On The Effect of Spatial Variability and Support on Validation of Remote Sensing Observations of CO₂

Jovan M. Tadić¹ and Anna M. Michalak¹

[1]{Department of Global Ecology, Carnegie Institution for Science, Stanford, CA 94305, USA}

Abstract

Validation of ground-based and satellite remote sensing CO₂ observations involves comparisons among platforms and with in situ airborne measurements. Several factors unrelated to observational errors can lead to mismatches between measurements, and must be assessed to avoid misinterpreting actual differences in observed values as errors. Here we explore the impact of CO₂ horizontal variability and differences in the spatial support of measurements. Case studies based on flights over Walnut Grove and Petaluma, California, are used to compare hypothetical airborne, TCCON, GOSAT, and OCO-2 measurements. We find that high CO₂ variability can lead to differences in inferred XCO₂ (1) of over 0.5 ppm between airborne and remote sensing observations, due to the spatial mismatch between spiral flight trajectories and atmospheric columns, and (2) of up to 0.3 ppm among remote sensing platforms, due to differences in the spatial support of observations. Horizontal CO₂ variability must therefore be considered in intercomparisons aimed at validation of remote sensing observations.

Keywords: column-integrated atmospheric CO₂ mole fraction, validation, TCCON, GOSAT, OCO-2, remote sensing

* Corresponding author at: Department of Global Ecology, Carnegie Institution for Science,
Stanford, CA 94305, USA.
E-mail address: jtadic@carnegiescience.edu (J. Tadić).

1. Introduction

Tying column measurements to WMO standards is required for ensuring the comparability of remote sensing measurements of atmospheric CO₂ from ground-based and space-based platforms (i.e. Messerschmidt et al., 2011). The most common way to achieve this is through comparisons with a series of calibrated airborne *in situ* instruments that measure atmospheric profiles, flown coincidentally with remote sensing measurements. This approach has been successfully deployed to validate column-integrated atmospheric CO₂ mole fractions (XCO₂) from ground-based TCCON measurements (Washenfelder et al., 2006; Wunch et al., 2010; Messerschmidt et al., 2011; Deutscher et al., 2010; Tanaka et al., 2012), and measurements derived from SWIR spectra of the GOSAT TANSO-FTS (Miyamoto et al., 2013; Tadić et al., 2014; Tanaka et al., 2012, Inoue et al., 2013), among others.

Due to the inherent maneuverability constraints of aircraft, vertical transects through the atmosphere are usually approximated by irregularly shaped trajectories, often spirals (Messerschmidt et al., 2011; Tadić et al., 2014; Deutscher et al., 2010), and thereby inevitably incorporate the impact of horizontal variability in CO₂ concentrations into the reported vertical profiles. The resulting profile represents a composite of individual measurements collected at increasing or decreasing heights, regardless of the horizontal coordinates of individual points in the trajectory. In previous remote sensing validation studies (e.g., Tadić et al., 2014), details

regarding the exact flight patterns used to approximate vertical profiles are often omitted, and the impact of horizontal CO₂ variability on the inferred vertical profile is not assessed.

Evidence of such horizontal variability has been documented both within and outside the planetary boundary layer (PBL). For example, a multi-year record of airborne CO₂ observations in the US Southern Great Plains (Biraud et al., 2013) provides extensive evidence of horizontal CO₂ variability, including gradients of up to 10 ppm within the PBL on 5 to 10 minute level flight legs, which, at aircraft speeds of up to 100 m/s, translates to distances of 20-40 km, and gradients of 0.5 ppm/km within the PBL. Small-scale variability within these 20-40 km transects may be even higher. Biraud et al. (2013) even stated that the existence of horizontal variability prevented a direct comparison to coincident, continuous cavity ring-down spectroscopy and discrete flask measurements due to the measurement time mismatch between platforms.

Conversely, the major source of uncertainty typically reported in remote sensing validation studies is the uncertainty resulting from the extrapolation of profiles above and below the sampled fraction of the full atmospheric column, due to constraints in instrument and aircraft flight capabilities (e.g., Tadić et al., 2014) or to constraints resulting from air-traffic control issues (e.g., Tanaka et al., 2012). This uncertainty has been estimated as being below 1 ppm (Miyamoto et al., 2013).

Another confounding factor is the fact that *in situ* airborne measurements and remote sensing observations have substantially different spatial supports, i.e., represent an average quantity over different volumes. For example, *in situ* aircraft sampling yields a series of point measurements (both vertically and horizontally), TCCON reports column-integrated observations at a point location (Wunch et al., 2011), and observations from space-based platforms not only represent a vertical integration but also an average over some horizontal footprint. The size of such

footprints can be of the same order of magnitude as distances over which significant horizontal gradients occur. For example, GOSAT has a circular nadir footprint of about 10.5 km diameter at sea level (Kuze et al., 2009) and OCO-2 has a footprint size of $2.25 \text{ km} \times 0.1\text{-}1.3 \text{ km}$ (Crisp et al, 2007; Crisp et al, 2008). The diameters of spiral-shaped aircraft profiles used for validation are also of the same order of magnitude. In Tadić et al. (2014), for example, a typical radius was 3 to 5 km, while radii of $\sim 3\text{-}8 \text{ km}$ (Messerschmidt et al., 2011), and $\sim 15 \text{ km}$ (Deutscher et al., 2010), have been reconstructed from published flight trajectory data for other studies. The differences in spatial support therefore raise further questions about direct comparisons across platforms.

An empirical analysis of the spatial variability of atmospheric CO_2 and its implications for inverse analyses and space-borne sensors was given in Lin et al. (2004). Authors concluded that the horizontal CO_2 variability causes deviations between values at a point location and spatial averages measured by space-borne sensors or represented in models. However, they quantified only representation errors resulting from mismatch between point measurements, or total column CO_2 measurements using expected OCO satellite with relatively small footprint, and spatial averages at the scale of grid cell size of 200-400 km.

In this paper we explore whether the impact of horizontal variability in CO_2 concentration and of differences in the spatial support of observations can be sufficiently large as to hinder the intercomparability of observations from *in situ* airborne, ground-and two popular space-based CO_2 remote-sensing platforms, GOSAT and OCO-2. We do so by creating realistic representations of 3D CO_2 fields for two sites having particularly high horizontal CO_2 variability (Walnut Grove and Petaluma, California), to capture an upper bound of the impact of horizontal variability and differences in spatial support. The 3D fields are sampled in a manner consistent

with *in situ* airborne, TCCON, GOSAT and OCO-2 observations, and these synthetic observations are compared to assess the expected magnitude of the differences between observations, resulting purely from horizontal variability and the differences in the spatial support of the observations.

2. Airborne observations

Two flights starting from NASA Ames Research Center with especially high encountered horizontal CO₂ variability were selected from nearly 100 analyzed flights, and used as a basis for the 3D CO₂ fields. The first one was flown to Walnut Grove, California, on July 20th 2011, and the second one to Petaluma, California, on February 4th, 2013. These flights represented a qualitative upper bound on the expected impacts of horizontal variability and differences in the spatial support of observations. Both flights were flown at near-uniform altitude, and thus did not capture vertical information. We therefore use information from two other flights, to Railroad Valley, Nevada, (conducted on June 25th and 26th, 2011) to represent realistic, sampled vertical variability.

In this way two complementary datasets - one with substantial horizontal information and one with substantial vertical information - were used to generate realistic hypothetical 3D CO₂ fields subsequently used to assess the magnitude of errors that are possible due to horizontal CO₂ variability, as detailed in Section 3. This approach is reasonable, because vertical and horizontal direction represents two orthogonal spatial subdomains, and variability of the CO₂ field in one of those does not necessarily imply any structure in the other one. The field created in this manner are not meant to represent “real” CO₂ variability at a given site and time; they instead exemplify a realistic representation of a possible atmospheric structure.

Measurements for all flights were taken by a tactical fighter aircraft (Alpha Jet) equipped with GPS and inertial navigation systems that provide position information, altitude and temperature. The aircraft carries a reconfigured Picarro 2301-m cavity ring-down instrument for CO₂ and other measurements. Additional details are provided in Tadić et al. (2014).

All calculations were conducted in a Euclidean coordinate system, while GPS instruments recorded the data in Lat/Long/Alt format based on WGS84 spheroid earth model (Carlson and Clay, 1999). The coordinates were converted into the Universal Transverse Mercator coordinate system, so that distances and angles could be computed using Euclidean geometry over short distances (Li and Heap, 2008).

3. Methodology

In this work, we quantify mismatches between XCO₂ values reported by different sampling platforms resulting from horizontal CO₂ variability and differences in the spatial support of observations. We do so by sampling a realistic 3D CO₂ field in a manner consistent with real airborne profiles, TCCON, GOSAT and OCO-2. Thus, the work is composed of two steps: a) generating the realistic 3D CO₂ fields, and b) sampling the fields in a manner consistent with different observation platforms.

3.1 Conditional realizations on horizontal planes

Conditional realizations, a.k.a. spatially-consistent Monte Carlo simulations, of CO₂ variability were generated on a horizontal plane for the two sites described in Section 2. These planes were at 578 m elevation for the Walnut Grove site, and 300 m elevation for the Petaluma site, to coincide with the average height of the flights (Figure 1). The planes were 28 km and 20

km square for the two sites, respectively. Conditional simulations represent equally probable realizations of a spatial random function, where each realization honors observed values at measurement locations, i.e. is ‘conditioned’ on the data (e.g., Chiles and Delfiner, 1999). Conditional realizations are thus consistent both with observed values and with the degree of variability expected of the true (unknown) distribution, where the degree of variability expected of the true distribution is assessed through a variogram analysis of the available observations and the variogram is converted to a covariance function of the form listed in eq. 4 below. In this way they differ from kriging, which, by design, provides point-wise “best estimates,” but that are spatially smoother than reality. The average across conditional realizations asymptotically approaches the kriging estimate as the number of realizations tends to infinity; while individually the conditional realizations provide an accurate representation of the degree of variability in the unknown “true” field.

An ensemble of 1000 conditional simulations was generated for each site, based on observations from the Walnut Grove (Figure 1a) and Petaluma (Figure 1b) flights. Each realization (s_{ci} , $m \times 1$) was calculated as (for example, see Zhou et al., 2013):

$$s_{ci} = \Lambda(\mathbf{z} - \mathbf{z}_{ui}) + \mathbf{s}_{ui} \quad (1)$$

where Λ is the $m \times n$ matrix of weights defined in eq 5, $\mathbf{z}$ ($n \times 1$) are the observations, and $\mathbf{z}_{ui}$ ($n \times 1$) and $\mathbf{s}_{ui}$ ($m \times 1$) are unconditional realizations at measurement (n) and estimation (m) locations, respectively, obtained from:

$$\mathbf{z}_{ui} = \mathbf{C}^T \mathbf{u} \quad (2)$$

where $\mathbf{u}$ is an $(n + m) \times 1$ vector of normally distributed random values with zero mean and unit variance (a new vector $\mathbf{u}$ is generated for each realization), and $\mathbf{C}$ is the $(n+m) \times (n+m)$ matrix resulting from the Cholesky decomposition of the covariance matrix:

$$\mathbf{Q}_{zz}\mathbf{Q}_{zs}\mathbf{Q}_{sz}\mathbf{Q}_{ss}=\mathbf{C}\mathbf{C}^T \quad (3)$$

where the components of $\mathbf{Q}$ are the covariance matrices between the measurements ($\mathbf{z}$) and estimation ($\mathbf{s}$) locations. All components of $\mathbf{Q}$ are defined as:

$$Q_{ijh}=\sigma^2+\sigma_Q^2, \text{ \&for } h_{ij}=0 \sigma^2 \exp(-h_{ij}/l), \text{ \&for } h_{ij}>0 \quad (4)$$

where σ^2 is the variance of the portion of the residual CO_2 variability that is spatially correlated, $3l$ is the practical correlation range, and σ_Q^2 is the variance of the portion of the variability that is not spatially correlated. The three parameters were obtained by fitting theoretical variogram to the experimental variograms of the measurements using standard procedures (e.g., Chiles and Delfiner, 1999).

The vector $\mathbf{\Lambda}$ used in eq 1, and $\mathbf{M}$, a matrix of Lagrange multipliers are obtained by solving the universal kriging system of linear equations:

$$\mathbf{Q}_{zz}\mathbf{X}_z\mathbf{X}_z^T\mathbf{\Lambda}+\mathbf{M}=\mathbf{Q}_{zs}\mathbf{X}_s^T \quad (5)$$

where $\mathbf{X}_z$ (and $\mathbf{X}_s$, analogously) are defined by:

$$\mathbf{X}_z=\begin{bmatrix} 1 & H_1 & \dots & 1 & H_n \end{bmatrix} \quad (6)$$

where $\mathbf{X}_z$ is an $n \times 2$ matrix of auxiliary variables, representing the spatial trend. In this study, the model of the trend included only an intercept and a trend with height (H). Additional details on the generation of conditional realizations as implemented here can be found in e.g. Zhou et al. (2013). Examples of the generated conditional realizations are shown in Figure 1c, d.

3.2 Modeling vertical variability

The Railroad Valley profiles (see Section 2) were used to model the vertical variability. To do so, these profiles were spatially superimposed on the horizontal planes by matching the altitude of the horizontal plane (578 m for Walnut Grove, 300 m for Petaluma) to the corresponding point in the profile, shifting all values in the profile by a constant such that the value at 578 m (Walnut Grove) / 300 m (Petaluma) matched the conditionally simulated concentration at the central grid cell of the horizontal plane, and finally using this adjusted vertical profile to represent the vertical variability at the central grid cell of the horizontal plane (Figure 2).

The conditional realizations of the CO₂ variability were then extended vertically down to the ground level and up to a height of 7200 m based on the measured vertical CO₂ profiles taken from the two flights over Railroad Valley while also incrementally suppressing horizontal variability with increasing height (equation 1). The profiles were further extrapolated to the top of the atmosphere according to Method 3 described in Tadić et al. (2014). The profiles observed by the different platforms converge by 7200 m (Fig 3c,d), such that the selection of the extrapolation method above that height would not play a role in assessing differences between platforms. The vertical extensions were done assuming that CO₂ concentrations converge with height to reach the reported 1 ppm of horizontal gradient at 5500 m observed in Biraud et al. (2013). The exponential convergence function was used:

$$C(x, y, H) = C_p(x_p, y_p, H) - C_p(x_p, y_p, H_{ref}) - C_{sci} \times e^{-abs(H - H_{ref})/K} \quad (7)$$

where C represents the concentration at a given location (x,y) and height (H) within the 3D CO₂ field, C_p is the concentration of the adjusted vertical profile at the same height H , or at the

height of the horizontal reference grid H_{ref} (578m for Walnut Grove, 300m for Petaluma), C_{sci} is the concentration of the conditional realization at location (x,y) , and K is a constant selected to impose the desired vertical convergence rate (2137.6 m in our cases).

3.3. Sampling the 3D CO₂ fields

A simple thought experiment was conducted to reveal the impact of horizontal variability and differences in spatial supports of the coincident airborne, TCCON, GOSAT, and OCO-2 observations. It was assumed that there were TCCON stations at the geometric centers of the flights (Fig 1a and b) near Walnut Grove and near Petaluma, and that satellites soundings with a circular nadir footprint of about ~11 km (as it is for GOSAT) and a rectangular nadir footprint of about ~1.3×2.25 km (as OCO-2) were collocated with the sites of the fictional TCCON stations. All remote sensing platforms were assumed to sample a vertical column of air, although accounting for deviations from a vertical profile would further increase the impact of horizontal variability. Two real aircraft trajectories, corresponding to adjusted profiles (See Section 3.2), were also taken from our previous study (Tadić et al., 2014), centered around the hypothetical TCCON locations, and used to simulate airborne measurements. Figure 3a and 3b show the hypothetical observation locations and the spatial supports for the airborne, TCCON, and satellite measurements for both sites. Column average values obtained from simulated profiles were calculated by pressure-weighting the simulated concentrations, a methodology described in details in Tadić et al. (2014). For the GOSAT and OCO-2 observations, the XCO₂ values were then averaged within the footprints of the soundings. A simple pressure-weighting was implemented for deriving column observations rather than using instrument-specific averaging kernels, because the emphasis here was to examine the profiles and columns that the different

instruments encounter, rather than what they report, due to horizontal CO₂ variability and differences in spatial support. Implementing instrument-specific averaging kernel would have instead yielded differences that were representative not only of these two factors, the influence of which we wish to isolate here, but also of differences in averaging kernels. In other words, the impact of differences in averaging kernels would represent another, but distinct, source of differences among observations reported by different instruments in a real measurement campaign. Within the defined scope of this study, the differences between platforms are driven exclusively by horizontal CO₂ variability.

3.4 Limitations

While both the horizontal and vertical variability used to create the 3D CO₂ fields were taken from actual flights, the suppression of the horizontal variability with height implemented here is not fully realistic. For example, expected discontinuities at the boundary between the PBL and the free troposphere are not present, and the assumption that horizontal variability decreases with height at the same rate throughout the examined domain is also questionable.

In addition, the sensitivity of the satellite observations was assumed to be uniform within the sounding footprint, whereas the point spread function is more complex in reality. A further simplification is that the analysis assumes that satellite soundings reflect a perfectly vertical profile, which is not the case. In reality, the fact that footprints are not always perfectly vertical allows additional horizontal variability to be captured by such footprints, which suggests that differences between platforms may be even higher.

Overall, these limitations would tend to decrease the estimated impact of horizontal variability and differences in spatial support, and the results presented here are therefore conservative estimates of errors expected in reality under conditions observed at the two sites.

4. Results and Discussion

4.1 Observed horizontal variability

A CO₂ gradient of 8 ppm was encountered over a 7 km flight path above Walnut Grove, CA at ~580 m altitude (Figure 1a), which translates into a ~1 ppm/km horizontal gradient. Over Petaluma, CA, at 300 m altitude, an even sharper gradient of ~9 ppm was observed over a 7 km distance (Figure 1b). Both flights took place well within the PBL and show that the magnitude of horizontal variability can be considerable, and could potentially impact validation of remote sensing platforms using aircraft measurements.

4.2 Effects of horizontal variability and differences in spatial supports

Examples of the vertical profiles encountered by aircraft, TCCON, GOSAT, and OCO-2 sampling within the 3D CO₂ field from a specific conditional realization at each site are presented in Figure 3c,d, together with the corresponding value of XCO₂. The difference between the vertical profiles along the aircraft flight path, at the location of a TCCON site, and within the footprints of GOSAT and OCO-2 are most pronounced within the PBL, where the magnitude of the horizontal variability is the highest. Differences between the profiles exceed 1 ppm at some heights.

The difference between XCO_2 derived from the airborne *in-situ* CO_2 profile and the “true” XCO_2 at that spatial location (as represented here by XCO_2 at the location of the hypothetical TCCON site) has two potential causes. The first is a difference between the XCO_2 at the TCCON site and the average XCO_2 at radius R around the TCCON site, where R is the radius of the spiral flight trajectory. The second is the inability of the aircraft to perfectly capture the average XCO_2 at radius R , which results from the vertical lag between individual loops in the spiral flight trajectory. To distinguish between these two potential causes, we calculated the average XCO_2 at radius R directly, but the mismatch between the XCO_2 values was not reduced. This result indicates that the mismatch is due primarily to the difference between the actual XCO_2 at the TCCON location versus at a radius R around it.

In the Walnut Grove case, the mean XCO_2 averaged across the XCO_2 from 1000 individual conditional realizations, was 388.76, 388.92, 388.92, and 388.98 ppm for aircraft, TCCON, aircraft, OCO-2 and GOSAT, respectively, showing no substantial systematic bias across platforms. In the Petaluma case, on the other hand, the derived column average values were 396.07, 395.53, 395.57, and 395.76 ppm, for aircraft, TCCON, OCO-2 and GOSAT, respectively, showing a consistent over-estimate from the airborne observations as the simulated flight path preferentially sampled an area with higher CO_2 concentrations relative to the centerpoint of the spiral (as represented by TCCON). The difference between XCO_2 inferred from the airborne observations and TCCON was 0.54 ± 0.39 ppm for the 1000 realizations. The differences between the true XCO_2 at the centerpoint of the domain (represented here by the hypothetical TCCON observation) and each of the other platforms are summarized in Figure 3e,f across the 1000 realizations for both sites.

The mean absolute difference between pairs of XCO_2 values estimated using all of the examined platforms (Table 1) gives an indication of the expected differences in XCO_2 due entirely to horizontal variability and differences in support when comparing XCO_2 as estimated using various platforms. These numbers are representative of conditions observed during the flights used to create the 3D CO_2 field used in this analysis, and generally represent an upper range in errors as the flights used here were selected specifically because of the observed high level of horizontal variability.

The largest differences occur between XCO_2 inferred from airborne observations and the other platforms for both sites. Comparing airborne observations to TCCON gives the clearest indication of the impact of horizontal variability (0.39 ppm for Walnut Grove; 0.57 ppm for Petaluma). This expected difference from airborne observations is actually lower for OCO-2 and especially for GOSAT, which represent average XCO_2 over a larger area, thereby reducing the impact of horizontal variability. In other words, the larger spatial support of OCO-2 and GOSAT partially offset the impact of horizontal variability on intercomparisons with airborne observations.

Differences among the remote sensing platforms are smaller than those with airborne observations, and are representative of the impact of spatial support in the presence of horizontal variability. As expected, TCCON and OCO-2 are more comparable to each other than either is with GOSAT, as the spatial support of GOSAT observations is substantially larger. The biggest observed difference is between TCCON and GOSAT for the Petaluma case (0.30 ppm).

Results indicate that even for co-located observations, the impact of horizontal variability and differences in support can exceed 0.5 ppm for XCO_2 . Although this number is below the single shot accuracy of instruments such as OCO-2 and GOSAT, it is nevertheless substantial

given typical gradients in XCO₂ on regional scales (Miller et al., 2007), and the effect of horizontal variability on the uncertainty of intercomparisons must therefore be taken into account. Minimizing the radius of the spiral airborne trajectories is thus one strategy for decreasing the impact of horizontal variability on validation studies. Differences due to horizontal variability and spatial support between GOSAT and OCO-2 are relatively small even for the two highly-variable cases examined here (≤ 0.20 ppm), but this is only if the two soundings are perfectly co-located which is generally not the case. Finally, intercomparisons between (collocated) TCCON and OCO-2 observations are least affected by horizontal variability and differences in spatial support (≤ 0.18 ppm), due to the small footprint of these observations.

4.3 Realism of implemented case studies

There is evidence that variability on the scale examined here has been observed in validation studies and has had a real impact on studies involving more than one measurement platform. For example, Figure 4 shows a vertical CO₂ profile captured on August 3rd, 2011, at Railroad Valley by the NASA Langley DC-8. CO₂ measurements were made by AVOCET, a modified LI-COR model 6252 non-dispersive infrared spectrometer (NDIR) used in hundreds of science flights and recognized as a benchmark for evaluating remote sensing observations of CO₂ (Choi et al., 2008; Vay et al., 2011). The flight coincided with a GOSAT satellite overpass. The aircraft appeared to capture very high ‘vertical’ variability of CO₂, in excess of 2 ppm even above 8000 m. In reality, however, it appears to have captured high horizontal variability superimposed on the reported vertical profile, as the highest values in each loop were found consistently in the Northwest portion of the spiral trajectory.

5. Conclusions

Results presented in this study evaluate the potential effects of horizontal variability and differences in spatial support on intercomparisons between airborne and remote sensing observations of XCO_2 , and among remote sensing observations from different platforms (TCCON, GOSAT, and OCO-2).

We find that comparisons between airborne and remote sensing observations must be evaluated with caution, unless the presence of horizontal variability can be ruled out. The impact of horizontal variability is strongest when comparing airborne observations to TCCON measurements, which have the smallest spatial support, with differences exceeding 0.5 ppm for conditions such as those examined here. The mismatch in observed XCO_2 values was attributed to real differences between XCO_2 at a TCCON site relative to that at a radius representative of an airborne spiral flight pattern, rather than to the inability of the aircraft to capture the vertical variability in CO_2 . The larger spatial supports of OCO-2, and especially GOSAT, observations help to mitigate the impact of horizontal variability, but only partially so.

In the presence of horizontal CO_2 variability, differences in spatial support among remote sensing platforms also lead to differences in observed XCO_2 values of up to 0.30 ppm. Although these are smaller than the expected differences between airborne observations and remote sensing measurements, they are nonetheless on the same order of magnitude as other sources of errors and should therefore be accounted for in intercomparisons. For example, in Tadić et al. (2014), airborne instrument uncertainty was 0.30-0.40 ppm, the uncertainty coming from incomplete coverage of the total column was 0.50 ppm, and the uncertainty of individual GOSAT soundings was $\sim 0.5\%$ or ~ 2 ppm.

Acknowledgment. This work was supported by the National Aeronautics and Space Administration (NASA) through grant no. NNX12AB90G and NNX13AC48G, and the National Science Foundation (NSF) through grant no. 1342076. Authors would like to thank members of the AJAX team from NASA Ames Research Center: Max Loewenstein, Emma Yates, Laura Iraci, Warren Gore and Antonio Trias for their participation in airborne campaigns and data acquisitions. We would also like to thank Max Loewenstein and Yoichi Shiga (Stanford University) for helpful discussions, Christian Frankenberg (NASA JPL) for help with the conversion of aircraft profiles into total column values (XCO₂), and Yonghoon Choi (National Institute for Aerospace, USA) for making available measurements used in Figure 4. All data used in this study are available upon request.

References

Biraud, S. C., M. S. Torn, J. R. Smith, C. Sweeney, W. J. Riley, and P. P. Tans (2013), A multi-year record of airborne CO₂ observations in the US Southern Great Plains. *Atmos. Meas. Tech.*, 6, 751-763, doi:10.5194/amt-6-751-2013.

Carlson, C. G. and D. E. Clay (1999), The Earth Model – Calculating Field Size and Distances between Points using GPS Coordinates, Site-Specific Management Guidelines series-11, Potash & Phosphate Institute (PPI).

Chiles, J., and P. Delfiner (1999), *Geostatistics: Modeling Spatial Uncertainty*, 695 pp., John Wiley, Hoboken, N. J.

Choi, Y., S. A. Vay, K. P. Vadrevu, A. J. Soya, J.-H. Woo, S. R. Nolf, G. W. Saxe, G. S. Diskin, D. R. Blake, N. J. Blake, H. B. Singh, M. A. Avery, A. Fried, L. Pfister, and H. E. Fuelberg (2008), Characteristics of the atmospheric CO₂ signal as observed over the conterminous United States during INTEx-NA. *J. Geophys. Res.* *113*, D07301, doi: 10.1029/2007JD008899.

Clerbaux, C., M. George, S. Turquety, K. A. Walker, B. Barret, P. Bernath, C. Boone, T. Borsdorff, J. P. Cammas, V. Catoire, M. Coffey, P.-F. Coheur, M. Deeter, M. De Mazière, J. Drummond, P. Duchatelet, E. Dupuy, R. de Zafra, F. Eddounia, D. P. Edwards, L. Emmons, B. Funke, J. Gille, D. W. T. Griffith, J. Hannigan, F. Hase, M. Höpfner, N. Jones, A. Kagawa, Y. Kasai, I. Kramer, E. Le Flochmœen, N. J. Livesey, M. López-Puertas, M. Luo, E. Mahieu, D. Murtagh, P. Nédélec, A. Pazmino, H. Pumphrey, P. Ricaud, C. P. Rinsland, C. Robert, M. Schneider, C. Senten, G. Stiller, A. Strandberg, K. Strong, R. Sussmann, V. Thouret, J. Urban, and A. Wiacek (2008), CO measurements from the ACE-FTS satellite instrument: data analysis and validation using ground-based, airborne and spaceborne observations. *Atmos. Chem. Phys.*, *8*, 2569-2594, doi:10.5194/acp-8-2569-2008.

Crisp, D., C. E. Miller, and P. L. DeCola (2007), NASA Orbiting Carbon Observatory: Measuring the column averaged carbon dioxide mole fraction from space, JARS, 12/1/2007.

Crisp, D., Miller, C. E., & DeCola, P. L. (2008). NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space. *J. Appl. Rem. Sens.* *2(1)*, 023508, doi:10.1117/1.2898457.

393

394 Deutscher, N. M., D. W. T. Griffith, G. W. Bryant, P. O. Wennberg, G. C. Toon, R. A.
395 Washenfelter, G. Keppel-Aleks, D. Wunch, Y. Yavin, N. T. Allen, J.-F. Blavier, R. Jimnez, B.
396 C. Daube, A. V. Bright, D. M. Matross, S. C. Wofsy, and S. Park (2010), Total column CO₂
397 measurements at Darwin, Australia – site description and calibration against in situ aircraft
398 profiles. *Atmos. Meas. Tech.*, 3, 947–958, doi:10.5194/amt-3-947-2010.

399

400 [http://www.ipni.net/publication/ssmg.nsf/0/0BDF314BF75B9FC1852579E5007691F9/\\$FIL](http://www.ipni.net/publication/ssmg.nsf/0/0BDF314BF75B9FC1852579E5007691F9/$FILE/SSMG-11.pdf)
401 [E/SSMG-11.pdf](http://www.ipni.net/publication/ssmg.nsf/0/0BDF314BF75B9FC1852579E5007691F9/$FILE/SSMG-11.pdf) (last accessed 01/09/2015).

402

403 Inoue, M., I. Morino, O. Uchino, Y. Miyamoto, Y. Yoshida, T. Yokota, T. Machida, Y.
404 Sawa, H. Matsueda, C. Sweeney, P. P. Tans, A. E. Andrews, S. C. Biraud, T. Tanaka, S.
405 Kawakami, and P. K. Patra (2013), Validation of XCO₂ derived from SWIR spectra of GOSAT
406 TANSO-FTS with aircraft measurement data. *Atmos. Chem. Phys.*, 13, 9771-9788,
407 doi:10.5194/acp-13-9771-2013.

408

409 Kuze, A., H. Suto, M. Nakajima, and T. Hamazaki (2009), Thermal and near infrared sensor
410 for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing
411 Satellite for greenhouse gases monitoring. *Appl. Opt.*, 48, 6716–6733.

412

413 Kuze, A., D. M. O'Brien, T. E. Taylor, J. O. Day, C. O'Dell, F. Kataoka, M. Yoshida, Y.
414 Mitomi, C. Bruegge, H. Pollock, R. Basilio, M. Helmlinger, T. Matsunaga, S. Kawakami, K.
415 Shiomi, T. Urabe, and H. Suto (2011), Vicarious calibration of the GOSAT sensors using the

416 Railroad Valley desert playa. *IEEE Trans. Geosci. Remote Sens.*, 49(5). 1781–1795, doi:
417 10.1109/TGRS.2010.2089527.

418

419 Li, J., and A. Heap (2008), A review of spatial interpolation methods for environmental
420 scientists. Geoscience Australia, Canberra, p. 137.

421

422 Lin, J. C., C. Gerbig, B. C. Daube, S. C. Wofsy, A. E. Andrews, S. A. Vay, and B. E.
423 Anderson (2004), An empirical analysis of the spatial variability of atmospheric CO₂:
424 Implications for inverse analyses and space-borne sensors. *Geophys. Res. Lett.*, 31(23), L23104,
425 doi:10.1029/2004GL020957.

426

427 Messerschmidt, J., M. C. Geibel, T. Blumenstock, H. Chen, N. M. Deutscher, A. Engel, D.
428 G. Feist, C. Gerbig, M. Gisi, F. Hase, K. Katrynski, O. Kolle, J. V. Lavric, J. Notholt, M. Palm,
429 M. Ramonet, M. Rettinger, M. Schmidt, R. Sussmann, G. C. Toon, F. Truong, T. Warneke, P. O.
430 Wennberg, D. Wunch, and I. Xueref-Remy (2011), Calibration of TCCON column-averaged
431 CO₂: the first aircraft campaign over European TCCON sites. *Atmos. Chem. Phys.*, 11, 10765–
432 10777, doi:10.5194/acp-11-10765-2011.

433

434 Miller, C. E., D. Crisp, P. L. DeCola, S. C. Olsen, J. T. Randerson, A. M. Michalak, A.
435 Alkhaled, P. Rayner, D. J. Jacob, P. Suntharalingam, D. B. A. Jones, A. S. Denning, M. E.
436 Nicholls, S. C. Doney, S. Pawson, H. Boesch, B. J. Connor, I. Y. Fung, D. O'Brien, R. J.
437 Salawitch, S. P. Sander, B. Sen, P. Tans, G. C. Toon, P. O. Wennberg, S. C. Wofsy, Y. L. Yung,

and R. M. Law (2007), Precision requirements for space-based XCO₂ data. *J. Geophys. Res.*,
112, D10314, doi: 10.1029/2006JD007659.

Miyamoto, Y., M. Inoue, I. Morino, O. Uchino, T. Yokota, T. Machida, Y. Sawa, H.
Matsueda, C. Sweeney, P. P. Tans, A. E. Andrews, S. C. Biraud, and P. K. Patra (2013),
Atmospheric column-averaged mole fractions of carbon dioxide at 53 aircraft measurement sites.
Atmos. Chem. Phys., 13, 5265-5275, doi:10.5194/acp-13-5265-2013.

Reuter, M., H. Bösch, H. Bovensmann, A. Bril, M. Buchwitz, A. Butz, J. P. Burrows, C. W.
O'Dell, S. Guerlet, O. Hasekamp, J. Heymann, N. Kikuchi, S. Oshchepkov, R. Parker, S. Pfeifer,
O. Schneising, T. Yokota, and Y. Yoshida (2013), A joint effort to deliver satellite retrieved
atmospheric CO₂ concentrations for surface flux inversions: the ensemble median algorithm
EMMA. *Atmos. Chem. Phys.*, 13, 1771-1780, doi:10.5194/acp-13-1771-2013.

Tadić, J. M., M. Loewenstein, C. Frankenberg, A. Butz, M. Roby, L.T. Iraci, E.L. Yates, W.
Gore, and A. Kuze (2014), A comparison of in-situ aircraft measurements of carbon dioxide and
methane to GOSAT data measured over Railroad Valley playa, Nevada, USA. *IEEE Trans.*
Geosci. Remote Sens., 52(12), doi: 10.1109/TGRS.2014.2318201.

Tanaka, T., Y. Miyamoto, I. Morino, T. Machida, T. Nagahama, Y. Sawa, H. Matsueda, D.
Wunch, S. Kawakami, and O. Uchino (2012), Aircraft measurements of carbon dioxide and
methane for the calibration of ground-based high-resolution Fourier Transform Spectrometers

and a comparison to GOSAT data measured over Tsukuba and Moshiri. *Atmos. Meas. Tech.*, *5*, 2003-2012, doi:10.5194/amt-5-2003-2012.

United States National Research Council, “Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond,” Available from <http://www.nap.edu/> (2007).

Vay, S. A., Y. Choi, K. P. Vadrevu, D. R. Blake, S. C. Tyler, A. Wisthaler, A. Hecobian, Y. Kondo, G. S. Diskin, G. W. Sachse, J-H. Woo, A. J. Weinheimer, E. T. Olsen, A. Stohl, J. F. Burkhardt, and P. O. Wennberg (2011), Patterns of CO₂ and radiocarbon across high Northern latitudes during International Polar Year 2008. *J. Geophys. Res. Atmos.*, *116*, D14301, doi: 10.1029/2011JD015643.

Washenfelder, R., G. Toon, J. Blavier, Z. Yang, N. Allen, P. Wennberg, S. Vay, D. Matross, and B. Daube (2006), Carbon dioxide column abundances at the Wisconsin Tall Tower site, *J. Geophys. Res.*, *111*(D22), 1-11, doi:10.1029/2006JD007154.

Wunch, D., G. C. Toon, P. O. Wennberg, S. C. Wofsy, B. B. Stephens, M. L. Fischer, O. Uchino, J. B. Abshire, P. Bernath, S. C. Biraud, J. F. L. Blavier, C. Boone, K. P. Bowman, E. V. Browell, T. Campos, B. J. Connor, B. C. Daube, N. M. Deutscher, M. Diao, J. W. Elkins, C. Gerbig, E. Gottlieb, D. W. T. Griffith, D. F. Hurst, R. Jiménez, G. Keppel-Aleks, E. A. Kort, R. Macatangay, T. Machida, H. Matsueda, F. Moore, I. Morino, S. Park, J. Robinson, C. M. Roehl, Y. Sawa, V. Sherlock, C. Sweeney, T. Tanaka, and M. A. Zondlo (2010), Calibration of the

Total Carbon Column Observing Network using aircraft profile data, *Atmos. Meas. Tech.*, 3(5),
1351-1362, doi:10.5194/amt-3-1351-2010.

Wunch, D., G. C. Toon, J.-F. L. Blavier, R. Washenfelder, J. Notholt, B. J. Connor, D. W. T.
Griffith, V. Sherlock, and P. O. Wennberg (2011), The total carbon column observing network,
Philos. Trans. R. Soc. - Ser. A Math. Phys. Eng. Sci., 369(1943), 2087-2112,
doi:10.1098/rsta.2010.0240.

Zhou, Y., D. R. Obenour, D. Scavia, T. H. Johengen, and A. M. Michalak (2013), Spatial
and temporal trends in Lake Erie hypoxia, 1987–2007. *Environ. Sci. Technol.*, 47 (2), 899, doi:
10.1021/es303401b.