GEOPHYSICAL RESEARCH LETTERS, VOL. 33, L12810, doi:10.1029/2006GL025910, 2006

Using continuous data to estimate clear-sky errors in inversions of satellite CO_2 measurements

Katherine D. Corbin[1] and A. Scott Denning[1]

Received 31 January 2006; revised 28 March 2006; accepted 12 May 2006; published 24 June 2006.

[1] We used continuous measurements of atmospheric CO_2 at two stations to investigate potential errors in inversions of temporal averages of satellite clear-sky column retrievals. Compared to the complete data sets, the mid-day CO_2 on clear days was systematically lower with a larger winter difference. Net ecosystem exchange (NEE) of CO_2 was enhanced on clear vs. all days, the summer boundary layer was deeper, and the CO concentration was systematically lower. During winter these differences cannot account for the CO_2 bias, which must be caused by advection. Summertime errors reflect a tradeoff between deeper mixing and enhanced NEE on clear days. If these sites represent mid-latitude forests and if the CO_2 difference is confined to the bottom 15% column mass, then inversions of temporally-averaged satellite column data products will incur a −0.2 to −0.4 ppm bias. CO_2 concentrations must therefore be assimilated at the place and time observed.

Citation: Corbin, K. D., and A. S. Denning (2006), Using continuous data to estimate clear-sky errors in inversions of satellite CO_2 measurements, *Geophys. Res. Lett.*, *33*, L12810, doi:10.1029/2006GL025910.

1. Introduction

[2] An important method to help quantify the large-scale surface exchanges of carbon is by tracer transport inversion, which uses atmospheric CO_2 concentrations and a transport model to infer information about surface sources and sinks [*Gurney et al.*, 2002; *Rodenbeck et al.*, 2003; *Baker et al.*, 2006]; however, flux estimates are still highly uncertain in many regions due to sparse data coverage [*Gurney et al.*, 2003]. Due to their global spatial sampling and data volume, satellite CO_2 measurements may help improve the inverse modeling constraint, particularly in regions that are poorly sampled by existing ground-based CO_2 monitoring networks. Global simulations with source-sink synthesis inversion models indicate that uncertainties in the atmospheric CO_2 balance could be reduced substantially if data from the existing in situ network were augmented by spatially-resolved, global measurements of the column-integrated dry air mole fraction (X_{CO2}) with precisions of ~1 ppm [*Rayner and O'Brien*, 2001; *Houweling et al.*, 2004].

[3] The Orbiting Carbon Observatory (OCO), scheduled to launch in 2008, is designed specifically to observe X_{CO2} with ~0.3% (1 ppm) precision on regional scales [*Crisp et al.*, 2004]. OCO will fly in a polar, sun-synchronous orbit just ahead of the Earth Observing System (EOS) Aqua Platform with a 13:15 equator crossing time and a 16-day repeat cycle; and it will collect high-resolution spectra of reflected sunlight in the 0.76 μm O_2 A-band and the CO_2 bands at 1.61 μm and 2.06 μm. To maintain an adequate number of soundings even in the presence of patchy clouds, OCO will have a 10 km-wide cross-track field of view that is divided into eight 1.25 km-wide samples with a 2.25 km down-track resolution at nadir.

[4] To obtain near-surface information, retrievals of total column CO_2 concentrations from near-IR spectra measured by space-borne instruments will require clear-sky conditions. Systematic differences in atmospheric CO_2 in clear vs. cloudy conditions might be expected because of the dependence of the photosynthesis rate on the directional character of solar radiation. NEE is strongest on slightly cloudy days due to greater light-use efficiency for diffuse relative to direct beam radiation, which may lead to lower than average CO_2 mixing ratios on partly cloudy days [e.g., *Freedman et al.*, 2001; *Gu et al.*, 2002]. Differences in atmospheric concentrations arising from differences in NEE depend on the spatial scale of the differences in radiative forcing: small-scale cloudy patches are expected to have less effect on concentrations than large-scale perturbations because of horizontal mixing by winds. In winter, since vegetation is not actively photosynthesizing, the a priori expectation is that CO_2 mixing ratios would not depend on cloud conditions. In addition to differences arising from biology, clouds are frequently associated with fronts, changes of air masses and convection with strong vertical motion, so atmospheric transport may be systematically different on clear vs. cloudy days.

[5] Systematic differences in atmospheric CO_2 concentrations between clear and cloudy conditions would introduce sampling errors into tracer transport inversions that use satellite CO_2 products to represent temporal averages. Satellite retrievals of only clear pixels might overestimate spatial or temporal averages of CO_2 because they will not see conditions with enhanced CO_2 uptake. Alternatively, heavy overcast conditions are expected to suppress NEE due to strongly reduced radiation and could lead to systematic underestimation from space-borne measurements during the growing season. Sampling errors could also be caused by advection associated with cloud cover. Depending on the treatment of the observations in the models, this sampling error could potentially introduce a bias; however, if modelers use satellite data at the same time and location and with the same atmospheric situation as the retrievals, these sampling errors would be eliminated. This study investigates clear-sky effects using continuous

---

[1] Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA.

0094-8276/06/2006GL025910$05.00

measurements of CO_2 concentrations from two tower sites over a period of several years.

2. Methods

[6] We analyzed continuous data from two towers: a tall television tower near Park Falls, WI (WLEF 45.95°N, 90.27°W) and the Environmental Monitoring Site at Harvard Forest, located in north-central MA (HF 42.54°N, 72.18°W). The WLEF tower is in a heavily forested zone of low relief, and mixed evergreen and deciduous forests dominate the area surrounding the tower (see *Davis et al.* [1997, 2003] for a description of the site and measurements). The CO_2 concentration is measured at 396 m with two independent Licor CO_2 gas analyzers, which have a mean absolute value difference of 0.25 ppm. To reduce data gaps, we used the average between the two measurements when available and a single analyzer when one had missing data. Photosynthetically active radiation (PAR) is also measured; and the net ecosystem exchange of CO_2 (NEE, defined as the net flux out of the ecosystem) has been computed using eddy covariance methods. The WLEF CO_2 and PAR measurements are available from 1995 through 2003, and we used NEE values from 1997 through 2001. The HF tower is also in a mixed forest that contains oak, maple, hemlock, and spruce (see *Wofsy et al.* [1993] and *Goulden et al.* [1996] for further details). Groups from the Atmospheric Sciences Research Center (ASRC) and Harvard University measure nearly continuous CO_2 concentrations, CO concentrations, PAR, turbulent CO_2 flux at 29 m, and the rate of change in canopy carbon storage below 29 m. All variables are available from 1993 through 2002. We calculated NEE at HF by subtracting the storage measurements from the turbulent CO_2 flux.

[7] We sampled the continuous record of near-surface CO_2 at mid-day corresponding to the OCO planned overpass time. We analyzed two time periods: measurements at 1300 local time and the average value measured from 1100 through 1600 (the mean of six hours). The first represents individual nadir pixels and the second represents the average of retrievals across an atmospheric transport model grid cell, which will be the basis of inversions using satellite CO_2 products. At a mean wind speed of 10 m/s, a six-hour average is equivalent to a 216 km swath of retrievals and comparable to global transport model grid scale. We chose to average six hours each day from 1100 to 1600 LST to avoid rapid variations in concentration associated with the morning and evening transitions between stable and mixed conditions.

[8] Since long-term boundary layer (PBL) depth data is not available at either tower, we analyzed PBL heights from the European Centre for Medium Range Forecasts (ECMWF) 40-year Re-Analysis (ERA-40), which has a six-hour time-step. To capture the daytime PBL depth, we used values at 1800 UTC from the grid cells that included the towers.

[9] We estimated the average difference in CO_2, NEE, CO, and PBL depth between clear and cloudy days by (1) creating clear-sky subsets of the time-series of each variable, (2) fitting separate analytical (harmonic) functions to the clear-sky subset and to the entire time-series, and (3) subtracting the two analytical functions to obtain a seasonal climatology of the clear-sky minus all-sky difference in each variable. Clear-sky subsets were defined by selecting the mid-day values of each variable for days on which measured PAR was greater than a threshold value defined by month for each site. The threshold PAR values were set by ranking measurements from all years at each site, then selecting the value corresponding to the percentage of clear days for each month at the nearest city recorded by the National Climatic Data Center (NCDC). The NCDC monthly climatology of clear days is based on at least 40 years of data and is determined by human observers who categorize each daytime hour as clear if the average cloud cover was less than 30%. For WLEF, the nearest stations in the NCDC database are Green Bay, WI (232 km away); Duluth, MN (175 km); and Minneapolis/St. Paul, MN (262 km). We used an average of the monthly clear-sky days from all three stations. At HF, the closest station is Worcester, MA (45 km). Since the NCDC clear-sky criteria is likely less stringent than satellite requirements, we decreased the reported percentages of clear days by 5% to ensure that the clear-sky differences are not overestimated by including partly cloudy days with enhanced NEE that will not be captured by satellites. The PBL depth clear-sky subsets included the same days as the clear-sky CO_2 subsets.

[10] We separately fit seasonally-varying harmonic functions of each variable to the entire time-series and to the clear-sky subset using a linear least squares method. We removed data for February 29, de-trended the CO_2 concentration, and required both the variable being investigated and the PAR measurement to be valid at each hour. We found that two harmonics per year fit seasonal variations adequately, without introducing spurious noise. Differences between the harmonic fits to the clear-sky subsets and to the corresponding complete data sets are presented below and interpreted as the seasonal sampling error expected to occur in an average year by a satellite which only observes the atmosphere in clear conditions.

3. Results

[11] Sampling the CO_2 concentration only on clear days resulted in underestimation of the mean concentrations at both towers at all times of year (Figure 1). The seasonal cycle of the sampling error is similar for all cases, with a greater near-surface difference in winter than during the summer months. At WLEF, the mean winter bias is -1.5 ppm and the mean summer bias is -0.8 ppm; and at HF the mean biases for winter and summer are -3.2 ppm and -1.5 ppm, respectively. The biases at the WLEF tower are smaller than at HF, which could reflect differences in vegetation or transport.

[12] To explain the clear-sky CO_2 bias, we analyzed the clear-sky NEE bias (Figure 2). Both towers have a large negative summer bias due to increased photosynthesis on clear days and negligible to slightly positive differences in the winter. Meteorological factors such as increased temperature and water stress may contribute to the changes in magnitude and timing. We investigated the clear-sky temperature bias and found that the HF temperatures are greater on clear days than on average and that the summertime temperature bias is ~0.4°C greater at 1300 than from 1100–

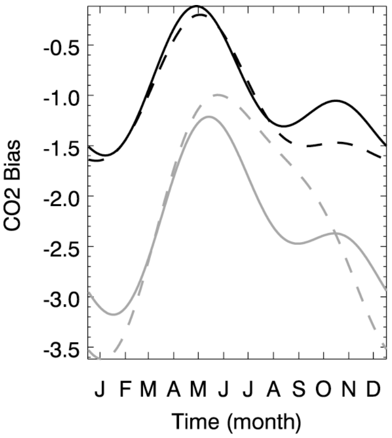

Figure 1. The clear-sky CO_2 sampling bias at WLEF (black) and at HF (gray), in ppm. Solid lines depict the 1100–1600 bias and dashed lines show the 1300 bias.

1600. The increased temperatures at 1300 could lead to increased respiration and decreased NEE, and these higher temperatures combined with the low solar zenith angle at 1300 may increase the water vapor pressure deficit, causing more stress on the vegetation and less CO_2 uptake.

[13] Since the surface CO_2 concentration is dependent on both the surface fluxes and vertical mixing, we analyzed the PBL depth clear-sky bias (Figure 3). The clear-sky bias is positive in the summertime at both towers, with the PBL ~200 m deeper on clear days than on average. During the winter the magnitude of the bias is smaller and the PBL is slightly shallower on clear days.

[14] We estimated the expected CO_2 bias from the mean differences in NEE and PBL height using a simple box model. We calculate a summer and winter estimate of the clear-sky effect on mixed-layer CO_2 concentration at both towers as

$$\Delta C = \Delta\left(\frac{\mathrm{NEE}\Delta t}{\rho z_i / M_{air}}\right), \qquad (1)$$

where ρ is the mean density of the mixed layer, z_i is the mean depth of the daytime mixed layer, M_{air} is the molecular weight of dry air, and Δt = 10 hours is the duration over which the NEE difference was assumed to act. The NEE and z_i values and the resulting biases are summarized in Table 1. Although the box model is sensitive to the parameters used, it indicates that the summertime CO_2 bias is weak at the towers because the lower concentrations from enhanced photosynthesis are mixed into a deeper boundary layer, diluting the effect of the larger flux on clear days. In winter, the CO_2 biases are also weak, which is not surprising since the PBL depth and NEE are nearly the same on clear days as they are on average. The box model suggests that the large winter CO_2 bias observed at both towers cannot be explained by differences in surface fluxes or vertical mixing, but instead likely results from non-local processes such as advection.

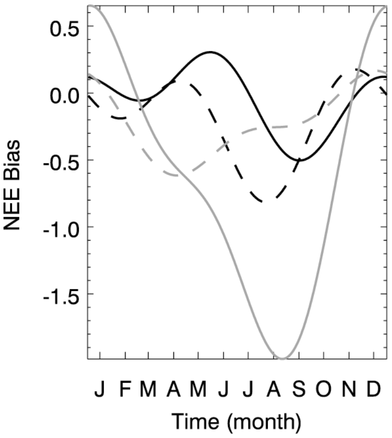

Figure 2. The NEE clear-sky bias, in μmol m^{-2} s^{-1}.

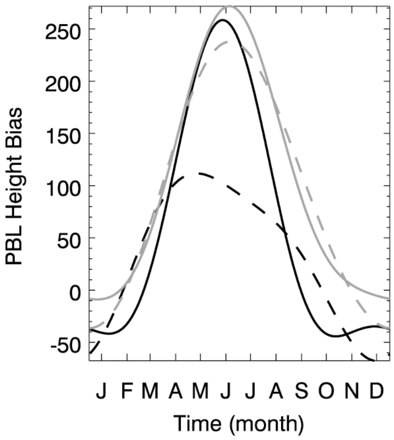

Figure 3. PBL depth clear-sky bias, in m.

[15] Finally, we calculated the bias in CO concentrations at HF. Since CO is a ubiquitous by-product of the same combustion processes as CO_2 and has an average lifetime of only 3 months, CO measurements can provide information on the intensity of anthropogenic activities [*Palmer et al.*, 2003; *Bakwin et al.*, 2004; *Suntharalingam et al.*, 2004]. The CO bias has a similar seasonal cycle to the CO_2 difference; and in both seasons the CO concentration is lower on clear days, indicating that the fossil fuel contribution is less. The mean clear-sky bias is −44 ppb and −18 ppb for January and July, respectively. Assuming that the primary source of CO is fossil fuel combustion and that the anthropogenic fluxes in the immediate vicinity of HF are negligible, the CO results indicate that part of the CO_2 bias is due to less advection of anthropogenic CO_2 on clear days. Using an average combustion efficiency of 95% [*Miller et*

Table 1. NEE and PBL Height (z_i) Values Used in the Box Model and the Resulting CO_2 Biases

	January/July		
	NEE, μmol/m^2/s	z_i, m	Bias, ppm
		WLEF	
Clear	1./−8.	650/2100	0.1/−0.1
Total	1./−7.	700/1900	
		HF	
Clear	1.4/−13.5	925/2000	0.3/0.1
Total	1./−12.	950/1750	

al., 2003; *Bakwin et al.*, 2004], the CO_2 bias resulting from reduced fossil fuel contributions on clear-sky days is ~−0.5 ppm in the summer and ~−1.2 ppm in the winter, or less than half the observed CO_2 difference at this site.

4. Conclusions

[16] This study indicates that sampling only in clear conditions leads to a systematic underestimation of the mean CO_2 concentration at both WLEF and HF. In summer, the mean clear-sky bias in mixed-layer CO_2 is ~−1.5 ppm at HF and −0.8 ppm at WLEF. A simple box model suggests that enhanced photosynthesis on clear days may be offset by a deeper boundary layer, mitigating some of the difference. During the winter, the clear-sky effects on NEE and boundary layer depth is weak, and the large observed CO_2 difference (~ −3 ppm at HF and −1.5 ppm at WLEF) cannot be explained in terms of local forcing. Seasonal patterns of mid-day differences in CO concentration at HF are similar to those of CO_2, with a greater difference in winter than summer, but are not sufficient to explain the CO_2 difference. Much of the clear-sky sampling error in CO_2 at these sites may be attributed to differential advection on clear vs. average days. Satellite retrievals of total column CO_2 concentrations are expected to be less affected by clear-sky sampling error than mixed-layer measurements. If these two sites are broadly representative of mid-latitude forested regions and if the CO_2 difference is confined to a PBL occupying 15% of the column mass, then inversions of temporally-averaged satellite column data products will incur a −0.2 to −0.4 ppm bias. Therefore, satellite total-column CO_2 retrievals must be assimilated at the time and location of the observations.

[17] **Acknowledgments.** We thank Steven Wofsy for the Harvard Forest data and Kenneth Davis for the data from the WLEF tower. This research was funded by NASA grant NCC5-621 and by NASA Earth System Science fellowship 53-1970. We gratefully acknowledge the constructive comments by two anonymous reviewers, which improved the quality of the manuscript.

References

Baker, D. F., et al. (2006), TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, *Global Biogeochem. Cycles*, *20*, GB1002, doi:10.1029/2004GB002439.

Bakwin, P. S., K. J. Davis, C. Yi, S. C. Wofsy, J. W. Munger, L. Haszpra, and Z. Barcza (2004), Regional carbon dioxide fluxes from mixing ratio data, *Tellus, Ser. B*, *56*, 301–311.

Crisp, D., et al. (2004), The Orbiting Carbon Observatory (OCO) mission, *Adv. Space Res.*, *34*, 700–709.

Davis, K. J., D. H. Lenschow, S. P. Oncley, C. Kiemle, G. Ehret, A. Giez, and J. Mann (1997), Role of entrainment in surface-atmosphere interactions over the boreal forest, *J. Geophys. Res.*, *102*, 29,219–29,230.

Davis, K. J., P. S. Bakwin, C. Yi, B. W. Berger, C. Zhao, R. M. Teclaw, and J. G. Isebrands (2003), The annual cycle of CO_2 and H_2O exchange over a northern mixed forest as observed from a very tall tower, *Global Change Biol.*, *9*, 1278–1293.

Freedman, J. M., D. R. Fitzjarrald, K. E. Moore, and R. K. Sakai (2001), Boundary layer clouds and vegetation-atmosphere feedbacks, *J. Clim.*, *14*, 180–197.

Goulden, M. L., J. M. Munger, S. M. Fan, B. C. Daube, and S. C. Wofsy (1996), Exchange of carbon dioxide by a deciduous forest: Response to interannual climate variability, *Science*, *271*(5255), 1576–1578.

Gurney, K. R., et al. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, *415*, 626–630.

Gurney, K. R., et al. (2003), TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus, Ser. B*, *55*, 555–579.

Gu, L., D. Baldocchi, S. B. Verma, T. A. Black, T. Vesala, E. M. Falge, and P. R. Dowty (2002), Advantages of diffuse radiation for terrestrial ecosystem productivity, *J. Geophys. Res.*, *107*(D6), 4050, doi:10.1029/2001JD001242.

Houweling, S., F. M. Breon, I. Aben, C. Rödenbeck, M. Gloor, M. Heimann, and P. Ciais (2004), Inverse modeling of CO_2 sources and sinks using satellite data: A synthetic inter-comparison of measurement techniques and their performance as a function of space and time, *Atmos. Chem. Phys.*, *4*, 523–538.

Miller, J. B., P. P. Tans, J. W. C. White, and T. J. Conway (2003), The atmospheric signal of terrestrial carbon isotopic discrimination and its implication for partitioning carbon fluxes, *Tellus, Ser. B*, *55*, 197–206.

Palmer, P. I., D. J. Jacob, D. B. A. Jones, C. L. Heald, R. M. Yantosca, J. A. Logan, G. W. Sachse, and D. G. Streets (2003), Inverting for emissions of carbon monoxide from Asia using aircraft observations over the western Pacific, *J. Geophys. Res.*, *108*(D21), 8828, doi:10.1029/2003JD003397.

Rayner, P. J., and D. M. O'Brien (2001), The utility of remotely sensed CO_2 concentration data in surface source inversions, *Geophys. Res. Lett.*, *28*(1), 175–178.

Rodenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003), CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem. Phys.*, *3*, 1919–1964.

Suntharalingam, P., D. J. Jacob, P. I. Palmer, J. A. Logan, R. M. Yantosca, Y. Xiao, M. J. Evans, D. G. Streets, S. L. Vay, and G. W. Sachse (2004), Improved quantification of Chinese carbon fluxes using CO_2/CO correlations in Asian outflow, *J. Geophys. Res.*, *109*, D18S18, doi:10.1029/2003JD004362.

Wofsy, S. C., M. L. Goulden, J. W. Munger, S. M. Fan, P. S. Bakwin, B. C. Daube, S. L. Bassow, and F. A. Bazzaz (1993), Net exchange of CO_2 in a midlatitude forest, *Science*, *260*(5112), 1314–1317.

K. D. Corbin and A. S. Denning, Department of Atmospheric Science, Colorado State University, 1371 Campus Delivery, Fort Collins, CO 80523-1371, USA. (kdcorbin@atmos.colostate.edu)