetary boundary layer height ($\mathbb{R}^2 = 0.33$), and low vertical diffusion diffusivity . Furthermore, many of the ($\mathbb{R}^2 = 0.31$). Note that many of these variables are closely correlated with one another, so the individual correlations listed above are all interrelated.

In addition, several of the meteorological variables exhibit a nonlinear relationship with the tracer CV; the, and the potential for bias in modeled atmospheric transport increases more quickly in stable atmospheric conditions. For example, the CV increases more quickly when net radiation and planetary boundary heights are low.

In contrast to land regions, the tracer CV over the oceans is most closely associated with low zonal wind speeds (R² = 0.29, Fig. 7). Over land regions, that correlation is zero. Uncertainties in atmospheric transport over the oceans are also associated with low PBL heights (R² = 0.25). These two meteorological variables explain different patterns in the tracer CV; PBL heights and zonal wind speeds over the ocean are not correlated with one another (R² = 0), so these two parameters may indicate different processes underlying the atmospheric transport errors.

These differences between land and ocean regions may reflect differences in synoptic-scale circulation. Over the oceans, high CV values are clustered in zonal bands, and these clusters often occur at the transition between distinctive synoptic flow patterns. Modeled atmospheric tracer transport is more uncertain in these transition regions – at the transition between southern westerlies and southern trade winds and at the transition between the North Atlantic trade winds and Gulf Stream. Zonal winds over the continents are often more variable than over the oceans (Fig. S17), and atmospheric transport uncertainties do not cluster into the same distinctive, zonal

²⁵ (Fig. S17), and atmospheric transport uncertainties do not cluster into the same distinctive, zonal bands.

The results of this synthetic tracer experiment hold a number of potential applications to top-down CO_2 - CO_2 flux estimation. The danger of obtaining abiased CO_2 -biased CO_2 budget is likely higher in regions with consistent low energy , and limited vertical mixing, and /or high albedo. These biases are unlikely to be represented by most existing inversion uncertainty calculations, as explained in the previous sections. Furthermore, the

. A number of existing studies indicate that uncertainties in PBLH and vertical mixing are closely tied to uncertainties in estimated trace gas transport or in estimated trace gas fluxes (e.g., Stephens et al., 2007; Williams et al., 2011; Miller et al., 2012; Pino et al., 2012; Kretschmer et

- 5 (e.g., Stephens et al., 2007; Williams et al., 2011; Miller et al., 2012; Pino et al., 2012; Kretschm This study further suggests that sustained transport errors due to PBLH are more likely in regions or at times when PBL heights and mixing are consistently low. The meteorological model ensemble is not necessarily more uncertain in these regions (see Fig.S16-17). Note that month-long CO₂ transport biases did not correlate as strongly with meteorology uncertainties.
- ¹⁰ Figs. S15-S16). Rather, the extent to which meteorological uncertainties translate into tracer transport uncertainties appears to depend, at least in part, on the stability and net energy input associated with the boundary layer later.

In summary, boundary layer energy and height explain some of the patterns in the estimated transport errors, but other patterns are associated with uncertainties in synoptic flow and are not

- related to a single meteorological parameter. In fact, over both terrestrial and oceanic regions, individual meteorological parameters only explain a maximum of 29-45% of the variability in the tracer CV. This result stresses the utility of a meteorological model to calculate the variances and covariances in atmospheric transport errors rather than relying on a single, meteorological proxy.
- Note that this study does not account for uncertainties in bottom-up, biogeochemical flux models due to uncertainties in driving meteorological variables. For example, process-based, biogeochemical models of CO₂ typically require estimates of meteorological variables like humidity, temperature, or precipitation to compute the surface fluxes. A number of existing studies have used atmospheric data and/or atmospheric models to explore the meteorological variables that drive CO₂ flux models. For example, Lin et al. (2011) explored how uncertainties
- in flux model drivers affected fluxes estimated for Canadian boreal forests. They found that uncertainties in downward shortwave radiation contributed to the largest uncertainties in the simulated fluxes. Similarly, numerous studies indicate that both air temperature and humidity are drivers of CO_2 fluxes (e.g., Law et al., 2002; Gourdji et al., 2012). These meteorological variables (e.g., downward shortwave radiation, temperature, and specific humidity) correlate with the persistent atmospheric transport uncertainties discussed earlier in this section. A future

Discussion Paper

In this paper, we use two case studies to investigate the potential for bias in top-down

4 Conclusions

- We use CAM-LETKF to explore the characteristics of persistent, covarying atmospheric CO₂ transport errors and the implications of those errors for CO₂ flux estimatesdue to errors in modeled atmospheric CO₂ transportflux estimates. The first case study examines the ability of in situ atmospheric observations to detect bias in estimated explores the relative magnitude of these errors at the monthly time scale and the effects of error covariances. These covariances
 play a critical role in the uncertainties in modeled atmospheric CO₂ fluxes. Among other results, the first case in end to be biased by 20 perturbations in grants by a factor of 5, 20
- ; we find that CT would need to be biased by 29uncertainties increase by a factor of 5–20 at individual observation sites when we include the error covariances in the analysis. These monthly-scale errors correspond to 13-150%, on average, before that bias were detectable above CO_2 transport uncertainties at terrestrial, atmospheric observation sites. These results
- are strongly influenced by temporal correlations in the transport uncertainties. In other words, atmospheric CO_2 of the afternoon CO_2 enhancement, depending on the site in question.

Existing top-down studies often overlook these covariances, and these results imply that atmospheric CO_2 measurements contain less information about the fluxes than is usually assumed by top-down studies that ignore transport error covariances often assumed. As aresult,

²⁵ most existing inversions are likely to result, existing inversions may underestimate the uncertainties in estimated CO_2 - CO_2 fluxes and/or may be vulnerable to unforeseen biases in the estimated fluxes. Accounting for these correlated errors can be as simple as modifying one of the covariance matrix inputs in a Bayesian inversion. Accordingly, this study provides information to improve the setup of future top-down inverse modeling studies – an improvement that will widen the confidence interval on the estimated fluxes. In a subsequent case study, we investigate the factors associated with month-long biases in atmospheric transport. The largest short-term CO₂ CO₂ transport errors correlate strongly with the location of the largest surface fluxes, but month-long biases in atmospheric transport are not only localized to regions with large fluxes. Rather, these biases may be more likely to occur at observation sites that are far from large fluxes and in regions with high atmospheric stability and low net radiation. Over the oceans, biases in atmospheric transport are also associated with weak zonal winds. Existing top-down flux studies may be more likely to estimate inaccurate regional fluxes under those conditions. However, a large fraction of the estimated atmospheric transport errors cannot be described by a single meteorological parameter. This result indicates the utility of a meteorological modeling system, like CAM-LETKF, to estimate errors in atmospheric CO₂ transport. Through this framework, we can better understand the connections between uncertain

atmospheric transport and uncertainties in CO₂ budgets estimated from atmospheric data.

The Supplement related to this article is available online at doi:10.5194/acpd-0-1-2015-supplement.

Acknowledgements. This work was conducted at the Department of Energy's (DOE) Lawrence Berkeley National Laboratory as part of a DOE Computational Science Graduate Fellowship. The research used resources of the National Energy Research Scientific Computing Center, which is supported by the DOE's Office of Science under Contract No. DE-AC02-05CH11231. CarbonTracker CT2011oi CT2011 SUBSCRIPTNBoi results are provided by NOAA ESRL, Boulder, Colorado, USA from the website at http://carbontracker.noaa.gov. We thank Steven Wofsy for his feedback on the manuscript and thank Ed Dlugokencky of NOAA.

25 **References**

20

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO₂ fluxes, 1988–2003, Global Biogeochemical Cycles, 20, doi:10.1029/2004GB002439, 2006.

- Chen, H., Zhou, T., Neale, R. B., Wu, X., and Zhang, G. J.: Performance of the New NCAR CAM3.5 in East Asian Summer Monsoon Simulations: Sensitivity to Modifications of the Convection Scheme,
- Journal of Climate, 23, 3657–3675, doi:10.1175/2010JCLI3022.1, 2010.
 Ciais, P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., and Ramonet, M.: Atmospheric inversions for estimating CO₂ fluxes: methods and perspectives, in: Greenhouse Gas Inventories, edited by Jonas, M., Nahorski, Z., Nilsson, S., and Whiter, T., pp. 69–92, Springer Netherlands, doi:10.1007/978-94-007-1670-4_6, 2011.
- ¹⁰ Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., Chang, P., Doney, S. C., Hack, J. J., Henderson, T. B., Kiehl, J. T., Large, W. G., McKenna, D. S., Santer, B. D., and Smith, R. D.: The Community Climate System Model Version 3 (CCSM3)., Journal of Climate, 19, 2122 – 2143, doi:10.1175/JCLI3761.1, 2006.

Enting, I.: Inverse Problems in Atmospheric Constituent Transport, Cambridge Atmospheric and Space Science Series, Cambridge University Press, Cambridge, 2002.

- Science Series, Cambridge University Press, Cambridge, 2002.
 Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO₂ with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, J. Geophys. Res.: Atmos., 108, doi:10.1029/2003jd003770, 2003.
- 20 Gerbig, C., Körner, S., and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, Atmospheric Chemistry and Physics, 8, 591–602, doi:10.5194/acp-8-591-2008, 2008.
 - Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and
- Michalak, A. M.: North American CO₂ exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, doi:10.5194/bg-9-457-2012, 2012.
 - Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y. H., Cials, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyu-
- tov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C. W.: Towards robust regional estimates of CO₂ sources and sinks using atmospheric transport models, Nature, 415, 626–630, doi:10.1038/415626a, 2002.

- Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J., Sauer, T., Szunyogh, I., Yorke, J. A., and Zimin, A. V.: Four-dimensional ensemble Kalman filtering, Tellus A, 56, 273–277, doi:10.1111/j.1600-0870.2004.00066.x, 2004.
- Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos:
- ⁵ A local ensemble transform Kalman filter, Physica D: Nonlinear Phenomena, 230, 112 126, doi:10.1016/j.physd.2006.11.008, 2007.
 - Huntzinger, D., Post, W., Wei, Y., Michalak, A., West, T., Jacobson, A., Baker, I., Chen, J., Davis, K., Hayes, D., Hoffman, F., Jain, A., Liu, S., McGuire, A., Neilson, R., Potter, C., Poulter, B., Price, D., Raczka, B., Tian, H., Thornton, P., Tomelleri, E., Viovy, N., Xiao, J., Yuan, W., Zeng, N., Zhao, M., and
- 10 Cook, R.: North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison, Ecological Modelling, 232, 144 – 157, doi:10.1016/j.ecolmodel.2012.02.004, 2012.
 - Huntzinger, D. N., Gourdji, S. M., Mueller, K. L., and Michalak, A. M.: The utility of continuous atmospheric measurements for identifying biospheric CO₂ flux variability, Journal of Geophysical
- ¹⁵ Research: Atmospheres, 116, doi:10.1029/2010JD015048, http://dx.doi.org/10.1029/2010JD015048, 2011.
 - Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP–DOE AMIP-II Reanalysis (R-2), Bulletin of the American Meteorological Society, 83, 1631– 1643, doi:10.1175/BAMS-83-11-1631, 2002.
- 20 Kitanidis, P. K.: Introduction to Geostatistics: Applications in Hydrogeology, Stanford-Cambridge program, Cambridge University Press, Cambridge, 1997.
 - Kretschmer, R., Gerbig, C., Karstens, U., and Koch, F.-T.: Error characterization of CO₂ vertical mixing in the atmospheric transport model WRF-VPRM, Atmospheric Chemistry and Physics, 12, 2441–2458, doi:10.5194/acp-12-2441-2012, 2012.
- Kretschmer, R., Gerbig, C., Karstens, U., Biavati, G., Vermeulen, A., Vogel, F., Hammer, S., and Totsche, K. U.: Impact of optimized mixing heights on simulated regional atmospheric transport of CO₂, Atmospheric Chemistry and Physics, 14, 7149–7172, doi:10.5194/acp-14-7149-2014, 2014.
 - Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., and Rayner, P. J.: Structure of the transport uncertainty in mesoscale inversions of CO₂ sources and sinks using ensemble model
- simulations, Biogeosciences, 6, 1089–1102, doi:10.5194/bg-6-1089-2009, 2009.
 - Law, B., Falge, E., Gu, L., Baldocchi, D., Bakwin, P., Berbigier, P., Davis, K., Dolman, A., Falk, M., Fuentes, J., Goldstein, A., Granier, A., Grelle, A., Hollinger, D., Janssens, I., Jarvis, P., Jensen, N., Katul, G., Mahli, Y., Matteucci, G., Meyers, T., Monson, R., Munger, W., Oechel, W., Olson, R.,

Pilegaard, K., U, K. P., Thorgeirsson, H., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation, Agricultural and Forest Meteorology, 113, 97 – 120, doi:10.1016/S0168-1923(02)00104-1, 2002.

- Li, H., Kalnay, E., and Miyoshi, T.: Simultaneous estimation of covariance inflation and observation
- 5 errors within an ensemble Kalman filter, Quarterly Journal of the Royal Meteorological Society, 135, 523–533, doi:10.1002/qj.371, 2009.
 - Lin, J. C. and Gerbig, C.: Accounting for the effect of transport errors on tracer inversions, Geophysical Research Letters, 32, doi:10.1029/2004GL021127, 2005.

Lin, J. C., Pejam, M. R., Chan, E., Wofsy, S. C., Gottlieb, E. W., Margolis, H. A., and McCaughey,

- J. H.: Attributing uncertainties in simulated biospheric carbon fluxes to different error sources, Global Biogeochemical Cycles, 25, doi:10.1029/2010GB003884, 2011.
 - Liu, J., Fung, I., Kalnay, E., and Kang, J.-S.: CO₂ transport uncertainties from the uncertainties in meteorological fields, Geophysical Research Letters, 38, doi:10.1029/2011GL047213, 2011.
- Liu, J., Fung, I., Kalnay, E., Kang, J.-S., Olsen, E. T., and Chen, L.: Simultaneous assimilation of AIRS
 XCO₂ and meteorological observations in a carbon climate model with an ensemble Kalman filter, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2011JD016642, 2012.
 - Masarie, K. A., Pétron, G., Andrews, A., Bruhwiler, L., Conway, T. J., Jacobson, A. R., Miller, J. B., Tans, P. P., Worthy, D. E., and Peters, W.: Impact of CO2 measurement bias on CarbonTracker surface flux estimates, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2011JD016270, 2011.
- ²⁰ Michalak, A., Bruhwiler, L., and Tans, P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, J. Geophys. Res.: Atmos., 109, doi:10.1029/2003JD004422, 2004.
 - Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, Journal of Geophysical Research: Atmospheres, 110, doi:10.1029/2005JD005970, 2005.
- Miller, S. M., Kort, E. A., Hirsch, A. I., Dlugokencky, E. J., Andrews, A. E., Xu, X., Tian, H., Nehrkorn, T., Eluszkiewicz, J., Michalak, A. M., and Wofsy, S. C.: Regional sources of nitrous oxide over the United States: Seasonal variation and spatial distribution, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2011JD016951, 2012.
 - Miyoshi, T.: The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation
- with the Local Ensemble Transform Kalman Filter, Monthly Weather Review, 139, 1519–1535, doi:10.1175/2010MWR3570.1, 2011.
 - Oleson, K. W., Niu, G.-Y., Yang, Z.-L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stöckli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., and Qian, T.: Improvements to the Community Land

Model and their impact on the hydrological cycle, Journal of Geophysical Research: Biogeosciences, 113, doi:10.1029/2007JG000563, 2008.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood,

- E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., and Kassem, K. R.: Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity, BioScience, 51, 933–938, doi:10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2, 2001.
- Parazoo, N. C., Denning, A. S., Kawa, S. R., Pawson, S., and Lokupitiya, R.: CO₂ flux estimation errors associated with moist atmospheric processes, Atmospheric Chemistry and Physics, 12, 6405–6416, doi:10.5194/acp-12-6405-2012, 2012.
 - Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., Werf, G. R. v. d., Randerson, J. T., Wennberg,
- P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proceedings of the National Academy of Sciences, with updates documented at http://carbontracker.noaa.gov, 104, 18 925–18 930, doi:10.1073/pnas.0708986104, 2007.
 - Pino, D., Vilà-Guerau de Arellano, J., Peters, W., Schröter, J., van Heerwaarden, C. C., and Krol, M. C.: A conceptual framework to quantify the influence of convective boundary layer development on car-
- bon dioxide mixing ratios, Atmospheric Chemistry and Physics, 12, 2969–2985, doi:10.5194/acp-12-2969-2012, 2012.
 - Raczka, B. M., Davis, K. J., Huntzinger, D., Neilson, R. P., Poulter, B., Richardson, A. D., Xiao, J., Baker, I., Ciais, P., Keenan, T. F., Law, B., Post, W. M., Ricciuto, D., Schaefer, K., Tian, H., Tomelleri, E.,
- Verbeeck, H., and Viovy, N.: Evaluation of continental carbon cycle simulations with North American flux tower observations, Ecological Monographs, 83, 531–556, doi:10.1890/12-0893.1, 2013.
 - Schuh, A. E., Denning, A. S., Corbin, K. D., Baker, I. T., Uliasz, M., Parazoo, N., Andrews, A. E., and Worthy, D. E. J.: A regional high-resolution carbon flux inversion of North America for 2004, Biogeosciences, 7, 1625–1644, doi:10.5194/bg-7-1625-2010, 2010.
- Sheskin, D.: Handbook of Parametric and Nonparametric Statistical Procedures: Third Edition, Taylor & Francis, Boca Raton, Florida, 2003.
 - Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.:

- Weak Northern and Strong Tropical Land Carbon Uptake from Vertical Profiles of Atmospheric CO₂, Science, 316, 1732–1735, doi:10.1126/science.1137004, 2007.
 - Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, 2005.
 - Williams, I. N., Riley, W. J., Torn, M. S., Berry, J. A., and Biraud, S. C.: Using boundary layer equilibrium to reduce uncertainties in transport models and CO₂ flux inversions, Atmospheric Chemistry and
- to reduce uncertainties in transport models and CO₂ flux inversions, Atmospheric Chemistry an Physics, 11, 9631–9641, doi:10.5194/acp-11-9631-2011, 2011.



Figure 1. Average $\frac{\text{CT-CO}_2 \cdot \text{Carbon Tracker}}{\text{CO}_2}$ fluxes (version 2011oi) for (a) (a) Feburary and (b) (b) July 2009. The fluxes include biosphere, ocean, fossil fuel, and biomass burning fluxes (http://www.esrl. noaa.gov/gmd/ccgg/carbontracker/CT2011_oi).

Discussion Paper

Discussion Paper







Discussion Paper

Discussion Paper

Figure 2. The top panels display average 6-hourly CO_2 6 hourly CO_2 transport uncertainties estimated by CAM-LETKF. The uncertainties (standard deviations 95% confidence intervals) are for the surface model layer for (a) (a) February and (b) (b) July 2009. To create these plots The bottom-panels (c and d), we calculate in contrast, display the ensemble variance at each time step and subsequently average the variances across all time stepsuncertainties in month-long averaged surface CO_2 concentrations. These standard deviations are the square root Note that these meaned 6-hourly variances. Furthermore, these plots include model output from all 24hours h of each day. The supplement Supplement provides analogous figures for daytime- or nighttime-only model output. The bottom-panels (c and d), in contrast, display the standard deviation in month-long averaged surface CO_2 concentrations.

Discussion Paper



Figure 3. Hourly averaged measured CO_2 CO_2 measurements at (a) (a) Moody, Texas, and (b) (b) Argyle, Maine, compared against the CAM–LETKF model ensemble. The Measurements are from the top inlet height at each location. In this figure, the model ensemble represents uncertainties due to atmospheric transport but not other errors (e.g., due to the fluxes, model resolution, etc.).



Figure 4. The uncertainties in monthly-averaged, afternoon atmospheric CO_2 (Sects. 2.4 and 3.3) at a selection of representative, global CO_2 observation sites. Panels (a) and (b) show the results at each site for February and July 2009, respectively. Dark blue bars indicate the difference between the top and bottom of the 95% confidence interval when we include error covariances. The light blue bars indicate the results when we remove these covariances in atmospheric transport errors. Observation sites in the figure include Ryori, Japan (RYO); Ochsenkopf, Germany (OXK); Talk Tower Angus, UK (TTA); East Trout Lake, Saskatchewan, Canada (ETL), Fraserdale, Ontario, Canada (FSD); and West Branch, Iowa, USA (WBI). For more information on these observation sites, refer to Table S1.



Figure 5. Results of the hypothesis test (sections 2.4 and 3.3) at Uncertainty in monthly-averaged afternoon CO_2 concentrations as a selection percentage of global CO_2 observation sites the average afternoon CO_2 boundary layer enhancement. Panel (a) shows This figure places the location, name, and type of each observation site examined in the hypothesis testuncertainties from Fig. Panels 4 (bdark blue bars) and (c) show in context of the test results for February and July 2009. The test asks the following question: how biased would CT fluxes need to be before afternoon CO_2 observation sites would detect that increment from surface fluxes. Larger percentages indicate greater potential for bias above the in monthly CO_2 budgets estimated CO_2 transport uncertainties? from atmospheric data.



Figure 6. The coefficient of variation (CV, unitless) for the monthly-averaged model surface layer. The results plotted here are for the synthetic tracer simulation (sections Sects. 2.5 and 3.4). In that simulation, the synthetic fluxes have a constant spatial distribution. The resulting $CV(\sigma/\mu)$ shows the distribution of month-long, surface-level transport uncertainties independent of the spatial distribution in the fluxes. Note that this plot displays the results from land regions only.



Figure 7. Each panel shows the correlation relationship between the synthetic tracer CV (Fig. 6) and various monthly-averaged meteorological parameters estimated by CAM-LETKF. The top row (a) shows the results for terrestrial regions while the bottom row (b) displays the results for ocean/marine regions. Darker colors in each panel indicate a higher density of points. We test the correlation with 60 different parameters (Table S2S2) and plot the relationships for which $R^2 \ge 0.3$ two parameters that correlate most closely with the tracer CV over terrestrial and marine regions, respectively. In all cases, we fit both a standard major axis regression and nonlinear least squares $(\frac{1}{[\beta_1 \times \text{parameter} + \beta_2]}, \frac{1}{[\beta_1 \times \text{parameter} + \beta_2]})$ and plot the regression with the higher correlation coefficient.