Inferring the vertical distribution of CO and CO_2 from TCCON total column values using the TARDISS algorithm

Harrison A. Parker[1], Joshua L. Laughner[2], Geoffrey C. Toon[2], Debra Wunch[3], Coleen M. Roehl[1], Laura T. Iraci[4], James R. Podolske[4], Kathryn McKain[5,6], Bianca C. Baier[5,~~7~~6], Paul O. Wennberg[1,~~8~~7]

[1] Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA
[2] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[3] Department of Physics, University of Toronto, Toronto, ON, Canada
[4] Atmospheric Science Branch, NASA Ames Research Center, Moffett Field, CA, USA
[5] Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, CO, USA
~~[6] Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA~~
[~~7~~6] Global Monitoring Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA
[~~8~~7] Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA, USA

Correspondence to: Paul O. Wennberg (wennberg@caltech.edu) and Joshua L. Laughner (josh.laughner@jpl.nasa.gov).

Abstract

We describe an approach for determining limited information about the vertical distribution of carbon monoxide (CO) and carbon dioxide (CO_2) from total column ~~observations from,~~ ground-based TCCON observations. For ~~long-lived trace gases, such as~~ CO and CO_2, it has been difficult to retrieve information about their vertical distribution from spectral line shapes ~~in the shortwave infrared (SWIR) spectra~~ because of the ~~large doppler widths at 6000 cm^{-1}, and~~ errors in the spectroscopy and ~~in~~ the atmospheric temperature profile ~~which~~that mask the effects of variations in their mixing ratio with altitude ~~in the troposphere~~. For CO_2 the challenge is especially difficult given that these variations are typically 2% or less. Nevertheless, if sufficient accuracy can be obtained, such information would be highly valuable for evaluation of retrievals from satellites and more generally for improving the estimate of surface sources and sinks of these trace gases.

We present here the Temporal Atmospheric Retrieval Determining Information from Secondary Scaling (TARDISS) retrieval algorithm. TARDISS uses several, simultaneously ~~-~~obtained total column observations of the same gas from different absorption bands with distinctly different vertical averaging kernels. ~~Since TARDISS avoids spectral re-fitting by ingesting retrieved column abundances, it is very fast and processes years of data in minutes.~~ The different total column retrievals are combined in TARDISS using a Bayesian approach where the weights and temporal covariance applied to the different retrievals include additional constraints on the diurnal variation in the vertical distribution for these gases. We assume that ~~only~~ the near ~~-~~surface ~~is influenced by local~~part of the column varies rapidly over the course of a day (from surface sources and sinks, ~~while variations in the distribution in the middle~~for example) and the upper ~~troposphere result primarily from advection that can be independently constrained using reanalysis data about the variation in mid-tropospheric potential temperature.~~ part of the column has a larger temporal covariance over the course of a day.

Using measurements from the five North American TCCON sites, we find that the retrieved lower partial column (between the surface and ~800 hPa) of the CO and CO_2 dry mole fractions (DMF) have slopes of ~~1.001~~0.999 ±0.002 and 1.~~007~~001±0.~~002~~003 with respect to lower column DMF from integrated in situ data measured ~~by~~ directly from aircraft and ~~AirCore~~in AirCores. The average error for our lower column CO retrieval is ~~0.857~~1.51 ppb (~1.5%) while the average error for our CO_2 retrieval is ~~3.55~~5.09 ppm (~~0.8%). We calculate degrees of freedom from signal of 0.218 per measurement for lower partial column CO on average and of 0.353 per measurement for lower partial column CO_2 on average.~~1.25%). Compared with classical line-shape-derived vertical profile retrievals, our algorithm

reduces the influence of forward model errors such as imprecision in spectroscopy (line shapes and intensities) and in the instrument line shape. In addition, because TARDISS uses the existing retrieved column abundances from TCCON (which themselves are computationally much less intensive than profile retrieval algorithms), it is very fast and processes years of data in minutes. We anticipate that this approach will find broad application for use in carbon cycle science.

1 Introduction

Remote sensing ~~measurements~~retrievals of atmospheric ~~gases~~gas abundances are ~~made around the world in an attempt~~used to ~~better understand~~diagnose the sources, sinks, and fluxes at the local, regional, and global scales (Connor et al., 2008, p.2; Deeter, 2004; Kerzenmacher et al., 2012; Wunch et al., 2011). Compared with in situ measurements, these retrievals, which are used in carbon cycle science investigations, are less influenced by nearby point sources or sinks and rapidly changing meteorological conditions that would lead to erroneous flux calculations (Keppel-Aleks et al., 2012) ~~which is both a strength and a weakness for use in carbon cycle science investigations. Additionally, because~~. Because the column represents the integral of a gas from the surface to the top of the atmosphere, flux estimates from column amounts are less sensitive to errors in the assumed vertical transport than those using surface measurements (Keppel-Aleks et al., 2011, 2012).

~~Contrastingly, column measurements have their own drawbacks for estimating surface fluxes. Total column concentrations are much less sensitive to the local emissions and flux estimation can be influenced by variation in the mixing ratio at higher altitudes. Signals~~ In contrast, since signals of CO_2 and CO ~~from~~fluxes at the surface are muted in the total column (due to the dilution of signals from the surface being integrated ~~across~~throughout an entire column~~.~~), they are less useful in diagnosing local emissions than in situ measurements. For CO_2, the total columns are strongly influenced by ~~mesoscale flux patterns~~synoptic scale transport in the troposphere ~~above the boundary layer~~ making it even more difficult to discern the influences of surface fluxes (Keppel-Aleks et al., 2011, 2012). For CO, its several-week lifetime in the free troposphere results in regional transport influences that can dampen the surface signals in the total column values (Deeter, 2004; Zhou et al., 2019). These issues ~~can~~ limit the effectiveness of total column measurements in surface flux analysis – particularly for local sources.

Profile retrievals can, in principle, ameliorate these issues and ~~provide clearer~~thereby enable more direct information on surface processes. ~~-~~ Theoretical analysis shows that two to three vertical degrees of freedom (DoF) can be achieved in CO_2 retrievals from near-IR (NIR) and mid-IR (MIR) spectra from high-resolution Fourier transform spectrometers (Connor et al., 2016; Kuai et al., 2012; Roche et al., 2021; Shan et al., 2021). In practice, however, Connor et al. (2016) and Roche et al. (2021) showed that the precision of retrieved CO_2 profiles using spectral windows in the NIR was much lower than the theoretical estimate due to ~~errors~~uncertainty in the ~~a priori~~ temperature profile and in the

forward radiative transfer model. Likewise, Shan et al. (2021) retrieve CO_2 profiles using spectral windows in the MIR. They use an a posteriori optimization method to improve the tropospheric CO_2 signal and they report errors near 2%. Although both of these methods retrieve profiles with sufficient degrees of freedom to observe some signals of the variation in the vertical distribution, they report errors ~~which~~sufficiently large enough to limit their use for carbon cycle studies.

Several operational CO profile retrievals exist, but these products still face the issues of column dilution or larger sensitivity to the free troposphere compared to the surface. ~~For example, spectra recorded from the MOPITT instrument aboard the Terra satellite in both MIR and NIR have been used to provide limited information on the vertical distribution of CO (between one and two degrees of freedom in the troposphere) (Deeter, 2004; Turquety et al., 2008).~~ The Network for the Detection of Atmospheric Composition Change (NDACC) retrieves profiles of CO in the atmosphere (Buchholz et al., 2017) with ~2 degrees of freedom for the signal providing information of a lower (surface-8km) layer sensitive to the boundary layer and an upper (8-~~20km~~20 km) layer with ~1-3% uncertainty in the total column (Zhou et al., 2018, 2019). These ground-based measurements require higher spectral resolution ~~and therefore a~~than those typically available in the TCCON interferometers. The higher resolution also requires longer measurement time, resulting in fewer observations per day. This limits their ability to capture diurnal changes and makes the measurements more susceptible to ~~cloud interruption.~~variations in solar viewing during acquisition of the interferograms. These measurements also require highly accurate knowledge of the spectral line widths, their temperature dependence, the instrument line shape (ILS), and the solar spectrum. These limitations motivate our work to develop a new product with better sensitivity to surface processes and higher temporal resolution from the existing TCCON retrievals.

~~Profile retrievals that fit measured spectra and exploit the profile information given by pressure broadening of spectral lines require high resolution data to obtain information about different levels of the atmosphere (Sepúlveda et al., 2014).~~ In ~~the~~our approach ~~described here~~, we do not retrieve profile information directly from the spectra. Instead, we utilize ~~information from~~ the vertical~~,~~ and temporal~~, and a priori vertical profile~~ domains to infer partial column dry mole fraction (DMF) values. We fit partial column scalar values to match TCCON retrieved total column ~~dry mole fraction measurements~~DMF that are 1) quality controlled and 2) individually tied to ~~the~~ World Meteorological Organization (WMO) trace gas standard ~~scale~~scales which mitigates a number of errors in the forward ~~spectroscopic~~radiative transfer model~~. We use~~, including those arising from errors in the spectroscopy. We use the extant multiple total column measurements from spectral windows with different line intensities~~,~~ and hence different shapes of the column averaging kernel. We extract the vertical information from the diurnally-varying differences in these total column values ~~between the different windows by fitting over an entire day of measurement in order to make use of the information from the temporal dimension. We optimize the separation between the near surface and the rest of the atmosphere using~~and additional a priori information about the expected temporal covariance in the

different partial columns based on known atmospheric behavior. This method allows us to ~~use the algorithm to~~ extract information focused on the lower atmosphere where the ~~concentrations~~trace gas DMF are most sensitive to surface exchange.

The ~~accuracy~~uncertainty of this new method for retrieving partial column values is evaluated using comparisons with in situ vertical ~~profiles.~~profile measurements. Section 2 describes the theory and parameters chosen for our retrieval, and the data used for the retrieval, validation, and comparison. Sections 3.1 ~~to~~through 3.~~2~~3 present our validation data~~,~~ and a sensitivity study of the retrieval parameters~~, and~~ . Section 3.4 presents an error and information content analysis. Finally, Sect. 3.~~3 and 3.4 give~~5 gives examples of the data retrieved using this approach ~~and provide evidence for the utility of this approach in flux estimation~~.

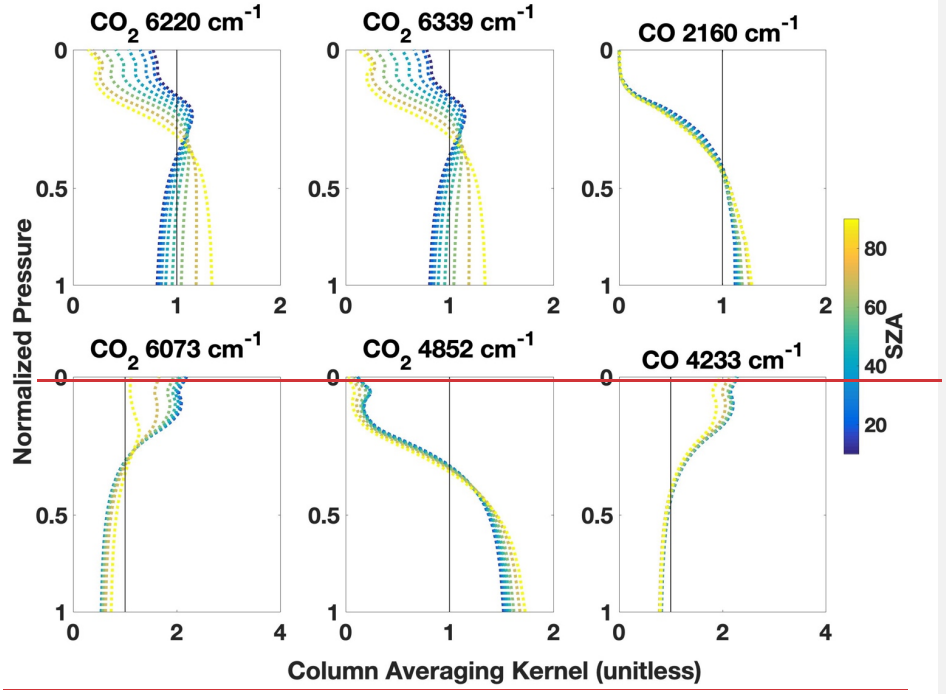
2 Methods

2.1 Total Carbon Column Observing Network

The Total Carbon Column Observing Network (TCCON) is ~~described by Wunch et al. (2011), although we will give a brief overview here to include aspects of the retrieval algorithm and observation scheme that have evolved since 2011. TCCON is a network of sites that use ground-based Fourier transform spectrometers with InGaAs and Si detectors to gather spectra for the 3900 to 15500 cm^{-1} spectral region.~~a ground-based network of solar viewing Fourier transform spectrometers equipped with InGaAs and Si detectors that gather spectra for the 3900 to 15500 cm^{-1} spectral region (Wunch et al., 2011). Importantly for our work here on CO, some sites are now equipped with an InSb detector that simultaneously allows spectral measurement down to 2000 cm^{-1} at the expense of simultaneous observations using the Si detector. CO_2 and CO are retrieved simultaneously over several spectral windows (independent spectral bands). These windows are chosen to provide high sensitivity to the gas of interest while limiting interference from other atmospheric absorbers.

Column abundances of atmospheric species are computed from the measured spectra using a nonlinear least-squares fitting algorithm, GFIT, which minimizes the residuals between a measured spectrum and one calculated by uniformly scaling a priori vertical profiles for the fitted atmospheric species, yielding the optimal VMR (volume mixing ratio) scaling factors (VSF) of the fitted gases. The ~~prior~~a priori profiles scaled by the VSF are integrated to calculate the total column abundance of a species. The retrieved scaled column abundances are converted to column dry mole fraction (DMF) by multiplying by 0.2095 and dividing by the column of O_2, retrieved from a different spectral window of the same spectrum. These retrievals are then quality-controlled and scaled to minimize both ~~any~~ airmass dependence and the difference with simultaneously measured in situ profiles. ~~These outputs from standard TCCON processing are used as input for TARDISS.~~

For each window and for each spectrum fit by GFIT, an associated column averaging kernel is computed that describes the sensitivity of the ~~VSF~~total column to changes in species abundance at each altitude~~.~~ (shown in Fig.1). A perfect column averaging kernel would have values of one for all altitudes. More commonly, the kernels will ~~be greater than 1 at lower/higher altitudes and less than 1 at higher/lower altitudes.~~vary slowly with altitude with a pressure weighted average value close to one. Values higher (lower) than 1 mean that the retrieval is more (less) sensitive to trace gas changes at that altitude. These sensitivities ~~also~~ vary with solar zenith angle (SZA) as the spectral absorption ~~deepen.~~deepens at higher SZA. The vertical sensitivity of each window is a result of its spectral properties. Optically thin spectral regions (windows) tend to be more sensitive to the upper troposphere and the stratosphere while optically thick windows tend to be more sensitive to the lower troposphere. Since information about the stratosphere comes only from near the line center as a result of diminished collisional broadening, if the absorption at the line center is saturated (nearly zero transmission), the spectrum will contain little information about the stratosphere and hence the kernel will be ~~small~~low there. The differences in column averaging kernel shapes are the main source of information used in ~~our~~the TARDISS algorithm. The outputs of the VSF values, a priori profiles, total column DMF values, and vertical averaging kernels from standard TCCON processing are used as input for the TARDISS algorithm.



We will refer to the spectral retrievals as being the TCCON retrievals and the temporal partial column retrievals as the TARDISS fit. We also use the terms retrieval and fit interchangeably to refer to the TCCON or TARDISS methodology.

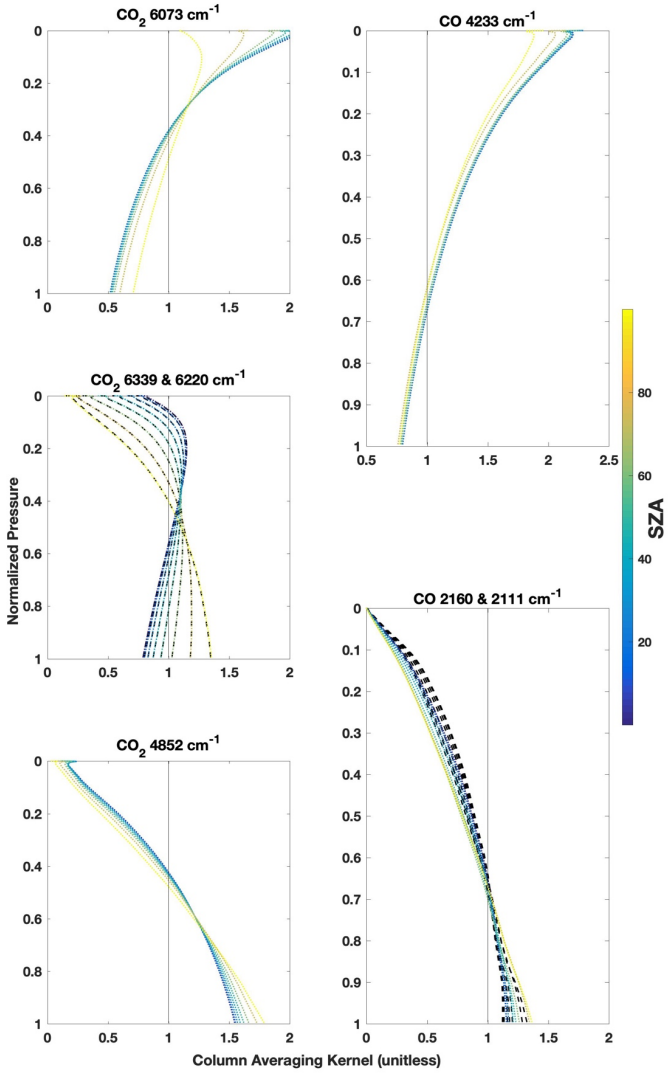

CO2 6073 cm-1
CO 4233 cm-1
CO2 6339 & 6220 cm-1
CO 2160 & 2111 cm-1
CO2 4852 cm-1
Normalized Pressure
SZA
Column Averaging Kernel (unitless)

Figure 1. Vertical sensitivities of the total column retrievals from GFIT ~~and~~ used in our algorithm for both CO_2 (left ~~and middle~~ column) and CO (right column) plotted against pressure normalized ~~pressure~~ to the surface and color coded by the solar zenith angle (SZA). ~~We also use a window centered at 2111 cm^{-1} for CO which has vertical sensitivities that are nearly identical to the 2160 cm^{-1} window.~~ A column averaging kernel greater than 1 means that the total column is more sensitive to molecules at this pressure level than the average sensitivity. For example, if we move some of the CO_2 molecules from 200 hPa to the surface in our a priori profile, the retrieved total column and scale factor (VSF) will decrease for the 6073 cm^{-1} window and increase for the 4852 cm^{-1} window while the true and a priori total ~~column remains~~columns remain unchanged. The 6220 and 6339 cm^{-1} CO_2 and 2160 and 2111 cm^{-1} CO windows have near-identical kernels due to the CO_2 bands being almost identical in their line strengths, separations, widths, and temperature dependences. The 6339 cm^{-1} CO_2 is represented by black dashed lines behind the dotted lines representing the 6220 cm^{-1} sensitivities and the 2111 cm^{-1} CO is represented by black dashed lines behind the dotted lines representing the 2160 cm^{-1} sensitivities.

2.2 The TARDISS Algorithm

Traditional profile retrievals fit spectra by adjusting the abundance of the trace gases at multiple vertical levels to determine the vertical distribution of a specific atmospheric species (Pougatchev et al., 1995; Roche et al., 2021). Here, we describe the Temporal Atmospheric Retrieval Determining Information from Secondary Scaling (TARDISS) algorithm that optimizes the scaling of the profile of our target gas separated into two layers, one near the surface and the other at and above the typical well-mixed surface boundary layer. This is illustrated by the flowchart in Fig. 2. The algorithm minimizes the cost function (Equation 1) by finding the maximum a posteriori solution for a state vector containing upper and lower column scale factors for all TCCON observations in a given day. That is, if a day has n_s observations, the state vector will have n_s lower column scale factors and n_s upper column scale factors, for $2n_s$ elements total. These are constrained by TCCON column average mole fractions and an assumed temporal covariance. The Jacobian matrix for TARDISS combines the TCCON averaging kernels and the TCCON assumed vertical CO or CO_2 profiles in an operator which maps the upper and lower scale factors back to column average mole fractions. We define our cost function as:

$$\chi^2 = \left(\boldsymbol{y} - \boldsymbol{K}(\boldsymbol{x}_\gamma - \boldsymbol{x}_{a,\gamma})\right)^T \boldsymbol{S}_\epsilon^{-1}\left(\boldsymbol{y} - \boldsymbol{K}(\boldsymbol{x}_\gamma - \boldsymbol{x}_{a,\gamma})\right) + (\boldsymbol{x}_\gamma - \boldsymbol{x}_{a,\gamma})\boldsymbol{S}_a^{-1}(\boldsymbol{x}_\gamma - \boldsymbol{x}_{a,\gamma}), \quad (1)$$

where $\mathbf{y}$ is the measurement vector, $\mathbf{K}$ is the Jacobian matrix, $\boldsymbol{x}_\gamma$ is the retrieved state vector, $\boldsymbol{x}_{a,\gamma}$ is the a priori state vector, $\boldsymbol{S}_\epsilon$ is the model covariance matrix, and $\boldsymbol{S}_a$ is the prior covariance matrix. In the following sections, we will derive the necessary equations for the construction of the components of the cost function in detail. Table S1 lists all the variable names in this work and their descriptions.

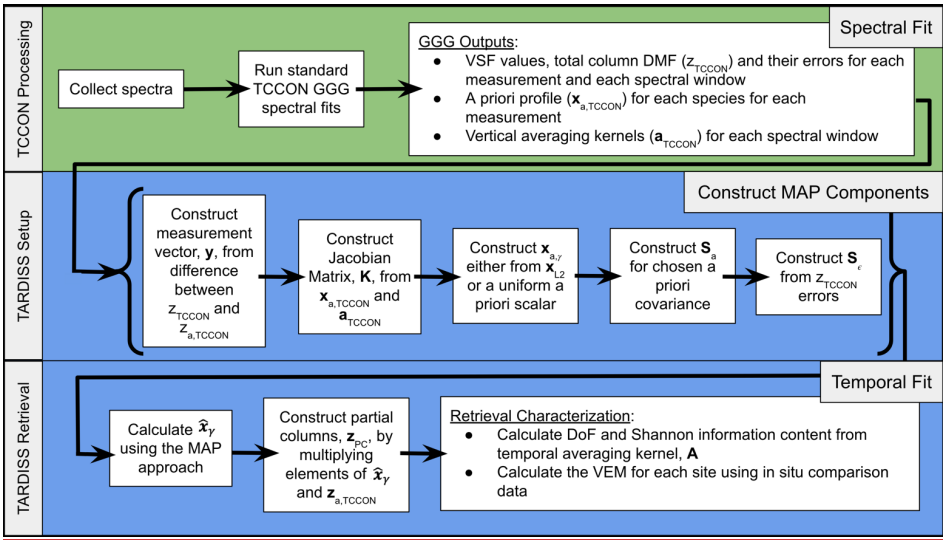


Figure 2. Flowchart illustrating the steps performed by of the TARDISS retrieval. The input to the TARDISS retrieval is the output of the spectral fitting done by the GGG2020 software suite represented by the green row. The setup of the components of the TARDISS algorithm from the output of the TCCON spectral fits is shown in Equations 11 through 14 and in the middle row. The TARDISS retrieval is performed using Equation 16, the output partial column DMF values are calculated using Equation 17, and the information content is calculated by Equation 18 and 19 as shown in the bottom row.

2.2.1 Derivation of the TARDISS Jacobian Matrix Components

We use the notation and concepts of Rodgers and Connor (2003) with vectors represented with bolded lower-case letters and matrices represented with bolded upper-case letters. We start in the vertical domain where Equations 3 through 9 are used for each spectral window, each TCCON measurement, and each species retrieved (CO and CO_2 in this work) in the TCCON fit. These equations are used to calculate the weights in the Jacobian matrix and values in the measurement vector for the temporal calculations in

Equation 10 and beyond (represented by the middle and bottom row of Fig. 2). We will therefore keep Equations 3 through 9 agnostic of species and window for this description.

To derive the values used in the Jacobian matrix, **K**, we start by relating the atmospheric profile of CO or CO2 to the column average mole fractions observed by TCCON. For TARDISS, we assume that the a posteriori atmospheric profile can be described as the profile output by the TCCON retrieval with the bottom q levels scaled separately from the top $n_l - q$ levels, where q is a chosen level index and n_l is the number of vertical levels in the profile:

$$\mathbf{x}_{\text{part}} = \begin{bmatrix} \gamma_L \cdot x_{a,TCCON,1} \\ \vdots \\ \gamma_L \cdot x_{a,TCCON,q} \\ \gamma_U \cdot x_{a,TCCON,q+1} \\ \vdots \\ \gamma_U \cdot x_{a,TCCON,n_l} \end{bmatrix}. \quad (2)$$

Here, $x_{a,TCCON}$ is the TCCON a priori profile scaled by the median TCCON retrieved VSF across all the TCCON spectral windows for this gas, and the γ_L and γ_U values are the lower and upper column scale factors, respectively, which our algorithm retrieves. We relate this to the TCCON total column value using the standard equation from Rodgers and Connor (2003):

$$z_{TCCON} = z_{a,TCCON} + \boldsymbol{a}_{TCCON}^{\xi T} (\boldsymbol{x}_{part} - \boldsymbol{x}_{a,TCCON}), \quad (3)$$

where z_{TCCON} is the total column DMF output of a chosen species in a particular window from the TCCON fit, $z_{a,TCCON}$ is the original vertical column DMF calculated from the a priori profile scaled by the median VSF of the windows used, and $\boldsymbol{a}_{TCCON}^{\xi}$ is the vector of column averaging kernel values output from the TCCON processing weighted by the pressure thickness of each atmospheric layer. All components in Equation 3 are dry mole fractions except for the averaging kernel which is unitless. Equation 3 tells us how the retrieved DMF would change if the profile constructed from the two partial columns differed from $x_{a,TCCON}$.

The next step is to rearrange this equation so that our observed quantity is on the left-hand side, and the right-hand side is a linear combination of the two scaling factors. Subtracting $z_{a,TCCON}$ from both sides and focusing on the rightmost term of Equation 3, the averaging kernel is multiplied by the difference of the a priori and scaled DMF profiles summed for each of the n_l levels of the atmosphere.

$$z_{TCCON} - z_{a,TCCON} = \mathbf{a}_{TCCON}^{\xi T} \left(\mathbf{x}_{part} - \mathbf{x}_{a,TCCON}\right) = \sum_{i=1}^{n_l} a_{TCCON,i} \left(x_{part,i} - x_{a,TCCON,i}\right). \quad (4)$$

Here, we assign $\boldsymbol{x}_{part}$ to be the TCCON a priori profile scaled by two independent values, one for the lower partial column and one for the upper partial column. To designate the partial columns, our method splits the total column at a specified altitude level index, q, and scales the a priori profile below and above the level q independently by the scalar values γ_L and γ_U such that:

$$z_{TCCON} - z_{a,TCCON} = \sum_{i=1}^{q} a_{TCCON,i}(\gamma_L x_{a,TCCON,i} - x_{a,TCCON,i}) + \sum_{i=q+1}^{n_l} a_{TCCON,i}(\gamma_U x_{a,TCCON,i} - x_{a,TCCON,i}) \quad (5)$$

Since Equation 5 is linear, we then group terms reducing the right side of Equation 5 to:

$$z_{TCCON} - z_{a,TCCON} = (\gamma_L - 1)\sum_{i=1}^{q} a_{TCCON,i}\, x_{a,TCCON,i} + (\gamma_U - 1)\sum_{i=q+1}^{n_l} a_{TCCON,i}\, x_{a,TCCON,i} \quad (6)$$

Defining two new variables, k_L and k_U, we can write this as:

$$z_{TCCON} - z_{a,TCCON} = (\gamma_L - 1)\, k_L + (\gamma_U - 1)\, k_U \quad (7)$$

where,

$$k_L = \sum_{i=1}^{q} a_{TCCON,i}\, x_{a,TCCON,i} \quad (8)$$

and

$$k_U = \sum_{i=q+1}^{n} a_{TCCON,i}\, x_{a,TCCON,i} \quad (9)$$

and k_L and k_U are both scalar values.

Equation 7 is applicable to all spectral windows for each spectrum measured. For example, for our CO_2 retrieval, we use four separate spectral windows per measured spectrum and often have a few hundred spectra measured per day. In the case of the CO_2 retrieval, the left-hand side of Equation 7 and the k_L and k_U values will be calculated for each of the four spectral windows used for each spectrum fit by TCCON. These values are aggregated into the vectors and matrices described by Equations 10 - 14 in order to fit the spectra measured over an entire day at one time.

2.2.2 Deriving the Maximum A Posteriori Equation and Solution

While Equation 7 can be set up and solved for each spectrum using the total column value from each spectral window used in the TCCON fit, the TARDISS retrieval uses an entire day's worth of TCCON

retrievals in order to increase the signal-to-noise and to utilize the information from the temporal variation in the kernels. Fitting over an entire day of TCCON retrievals reduces the retrieved partial column error values compared to fitting individual measurements using Equation 7. Section S1 shows the influence of including multiple observations on the retrieved partial column errors. Let n_w denote the number of windows and n_s the number of spectra over a day and w_i and s_i denote the ith window and spectrum. We combine the above equations into a matrix form:

$$\boldsymbol{y} = \boldsymbol{K}(\boldsymbol{x}_\gamma - \boldsymbol{x}_{a,\gamma}) + \boldsymbol{\epsilon} \quad . \quad (10)$$

where, $\boldsymbol{y}$ is the measurement vector composed of values from the left side of Equation 7

$$\boldsymbol{y} = \begin{bmatrix} z_{TCCON,1,1} - z_{a,TCCON,1} \\ \vdots \\ z_{TCCON,w_i,s_i} - z_{a,TCCON,s_i} \\ \vdots \\ z_{TCCON,n_w,n_s} - z_{a,TCCON,n_s} \end{bmatrix}, \quad (11)$$

K is the Jacobian matrix of the k_L and k_U values over a day,

$$\boldsymbol{K} = \begin{bmatrix} k_{L,1,1} & \ddots & 0 & k_{U,1,1} & \ddots & 0 \\ 0 & & k_{L,1,n_s} & 0 & & k_{U,1,n_s} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_{L,n_w,1} & & 0 & k_{U,n_w,1} & & 0 \\ & \ddots & & & \ddots & \\ 0 & & k_{L,n_w,n_s} & 0 & & k_{U,n_w,n_s} \end{bmatrix}, \quad (12)$$

$\boldsymbol{x}_\gamma$ is our state vector of partial column scalars which are the same for all windows in each measured spectrum,

$$\boldsymbol{x}_\gamma = \begin{bmatrix} (\gamma_L - 1)_1 \\ \vdots \\ (\gamma_L - 1)_{n_s} \\ (\gamma_U - 1)_1 \\ \vdots \\ (\gamma_U - 1)_{n_s} \end{bmatrix}, \quad (13)$$

and $\boldsymbol{x}_{a,\gamma}$ is our vector of a priori partial column scalars,

$$\boldsymbol{x}_{a,\gamma} = \begin{bmatrix} (\gamma_{a,L} - 1)_1 \\ \vdots \\ (\gamma_{a,L} - 1)_{n_s} \\ (\gamma_{a,U} - 1)_1 \\ \vdots \\ (\gamma_{a,U} - 1)_{n_s} \end{bmatrix} . \quad (14)$$

With n_s measurements made in a day, n_w spectral windows, and two partial columns, the y vector is of the size 1 by $n_w n_s$, the $\boldsymbol{K}$ matrix is of the size $n_w n_s$ by $2n_s$ and the $\boldsymbol{x}_\gamma$ and $\boldsymbol{x}_{a,\gamma}$ vectors are of the size $2n_s$ by 1. So, for each spectrum, there is one γ_L value and one γ_U value, representing the partial column scale factors aggregated over the windows.

Since Equation 10 is linear, we can apply a basic linear least-squares method to solve for the partial column scalars:

$$\boldsymbol{x}_{L2} = (\boldsymbol{K}^T\boldsymbol{K})^{-1}\boldsymbol{K}^T\boldsymbol{y} . \quad (15)$$

While the linear least-squares method provides a useable solution to our retrieval, it also has partial column error values on the order of 10 ppm, due to the strong anti-correlation of the lower and upper partial columns, which render the solutions unsuitable for carbon cycle science. Including constraints through a Bayesian approach reduces the retrieved partial column error values as shown in Fig. S1. In addition, the least-squares method does not allow us to utilize additional a priori information in the covariance of the partial columns.

We use the maximum a posteriori (MAP) approach (Rodgers, 2008) to calculate the most probable state vector from the given models and a priori information. In line with the assumptions of the MAP approach, we assume our problem is linear and follows a gaussian distribution. The MAP solution can take a few equivalent forms. In this work we use:

$$\boldsymbol{x}_\gamma = \boldsymbol{x}_{a,\gamma} + \boldsymbol{S}_a\boldsymbol{K}^T(\boldsymbol{K}\boldsymbol{S}_a\boldsymbol{K}^T + \boldsymbol{S}_\epsilon)^{-1}(\boldsymbol{y} - \boldsymbol{K}\boldsymbol{x}_{a,\gamma}) , \quad (16)$$

where ~~Since the terminology for the fitting done in a standard TCCON retrieval~~$\boldsymbol{x}_{a,\gamma}$ is ~~similar to that used in~~ the a priori partial column scalar values, $\boldsymbol{S}_a$ is the a priori covariance matrix, $\boldsymbol{K}$ is the forward

mapping matrix, $\boldsymbol{S}_\epsilon$ is the model covariance matrix, $\boldsymbol{y}$ is the measurement vector, and $\boldsymbol{x}_\gamma$ is the output solution vector. The input components ($\boldsymbol{x}_{a,\gamma}$, $\boldsymbol{S}_a$, and $\boldsymbol{S}_\epsilon$) are described in Sect. 2.3.2.

Once we have calculated the most likely solution for the partial column scalars as a vector in temporal space, $\boldsymbol{x}_\gamma$, we reconstruct the partial column DMF ~~retrieval discussed in this work, we will refer to~~values for the day for the lower and upper partial columns as:

$$\boldsymbol{z}_{PC} = \begin{bmatrix} z_{PC,L,1} \\ \vdots \\ z_{PC,L,n_s} \\ z_{PC,U,1} \\ \vdots \\ z_{PC,U,n_s} \end{bmatrix} = \begin{bmatrix} (x_{\gamma L,1} + 1) \cdot z_{a,L,TCCON,1} \\ \vdots \\ (x_{\gamma L,n_s} + 1) \cdot z_{a,L,TCCON,n_s} \\ (x_{\gamma U,1} + 1) \cdot z_{a,U,TCCON,1} \\ \vdots \\ (x_{\gamma U,n_s} + 1) \cdot z_{a,U,TCCON,n_s} \end{bmatrix} \quad (17)$$

where $z_{a,L,TCCON}$ and $z_{a,U,TCCON}$ are the values of the a priori partial column DMFs calculated by integrating the median TCCON a posteriori profiles for the measurements in a day using the same method as the standard TCCON full column retrievals (Wunch et al., 2011).

2.2.3 Calculating Informational Content

The MAP retrieval allows us to calculate the information content of the retrieval. In particular, we compare the degrees of freedom for our retrieval calculated by taking the trace of the averaging kernel of the fit, calculated as the following:

$$DoF = tr(\boldsymbol{A}) = tr((\boldsymbol{K}^T \boldsymbol{S}_\epsilon^{-1} \boldsymbol{K} + \boldsymbol{S}_a^{-1})^{-1} \boldsymbol{K}^T \boldsymbol{S}_\epsilon^{-1} \boldsymbol{K}) \quad (18)$$

as well as the Shannon information content derived from the natural log of the determinant of the difference between the averaging kernel and an identity matrix:

$$H = -\frac{1}{2} \ln(|\boldsymbol{I} - \boldsymbol{A}|) \quad (19)$$

Generally, profile retrieval averaging kernels represent the sensitivity of a specific level of a profile to the rest of the levels in the profile. The averaging kernel for the TARDISS inversion is a temporal averaging kernel relating how each partial column calculation relates to every other measurement during a day. The DoF value for a day of the retrieval represents how many individual pieces of partial column information we can infer over the day of measurements. We either report the number of degrees of freedom from the fit over a day or normalize the degrees of freedom by the number of measurements in each day for a more comparative understanding of the TARDISS degrees of freedom

with respect to a traditional profile retrieval as well as between days with a large variation in the number of measurements.

2.2.4 In Situ Comparison Calculations

To evaluate the accuracy of our partial column retrieval, we use the smoothing calculation shown in Equation 3 of Wunch et al. (2010), altered to use the terminology of this work, to determine the value of the partial columns of the TCCON total columns used as input:

$$z_s = z_{a,TCCON} + \boldsymbol{a}_{TCCON}^{\xi T}\left(\boldsymbol{x}_{true} - \boldsymbol{x}_{a,TCCON}\right), \quad (20)$$

where ~~as being the phase 1 (P1) fit and~~ z_s is the smoothed column averaged DMF, $z_{a,TCCON}$ is the column averaged DMF of the scaled a priori profile, $\boldsymbol{a}_{TCCON}^{\xi}$ is the vertical averaging kernel for the specific spectral window dotted with an integration operator, $\boldsymbol{x}_{true}$ is the measured, in situ profile in DMF, and $\boldsymbol{x}_{a,TCCON}$ is the scaled a priori profile. We use this equation to create the smoothed partial column TCCON DMF values by integrating to the same split point, q, as in Equation 5. These values serve as a sort of null hypothesis to compare to the TARDISS retrieval to determine if the fits are effective in inferring partial column information.

In order to compare the partial column retrievals to in situ profiles for validation purposes, we calculate the vertical sensitivities of the TARDISS fit (shown in Fig. 8) using the gain matrix, **G**, from the TARDISS inversion and the averaging kernel profiles from the TCCON measurement windows as:

$$\boldsymbol{G} = (\boldsymbol{K}^T\boldsymbol{S}_\epsilon^{-1}\boldsymbol{K} + \boldsymbol{S}_a^{-1})^{-1}\boldsymbol{K}^T\boldsymbol{S}_\epsilon^{-1}. \quad (21)$$

$$\boldsymbol{A}_{vert} = \boldsymbol{G}\,\boldsymbol{\Xi}_{TCCON}, \quad (22)$$

where

$$\boldsymbol{\Xi}_{TCCON} = \begin{bmatrix} \boldsymbol{a}_{TCCON,1,1} \\ \vdots \\ \boldsymbol{a}_{TCCON,1,n_s} \\ \boldsymbol{a}_{TCCON,n_w,1} \\ \vdots \\ \boldsymbol{a}_{TCCON,n_w,n_s} \end{bmatrix}, \quad (23)$$

and $\boldsymbol{a}_{TCCON}$ is the same vector of column averaging kernels from Equation 3 without the integration operator for each window used and $\boldsymbol{A}_{vert}$ is the vertical sensitivity of the partial column related to the

profile. G has dimensions of $2n_s$ by $n_w n_s$, $\boldsymbol{\Xi}_{TCCON}$ has dimensions of $n_w n_s$ by 51, and $\boldsymbol{A}_{vert}$ has dimensions of $2n_s$ by 51. The gain matrix relates each measurement in a day to the upper and lower partial column calculation which is useful to calculate the temporal DoF but is not directly comparable to in situ vertical profiles. The $\boldsymbol{A}_{vert}$ term converts the temporal sensitivities of the gain matrix to vertical sensitivities using the TCCON vertical averaging kernel allowing us to compare with the in situ validation profiles. We apply the average vertical sensitivities for the measurements used in comparison with in situ profile measurements.

Since ~~as the TARDISS fit. We~~ $\boldsymbol{a}_{TCCON}$ represents the change in TCCON total column DMF (also called X_{gas}) per change in true DMF at each level ($\frac{\delta X_{gas,TCCON}}{\delta x_{true}}$) and the gain matrix represents the change in partial column scalar per change in TCCON total column DMF ($\frac{\delta \gamma}{\delta X_{gas,TCCON}}$), $\boldsymbol{A}_{vert}$ has units of change in partial column scalar per change in level DMF value ($\frac{\delta \gamma}{\delta x_{true}}$) and relies on the difference between a 'true' in situ profile and the ~~use the terms retrieval and fit interchangeably to refer to the P1 or TARDISS methodology~~a priori profile used in the inversion.

For our TARDISS comparisons, we use an adjusted version of Equation 20 to determine the value the inversion would return if it were using the true profile instead of the scaled TCCON priors:

$$\boldsymbol{z}_s = z_{a,TCCON} + \boldsymbol{A}_{vert}\left(\boldsymbol{x}_{true} - \boldsymbol{x}_{a,TCCON}\right), \quad (24)$$

where $\boldsymbol{x}_{a,TCCON}$ is the a priori profile used in Equation 3 and $\boldsymbol{x}_{true}$ is the measured in situ profile in DMF. The in situ profile is interpolated to the same vertical levels as the TCCON a priori profile as shown in Fig. 4. After calculating the smoothed in situ profile, we integrate the profile from the surface to the vertical level at which the partial columns are separated, q in Equation 5, for the lower column. For the upper partial column, we integrate from the level q+1 to the top of the atmosphere for the upper column using the method outlined in Wunch et al. (2010). We then compare the integrated, smoothed, in situ partial column DMFs directly with the reconstructed lower and upper partial columns calculated by Equation 17.

2.2.5 Error Calculations

Finally, the error for the retrieval is made up of model parameter error, smoothing error, and retrieval noise (Rodgers, 2008). There are no model parameters in the state vector of the TARDISS retrieval, so the model parameter error is zero. The smoothing error is the diagonal of the following matrix:

$$\boldsymbol{S_s} = (\boldsymbol{K}^T\boldsymbol{S_\epsilon}^{-1}\boldsymbol{K} + \boldsymbol{S_a}^{-1})^{-1}\boldsymbol{S_a}^{-1}(\boldsymbol{K}^T\boldsymbol{S_\epsilon}^{-1}\boldsymbol{K} + \boldsymbol{S_a}^{-1})^{-1} \quad , \quad (25)$$

and the retrieval noise is the diagonal of the matrix calculated by:

$$\boldsymbol{S_r} = (\boldsymbol{K}^T\boldsymbol{S_\epsilon}^{-1}\boldsymbol{K} + \boldsymbol{S_a}^{-1})^{-1}\boldsymbol{K}^T\boldsymbol{S_\epsilon}^{-1}\boldsymbol{K}(\boldsymbol{K}^T\boldsymbol{S_\epsilon}^{-1}\boldsymbol{K} + \boldsymbol{S_a}^{-1})^{-1} \quad , \quad (26)$$

and the sum of the two are the total error for the fit.

In order to report an error for our retrieval that reflects the performance of the retrieval in the validation comparisons in Section 3.1, the retrieval output errors are multiplied by a scalar calculated from the 1-to-1 comparisons. Using the multiplier ensures that we are reporting a conservative estimate of the error in the retrieval. We use the 1-to-1 comparisons to scale our error values to the point where at least 50% of the comparison points are within the one standard deviation error range of the 1-to-1 line. We calculate the scalar values as:

$$VEM = Median(\frac{|\boldsymbol{z}_{comp} - \boldsymbol{z}_s|}{\boldsymbol{\sigma}}) \quad (27)$$

where $\boldsymbol{z}_{comp}$ is the comparison partial column values, $\boldsymbol{z}_s$ is the integrated, smoothed, in situ partial column values, $\boldsymbol{\sigma}$ is the output retrieval errors, and VEM is the calculated validation error multiplier that is unitless. The VEM is calculated and applied to all retrieved errors for each site so that the retrieved dataset for a site reflects the best representative error values. If a calculated VEM is less than one, we use a VEM of one instead to avoid spuriously reducing error values. A complete discussion of the retrieval error is in Section 3.4.2.

2.3 Algorithm Setup and Choices

2.3.1 Pre-processing of the TCCON Data

We begin by preprocessing the TCCON fits. We take the TCCON a priori profile and scale it by the median value of the TCCON output scalar values for each spectrum from the windows used so that our TARDISS fit is centered around the median TCCON a posteriori profile for each measurement point. The a posteriori errors from each window are not included in this calculation but are included in the formation of the measurement covariance matrix. This assumes that the true column-averaged VMR of a species is some linear combination of the VMRs calculated from the windows used in the TARDISS fit. Then, we calculate the a priori partial columns by integrating the scaled a priori profiles over the respective pressure levels for the upper and lower partial column. Finally, we assemble the necessary matrices for the fit described by Equation 16.

2.3.2 Maximum a Posteriori Components

The different components of Equation 16 reflect where a priori information can be used in the algorithm and several additional choices can be made to improve the fit. The following describes our standard input for these components. We present tests of the retrieval's sensitivity to these choices in Sect. 3.2.

For the a priori covariance matrix, $\boldsymbol{S}_a$, we use an identity matrix for the lower partial column scalar portion of the covariance matrix, and we use an exponential decay over the day of measurements from the diagonal for the upper partial column scalar portion of the covariance matrix. This requires that upper column scalar values shift in relation to one another and prevents the upper partial column scalars changing too rapidly in time. The off-diagonal values of the upper partial column portion of the a priori covariance matrix decay with respect to the measurements made before and after them over the course of one-third of a day of measurement. We assume no correlation between the upper and lower partial columns, although this is a place for future study. Since the a priori covariance matrix is inverted in the calculations, decreasing the magnitude of the a priori covariance matrix scalar increases the constraints imposed during the calculations so that a scalar of 10^{-5} is a more strict constraint than a scalar of 10^{-4}. A discussion of the influence of the temporal covariance is in Sect. 3.4.1.

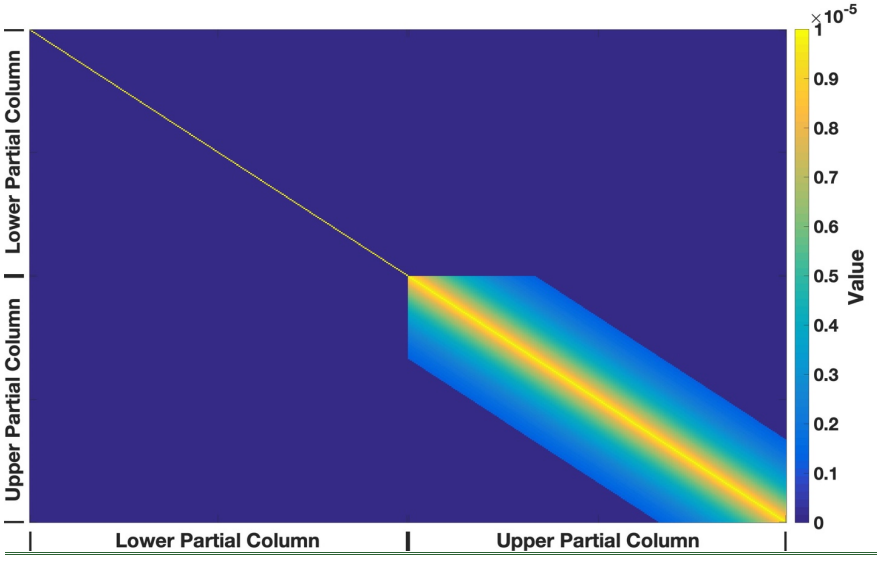


Figure 3. Example of an a priori covariance matrix color coded by the magnitude of the value. The axes represent the relationship of the contribution of each measurement to each partial column and each other measurement. The upper right and lower left quadrants are dark blue and represents zero assumed correlation between the upper and lower partial columns over a day of measurements. The diagonal is scaled to constrain the fit and the lower right quadrant shows the assumed correlation between upper partial column scalar values over a day of measurement. The lower partial column has an a priori covariance that is a scaled identity matrix, the upper partial column has an a priori covariance that decays over one third of the measurement day, and the cross covariances between the upper and lower partial columns are assumed to be zero.

The measurement error covariance matrix, $\boldsymbol{S}_{\epsilon}$, is a diagonal matrix composed of the squares of the TCCON errors for each spectral window so that measurements with smaller errors are weighted more heavily than those with larger errors.

CO_2 and CO use different values for the a priori TARDISS scale factors ($\boldsymbol{x}_{a,\gamma}$). For CO, we assume a uniform a priori scale factors of one for all observations. For CO_2 we use the solution to the least-squares method, $\boldsymbol{x}_{L2}$ from Equation 15 as $\boldsymbol{x}_{a,\gamma}$ in Equation 16. We adopted different approaches for these two gases since using a static a priori partial column scalar of one for the CO_2 retrievals worsened the comparison to in situ data but improved the validation comparison for the CO retrievals (shown in Sect. 3.2).

2.3.3 Choosing Spectral Windows for the TARDISS Fit

The primary information content used in our algorithm is derived from the fact that the total column abundances retrieved from different spectral windows of the same species will differ due to differences in their kernels unless the shape of the a priori profile is perfect. Accordingly, for this method to have sufficient information, windows with different vertical averaging kernels are needed, such as those shown in Fig. 1. Preferably, the TARDISS retrieval would use a window that is more sensitive to the lower atmosphere and a window that is more sensitive to the upper atmosphere so that a larger amount of information is contained between them. While it is imperative to use windows that have differing averaging kernel profiles, it is also necessary to use windows that have sufficiently low error in the TCCON fit.

For the partial column CO_2 calculations, we use four spectral windows in the TCCON process centered at 6339, 6220, 4852, and 6073 cm^{-1} which were suggested for profile retrieval exploration by Connor et al. (2016). The 6339 cm^{-1} and 6220 cm^{-1} windows are spectroscopically similar and have column averaging kernel profiles that vary with solar zenith angle providing some vertical information over the course of a day (see Fig. 1). The 4852 cm^{-1} window has an averaging kernel profile that is

largest at the surface and smallest at the upper troposphere and the 6073 cm^{-1} window has an averaging kernel profile that is effectively the opposite of the 4852 cm^{-1} window. Both the 4852 cm^{-1} and 6073 cm^{-1} window averaging kernels are largely independent of solar zenith angle with the exception of the highest levels in the 6073 cm^{-1} window profile.

For the partial column CO calculations, we use three spectral windows fit during the TCCON process. There is one window in the NIR region centered at 4233 cm^{-1} and two windows in the MIR region centered at 2111 and 2160 cm^{-1}. The two MIR windows have similar averaging kernel profiles that maximize at the surface and drop to nearly zero at upper levels. The NIR window averaging kernel profile has a minimum at the surface and a maximum at the upper levels.

Unlike the CO_2 windows that are all observed by the InGaAs detector, the MIR CO windows are measured by a liquid nitrogen cooled InSb detector. For this reason, we only have results of the CO partial column fits at the Caltech, Lamont, and East Trout Lake TCCON sites and, due to the lack of in situ profiling data in Pasadena, we only have direct vertical profile comparison results from the Lamont and East Trout Lake TCCON site.

Other windows output by TCCON retrievals were considered for the partial column calculations for both species. However, they had high levels of error in the TCCON fit or had fits that were particularly sensitive to changes in temperature.

2.3.4 Choice of Partial Column Height

We chose the lower partial column to integrate from the surface through the first five vertical layers of the GEOS meteorological fields. Using this criterion, a site at sea level has a lower partial column from sea level to 2 km and the upper partial column from 2 to 70 km. While somewhat arbitrary, the choice of 2 km was made to have the lower partial column encompass the surface mixed layer at most locations while minimizing the dilution of surface exchange signals that would result from integrating over a larger partial column. If there are known significant enhancements species enhancement near the 2 km level (such as CO during wildfire events), the retrieval performance may be degraded, and a different partial column height may be a more appropriate choice.

2.4 Sites Used in this Work

In this study, we use data from the five TCCON sites located across ~~the United States~~North America. The data record extends from as early as ~~2004~~2011 to as recent as 2021~~.~~ (Table 1). These sites are located at Park Falls, Wisconsin; NASA Armstrong, Edwards Air Force Base, California; Lamont, Oklahoma (the ~~DOE~~Department of Energy Southern Great Plains ~~ARM~~Atmospheric Radiation Measurement site), the California Institute of Technology (Caltech), in Pasadena, California, and East Trout Lake, Saskatchewan~~, Canada. Table 1 presents a summary of the sites used in this work.~~.

Park Falls, WI hosts the first operational TCCON site (July 2004-present). The site is in a rural, heavily forested area and generally far from anthropogenic influence. The FTS does not have an InSb detector, so we are able to only retrieve partial column values for CO_2. We focus on data obtained since 2012, when the alignment of the instrument has been more consistent. The increased variance of the TARDISS retrieval for data before 2012 likely reflects the inconsistent alignment of the FTS.

We use similar data from the TCCON site located at NASA's Armstrong Flight Research Center (formerly the Dryden Flight Research Center) in California which has been operational since July 2013. We report CO_2 partial column values for the 2013 to 2021 time period. The Armstrong site is on the northwest edge of Rogers Dry Lake within the Edwards Air Force Base in the Mojave Desert.

The Lamont, OK TCCON site is surrounded by farmland. It has been operational since July 2008, and an InSb detector was installed in October 2016. We focus on data from Lamont obtained after 2011 after an issue with the instrument laser was resolved. We report CO_2 partial column values from ~~2008~~2011 to 2021 and CO partial column values from 2017 to 2021.

The TCCON site on the Caltech campus in Pasadena, CA has been operational since July 2012 with an InSb detector measuring since October 2016. We report CO_2 partial column values from 2012 to 2021 and CO partial column values from 2017 to 2021.

The East Trout Lake, ~~Sask., CA~~SK, TCCON site is located in a remote, heavily forested area in the middle of the Saskatchewan Province. The instrument uses an InSb detector ~~so~~ allowing us to retrieve partial column CO values. It has been operational since October 2016, and we report partial column values for CO and CO_2 from 2017 to 2021.

Site	Location	Dates of Measurements Used	Data DOI
Park Falls, WI	45.945N, 90.273W	CO_2: 2012 - 2021	10.14291/tccon.ggg2020.parkfalls01.R0
NASA Armstrong, Edwards Air Force Base, CA	34.958N, 117.882W	CO_2: 2013 - 2021	10.14291/tccon.ggg2020.edwards01.R0
Lamont, OK	36.604N, 97.486W	CO_2: ~~2008~~2011 - 2021 CO: 2017- 2021	10.14291/tccon.ggg2020.lamont01.R0

Caltech, Pasadena, CA	34.1362N, 118.126W	CO_2: 2012 - 2021 CO: 2017 - 2021	10.14291/tccon.ggg2020.pasadena01.R0
East Trout Lake, ~~Sask., CA~~SK	54.354 N, 104.987W	CO_2: 2017 – 2021 CO: 2017 – 2021	10.14291/tccon.ggg2020.easttroutlake01.R0

Table 1. Location, dates of measurement, and DOIs of the TCCON sites used in this work. CO measurements require an InSb detector to cover the 2160 and 2111 cm^{-1} windows, which has only been available since 2017 at Caltech, Lamont, and East Trout Lake.

~~2.2 The TARDISS Algorithm~~

~~Traditional profile retrievals fit spectra by adjusting the abundance of the trace gases at multiple vertical levels to determine the vertical distribution of a specific atmospheric species. Here, we focus on developing an algorithm that we are calling the Temporal Atmospheric Retrieval Determining Information from Secondary Scaling (TARDISS), that optimizes separating the profile of our target trace gas into two layers, one near the surface and the other at and above the middle troposphere (two 'partial columns', here scaled prior DMFs).~~

~~We use the notation and concepts of Rodgers and Connor (2003) with vectors represented with bolded lower-case letters and matrices represented with bolded upper-case letters. The following equations are used for each spectral window, each TCCON measurement, and each species retrieved (CO and CO_2 in this work) in the P1 fit. Equations 1 through 7 are used to calculate the weights and values that are used in Equation 8 and beyond as we shift the focus from one measurement of one spectral window to including all the measurements in a day of all the spectral windows used for a particular species. We will therefore keep the equations species and window agnostic for this description. We start with an equation expressing the calculation of the total column value:~~

$$z_{p1} = z_{a,p1} + \boldsymbol{a}_{p1}^{T}(\boldsymbol{x}_{part} - \boldsymbol{x}_{a,p1}) \quad (1)$$

~~Where $\boldsymbol{z_{p1}}$ is the total column DMF output of a chosen species in a particular window from the P1 fit, $z_{a,p1}$ is the original vertical column DMF calculated from the prior profile scaled by the median VSF of the windows used, $\boldsymbol{a_{p1}}$ is the vector of column averaging kernel values output from the P1 TCCON processing weighted by the pressure thickness of each atmospheric layer, $\boldsymbol{x_{a,p1}}$ is the prior profile for the chosen species also scaled by the median VSF of the windows used, and $\boldsymbol{x_{part}}$ is partial column DMF to be retrieved for the chosen species. All components in Equation 1 are in dry mole fractions and the~~

averaging kernel is unitless. Equation 1 tells us how the retrieved DMF would change if the profile constructed from the two partial columnns differed from $x_{a,p1}$.

Focusing on the rightmost term of Equation 1, the averaging kernel is multiplied by the difference of the prior and unknown DMF profiles summed for each level of the atmosphere.

$$z_{p1} - z_{a,p1} = \boldsymbol{a}_{p1}^{T}\left(\boldsymbol{x}_{part} - \boldsymbol{x}_{a,p1}\right) = \sum_{i=1}^{n} a_{p1,i}\left(x_{part,i} - x_{a,i}\right) \quad (2)$$

Our method splits the total column at a specified altitude, q, and scales the prior profile below and above q independently2. such that:

$$z_{p1} - z_{a,p1} = \sum_{i=1}^{q} a_{p1,i}\left(\gamma_L x_{a,p1,i} - x_{a,p1,i}\right) + \sum_{i=q+1}^{n} a_{p1,i}\left(\gamma_U x_{a,p1,i} - x_{a,p1,i}\right) \quad (3)$$

where $\boldsymbol{\gamma_L}$ and $\boldsymbol{\gamma_U}$ are the lower partial column scaling factor and upper partial column scaling factor, respectively. As this is linear, we group terms reducing the right side of Equation 3 to:

$$z_{p1} - z_{a,p1} = (\gamma_L - 1)\sum_{i=1}^{q} a_{p1,i}\, x_{a,i} + (\gamma_U - 1)\sum_{i=q+1}^{n} a_{p1,i}\, x_{a,i} \quad (4)$$

Or further to:

$$z_{p1} - z_{a,p1} = (\gamma_L - 1)\, J_L + (\gamma_U - 1)\, J_U \quad (5)$$

Where,

$$J_L = AW \sum_{i=1}^{q} a_{p1,i}\, x_{a,p1,i} \quad (6)$$

and

$$J_U = AW \sum_{i=q+1}^{n} a_{p1,i}\, x_{a,p1,i} \quad (7)$$

$\boldsymbol{J_L}$ and $\boldsymbol{J_U}$ both reduce to scalar values for each spectral window and prior profile. The AW term is a weighting term based on daily anomaly values between individual windows and is referred to as window

weighting from here on. We discuss the choices and reasoning for the AW term in Sect. 3.1.1 and Appendix A.

Equation 5 is applicable to all spectral windows for each spectrum measured. For example, for our CO_2 retrieval we use four separate spectral windows per measured spectrum and often have a few hundred spectra measured per day.

Our TARDISS retrieval uses an entire day's worth of TCCON retrievals in order to utilize the information in the temporal dimension. We combine the above equations into a matrix form:

$$\boldsymbol{y} = \boldsymbol{K}\boldsymbol{x}_{\boldsymbol{y}} + \boldsymbol{\epsilon} \quad (8)$$

Where $\boldsymbol{y}$ is a vector of values from the left side of Equation 5 for the α number of windows and k number of spectra over a day,

$$\boldsymbol{y} = \begin{bmatrix} (\boldsymbol{z_{p1}} - \boldsymbol{z_{a,p1}})_{w1,1} \\ \vdots \\ (\boldsymbol{z_{p1}} - \boldsymbol{z_{a,p1}})_{w1,k} \\ \vdots \\ (\boldsymbol{z_{p1}} - \boldsymbol{z_{a,p1}})_{w\alpha,1} \\ \vdots \\ (\boldsymbol{z_{p1}} - \boldsymbol{z_{a,p1}})_{w\alpha,k} \end{bmatrix} \quad (9)$$

K is the matrix of the $\boldsymbol{J_L}$ and $\boldsymbol{J_U}$ values for the α number of windows and k number of spectra over a day,

$$\boldsymbol{K} = \begin{bmatrix} J_{L,w1,1} & \cdots & 0 & J_{U,w1,1} & \cdots & 0 \\ \cdots & \ddots & \cdots & \cdots & \ddots & \cdots \\ 0 & \cdots & J_{L,w1,k} & 0 & \cdots & J_{U,w1,k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ J_{L,w\alpha,1} & \cdots & 0 & J_{U,w\alpha,1} & \cdots & 0 \\ \cdots & \ddots & \cdots & \cdots & \ddots & \cdots \\ 0 & \cdots & J_{L,w\alpha,k} & 0 & \cdots & J_{U,w\alpha,k} \end{bmatrix} \quad (10)$$

and x_γ is our state vector of partial column scalars which are the same for all windows in each measured spectrum.

$$x_\gamma = \begin{bmatrix} (\gamma_L - 1)_1 \\ \vdots \\ (\gamma_L - 1)_k \\ (\gamma_U - 1)_1 \\ \vdots \\ (\gamma_U - 1)_k \end{bmatrix} \quad (11)$$

With k measurements made in a day, four spectral windows, and two partial columns, the y vector is of the size 4k by 1, the x_γ vector is of the size 2k by 1, and the K matrix is of the size 4k by 2k.

Fitting over an entire day of TCCON retrievals reduces the retrieved partial column error values compared to fitting individual measurements using Equation 5. Appendix B shows the influence of including multiple observations on the retrieved partial column errors.

We use the maximum a posteriori (MAP) approach (Rodgers 2008) to calculate the most probable state vector from the given models and prior information. In line with the assumptions of the MAP approach, we assume our problem is linear and follows a gaussian distribution. The MAP solution can take a few equivalent forms. In this work we use:

$$x_\gamma = x_{a,\gamma} + S_a K^T (K S_a K^T + S_\epsilon)^{-1} (y - K x_{a,\gamma}) \quad (12)$$

Where $x_{a,\gamma}$ is the prior partial column scalar values, S_a is the prior covariance matrix, K is the forward mapping matrix, S_ϵ is the model covariance matrix, y is the measurement vector, and x_γ is the output solution vector. The input components ($x_{a,\gamma}$, S_a, and S_ϵ) are described in Sect. 2.3.2.

Once we have calculated the most likely solution for the partial column scalars, x_γ, we reconstruct the partial column DMF for the lower and upper partial columns as:

$$x_{pc} = x_\gamma x_{a,p1} + x_{a,p1} \quad (13)$$

where $x_{a,p1}$ is the prior partial column DMF calculated by integrating the median P1 posterior profile using the same method as the standard TCCON full column retrievals (Wunch et al., 2011).

The MAP retrieval allows us to calculate the information content of the retrieval. In particular, we compare the degrees of freedom for our retrieval calculated by taking the trace of the averaging kernel of the fit, calculated as the following:

$$DoFs = tr(A) = tr((K^T S_\epsilon^{-1} K + S_a^{-1})^{-1} K^T S_\epsilon^{-1} K) \quad (14)$$

as well as the Shannon information content derived from the natural log of the determinant of the difference between the averaging kernel and an identity matrix:

$$H = -\frac{1}{2}\ln(|\boldsymbol{I} - \boldsymbol{A}|) \quad (15)$$

Generally, profile retrieval averaging kernels represent the sensitivity of a specific level of a profile to the rest of the levels in the profile. The averaging kernel for the TARDISS inversion is primarily a temporal averaging kernel relating how each partial column calculation relates to every other measurement during a day. We normalize the degrees of freedom by the number of measurements in each day for a more comparative understanding of the TARDISS degrees of freedom with respect to a traditional profile retrieval as well as between days with a large variation in the number of measurements.

In order to compare the partial column retrievals to the in situ profiles for validation purposes, we calculate the vertical sensitivities of the TARDISS fit (shown in Fig. 4) using the gain matrix, **G**, from the TARDISS inversion and the averaging kernel profiles from the P1 measurement windows as:

$$\boldsymbol{G} = (\boldsymbol{K}^T\boldsymbol{S}_\epsilon^{-1}\boldsymbol{K} + \boldsymbol{S}_a^{-1})^{-1}\boldsymbol{K}^T\boldsymbol{S}_\epsilon^{-1} \quad (16)$$

$$\boldsymbol{A}_{vert} = \boldsymbol{x}_{a,P1} * \boldsymbol{G} * \boldsymbol{a}_{P1} \quad (17)$$

where $\boldsymbol{a}_{P1}$ is the same vector of column averaging kernels from Equation 1.

Since $\boldsymbol{a}_{P1}$ represents the change in TCCON X_{gas} DMF **per change in true DMF at each level** $\left(\frac{\delta x_{gas,TCCON}}{\delta x_{true}}\right)$ **and the gain matrix represents the change in partial column scalar per change in** TCCON X_{gas} **DMF** $\left(\frac{\delta y}{\delta x_{gas,TCCON}}\right)$**, $\boldsymbol{A}_{vert}$ has units of change in partial column scalar per change in level DMF value** $\left(\frac{\delta y}{\delta x_{true}}\right)$ **and relies on the difference between a 'true' in situ profile and the** prior profile used in the inversion.

2.3 Algorithm Setup and Choices

2.3.1 Pre-processing of the Phase 1 Data

We begin by preprocessing the P1 fits. We take the P1 model prior profile and scale it by the median value of the P1 output scalar values for each spectrum from the windows used so that our TARDISS fit is centered around the median P1 posterior profile for each measurement point. This assumes that the true column VMR of a species is some linear combination of the VMRs calculated

from the windows used in the TARDISS fit. Then, we calculate the partial column priors by integrating the scaled prior profiles over the respective pressure levels for each chosen partial column. Finally, we assemble the necessary matrices for the fit described by Equation 12.

2.3.2 Maximum a Posteriori Components

The different components of Equation 12 reflect where prior information can be used in the algorithm and several additional choices can be made to improve the fit. The following describes our standard input for these components. We present tests of the retrieval's sensitivity to these choices in Sect. 3.1.1.

For the prior covariance matrix, we use an identity matrix for the lower partial column scalar portion of the covariance matrix, and we use an exponential decay from the diagonal for the upper partial column scalar portion of the covariance matrix. This requires that upper column scalar values shift in relation to one another and theoretically imposes higher costs if the upper partial column scalars change rapidly in time. The off-diagonal values of the upper partial column portion of the prior covariance matrix decay with respect to the measurements made before and after them over the course of one-third of a day of measurement. Since the prior covariance matrix is inverted in the calculations, decreasing the magnitude of the prior covariance matrix scalar increases the constraints imposed during the calculations so that a scalar of 10^{-5} is a larger constraint than a scalar of 10^{-4}. A discussion of the influence of the temporal covariance is in Sect. 3.1.3.

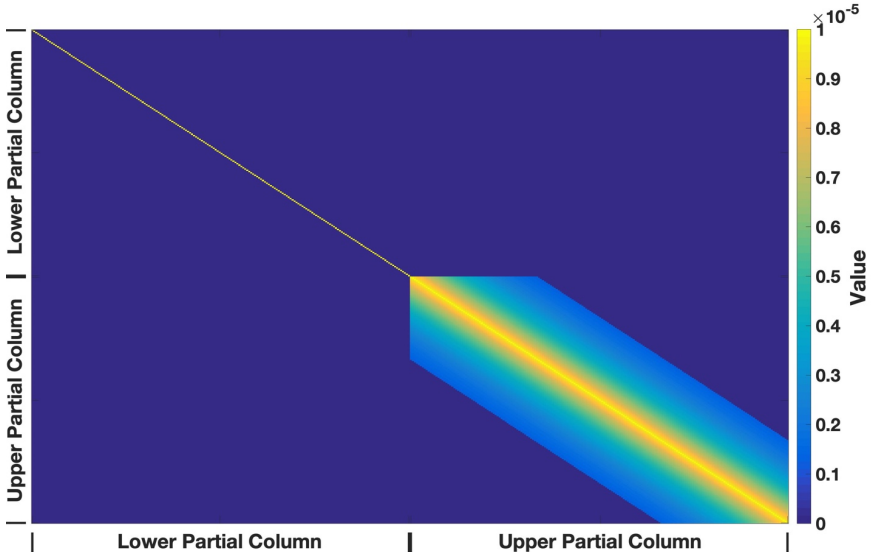


Figure 2. Example of a prior covariance matrix color coded by the magnitude of the value. The lower partial column has a prior covariance that is a scaled identity matrix, the upper partial column has a prior covariance that decays over one third of the measurement day, and the cross covariances between the upper and lower partial columns is assumed to be zero.

The measurement error covariance matrix is a diagonal matrix composed of the squares of the P1 errors for each spectral window so that measurements with smaller errors are weighted more heavily than those with larger errors.

CO_2 and CO use different values for the prior TARDISS scale factors ($\boldsymbol{x}_{a,\gamma}$). For CO, we assume a uniform prior scale factor of one for all observations. For CO_2 we solve Equation 12 using the linear least-squares method:

$$\boldsymbol{x}_{L2} = (\boldsymbol{K}^T\boldsymbol{K})^{-1}\boldsymbol{K}^T\boldsymbol{y} \quad (18)$$

and use the median daily value of $\boldsymbol{x}_{L2}$ as $\boldsymbol{x}_{a,\gamma}$ in Equation 12. While the linear least-squares method provides a solution to our retrieval, it does not allow us to utilize additional prior information in the

covariance of the partial columns or to specify prior partial column scale factors. Including these pieces of information reduces the retrieved partial column error values as shown in Fig. B. in Appendix B.

We adopted different approaches for these two gases since using a static prior of one for the CO_2 retrievals worsened the comparison to in situ data but improved the validation comparison for the CO retrievals (shown in Sect. 3.1.1).

2.3.3 Choosing Spectral Windows for the TARDISS Fit

The primary information content used in our algorithm is derived from the fact that the total column abundances retrieved from different spectral windows of the same species will differ due to differences in their kernels, unless the shape of the a priori profile is perfect. Thus, differences in the retrieved columns from different windows, together with their kernels, can be used to infer the errors in the a priori VMR profile, and hence derive a better VMR profile than one which is determined by simply scaling the a priori VMR profile. For this method to have sufficient information, windows with different vertical averaging kernels are needed, such as those shown in Fig. 1. Preferably, the windows used for the TARDISS retrieval would have a window that is more sensitive to the lower atmosphere and a window that is more sensitive to the upper atmosphere so that a larger amount of information is contained between them. While it is imperative to use windows that have differing averaging kernel profiles, it is also necessary to use windows that have low error in the P1 fit. The higher the error in a particular spectral window, the more uncertainty that that retrieval will add to the TARDISS results.

For the partial column CO_2 calculations, we use four spectral windows in the P1 process centered at 6339, 6220, 4852, and 6073 cm^{-1}. The 6339 cm^{-1} and 6220 cm^{-1} windows are spectroscopically similar and have column averaging kernel profiles that vary with solar zenith angle providing some vertical information over the course of a day. The 4852 cm^{-1} window has an averaging kernel profile that is largest at the surface and minimal at upper troposphere and lower stratosphere and the 6073 cm^{-1} window has an averaging kernel profile that is effectively the opposite of the 4852 cm^{-1} window. Both the 4852 cm^{-1} and 6073 cm^{-1} window averaging kernels are largely independent of solar zenith angle with the exception of the highest levels in the 6073 cm^{-1} window profile.

For the partial column CO calculations, we use three spectral windows fit during the P1 process. There is one window in the NIR region centered at 4233 cm^{-1} and two windows in the MIR region centered at 2111 and 2160 cm^{-1}. The two MIR windows have similar averaging kernel profiles that maximize at the surface and drop to nearly zero at upper levels. The NIR window averaging kernel profile has a minimum at the surface and a maximum at the upper levels.

Unlike the CO_2 windows that are all observed by the InGaAs detector, the MIR CO windows are measured by a liquid nitrogen cooled InSb detector. For this reason, we only have results of the CO partial column fits at the Caltech, Lamont, and East Trout Lake TCCON sites and, due to the lack of in

~~situ profiling data in Pasadena, we only have direct vertical profile comparison results from the Lamont and East Trout Lake TCCON site.~~

~~Other windows measured by TCCON instruments were considered for the partial column calculations for both species however they had high levels of error in the P1 fit, had fits that were particularly sensitive to changes in temperature, or their averaging kernels were similar enough to the other windows that they would not provide any extra information while still increasing the error.~~

~~2.3.4 Choice of Partial Column Height~~

~~We chose the lower partial column to integrate from the surface through the first five vertical layers of the GEOS meteorological fields. Using this criterion, a site at sea level has a lower column from sea level to 2 km and the upper partial column from 2 to 70 km. While somewhat arbitrary, the choice of 2 km was made to have the lower partial column encompass the surface mixed layer at most locations while minimizing the dilution of surface exchange signals that would result from integrating over a larger partial column. If there are known enhancements of significant species enhancement near the 2 km level (such as CO during wildfire season), the retrieval performance may be degraded and a different partial column height may be a more appropriate choice.~~

~~2.4~~ Comparison Data

~~2.4.1 In situ Vertical Profile Data~~

We use in situ data from multiple aircraft programs and AirCore ~~campaigns~~flights between 2008 and 2020 (Cooperative Global Atmospheric Data Integration Project; (2019~~)~~), Baier et al., 2021) to evaluate our partial column retrieval. ~~All in situ CO_2 data are reported on the WMO X2007 scale.~~

The aircraft ~~data are from the SEAC4RS campaign in 2013, the ATom and KORUS-AQ campaigns in 2016, and from measurements made by the Goddard Space Flight Center between 2014 and 2016. We use AirCore measurements taken in July of 2018 at the Armstrong, Lamont, and Park Falls sites. In addition, we use measurements from the NOAA Global Greenhouse Gas Reference Network's Aircraft Program to compare with our lower column measurements. Table S2 provides a summary of the in situ data used in this work~~CO_2 measurements are from the NASA Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) campaign (Toon et al., 2016) using an AVOCET instrument, from the 2016 Atmospheric Tomography Mission (ATom) (Wofsy et al., 2021; Thompson et al., 2022) using a Picarro cavity ringdown spectroscopy (CRDS) trace gas analyzer (Crosson, 2008), from the Korea-United States Air Quality Study (KORUS-AQ) campaign using a non-dispersive IR spectrometer, and from measurements made by the Goddard Space Flight Center using a Picarro CRDS trace gas analyzer.

The NASA Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) campaign used an AVOCET instrument on the ER-2 aircraft to sample in situ CO_2 in the atmosphere from the surface into the lower stratosphere (Toon et al., 2016). The SEAC4RS measurements align with TCCON measurements at the Armstrong site on 23 September 2013 where the ER-2 sampled from 1.5 km to 19 km altitude.

The Korea-United States Air Quality Study (KORUS-AQ) campaign used a multitude of instruments and platforms to measure and better understand air quality in South Korea and how to improve it (Crawford et al., 2021). One of the measurement platforms was the NASA DC-8 aircraft which made coincident measurements with the TCCON instrument at the Armstrong site on 18 June 2016. Measurements of CO_2 using a non-dispersive IR spectrometer during a flight that sampled from 0.68 km to 12 km are used here for comparison to retrieved partial columns.

The Atmospheric Tomography Mission (ATom) generated a global dataset to study the interactions of anthropogenic air pollution and greenhouse gases from Summer 2016 through Spring 2018 (Wofsy et al., 2021; Thompson et al., 2022). During this mission, the NASA DC-8 aircraft made measurements of trace gases worldwide. On 22 August 2016, their path coincided with the TCCON measurements at the Park Falls site, where they measured in situ CO_2 using a Picarro cavity ringdown spectroscopy (CRDS) trace gas analyzer (Crosson, 2008) from 0.79 km to 12 km.

We use Goddard Space Flight Center aircraft measurements that were made at the Armstrong site on 20 and 22 August 2014, 02 October 2015, and 10 February 2016. Picarro CRDS measurements were made on these flights up to ~13 km and down to 0.6 km. Multiple measurements were made on 20 and 22 August 2014.

The AirCore sampling system is composed of coiled stainless-steel tubing that is open on one end and samples ambient air as it descends from a balloon flight. This sample is then analyzed using a Picarro CRDS trace gas analyzer using an algorithm that accounts for the effect of longitudinal mixing on vertical resolution (Karion et al., 2010; Tans, 2009). We use AirCore data from measurements made on 16, 17, and 18 July 2018 at the Armstrong site; on 23, 25, and 27 July 2018 at the Lamont site, and on 30 and 31 July 2018 at the Park Falls site.

Finally, we use CO and CO_2 data measured at the Lamont site (site code SGP) and at the East Trout Lake site (site code ETL) as a part of the NOAA Global Greenhouse Gas Reference Network's Aircraft Program to measure the seasonal climatology of greenhouse gases in North America (Sweeney et al., 2015). We use data from 282 of the 399 flights made between 2008 and 2018 and all 34 flights for CO made between 2017 and 2020 at the SGP site. At the ETL site we use 26 of the 40 flights for CO_2 and 26 of the 39 flights for CO made between 2017 and 2021. Most of the measurements were made below 6 km altitude at the SGP site and below 7 km altitude at the ETL site, with some measurements as low as 0.17 km altitude. We use data from individual flights that have more than 2 measurements points within our 1.9 km lower partial column and at least one point at or below 1 km. Since these datasets do not include much data within the upper partial column, we compare with these

~~measurements only to our retrieved lower partial column values and exclude them from the validation discussion in Section 3.1.~~

~~2.4.2 Ground-based Data~~

~~In Section 3.4, we use ground-based data to compare to our lower partial column retrieval output and to begin discussions of the applications of our retrievals.~~

~~Directly next to the Park Falls TCCON site is a 400 m tall tower that is part of the NOAA GGGRN tall-tower and AmeriFlux networks and has multiple vertical levels altitudes along the tower at which measurements are made (Berger et al., 2001; Andrews et al., 2014). The tall tower uses a Licor LI-6262 to measure CO_2, and we use the measurements at the 396 m level to compare to our lower partial column retrievals.~~

~~A Cimel CE-318-N multiband photometer measures aerosol optical depth as a part of the AERONET network from the top of the Caltech Hall roughly 100 m from the Caltech TCCON site (Holben et al., 1998). It measures aerosol optical depth at 340, 380, 440, 500, 675, 870, and 1020 nm as well as atmospheric water vapor in centimeters. The photometer measures in 15-minute intervals and has been measuring at this location since 2010. We compare the level 1.5 cloud screened aerosol measurements to our lower partial column CO measurements to explore the connection between emissions, meteorology, and aerosols in Sect. 3.4.~~

~~3 Results and Discussion~~

~~To understand the effectiveness of our partial column retrieval, we use an adjusted version of the smoothing calculation shown in Equation 3 of Wunch et al. (2010) to determine the value the inversion would return if it were using the true profile instead of the scaled P1 priors:~~

~~$$z_s = z_{a,p1} + A_{vert}(x_{true} - x_{a,p1})$$~~ We use AirCore profiles from July and August of 2018 at the Armstrong, Lamont, and Park Falls sites (Baier et al., 2021). The AirCore sampling system is composed of coiled stainless-steel tubing that is open on one end while ascending on balloon to ~ 30 km, and passively samples ambient air as it descends to the ground on a parachute. This sample is then analyzed for CO_2, CH_4 and CO using a Picarro CRDS trace gas analyzer, and a fill dynamics model accounts for the effect of longitudinal mixing due to diffusion on vertical resolution (Karion et al., 2010; Tans, 2009; Tans, 2022).

Finally, we use CO and CO_2 data measured at the Lamont site (site code SGP) and at the East Trout Lake site (site code ETL) as a part of the NOAA Global Greenhouse Gas Reference Network (GGGRN) aircraft network in North America (Sweeney et al., 2015). Since these datasets do not include much data within the upper partial column, we compare with these measurements only to our

retrieved lower partial column values and exclude them from the validation discussion in Section 3.2. Table S2 provides a summary of the in situ data used in this work.

~~(19)~~

~~where x_a is the prior profile used in Equation 1 and x_{true} is the measured in situ profile converted to DMF. The in situ profile is interpolated to the same vertical levels as the TCCON prior profile as shown in Fig. 3. After calculating the smoothed in situ profile, we integrate from the surface of the profile to the same vertical level at which the partial column is separated, q in Equation 3, for the lower column or that level to the top of the atmosphere for the upper column using the method outlined in Wunch et al. (2010). We then compare the integrated, smoothed, in situ partial column DMF directly to the output from the reconstructed lower and upper partial columns calculated in Equation 13.~~

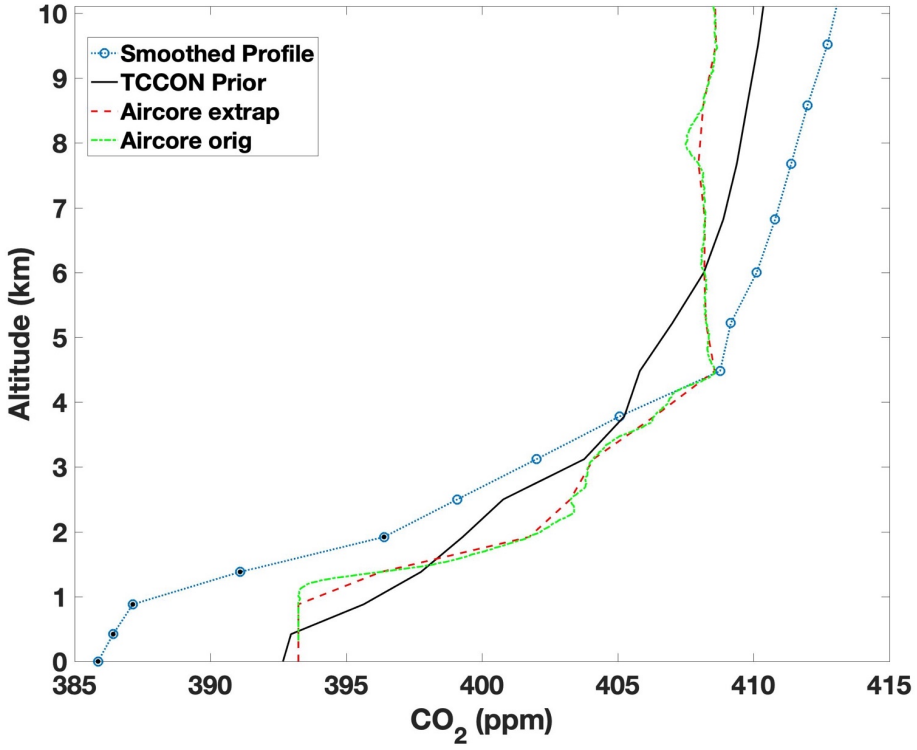


Figure ~~3~~4. An example of the profiles used in the direct comparison calculations using data from the Park Falls site on July 27, 2018. The profile above ~~6~~10 km is not shown. The solid black line is the ~~TARDISS prior~~TCCON a priori profile scaled by the median of the ~~TCCON~~ vertical scaling factors from the spectral windows used. The green dot-dashed line is the measured AirCore ~~mixing ratios.~~mole fraction. The red, dashed line is the AirCore measurements interpolated to the vertical spacing of the ~~TARDISS~~TCCON prior, and the blue, dotted line with circles is the smoothed, vertical sensitivity weighted profile that is integrated to calculate the partial column that ~~our phase 2 fitting~~the TARDISS retrieval would calculate if it had a 'true,' AirCore profile. The black dots within the blue circles represent the points of the profile that make up the lower partial column.

3 Results and Discussion

The TARDISS algorithm is very ~~efficient~~quick – taking only a minute of processing time per year of data for each species~~.~~ – because it does not repeat the spectral fitting. This ~~efficiency~~speed enables the validation comparisons to be performed using many different model choices. Thus, we evaluated the sensitivity of the TARDISS inversion by varying different forward model choices. The set of choices that we have designated as the operational setup for CO_2 inversion are:

- The covariance matrix, $\boldsymbol{S}_a$, is scaled by 10^{-5} to better constrain the fit
- The ~~prior~~value of the a priori scalar for the lower and upper partial column scalar ($\boldsymbol{x}_{a,\gamma}$ in Equation 16) is the ~~daily median of the~~ least squares solution for the respective column ($\boldsymbol{x}_{L2}$ in Equation 15).
- ~~The J values for each window are weighted by the square of the respective daily anomaly ratio of the individual columns (further explained in Appendix A).~~

For the CO inversion, the operational setup parameters are:

- A covariance matrix, $\boldsymbol{S}_a$, scaled by 10^{-5}~~4~~
- An ideal ~~prior~~a priori partial column scalar ($\boldsymbol{x}_{a,\gamma}$) of one~~.~~
- ~~The J values are left unweighted.~~

~~The daily anomaly ratios weightings are only applied to the CO_2 fits as changing the weightings in the CO inversion did not significantly affect the output. We vary three~~
We vary two aspects of the algorithm and observe the differences in the validation comparisons. The results of these tests are discussed in Sect. 3.~~1.1~~2 and represented in Table 2 and Table 3.

3.1 Validation Comparisons

We compare retrieved partial column values from three of the five sites presented in this work using measurements from the same set of in situ data used to evaluate and derive the ~~so-called 'in situ scaling factor' of the P1 TCCON retrievals.~~'in situ scaling factor' of the TCCON retrievals (Wunch et al., 2011). For CO_2, there are twenty-four points of comparison obtained between 2013 to 2018. Twelve of those comparisons are from the Armstrong TCCON site ~~spanning 2013 to 2018 and during different months of the year. Four~~, four profiles are available above the Park Falls TCCON site ~~(August 2016~~, and ~~July 2018). The~~the remaining eight profiles are from the Lamont TCCON site ~~and are all from July 2018~~. As the Lamont site is the only site in this work with an InSb detector and overlapping in situ measurements, the eight profiles measured at the Lamont site serve as the ~~entire~~totality of the CO comparison dataset.

We also compare the partial columns calculated from the TCCON individual windows to further contextualize the performance of the TARDISS algorithm in Sect 3.3.1 and summarized in Table 4. The comparisons of the TCCON individual windows are performed in the same way as the TARDISS comparisons using Equation 20 to calculate the smoothed, in situ partial columns.

The comparison profiles were measured by aircraft-based instruments or AirCore measurements~~,~~ ~~all~~ as described in Sect. 2.~~4.1.~~5 and Table S2. We revert to the TCCON priors for parts of the profile not measured by in situ methods. For the errors associated with the aircraft measurements, we use the reported measurement error for the measured parts of the profile, and, for the unmeasured parts of the profile, we use the average reported measurement error and, to account for the errors involved with estimating the parts of the profiles not measured by in situ methods, we add in quadrature twice the standard deviation of the measured profile in the respective partial column. For the errors associated with the AirCore measurements, we use the same approach as for the aircraft measurement and include an extra error term to conservatively account for atmospheric variability as captured by duplicate AirCores launched at approximately the same time. The error for AirCore from atmospheric variability is 0.6 ppm for CO_2 and 8 ppb for CO compared to the analyzer error of 0.05 ppm and 3 ppb. The partial column error values are calculated by integrating a profile shifted by the error values and subtracting it from the integration of the original smoothed profile. The difference between these two integrated, smoothed partial columns provides a conservative error value that represents the unlikely occurrence that the profile at every altitude has 100% error.

We compare the TARDISS retrievals from spectra obtained within one hour of the ~~of the in situ profile. One benefit of the A_{vert} term in Equation 19 is that it includes the temporal sensitivity of the inversion data which is not able to be taken into account in other remote sensing validations; however, the temporal sensitivity (represented by off-diagonal terms in A_{vert}) is often close to zero so that the retrieved partial column values have little influence on the data before or after them.~~in situ profile to the integrated, smoothed, in situ partial columns calculated using Equation 24. We report linear fits between the partial column retrievals and the integrated, smoothed, in situ partial columns ~~with y-intercepts forced through zero~~. Since our retrieval is designed to be linear, we use fits with y-intercepts forced through zero. As there are only scaling values in our retrieval, a non-zero y-intercept would introduce spurious error into our analysis. Since the reported coefficient of determination ~~(commonly referred to as the r^2 value)~~ for this linear fit would be spuriously high, we ~~report~~take the ~~mean~~ ratio of our retrieved partial column to the integrated, smoothed, in situ measurement and subtract one to quantify how much they deviate from each other. We report the mean of the absolute value of the ratio as it deviates from one as the mean ratio deviation. For example, a one percent difference in values would give a mean ratio deviation of 0.01. This mean ratio deviation value gives a more direct understanding of how the partial column values compare ~~as a lower value signifies a better comparison.~~.

We use these validation comparisons to perform sensitivity tests of our algorithm parameters and determine an operational set of parameters. We then use these optimal parameters for the CO_2 and CO retrievals to quantify the total error of our retrieval by calculating a validation error multiplier for each site. Validation error multipliers for each site and partial column are shown in Table 6.

~~**3.1.1 Sensitivity Analysis**~~

3.2 Choice of Operational Parameters from Validation Comparison

Several terms in our retrieval do not have ~~a single,~~ unambiguously correct ~~choice for their~~ values. To evaluate the sensitivity our retrieval to the choices made for these parameters, we have run our retrieval with alternate ~~choices~~values and report the degrees of freedom and comparison to in situ data (specifically, the retrieval comparison error, slope of the zero-forced linear fit, and the mean ratio deviation value of the linear fit) for each test. We tested changes to ~~three~~two terms: the TARDISS a priori scale ~~factor priors,~~ factors and the a priori covariance matrix scaling~~, and changes to the weightings on the individual windows~~.

To test the sensitivity of the retrieval to the partial column scalar prior, we compare the changes in the validation when using $\boldsymbol{x}_{L2}$ from Equation ~~18~~15 as the ~~prior, the daily median of $\boldsymbol{x}_{L2}$~~a priori partial column scalar (our operational choice for CO_2), the daily median of $\boldsymbol{x}_{L2}$, as well as the idealized scalar of unity (our operational choice for CO) to each other. In Tables 2 and 3, these are identified as "~~Xl2,~~" "~~Xl2~~$\boldsymbol{x}_{L2}$," "$\boldsymbol{x}_{L2}$ daily median," and "static ideal prior," respectively.

We also test the sensitivity of the retrieval to how the ~~prior~~a priori covariance matrix is scaled. ~~Doing so~~This term changes how strongly the retrieval is constrained to the prior. ~~Here,~~In Table 2 (CO_2) and Table 3 (CO), we ~~alter~~illustrate the ~~a priori covariance matrix by scaling the matrix uniformly to evaluate the optimal weighting with respect to the validation comparisons. We report the results~~influence of ~~using scaling values of~~choosing $1x10^{-4}$, $5x10^{-5}$, and $1x10^{-5}$ as an a priori covariance matrix scalar. While other scaling values were tested, the resulting errors were large enough or the resulting degrees of freedom were small enough, that the values were disregarded from further study.

TARDISS a priori Choice	A priori Covariance Matrix Scaling	DoF per measurement (overall)	Lower Column Error (ppm)	Lower Column Validation Slope	Lower Column Mean Ratio Deviation	Upper Column Error (ppm)	Upper Column Validation Slope	Upper Column Mean Ratio Deviation
	10^{-5}	0.046 (2.12)	1.146	1.004	0.008	0.497	0.999	0.003
$\boldsymbol{x}_{L2}$ daily median	10^{-4}	0.311 (15.1)	3.063	1.006	0.010	1.033	0.999	0.003

	$5x10^{-5}$	0.183 (8.48)	2.378	1.005	0.010	0.658	0.999	0.003
x_{L2}	10^{-5}*	0.046 (2.12)	1.146	1.001	0.011	0.497	0.999	0.002
	10^{-4}	0.311 (15.1)	3.063	1.004	0.011	1.033	1.000	0.002
	$5x10^{-5}$	0.183 (8.48)	2.378	1.003	0.009	0.658	1.000	0.002
Static ideal prior	10^{-5}	0.046 (2.12)	1.146	1.012	0.014	0.497	0.997	0.003
	10^{-4}	0.311 (15.1)	3.063	1.013	0.010	1.033	0.997	0.003
	$5x10^{-5}$	0.183 (8.48)	2.378	1.013	0.013	0.658	0.997	0.003

Table ~~Finally, we test the sensitivity to scaling the window weighting matrices defined by Equations 6 and 7. In our testing, we scale these matrices (the AW term in Equation 6 and 7) by one, the daily anomaly ratios discussed in Appendix A, and the square of the daily anomaly ratios. We multiply the matrices by these scalars since each spectral window is sensitive to different parts of the atmosphere and therefore returns a different total column X_{gas} value for the same atmospheric state. The ratio of the daily anomalies measured with the different windows used allows us to weight the windows in the inversion based on their observed relationships. Scaling the window weightings for the CO_2 retrieval by the square of the daily anomaly ratios gave the best comparison likely due to the increase of the contribution of the 4852 cm^{-1} window to the retrieval. Scaling the window weightings for the CO retrieval had little effect so we proceed with results retrieved with unscaled window weightings. We show the results of the scaled CO_2 retrieval comparisons in Table 2 and the results of the unscaled CO retrieval comparisons in Table 3.~~

~~For the CO_2 retrievals, the best performance and the operational set of parameters comes from using the daily median of the least squares solution as the prior, scaling the a priori covariance matrix by 10^{-5}, and scaling the window weights by the square of the daily anomaly ratios. The degrees of freedom per measurement changes with the scaling of the prior covariance matrix and varies between 0.364 (16.1 for the overall comparison) for the 10^{-5} scaling increasing to 0.800 (34.5 overall) for the 10^{-4} scaling. The same pattern holds for the errors with values of 0.586 ppm and 0.942 ppm for the 10^{-5} scaling increasing to 1.331 ppm and 1.661 ppm for the 10^{-4} scaling for the upper and lower columns, respectively. The validation slopes change with both the prior covariance matrix scaling and the prior scalar choice.~~ **2.** Variations in CO_2 retrieval upper and lower column validation slopes, upper and lower column mean ratio deviations, upper and lower column comparison errors, and DoF for different TARDISS a priori choices and a priori covariance matrix scaling values. The asterisk in the fourth row indicates that this is the operational set of parameter choices for the CO_2 retrieval.

The agreement between the in situ and TARDISS retrievals for CO and CO_2 change with both the a priori covariance matrix scaling and the a priori scalar choice. As we are trying to determine the parameters that give the best comparison results between the in situ and lower partial column

~~CO_2~~retrieval data specifically, we chose the parameters that resulted in the validation slope closest to one for the lower partial column. ~~The validation~~ For the lower partial column CO_2, the best result (slope ~~closest to one is 1.007 which is the results~~ of 1.001) comes from using the ~~daily median of the~~ $\boldsymbol{x}_{L2}$ values as an a ~~prior~~priori scalar and scaling the ~~prior~~a priori covariance matrix by 10^{-5}. The validation slope for the upper column comparison with these parameters (0.999) is ~~1.004 which is comparable~~similar to values from other parameter choices. ~~While~~ For the lower partial column CO, the best result for the lower column (slope of 0.999) results from the retrieval using a static a priori scalar of one and scaling the ~~prior~~a priori covariance matrix by ~~$5x10^{-5}$ and using a prior scalar of unity optimizes the upper partial column~~ 10^{-4}. Over the two hours of the comparison, ~~it leads to a poorer lower column comparison which is counter to the intention of the algorithm. The full comparison data~~the degrees of freedom are about 2.12 for CO_2 and 3.51 for CO – consistent with between ~~retrieval parameters is shown in Table 2 and the comparison data for the unscaled window weights are shown in Table S1. The poor lower column comparisons~~one and ~~excellent upper column comparisons for the retrievals using no window weightings shown in Table S1 suggest that the window weightings are acting to emphasize the information from the 4852 cm^{-1} window (and theoretically the surface) as intended.~~two DoF per hour of measurements. Since the largest variation in validation slopes in either partial column and ~~for different window weightings~~either species is driven by the change in the ~~prior~~a priori partial column scalar, we posit that the ~~prior~~a priori partial column scalar choice is the most significant parameter in the ~~CO_2~~ retrieval~~.~~ for determining validation slopes while the a priori covariance matrix scaling is the most significant parameter for determining the degrees of freedom of the fit and the retrieval errors.

TARDISS ~~Prior~~A priori Choice	~~Prior~~A priori Covariance Matrix Scaling	DoF per measurement (overall)	Lower Column Error (~~ppm~~ ppb)	Lower Column Validation Slope	Lower Column Mean Ratio Deviation	Upper Column Error (~~ppm~~ ppb)	Upper Column Validation Slope	Upper Column Mean Ratio Deviation
~~X12~~ $\boldsymbol{x}_{L2}$ daily median	10-5~~*~~	0.~~364 (16.1~~010 (0.402)	0.~~94 24~~40	~~1.007~~0 .935	0.~~0100~~ 55	0.~~58 61~~82	1.~~0040~~ 99	0.~~012~~ 100
	10-4	0.~~800 (34.5~~088 (3.51)	1.~~66 1~~334	~~1.012~~0 .938	0.~~0140~~ 52	~~1.33 1~~0.3 70	1.~~0051~~ 22	0.~~012~~ 128

Deleted Cells

	5x10-5			0.700 (30.4047 (1.88)	1.44 40.9 65	1.0100 .937	0.0120 53	1.04 20.3 03	1.00516	0.012 120
x_{L2}	10-5	0.364 (16.1)	0.94 2	1.008 0.010 (0.402)	0.58 6440	0.918	0.076	0.18 2	1.00413	0.012 115
	10-4			0.800 (34.5088 (3.51)	1.66 1334	1.0120 .921	0.0140 74	1.33 10.3 70	1.00413	0.011 142
	5x10-5	0.700 (30.4)		0.047 (1.44488)	1.01 00.9 65	0.0129 20	1.0420 .075	1.00 40.3 03	0.0121 .128	0.134
Static ideal prior	10-5			0.364 (16.1010 (0.402)	0.94 2440	1.0130 .996	0.0130 03	0.58 6182	1.0480 .998	0.005 050
	10-4*			0.800 (34.5088 (3.51)	1.66 1334	1.0160 .999	0.0160 05	1.33 10.3 70	1.00108 1	0.007 081
	5x10-5			0.700 (30.4047 (1.88)	1.44 40.9 65	0.9981 .014	0.0150 04	1.04 20.3 03	1.00007 3	0.006 075

Deleted Cells
Deleted Cells
Deleted Cells
Inserted Cells
Inserted Cells
Inserted Cells
Deleted Cells
Deleted Cells

Table 2. Variations in CO_2 retrieval upper and lower column validation slopes, upper and lower column mean ratio deviations, upper and lower column comparison errors, and DoF for different TARDISS prior choices and prior covariance matrix scaling values. All values shown here are for retrievals using the square of the anomaly weightings. The asterisk in the first row indicates that this is the operational set of parameter choices for the CO_2 retrieval.

For the CO retrievals, the best performance uses a static prior scalar of one, scaling the prior covariance matrix by 10^{-5}, and not scaling the window weights. The degrees of freedom per measurement increases from 0.170 (6.77 overall) to 0.638 (25.3 overall) and the errors increase from 0.174 ppb and 0.413 ppb to 0.402 ppb and 0.861 ppb as we change the scaling of the prior covariance matrix from 10^{-5} to 10^{-4} for the upper and lower columns, respectively. All the tests resulted in similar values of mean ratio deviation and errors under 1 ppb but using a static prior scalar of one and scaling the prior covariance matrix by 10^{-5} resulted in the lower column validation slope that was closest to one (1.001), so we chose this set of parameters as the operational parameters for the CO retrievals. These parameters also had an upper column validation slope of 1.031 which is closer to one than slopes from most other parameter choices.

TARDISS Prior Choice	Prior Covariance Matrix Scaling	DoF per measurement (overall)	Lower Column Error (ppb)	Lower Column Validation Slope	Lower Column Mean Ratio Deviation	Upper Column Error (ppb)	Upper Column Validation Slope	Upper Column Mean Ratio Deviation
Xl2 daily median	10^{-5}	0.170 (6.77)	0.413	0.977	0.024	0.174	1.079	0.078
	10^{-4}	0.638 (25.3)	0.861	0.987	0.014	0.402	1.095	0.094
	$5x10^{-5}$	0.485 (19.3)	0.727	0.983	0.017	0.310	1.092	0.091
Xl2	10^{-5}	0.170 (6.77)	0.413	0.973	0.028	0.174	1.088	0.087
	10^{-4}	0.638 (25.3)	0.861	0.985	0.017	0.402	1.101	0.100
	$5x10^{-5}$	0.485 (19.3)	0.727	0.980	0.021	0.310	1.098	0.098
Static ideal prior	10^{-5}*	0.170 (6.77)	0.413	1.001	0.005	0.174	1.031	0.033
	10^{-4}	0.638 (25.3)	0.861	1.007	0.012	0.402	1.053	0.055
	$5x10^{-5}$	0.485 (19.3)	0.727	1.005	0.009	0.310	1.046	0.048

Table 3.

Table 3. Variations in CO retrieval upper and lower column validation slopes, upper and lower column mean ratio values, upper and lower column comparison errors, and DoF for different TARDISS prior a priori choices and prior a priori covariance matrix scaling values. The asterisk in the second to last row indicates that this is the operational set of parameter choices for the CO retrieval.

3.1.2 Standard Output

Using the operational setup for our TARDISS fit, the comparison of CO_2 between our output lower partial column VMRs and the integrated, smoothed, in situ partial columns for all of the sites gives a slope of 1.007, a best fit standard error of 0.002 on the slope, and a corresponding mean ratio deviation value of 0.010. Figure 4 shows the 1-to-1 comparison plots for the upper and lower partial columns for CO_2 and CO. Across all twenty-four points, the average degrees of freedom of the daily inversions was 0.364 degrees of freedom per measurement for the lower column and the comparisons include between thirteen and sixty-seven inverted partial column data points. The error in the lower partial column values was 0.942 ppm on average across the comparison and the average reported error from the in situ profiles was between 0.03 and 0.156 ppm for the comparative partial column. The comparisons vary by site with the Lamont comparisons having the largest offset slope of 1.008. The

~~Park Falls and Armstrong comparisons have slopes of 1.000 and 1.006 respectively. Since neither the P1 or TARDISS retrieval can fully optimize the shape of the partial profile that they are scaling, the site to site differences are likely due to the variation of ability of the TCCON P1 priors to capture the source, sink, and transport complexities in the shape of the prior profile that directly translate to the DMFs used in the partial column scaling.~~

~~The operational comparison for the lower partial column CO fit gives a slope of 1.001, a best fit standard error of 0.002, and a corresponding mean ratio deviation value of 0.005. The inversion days have an average of 0.170 degrees of freedom per measurement with between 33 and 43 data points per comparison. The error in the lower partial column values was 0.413 ppb on average across the comparison.~~

~~To quantify the total error of our retrieval, we use the 1-to-1 comparisons to scale our error values to the point where at least 50% of the comparison points are within the error range of the 1-to-1 line. We calculate the scalar values as:~~**3.3 TARDISS Performance Using Operational Parameters**

3.3.1 Comparisons with Calculated TCCON Partial Columns

We compare the validation performance of the TARDISS partial column retrievals to the partial column validations of the TCCON individual windows used in the retrieval to demonstrate that TARDISS provides addition information about vertical distribution compared to the TCCON retrieval. We compute a partial column from the TCCON output by integrating the posterior TCCON CO or CO2 profile (i.e. the prior profile times the retrieved TCCON VSF) over the same pressure levels as the partial columns are calculated over for TARDISS. We compare the TCCON partial columns to the integrated, averaging kernel-smoothed, in situ partial columns calculated using Equation 20. The comparisons are shown in Table 4 and the slopes of the TCCON window partial column comparisons are shown as dotted lines in Fig. 5.

The comparisons show that the TARDISS retrieved partial columns for CO_2 have lower and upper partial columns slopes closer to one than the TCCON input windows. The mean ratio deviation for the lower column CO_2 is slightly larger than the mean ratio deviation for the TCCON input windows (0.011 compared to a TCCON average of 0.007) which is reflected in the error of the lower partial column CO_2 retrieval. The retrieved lower partial column for CO has a slope much closer to one than the slopes of the TCCON input and with a much smaller mean ratio deviation (0.002 compared to a TCCON average of 0.024). The retrieved upper partial column CO has a slope that is between the slopes of the TCCON input windows but still has a smaller mean ratio deviation suggesting increased precision.

These comparisons suggest that, for CO, the TARDISS algorithm is very effective at separately inferring the lower partial column CO values since the validation slope is closer to one and the mean ratio deviation is smaller than the individual windows. The algorithm is limited in its retrieval of the upper partial column CO which is shown by its direct comparisons and mean ratio deviation being similar to the TCCON input window partial column. The performance of the algorithm suggests that the

large variations in the CO vertical profile shapes benefit from the increased flexibility in the lower column but that there might be some spectroscopic biases to correct, particularly in the mid infrared windows.

For CO_2, the comparisons show that the algorithm can effectively infer upper partial column values but is less effective at retrieving the lower partial column CO_2 values. The lower partial columns benefit from the secondary scaling as the have less bias (a slope closer to one) than the individual windows but the slight increase in mean ratio deviation suggests that the retrieval cannot be as precise at adjusting for the surface errors in the a priori profile shape. The a priori profiles for CO_2 intentionally do not include variations of local sources or sinks at the surface but are quite accurate in the middle and upper troposphere. Accordingly, the secondary scaling of the upper partial column has improved accuracy and precision compared to the individual windows.

Finally, we compare the performance of the total column values calculated from the TARDISS scaled partial columns to the total column validations of the TCCON individual windows. The comparisons are shown in Fig. S3 and summarized in Table S3. The total column comparisons show similar trends as the upper column comparisons. This is likely due to the upper partial column vertical sensitivity being much larger than the lower partial column sensitivities as is discussed in Sect. 3.4.1.

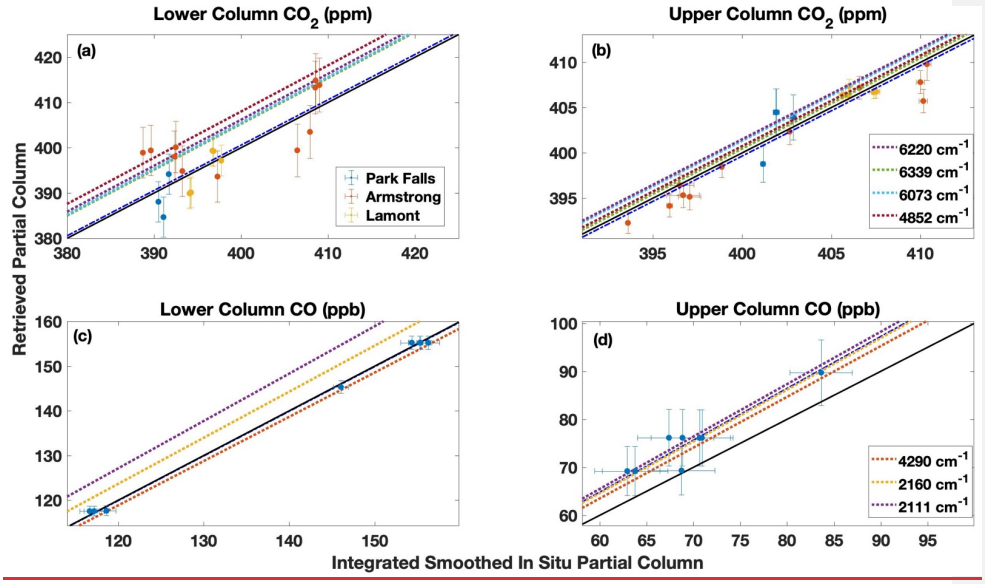
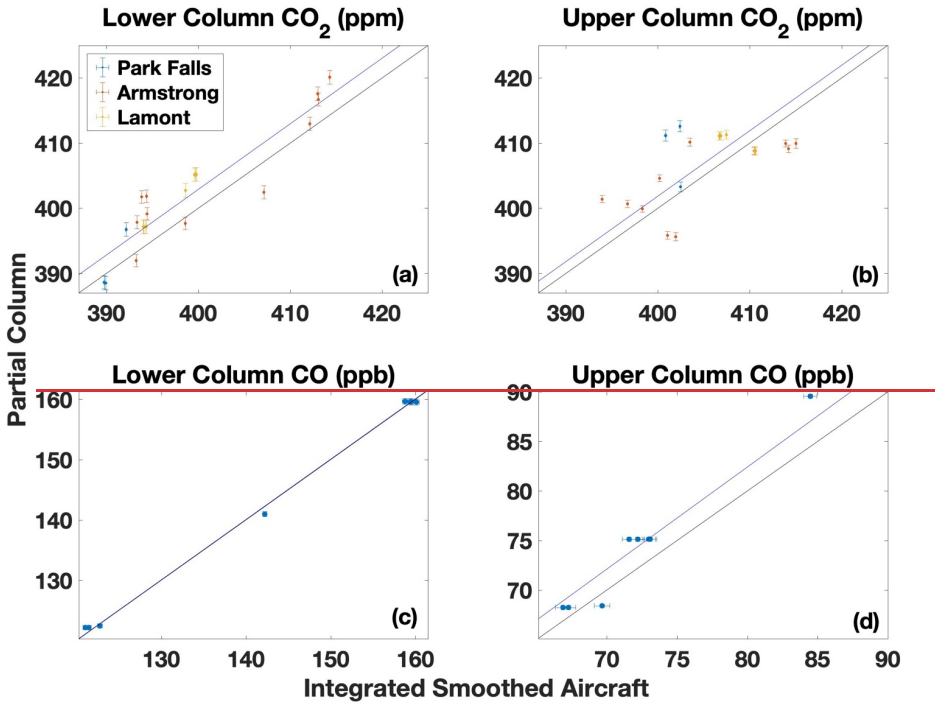


Figure 5

$$VEM = Median\left(\frac{|z_{comp} - z_s|}{\sigma}\right) \quad (18)$$

where z_{comp} is the comparison partial column value, z_s is the integrated, in situ partial column value, σ is the output total retrieval error, and VEM is the calculated validation error multiplier. For CO_2 at Park Falls, the lower and upper column VEM are 1.23 and 11.7; at Armstrong, the lower and upper column values are 4.48 and 8.33; and at Lamont the values are 3.43 and 6.29 for the lower and upper column, respectively. Since Caltech and East Trout Lake do not have comparison data, we apply error multiplier values of 4.48 and 11.7 as they are the largest multiplier values from among the other sites. For CO, the Lamont site the multiplier values are 1.57 and 12.5 which we use for the Caltech and East Trout Lake site CO retrieval data as well. The upper column VEM values are consistently larger than the associated lower column values; however, since the upper column errors are smaller than the lower column errors, the total errors for the lower and upper columns are closer in magnitude than the VEM values.



~~Figure 4~~**.** The direct comparisons between the partial column DMF values retrieved from the TARDISS fit and the integrated, smoothed ~~aircraft~~in situ partial columns for CO_2 (a,b) and the CO (c,d) for the lower (a,c) and upper (b,d) columns. The CO_2 comparisons are color coded by site and the CO comparisons are solely from the Lamont site. The error bars in the x-direction are the reported errors from the aircraft data smoothed the same way as the in situ measurements and the error bars in the y-direction are the output errors from the TARDISS fit~~.~~ scaled by the VEM values. The black solid line is the 1-1 line and the blue dot-dash line is the linear fit of the data with the y-intercept forced through zero. The blue dot-dash line for the lower partial column CO fit is overlapping with the solid black line. The slopes of the partial column validation of the TCCON spectral windows used in the retrieval are represented by dashed lines.

TCCON Window (cm^{-1})	Lower Column Validation Slope	Lower Column Validation Slope Error	Lower Column Mean Ratio Deviation	Upper Column Validation Slope	Upper Column Validation Slope Error	Upper Column Mean Ratio Deviation
CO_2						
6220	1.016	0.004	0.007	1.004	0.0010	0.003
6339	1.013	0.004	0.005	1.001	0.0009	0.003
6073	1.014	0.004	0.009	1.003	0.0011	0.003
4852	1.020	0.006	0.007	1.002	0.0011	0.004
TARDISS CO_2	1.001	0.003	0.011	0.999	0.0008	0.002
CO						
4290	0.990	0.034	0.041	1.058	0.077	0.106
2160	1.031	0.019	0.052	1.077	0.024	0.095
2111	1.059	0.020	0.061	1.092	0.023	0.108
TARDISS CO	0.999	0.002	0.005	1.081	0.012	0.081

Table 4. Comparisons of the TARDISS partial column retrieval to the partial column comparisons of the fits of the TCCON spectral windows from TCCON used as input for the TARDISS algorithm. The data in the TARDISS row uses the operational parameters for the fit that are identified in Table 2 and 3 by an asterisk.

3.3.2 Comparisons with Low Altitude In Situ Profiles

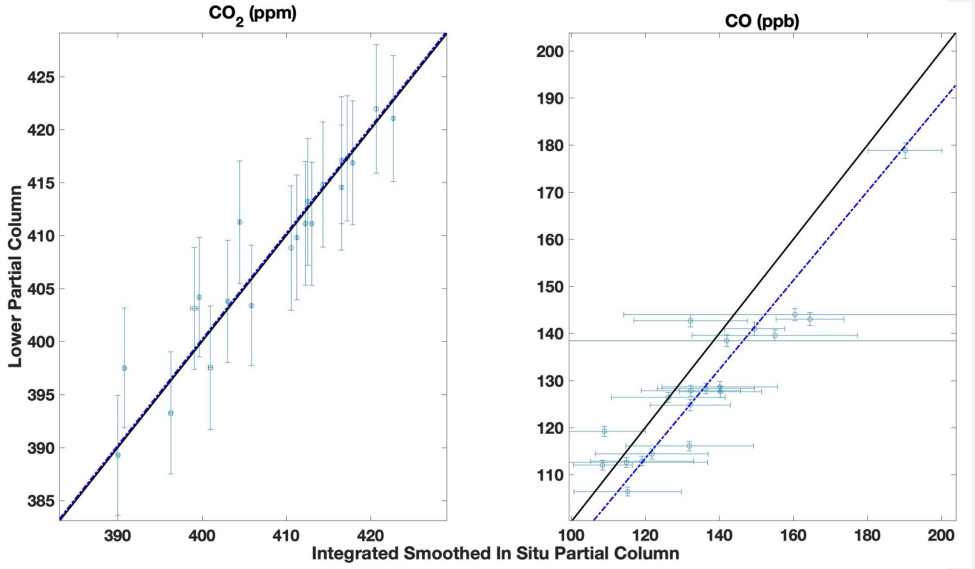


Figure 6. East Trout Lake site direct comparisons between the partial column DMF values retrieved from the TARDISS fit and the integrated, smoothed aircraft partial columns for lower column CO_2 and CO. The error bars in the x-direction are the integrated partial columns of the profile shifted by the error values and then subtracted from the original partial column integration. The error bars in the y-direction are the output errors from the TARDISS fit scaled by the VEM value for the site. The black solid line is the 1-1 line and the blue dot-dash line is the linear fit of the data with the y-intercept forced through zero. The slope for the fit is 1.001±0.002 for CO_2 and is 0.945±0.012 for CO.

In addition to the aircraft and AirCore validation data that include profile measurements at altitudes in the upper troposphere and lower stratosphere, we compare to aircraft data obtained as part of the NOAA GGGRN aircraft program at the Lamont and East Trout Lake sites. These measurements were made more frequently but do not include enough high-altitude measurements to compare with our retrieved upper partial column values, so we use them as an independent comparison to our validation data for our lower column CO_2 and CO retrievals. We use data obtained between the surface and 7 km from 26 of the 40 flights made between 2017 and 2020 at East Trout Lake. We also use data obtained between the surface and 6 km from 267 of the 399 flights performed at the Lamont site over the period of 2008 to 2018 and all 34 flights for CO made between 2017 and 2021. Figure 6 (East Trout Lake) and Fig. 7

(Lamont) show the retrieved lower partial column DMF plotted against the integrated, smoothed, in situ columns similar to Fig.

5.

Similar to the validation comparison, we revert to the a priori profile for altitudes not measured by in situ methods. To account for the errors in using the a priori profile, we add twice the standard deviation of the partial column that is measured to the average measurement error in quadrature. Given the lower altitudes measured by the GGGRN program, the errors associated with the parts of the profile that use the a priori profile are higher and, therefore, the errors in the long-term comparative measurements tend to be much higher than the validation measurements as shown in the CO comparisons in Fig 6.

Despite the larger error values, the consistency of the statistical parameters (summarized in Table S4) using the larger number of measurements in the long-term comparisons further motivates the use of the extended validation dataset. Some of the in situ profile comparisons occur during times with larger CO DMFs that suggest influences from sources not accounted for by the TCCON a priori profiles such as those from wildfires which likely resulted in the large VEM for the long-term CO comparisons. Although the comparisons with the long-term data are not used for validation, the long-term comparisons show that the validation comparisons are generally representative of the performance of the TARDISS algorithm overall.

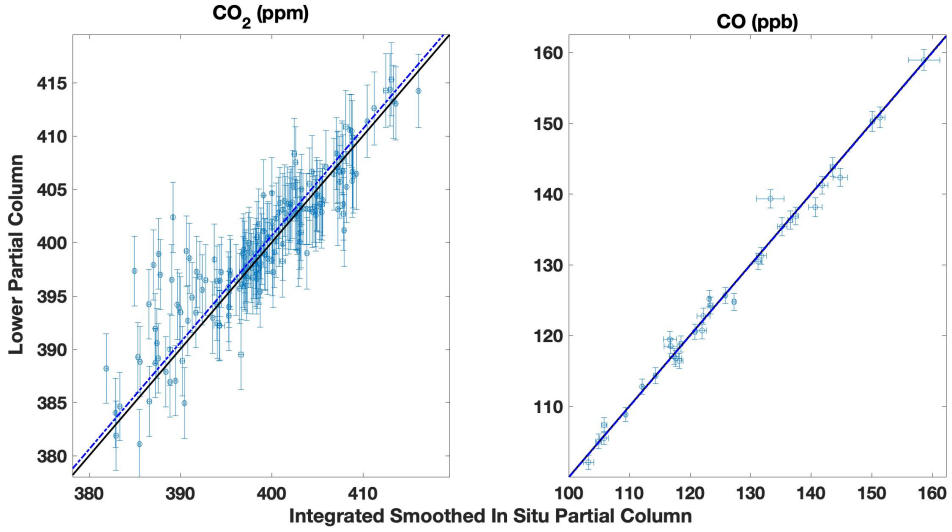


Figure 7. Lamont site direct comparisons between the partial column DMF values retrieved from the TARDISS fit and the integrated, smoothed airborne partial columns for lower column CO_2 and CO. The error bars in the x-direction are the integrated partial columns of the profile shifted by the error values and then subtracted from the original partial column integration. The error bars in the y-direction are the output errors from the TARDISS fit scaled by the VEM value for the site. The black solid line is the 1-1 line and the blue line is the linear fit of the data with the y-intercept forced through zero. The slope for the fit is 1.002±0.001 for CO_2 and is 1.000±0.002 for CO.

3.4 Retrieval Characterization

~~3.1.3~~ 3.4.1 TARDISS Vertical Sensitivity and Temporal Covariance

~~The analysis above implemented~~TARDISS uses an a priori covariance matrix with temporal covariance between upper partial column scalars over the course of a day of measurement, as shown in Fig. ~~2~~3. To determine how this ~~affects our~~constraint influences the retrievals, we compare the data above to the validation comparison from a CO_2 retrieval not constrained by a temporal covariance. The ~~prior~~a priori covariance matrix without the temporal covariance is simply a diagonal matrix of the 10^{-5}

scalar value. Table 45 shows that the retrievals without temporal constraints have a slightly poorer validation comparison overall, including ~~higher~~larger errors and fewer degrees of freedom. However, the site-by-site differences in validation data show that the ~~lower~~upper column VEM ~~for the Armstrong site is~~values are smaller when using a temporally unconstrained fit, whereas ~~both~~ the ~~Park Falls and Lamont VEMs~~lower column VEM values are improved when implementing the temporal constraints. ~~Similarly, the upper column VEM values at all sites improve without the temporal constraints.~~ While the purpose of this ~~publication~~study is to create ~~an~~a universally-applicable operational ~~product, the varying~~ algorithm, local differences in the sources and meteorology may alter the effects of the ~~prior~~a priori covariance matrix choice on the site VEMs ~~suggest~~. This suggests that ~~the~~ site-by-site parameter choices ~~could be individually determined in order to minimize error~~may enable smaller errors and ~~increase the partial column~~increased precision ~~if there were sufficient in situ validation profiles over the measurement site.~~.

~~Statistics~~	Statistics	Temporally Constrained Upper Column	Temporally Unconstrained Upper Column
Validation DoF (Overall)		0.~~364 (16.1~~0462 (2.12)	0.~~317 (13.9~~0352 (1.59)
Lower Column CO_2			
	Error (ppm)	~~0.942~~1.15	~~0.978~~1.15
	Validation Slope	1.~~007~~001	1.~~008~~002
	Mean Ratio Deviation	0.~~010~~011	0.~~012~~009
	Park Falls VEM	~~1.23~~3.25	~~1.52~~3.75
	Armstrong VEM	~~4.48~~2.98	4.~~35~~42
	Lamont VEM	~~3.43~~1.35	~~5.09~~2.50
Upper column CO_2			
	Error (ppm)	0.~~586~~497	~~1.053~~0.956
	Validation Slope	~~1.004~~0.999	~~1.002~~0.998
	Mean Ratio Deviation	0.~~012~~002	0.~~013~~003
	Park Falls VEM	~~11.7~~3.61	~~9.04~~1.92
	Armstrong VEM	~~8.334~~.63	~~5.26~~1.66
	Lamont VEM	~~6.29~~2.70	~~3.36~~1

Table 45. Validation comparison DoF, error, validation slope and mean ratio deviation and site VEM values for lower and upper column CO_2 for retrievals using a temporally constrained upper column and a temporally unconstrained upper column. The retrievals are performed with the operational parameters denoted by asterisks in Table 2.

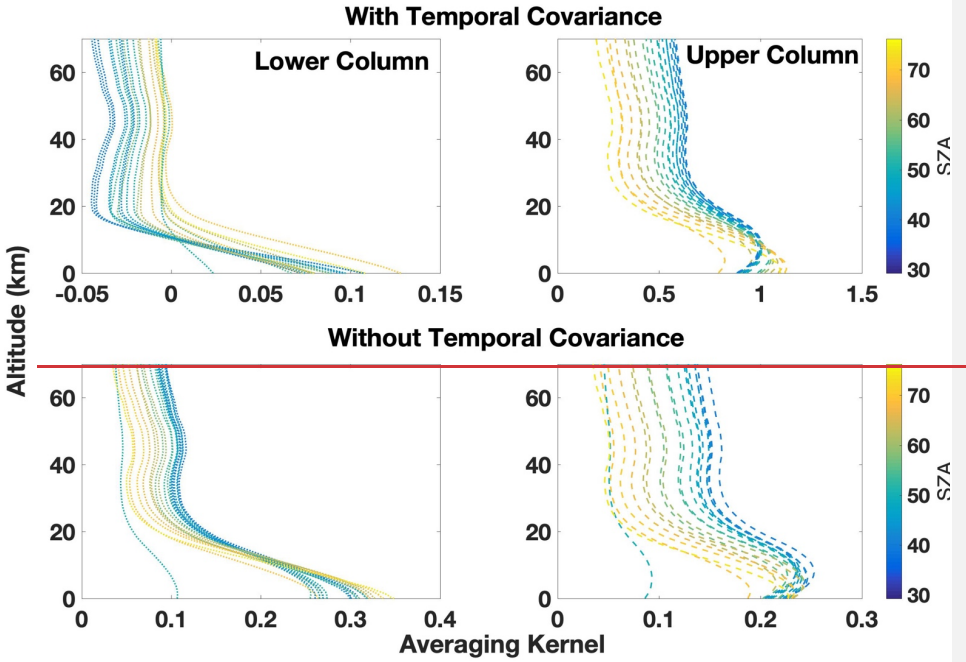
The ~~way that the~~ temporal covariance impacts our validation comparison ~~is~~ through the partial column vertical sensitivities described in Equation ~~17~~22 via the gain matrix (Equation ~~16~~21). To assess the importance of ~~our chosen prior~~the choice of a priori covariance matrix, we compare the

vertical sensitivities for a temporally constrained upper column and a temporally unconstrained upper column (shown in Fig. ~~5~~8) for a representative day (July 27th, 2018, at the Lamont site~~) using the operational parameters denoted by the asterisk in Table 2.~~).

Without the temporal constraint, the upper column sensitivities are on the same order as the lower column sensitivities with values between -0.05 and 0.~~12. For reference, a change of 1 ppm at a level with a sensitivity of 1 would result in a change in the partial column scalar of 0.025 and partial column scalars for CO_2 change on the order of a few percent O(0.01) per measurement.~~18. The upper column sensitivity peaks around the ~~10~~15 km level at low solar zenith angles and the peak moves toward the surface at higher solar zenith angles~~.~~ consistent with the changing kernel of the 6220 and 6339 cm^{-1} bands. The lower column sensitivities always peak ~~at~~near the surface ~~but~~(~2 km or below) and the sensitivity increases at higher solar zenith angles.

With the temporal constraint, the ~~pattern~~altitude of the ~~peaks~~ maximum sensitivities with respect to SZA remains similar but the upper column sensitivities are roughly ~~five times~~twice the value and the lower column sensitivities are half the value as the temporally unconstrained values. The imposed temporal covariance constrains the upper column to vary ~~together~~slowly over the span of a measurement day so that a change ~~at one level~~ in the column at one measurement point ~~would also induce~~induces changes at other measurement points ~~therefore~~thereby increasing the vertical sensitivities in the upper column over the entire ~~retrieval~~ day. This constraint is also stringent enough that it propagates into the sensitivity of the lower column scalar. Since our goal is to retrieve a lower partial column, it seems counterintuitive that using sensitivities with an order of magnitude difference provides a better validation comparison. However, for this method we assume that we know the shape and behavior of the upper column fairly well and that most of the change occurs near the surface. Given these assumptions, constraining the upper column more heavily by introducing expected daily patterns through the a priori covariance matrix allows for the lower column retrieval to have improved comparisons with in situ data despite the decreased vertical sensitivities.

While we test retrievals simply with and without temporal covariance, the possible choice of ~~prior~~a priori covariance matrix shape could be much more complex. Future study could include using model generated or back trajectory based temporal covariances to include outside information in the retrieval dynamically. For an operational retrieval product, we will include the temporal covariance in the ~~prior~~a priori covariance matrix as an operational parameter.


With Temporal Covariance
Lower Column
Upper Column
Without Temporal Covariance
Altitude (km)
Averaging Kernel
S7A

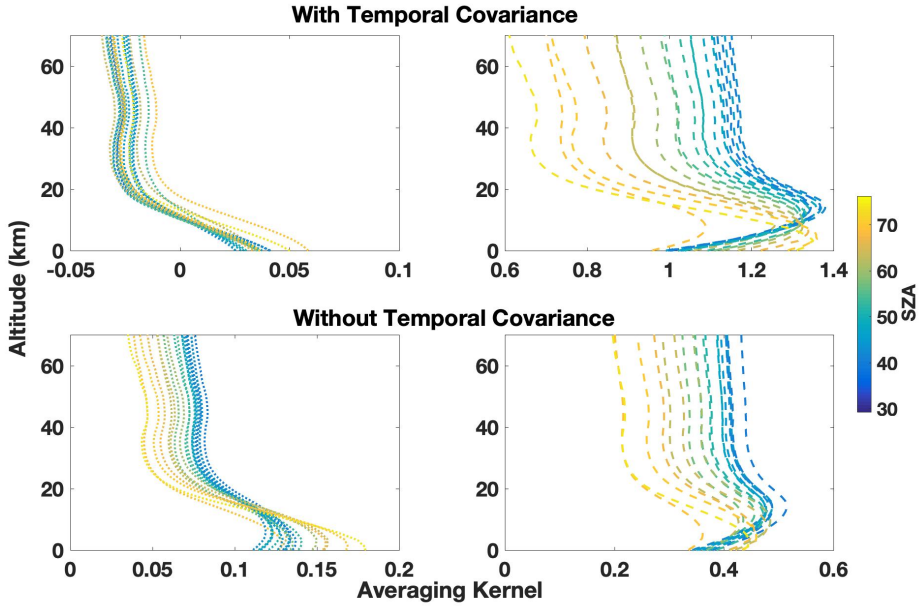


Figure ~~5~~8. Vertical sensitivities of the lower partial column (~~a,c~~left column) and upper partial column (~~b,d~~right column) scalars color coded by solar zenith angle in degrees. The sensitivities calculated when using a temporally covariant ~~prior~~a priori covariance matrix are shown in ~~(a) and (b)~~the top row and when using a non-temporally covariant ~~prior~~a priori covariance matrix are shown in ~~(c) and (d).~~the bottom row.

3.4.2 ~~Errors and Information Content~~

~~3.2.1~~ Error Analysis

Using the information from the validation comparison, we can ~~study~~evaluate the errors of the entire dataset from each of the five sites. The output of the retrieval is the partial column scalar and the error retrieved is the standard deviation of the partial column scalar calculated from the retrieval variance and represented as another scalar value. To convert our partial column scalar error to ~~parts per million~~a dry air mole fraction, we multiply the error scalar value by the ~~prior~~a priori partial column mixing ratio~~.~~ ($z_{a,TCCON}$ in Equation 17). Error varies from site to site due to variations in the ~~P1~~TCCON total column

errors that are input to the measurement covariance matrix and due to how well the ~~prior~~a priori partial column DMF matches the, generally unknown, actual partial column DMF. We report the total retrieval error, retrieval error components, and the error contribution from the validation comparison measurements in Table ~~5~~6.

~~Amongst all the sites, the~~The retrieval error values range from ~~0.977~~1.16 ppm to 1.~~068~~41 ppm for lower column CO_2 and from 0.~~605~~26 ppm to ~~0.737~~1.33 ppm for the upper column CO_2 ~~with the highest average error in Park Falls and the lowest average error in Armstrong for both columns.~~. For CO retrievals, the average total retrieval error ranges from 0.~~391~~48 ppb to 14.0~~.832~~ ppb for the lower column and 0.~~105~~032 ppb to ~~0.492~~2.23 ppb for the upper column. In general, the errors vary minimally ~~with season~~over the record, but ~~the Lamont site has~~there is a distinct seasonality for both lower column CO and CO_2 retrievals with the highest errors ~~are~~ during the summer perhaps ~~due to differences between the true and prior~~as a result of errors in the near surface a priori profiles ~~at this site during summer~~ (Fig. ~~S1~~S4). The absolute errors for CO_2 generally increase over time since ~~the prior partial column that is being scaled~~ simply because CO_2 is increasing ~~with continuing~~ due to anthropogenic emissions ~~even though~~. Fractionally, the ~~scalar values~~errors remain similar across the dataset for both CO and CO_2 (Fig. ~~S2).~~

~~The total retrieval error is the total of the model parameter error, the smoothing error, and the retrieval noise (Rodgers, 2008). In this retrieval, since there are no model parameters in the state vector, the model parameter error is zero. The variance due to model parameters is represented in the sensitivity analysis and becomes zero when we choose a particular set of model parameters. In future implementations, the model parameters could be included in the state vector and optimized within the retrieval.~~S5).

Because the model parameter error goes to zero in our implementation, ~~this means that~~ the current total retrieval error is the square root of the sum of the smoothing error (Equation 25) and the retrieval noise~~.~~ (Equation 26). The smoothing error is ~~59.3% to 71~~94.0% ~~of the total retrieval error on average for CO_2 and 40~~to 96.5% ~~to~~ of the total retrieval error on average for CO_2 and 81.~~0~~6% to 87.8% of the total retrieval error on average for CO depending on the site and is directly related to the scaling of the ~~prior~~a priori covariance matrix. ~~For example, if we did not scale our prior covariance matrix our smoothing error would be nearly zero.~~ While scaling the a priori covariance matrix by a higher value increases the smoothing error, it also results in a reduction to the total retrieval error. Furthermore, the fit of the lower partial column CO_2 benefits from a stronger constraint since the slope of the lower partial column CO_2 validation is closest to one when using the tightest covariance matrix as shown in Table 2. The retrieval noise is ~~29.0~~3.5% to ~~40.7~~6.0% of the total retrieval error on average for CO_2 and ~~19.0~~18.4% to ~~59.5~~12.2% of the total retrieval error on average for CO depending on the site and has the opposite relationship to the scaling of the a priori covariance matrix. The retrieval noise reflects the effect of the model covariance matrix that is composed of the ~~P1~~TCCON total column measurement errors and therefore reducing ~~the P1~~these errors would also reduce the retrieval noise.

~~The total retrieval~~Using the operational setup for our TARDISS fit, we calculate the site specific VEM values using Equation 27 (Tables 5 and 6). These values are used to scale the error of the TARDISS fit for all the comparisons in this work. The VEM scaled errors serve as a conservative estimate for the retrieval errors and should be reevaluated with additional in situ profile measurements as they become available. For CO_2 at Park Falls, the lower and upper column VEM are 3.61 and 3.25; at Armstrong, the lower and upper column values are 4.63 and 2.98; and at Lamont the values are 2.70 and 1.35 for the lower and upper column, respectively. Since Caltech and East Trout Lake do not have comparison data, we apply error multiplier values of 4.63 and 3.25 as they are the largest multiplier values from among the other sites. For CO, the Lamont site multiplier values are 1.00 and 15.4, which we use for the Caltech and East Trout Lake site CO retrieval data as well.

Since the TARDISS retrieval cannot fully optimize the shape of the partial profile, the site-to-site differences in VEM are likely due to the variation in the accuracy of the TCCON priors which by design do not capture the local source, sink, and transport complexities. For CO_2, the upper column VEM and retrieval error values are consistently smaller than the associated lower column values suggesting that these data support the assumption that the shape of the profile of the upper partial column is generally much more accurately captured by the TCCON priors.

The total error for each site is determined by the multiplying the retrieved errors by the site and partial column respective VEM values. After implementing the VEMs, the errors for the lower partial column CO_2 retrieval range from ~~1.31~~3.38 ppm to ~~4.66~~5.88 ppm and from ~~0.614~~1.22 ppb to 1.~~31~~96 ppb for CO across all sites and data. As the Caltech and East Trout Lake sites have no validation comparisons, we use the largest validation error multiplier (that of the lower column Armstrong and upper column Park Falls comparison) as a higher bound. ~~The comparisons in Fig. 4 are recreated with the scaled errors in Fig. S4.~~

~~The retrieval~~Since the overall biases are small with validation slopes close to one, the errors are sufficiently small ~~enough~~ that the TARDISS ~~results~~retrievals have ~~some power for~~skill in evaluating CO_2 fluxes at TCCON sites. The error compared to the overall lower partial column DMF is small, ~~0.80~~1.25% on average across the five sites for CO_2.

Site	~~Mean Lower/Upper Column~~ Retrieval ~~Error (ppm for CO_2; ppb for CO~~Noise (% of total)	~~Retrieval Noise~~Smoothing Error (% of total)	~~Smoothing Error (% of total)~~Mean Lower/Upper Column Retrieval Error (ppm for CO2; ppb for CO)	Lower/Upper Column Validation Error Multiplier (unitless)	Mean Total Lower/Upper Column Error (ppm for CO_2; ppb for CO)

CO_2 Retrievals					
Park Falls	1.068/0.7373.5	29.096.5	71.1.257/0.655	1.23/11.73.61/3.25	1.31/8.624.54/2.13
Armstrong	6.0.977/0.608	40.794.0	59.31.253/0.500	4.48/8.3363/2.98	4.12/5.0680/1.49
Lamont	1.035/0.6784.5	32.195.5	67.91.252/0.582	3.43/6.292.70/1.35	3.08/4.2638/0.786
Caltech	1.040/0.6704.5	35.095.5	65.1.271/0.568	4.48/11.763/3.25	4.66/7.845.88/1.85
East Trout Lake	1.025/0.6055.4	36.294.6	63.81.268/0.514	4.48/11.763/3.25	4.59/7.085.87/1.67
CO Retrievals					
Lamont	0.832/0.49212.2	59.587.8	40.51.34/0.447	1.57/12.500/15.4	1.3134/6.1588
Caltech	0.413/0.10518.4	81.619.0	81.1.96/0.318	1.57/12.500/15.4	0.648/1.3196/4.90
East Trout Lake	0.391/0.18215.7	27.184.3	72.91.22/0.355	1.57/12.500/15.4	0.614/2.281.22/5.47

Table 56. Errors in the CO and CO_2 lower partial column retrievals of each site shown as the average of the entire data time series and broken down into total retrieval error, retrieval noise, smoothing error, validation error multiplier, and total error. The values for total retrieval error and total error arerepresent one sigmastandard deviation.

3.2.24.3 Information Content Analysis

The information content of the retrieval is determined by the DoF and Shannon information content (H) of the retrieval, each calculated from the averaging kernel of the retrieval.. The DoF represent the independent pieces of information that can be retrieved from a measurement. We report our DoF values both normalized by the number of measurements made in a day, as well as the daily averageoverall DoF. Since the DoF are calculated as the trace of the averaging kernel, we isolate and report the DoF from the upper and lower column separately along with the total. The Shannon information content is a single value to represent the effectiveness of the retrieval to recover information from the model with respect to the variance in the data. Higher Shannon information content values correspond to a retrieval with a higher possible effectiveness.

The information content is summarized for each site in Table 67. The overall average lower column DoF per measurement across all sites and collected data is 0.353047 for CO_2 and 0.21815 for CO. The lowest DoF average value of 0.287034 is in Park Falls and the highest DoF average value of 0.425061 is in Armstrong for CO_2 and, between the three sites with CO retrievals, East Trout LakeCaltech has the highest average lower column DoF of 0.29218 compared to 0.21412 for Lamont and 0.19715 for CaltechEast Trout Lake. The upper columnretrievals of CO have much larger DoF are

~~significantly less than~~compared to CO_2 primarily since the ~~lower column DoF due to~~ CO_2 requires a stronger scaling constraint of the ~~constraints implemented in the prior~~a priori covariance matrix limiting the amount of information that can be inferred.

Ideally, DoF values greater than one are desired for traditional profile retrievals. However, the temporal aspect of our retrieval complicates the discussion. If we consider the CO_2 retrievals, the five sites used in this work made an average of ~~173~~172 measurements per day so that the DoF value average of 0.~~353~~0470 per measurement ~~still retrieve an average of 60.4~~yields 8.08 independent pieces of information about the lower partial column per day which provides significant information on the diurnal variation and the fluxes into and out of the lower column.

The information content shown in the DoF are mirrored in the Shannon information content. Similar to the DoF, Park Falls has the lowest and Armstrong has the highest Shannon information content on average for CO_2. These differences are likely driven by the combination of the ~~P1~~TCCON retrieval errors and how well the ~~chosen prior~~a priori covariance matrix matches the temporal aspects of local meteorology, such as cloud cover or upper tropospheric transport, or the magnitude and time scales of the local carbon fluxes in the boreal forest ~~and~~versus the lack of such fluxes in the Mojave Desert. For CO, the ~~East Trout Lake~~Caltech retrieval has the highest DoF and Shannon information content of the three sites. ~~Interestingly, the Lamont retrieval has a higher DoF but the Caltech retrieval has a higher average Shannon information content.~~ While the differences in Shannon information content and DoF between sites are not necessarily directly comparable, these differences also might be due to the ~~P1~~TCCON retrieval errors and how well the chosen ~~prior~~a priori covariance matrix constrains the solution~~. Since the Shannon information content includes the off diagonal terms of the averaging kernel matrix, the larger information content at the Caltech site suggests that the chosen model covariance matrix and prior covariance matrix are an effective constraint on the Caltech retrieval~~.

Site	Total Degrees of Freedom per ~~measurement~~Measurement (per day)	Lower Column DoF per ~~measurement~~Measurement (per day)	Upper Column DoF per ~~measurement~~Measurement (per day)	Average Measurements per day	Shannon Information Content per day
CO_2 Retrievals					
Park Falls	0.~~370 (43.1~~151 (14.0)	0.~~287 (35.9~~0338 (4.30)	0.~~0830 (7.19~~117 (9.72)	116	~~28.7~~9.96
Armstrong	0.~~495 (112~~165 (33.2)	0.~~425 (98.9~~0613 (14.3)	0.~~0700 (13.3~~104 (18.9)	227	~~81.0~~24.7
Lamont	0.~~415 (64.9~~163 (20.6)	0.~~337 (55.5~~0444 (7.22)	0.~~0782 (9.35~~119 (13.4)	~~159~~155	~~44.4~~15.0

Caltech	0.~~440 (77.8~~156 (23.1)	0.~~364 (67.40~~452 (8.45)	0.~~0760 (10.5~~111 (14.7)	180	~~54~~17.0
East Trout Lake	0.~~447 (~~181 (25.5~~81.6~~)	0.~~354 (700~~503 (10.2)	0.~~0925 (11.4~~131 (15.3)	~~182~~181	~~34.8~~19.0
Overall	0.~~433 (75.9~~163 (23.2)	0.~~353 (60.40~~470 (8.89)	0.~~0799 (10.3~~116 (14.4)	~~173~~172	~~48.6~~17.1
CO Retrievals					
Lamont	0.~~247 (30.5~~236 (26.1)	0.~~214 (27.0~~123 (15.7)	0.~~0333 (3.52~~113 (10.4)	~~119~~120	~~18~~17.5
Caltech	0.~~197 (39.0~~227 (43.6)	0.~~188 (37.5~~184 (36.9)	0.~~0091 (1.570~~431 (6.76)	~~189~~194	~~23.5~~26.8
East Trout Lake	0.~~292 (55.5~~263 (43.4)	0.~~253 (49.9~~146 (29.5)	0.~~0333 (5.59~~113 (13.8)	178	~~25.0~~26.2
Overall	0.~~246 (41~~242 (37.7)	0.~~218 (38.2~~151 (27.4)	0.~~0273 (0~~910 (10.3~~.56~~)	~~162~~164	~~22.3~~23.5

Table ~~6~~7. Degrees of freedom per measurement (and per day) for the lower column, upper column, and total retrieval, in addition to the Shannon information content separated by site for the CO and CO_2 retrievals.

410

~~3.2.3 Long-term Comparisons~~

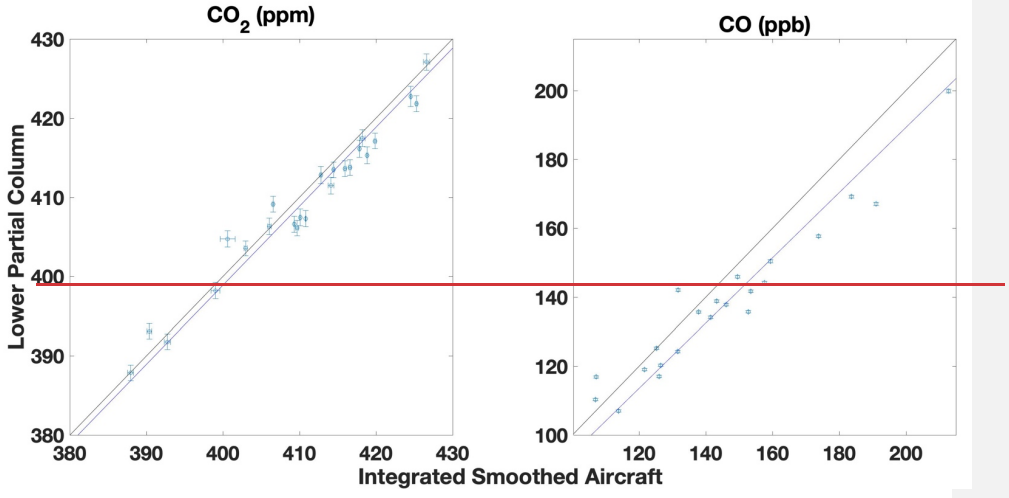
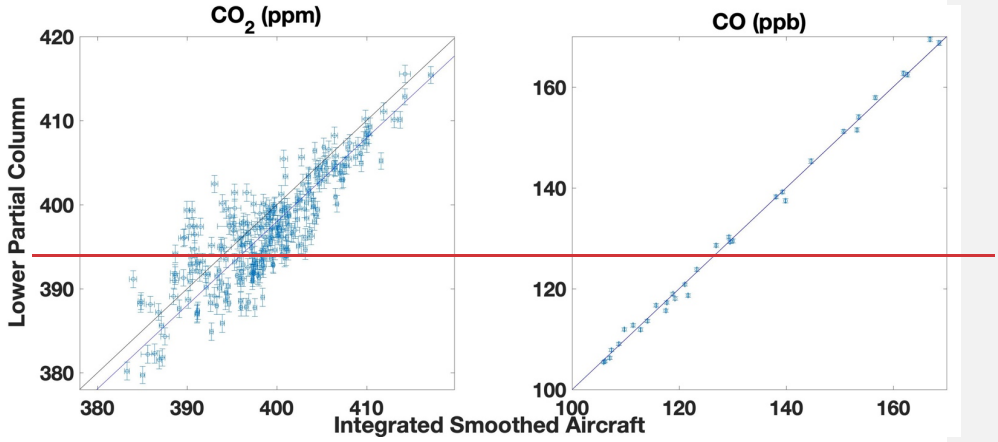


Figure 6. East Trout Lake site direct comparisons between the partial column DMF values retrieved from the TARDISS fit and the integrated, smoothed aircraft partial columns for lower column CO_2 and CO. The error bars in the x-direction are the reported errors from the aircraft data smoothed the same way as the in situ measurements and the error bars in the y-direction are the output errors from the TARDISS fit. The black solid line is the 1-1 line and the blue line is the linear fit of the data with the y-intercept forced through zero.

In addition to the aircraft and AirCore validation data that include profile measurements at altitudes in the upper troposphere and lower stratosphere, we compare to aircraft data obtained at the Lamont and East Trout Lake sites. We use data obtained between the surface and 7 km from 26 of the 40 flights for CO2 and 26 of the 39 flights for CO made between 2017 and 2020 at East Trout Lake. We also use data obtained between the surface and 6 km from 282 of the 399 flights performed at the Lamont site over the time period of 2008 to 2018 and all 34 flights for CO made between 2017 and 2021. Measurements at both sites were made as part of the NOAA GGGRN aircraft program. Figure 6 (East Trout Lake) and Figure 7 (Lamont) show the retrieved lower partial column DMF plotted against the integrated, smoothed, in situ columns similar to Fig. 4. These measurements were made more frequently but do not include enough high-altitude measurements to compare with our retrieved upper partial column values, so we use them as an independent comparison to our validation data for our lower column CO_2 and CO retrievals.

The consistency of the statistical parameters using the larger number of measurements in the long-term comparisons further reinforces the use of the validation comparison results across the entire retrieval dataset. Using the operational retrieval parameters, the long-term comparisons have similar informational and error statistics to the validation comparisons. For the Lamont site, the CO_2 retrievals have an average DoF per measurement of 0.334 and the comparison slope is 0.995. The overall VEM and mean total error calculated from the long-term comparisons is 3.15 and 3.20 ppm compared to 3.43 and 3.60 ppm from the Lamont, lower column validation comparisons. The Lamont CO retrievals have an average DoF per measurement of 0.215 and the comparison slope is 1.0003. The overall VEM and mean total error calculated from the long-term comparisons is 1.64 and 0.595 ppb compared to 1.57 and 1.31 ppb from the Lamont, lower column validation comparisons. For the East Trout Lake site, the CO_2 retrievals have an average DoF per measurement is 0.372 and the comparison slope is 0.997. The long-term VEM and mean total error for CO_2 are 2.42 and 2.48 ppm compared to the Armstrong VEM of 4.48 used for East Trout Lake and the resulting total error of 4.59 ppm. For long-term CO retrievals at East Trout Lake, the DoF per measurement is 0.263 and the comparison slope is 0.946. The overall VEM and mean total error calculated from the long-term comparisons is 22.5 and 8.34 ppb compared to the Lamont VEM of 1.57 used for East Trout Lake and the resulting total error of 0.614 ppb. Some of the in situ profile comparisons occur during times with larger CO DMFs that suggest influences from

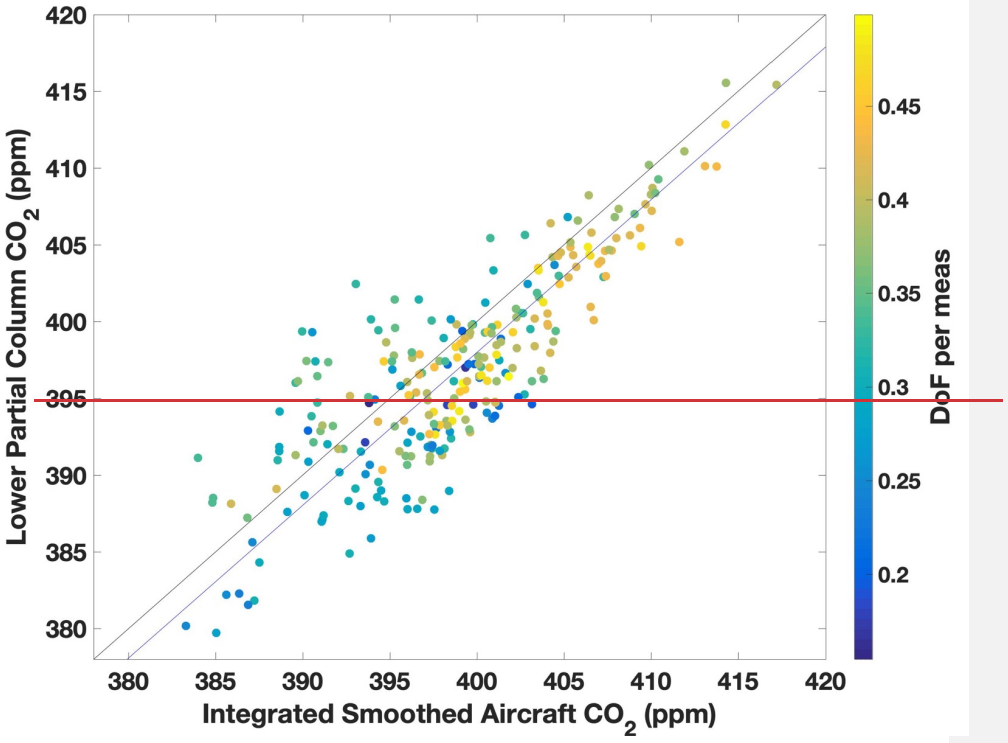
~~sources not accounted for by the P1 a prior profiles such as wild fires which likely resulted in the large VEM for the long-term CO comparisons.~~



~~Figure 7.~~ ~~Lamont site direct comparisons between the partial column DMF values retrieved from the TARDISS fit and the integrated, smoothed aircraft partial columns for lower column CO_2 and CO. The error bars in the x-direction are the reported errors from the aircraft data smoothed the same way as the in situ measurements and the error bars in the y-direction are the output errors from the TARDISS fit. The black solid line is the 1-1 line and the blue line is the linear fit of the data with the y-intercept forced through zero.~~

The informational content of the retrieval ~~helps us understand~~assists in evaluating the TARDISS algorithm ~~more thoroughly,~~ but ~~could~~ also ~~serve~~serves as a diagnostic ~~parameter to indicate~~of the effectiveness of the retrieval for ~~a particular~~each day of measurement. Figure ~~8~~9 shows the long-term comparisons between the retrieved lower partial column and the smoothed, integrated, in situ data at the Lamont site color-coded by the DoF per measurement for each point. The comparisons with higher DoF per measurement generally sit closer to the 1-to-1 line as ~~we would expect~~expected and suggest that days with higher DoF per measurement ~~would~~ have ~~a~~ lower associated VEM. Figure ~~S5~~S7 shows the VEM calculated ~~when filtering out measurement~~after removing days that have DoF per measurement values below a specific threshold. The ~~lowest~~ VEM ~~is~~ calculated ~~when filtering out days~~for the long-term comparison data decreases consistently with increasing DoF ~~per measurement lower than~~ filters until it reaches one at ~0.~~35 which~~07 DoF per measurement. This, however, this excludes roughly

~~half~~90% of the ~~282 flights used. A threshold higher than 0.35 reduces the number of measurements enough that it is no longer representative of the dataset.~~data. As a first step, the data could be filtered for low DoF or low Shannon information content. ~~Moving forward~~In the future, the information content could be used to create more dynamic VEM values for our datasets and provide more precise error values than the conservative, static VEM per site reported in Table ~~5. In addition to the DoF, the Shannon information content can be used to filter retrieval days~~ 6.~~where~~ the



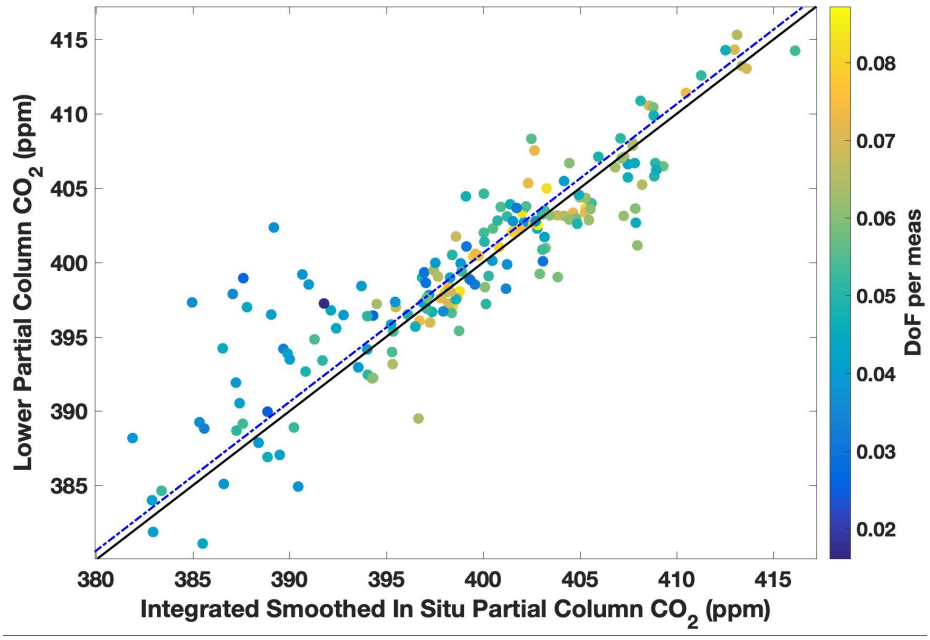


Figure ~~8~~9. The same comparison shown in Fig. 7 is shown here without error bars and color coded by the DoF per measurement for the comparison day retrieval. The blue dot-dash line ~~below~~above the black 1-to-1 line is the linear fit of the data with the y-intercept forced through zero with a slope of 1.002±0.001.

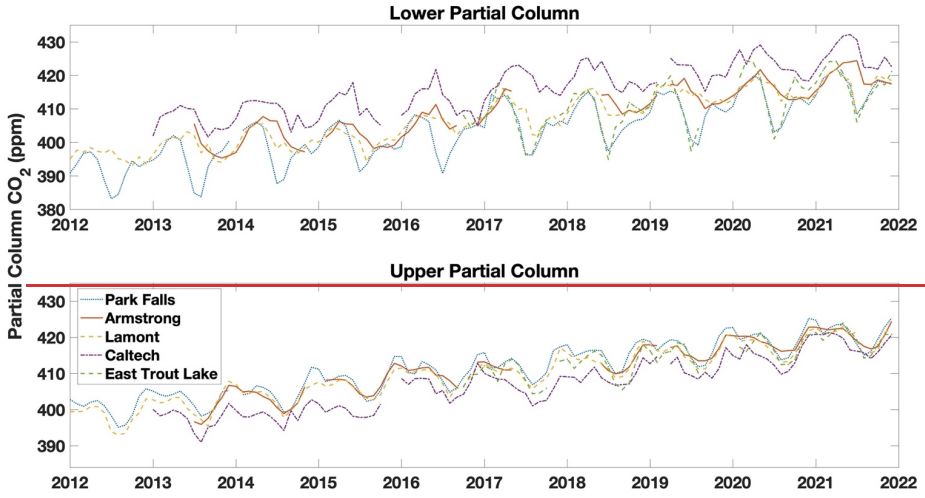
~~3.3 Current Products~~3.5 Time Series of the TARDISS Retrieval

The TARDISS algorithm is applicable to any spectra reported as TCCON data with the correct detector requirements (InGaAs for CO_2 and both InGaAs and InSb for CO). Overall, there are at least nine years of CO_2 data at each site in this work and approximately five years of CO data at the East Trout Lake, Lamont, and Caltech sites.

Figure ~~9~~10 shows the monthly mean lower and upper partial column data retrieved from spectra ~~taken~~obtained over the last decade ~~for~~at the ~~five~~North American TCCON sites ~~discussed in this work.~~. These ~~data show~~upper columns reflect the global seasonal patterns in CO_2 ~~in all sites with values in the Armstrong and Lamont traces being the most similar and consistent.~~. The lower column at Park Falls

and East Trout Lake ~~traces show~~reflect the local influences on CO_2 in the sharp ~~dips~~decline in ~~the summers~~surface CO_2 when the surrounding forest is most photosynthetically active. ~~The~~In contrast, the lower column Caltech trace shows a consistent urban enhancement over the global trends of ~5 ppm. All five upper column traces are generally consistent with one another and have a ~6 ppm seasonal fluctuation.

Figure ~~10~~11 shows the monthly median retrieved lower and upper partial column CO data from the East Trout Lake, Lamont, and Caltech site. We observe a slight seasonality at ~~both sites~~each site with maximums in the winter months and minimums in the summer months. The CO lower partial column data from the Caltech site tends to be larger than those from the Lamont site due to the urban enhancement despite the recent decreasing trend ~~but this is muted when using the monthly medians shown here. The East Trout Lake site show influences from the incomplete combustion of wildfires in both the upper partial column CO traces in both 2017 and 2021.~~. An example of effect of the urban enhancement on total and partial column values is shown in Fig. S8.



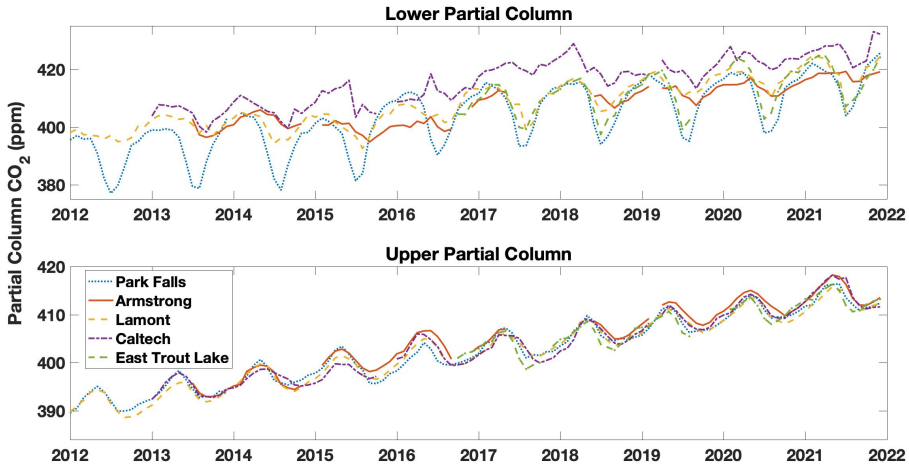
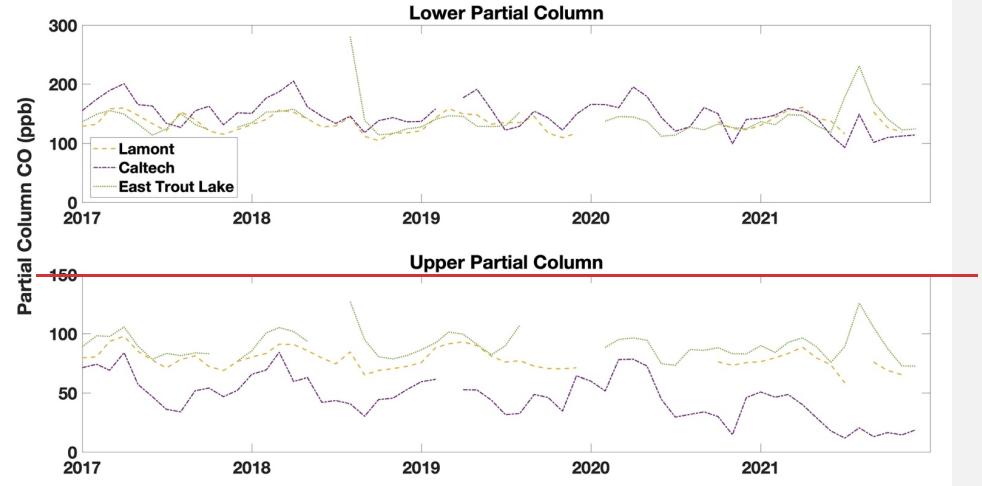


Figure 9<u>10</u>. Time series plot of the monthly median lower (top) and upper (bottom) partial column values of CO_2 in ppm for the five sites used in the work from 2012 (or the start of measurement) to the end of 2021. Data from before 2012 measured in Park Falls <u>and 2011 in Lamont</u> are not used due to instrument alignment <u>issues and laser</u> issues.


Lower Partial Column
Partial Column CO (ppb)
300
200
100
0
Lamont
Caltech
East Trout Lake
2017
2018
2019
2020
2021
Upper Partial Column
150
100
50
0
2017
2018
2019
2020
2021

520

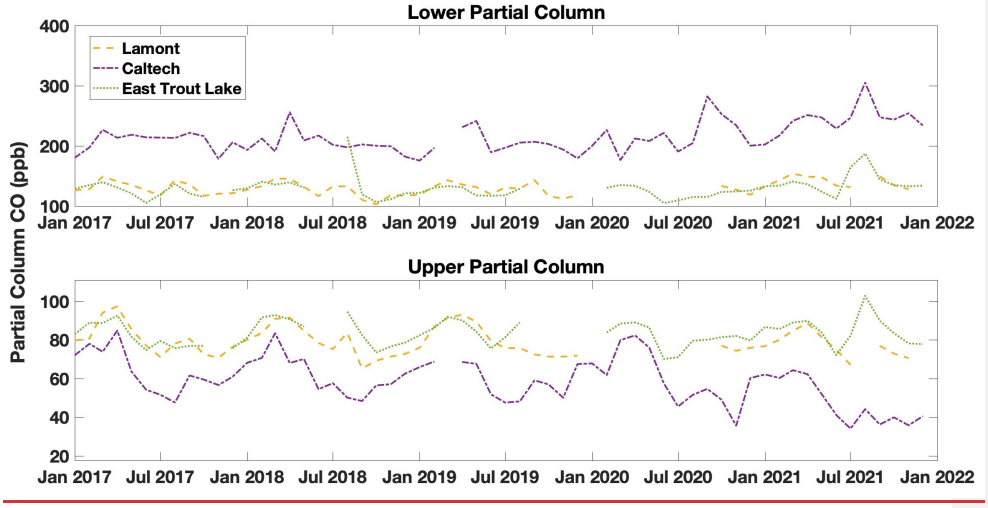

Lower Partial Column
Partial Column CO (ppb)
400
300
200
100
Lamont
Caltech
East Trout Lake
Jan 2017
Jul 2017
Jan 2018
Jul 2018
Jan 2019
Jul 2019
Jan 2020
Jul 2020
Jan 2021
Jul 2021
Jan 2022
Upper Partial Column
100
80
60
40
20
Jan 2017
Jul 2017
Jan 2018
Jul 2018
Jan 2019
Jul 2019
Jan 2020
Jul 2020
Jan 2021
Jul 2021
Jan 2022

Figure ~~10~~11. Time series plot of the monthly median lower (top) and upper (bottom) partial column values of CO in ppb for the three sites used in the work that have the InSb detector from 2017 to the end of 2021. CO has been declining in most of the US cities due to emissions control technologies.

~~3.4 Future Applications~~

~~Using the lower partial column data product, we can analyse carbon fluxes in novel ways. The data retrieved from the Park Falls site is collocated with in situ CO_2 data from a ~400-meter tall tower that measures continuously (Andrews et al., 2014) and also reports eddy covariance fluxes (Berger et al., 2001). Our lower partial column values are comparable to the in situ data in magnitude and general trend, particularly during the midday when turbulent mixing is often strong enough to create a more homogenous mixed layer CO_2 concentration (Xu et al., 2019). In future work, we will compare the tower CO_2 concentration and flux values with estimates of the diurnal variation in partial column CO_2.~~

~~The lower partial column CO are useful for comparison with other column averaged pollution tracers such as aerosol optical depth (AOD) and TROPOMI NO_2. For example, the chemical composition of the atmosphere in the South Coast Air Basin (SoCAB) continues to rapidly change (Parrish et al., 2016; Van Rooy et al., 2021) and it is of interest to diagnose whether the relationships between primary VOC emission sources, meteorology, and aerosols are also evolving. In future work, we will examine these relationships over the entire partial column data record using both the FTS data and observations from a nearby AERONET AOD measurement. Preliminary comparisons of afternoon average lower column CO values versus afternoon average 500 nm AOD values for days in which the afternoon average temperatures were above 25 degrees Celsius for the 2016 to 2021 time period show a strong correlation, particularly when decoupled from temperature and accounting for atmospheric water.~~

~~Appendix A: Daily Anomaly Calculation~~

~~Daily anomalies in this work mean the difference between the column values at a particular solar zenith angle in the afternoon and the column value at the same solar zenith angle in the morning. This approach removes air mass dependencies and allows for a direct comparison of the measured change in column values over a particular day. Due to the differences in the averaging kernel of each window, spectral windows sensitive to different parts of the atmosphere return different total column Xgas values and the ratio of the daily anomalies measured with the different windows used provides insight into how to weight the different inputs in the inversion. Since the sensitivities of each spectral window is determined by spectroscopy, the daily anomaly ratios are expected to be the same and independent of measurement location.~~

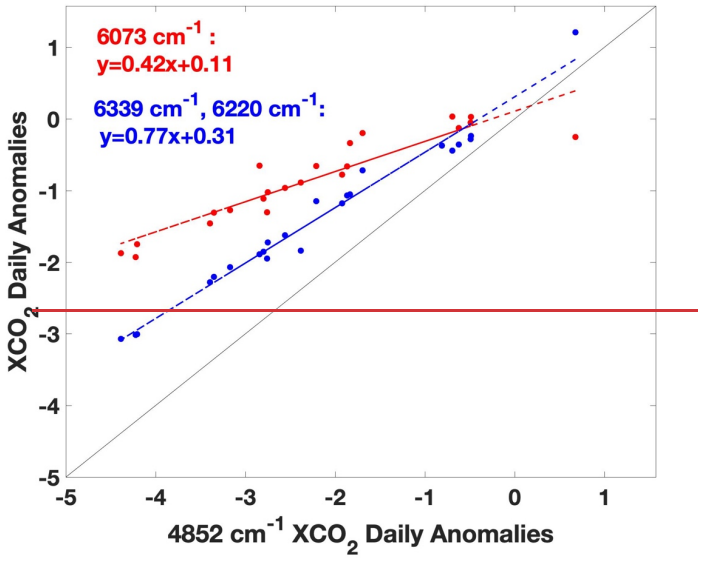


Figure A. Scatter plot of the daily anomaly values for the standard 6220 cm^{-1} and 6339 cm^{-1} TCCON windows and the 6073 cm^{-1} CO_2 window plotted against the daily anomaly values from the 4852 cm^{-1} CO2 window for days in July, 2018. The black line is the 1-to-1 line and the least squares linear fits for the respective windows are shown in the text on the plot. The slopes of the fit are used as weightings in the TARDISS retrieval.

Appendix B: Temporal Assimilation

We want to test the influence of the number of observations included in a single retrieval. To do this, we compare the retrieved error value for one measurement fit on its own and with an increasing number of observations until we retrieve with a full day. We take the midday observation from the Park Falls site on July 18, 2018 and retrieve the partial column error values using the least squares method and the maximum a posteriori method (using a static ideal prior scalar to avoid influences from the least squares approach). These values are represented by the points that correspond with zero on the x axis of Fig. B for both the lower and upper partial column errors. We then retrieve the errors of the midday measurement again including the observation before and after it which is represented by the points that correspond with 2 on the x axis of Fig. B. We repeat this method, expanding the number of observations included until we are fitting the entire day of observations.

The left-hand plot of Fig. B1 shows the decrease of the retrieved upper and lower partial column error of the midday point as the number of observations included in the retrieval increases. The upper partial column errors decrease more than the lower partial column errors partially due to the temporal constraints of the a priori covariance matrix. On the contrary, the right-hand plot of Fig. B shows that the inclusion of more observations in the least squares fit does not change the retrieved partial column errors of the midday measurement. Moreover, the partial column errors retrieved using the least squares method are at least six times larger than the partial column errors retrieved using the MAP method. This is due to the use the a priori covariance matrix in the MAP method that can improve upon the best estimate retrieval of the least squares method.

To understand the influence of the prior covariance matrix (overall scaling and temporal constraints), we compare the error values of the least squares method with the MAP method with an entirely uninformed prior covariance matrix. Shown in Fig. B2, the uninformed MAP approach returns errors of similar magnitude to the least squares method. This suggests that a main value of the MAP approach is the use of external information to improve and inform the retrieval.

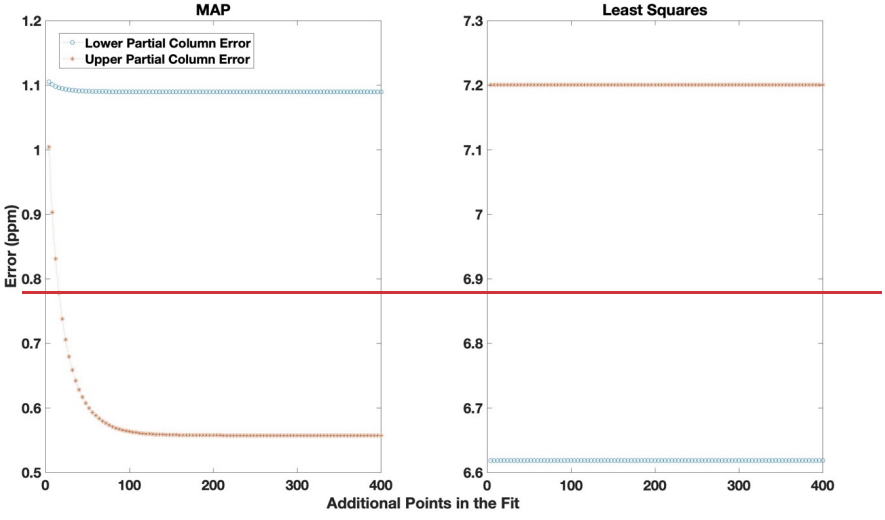
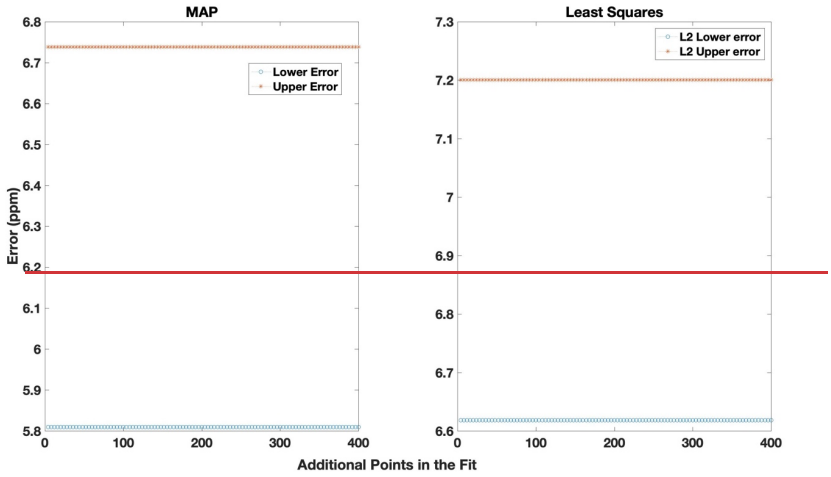


Figure B1. Errors in the retrieval of CO_2 from the midday total column measurement at the Park Falls site on July 18, 2018 using the MAP method outlined by equation 12 and the least squares method outlined by equation 18. The blue circles represent the error in the lower partial column and the orange asterisks represent the error in the upper partial column. Note the difference in the range of the y axis in

~~the left and right plots both of which are in parts per million. The x axis indicates the number of points included in the overall fit with zero additional points representing the retrieval of a single spectra.~~



~~**Figure B2. Same as Fig. B1, except the prior covariance is removed from the MAP retrieval.**~~**4 Conclusions**

The TARDISS retrieval algorithm enables partial column information to be derived from the TCCON total column observations of CO_2 and CO derived from different absorption bands with different vertical averaging kernels. Compared to traditional vertical retrieval approaches, the algorithm relaxes the requirement of very accurate meteorology knowledge, is less biased by spectroscopic errors, and is computationally inexpensive to run since it does fit spectra directly. By inferring information from the differences between total column DMF values from spectral windows that are quality controlled, the retrieval is restricted to imposing small changes to the partial and total columns. This effectively limits the amount of informational content that can be retrieved but also mitigates the issues of oscillation or large deviations in the retrieved vertical profile, partial columns in this case. Finally, this algorithm takes advantage of the temporal dimension by fitting over an entire day of measurements to retrieve enough information to infer temporal changes in the lower (surface to ~2 km) and upper (2 to 70 km) partial columns which also allows for the input of external, a priori, temporal information that is shown to improve the information content in the lower partial column fit.

Using measurements from the five North American TCCON sites, we compare our retrieved partial columns of CO and CO_2 DMF to the partial columns calculated from integrated, smoothed in situ data measured by aircraft and AirCore. We report slopes of 1.001±0.003 and 0.999±0.001 for the lower and upper partial column CO_2 comparisons, respectively, and slopes of 0.999±0.002 and 1.081±0.012 for the lower and upper partial column CO comparisons, respectively. The retrieved partial columns have improved direct comparisons and precision compared to the partial columns calculated from the original TCCON spectral windows.

We use the comparison data to calculate validation error multiplier (VEM) values to scale retrieved errors to be representative of the in situ comparisons. The average VEM scaled errors for the lower partial column CO and CO_2 retrievals are 1.51 ppb (~2%) and 5.09 ppm (~1.25%), respectively. The magnitudes of these error values suggest that the TARDISS retrieval will be useful in its current state for understanding surface fluxes of CO and will have some power for evaluating surface fluxes of CO_2.

The Bayesian TARDISS algorithm enables the informational content of the retrieval to be estimated. The average DoF for the lower partial column retrievals are 8.89 and 27.4 degrees of freedom so that ~9 and ~27 lower partial column values can be retrieved over a day of measurement for CO_2 and CO, respectively. The information content is affected by the parameters of the retrieval so that there is a tradeoff between retrieved error and the DoF of the retrieval. Furthermore, the daily DoF normalized by the number of measurements made in a day could serve as a quality control variable.

Future implementations of the retrieval could use the DoF values to create dynamic VEM to provide error values that are more precise than the static VEM. Similarly, future work could improve the effectiveness of the retrieval of lower partial column CO_2 using the TARDISS algorithm with the input of external information through the a priori covariance matrix, a priori partial column scalar, or the inclusion of the other parameters in the state vector.

Acknowledgements

We thank NASA via 80NSSC22K1066 for support of retrievals from the TCCON stations. A portion of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). BB is supported by NASA grant 80NSSC18K0898. The authors ~~would like to~~ thank the ObsPack ~~scientists~~team and data providers for the ~~use of the~~ in situ ~~profiles~~profile data used for validation ~~that were gathered during the various campaigns.~~. The data ~~are~~were downloaded from https://gml.noaa.gov/ccgg/obspack/data.php?id=obspack_co2_1_GLOBALVIEWplus_v5.0_2019-08-12 and were most recently accessed on September 2nd, 2022. In particular, we ~~would like to~~ thank the NASA LaRC AVOCET and DACOM groups for the KORUS-AQ CO_2 and CO data, respectively;

NASA Goddard for the Picarro CO_2 data at the Armstrong AFB; the SEAC4RS and ATom groups for the CO_2 data; the NOAA Global Monitoring ~~Division~~Laboratory for the AirCore CO_2 and CO data, and the long term aircraft CO_2 and CO at the SGP ~~ARM site~~ and ~~the~~ ETL ~~site. We thank Jochen Stutz for their effort in establishing and maintaining Caltech AERONET sites~~sites.

Data Availability

The data used in this study are made up of TARDISS retrieval products from five TCCON stations. The retrieval data are publicly available through CaltechDATA (https://doi.org/10.22002/pn9de-cry27) and the data input into the retrieval are publicly available via https://tccondata.org/. Retrieval code is currently available by request.

Author Contributions

HP wrote the TARDISS algorithm following an approach suggested by PW. HP retrieved the data with it and prepared the paper with thorough feedback from the coauthors. JL developed the theoretical framework for the TARDISS algorithm. CR retrieved the TCCON data using GGG for the Lamont, Caltech, and Park Falls sites. GCT gave input on the retrieval algorithm. DW gave input on the validation data and method. LTI and JRP maintain the Armstrong site. KM and BB provided insight and in situ data for the validation. All authors contributed to the review and editing of the work.

Competing Interests

The authors declare they have no conflicts of interest.

References

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., De Wekker, S. F. J., Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO2, CO, and CH4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmos. Meas. Tech., 7, 647–687, https://doi.org/10.5194/amt-7-647-2014, 2014.

Baier, B., Sweeney, C., Newberger, T., Higgs, J., Wolter, S., and NOAA Global Monitoring Laboratory: NOAA AirCore atmospheric sampling system profiles (20210813), https://doi.org/10.15138/6AV0-MY81, ~~2021~~2021a.

Baier, B., Sweeney, C., Newberger, T., Higgs, J., Wolter, S., and NOAA Global Monitoring Laboratory: NOAA AirCore atmospheric sampling system profiles (20210813), https://doi.org/10.15138/6AV0-MY81, 2021b.

Berger, B. W., Davis, K. J., Yi, C., and Bakwin, P. S.: Long-Term Carbon Dioxide Fluxes from a Very Tall Tower in a Northern Forest: Flux Measurement Methodology, JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY, 18, 14, 2001.

Buchholz, R., Deeter, M., Worden, H., Gille, J., Edwards, D., Hannigan, J., Jones, N., Paton-Walsh, C., Griffith, D., Smale, D., Robinson, J., Strong, K., Conway, S., Sussmann, R., Hase, F., Blumenstock, T., Mahieu, E., and Langerock, B.: Validation of MOPITT carbon monoxide using ground-based Fourier transform infrared spectrometer data from NDACC, Faculty of Science, Medicine and Health - Papers: part A, 1927–1956, https://doi.org/10.5194/amt-10-1927-2017, 2017.

Connor, B. J., Boesch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: Inverse method and prospective error analysis: OCO INVERSE METHOD, Journal of Geophysical Research: Atmospheres, 113, n/a-n/a, https://doi.org/10.1029/2006JD008336, 2008.

Connor, B. J., Sherlock, V., Toon, G., Wunch, D., and Wennberg, P. O.: GFIT2: an experimental algorithm for vertical profile retrieval from near-IR spectra, Atmospheric Measurement Techniques, 9, 3513–3525, https://doi.org/10.5194/amt-9-3513-2016, 2016.

Cooperative Global Atmospheric Data Integration Project: Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2017; obspack_co2_1_GLOBALVIEWplus_v4.2_2019-03-19, https://doi.org/10.25925/20190319, 2018.

Crawford, J. H., Ahn, J.-Y., Al-Saadi, J., Chang, L., Emmons, L. K., Kim, J., Lee, G., Park, J.-H., Park, R. J., Woo, J. H., Song, C.-K., Hong, J.-H., Hong, Y.-D., Lefer, B. L., Lee, M., Lee, T., Kim, S., Min, K.-E., Yum, S. S., Shin, H. J., Kim, Y.-W., Choi, J.-S., Park, J.-S., Szykman, J. J., Long, R. W., Jordan, C. E., Simpson, I. J., Fried, A., Dibb, J. E., Cho, S., and Kim, Y. P.: The Korea–United States Air Quality (KORUS-AQ) field study, Elementa: Science of the Anthropocene, 9, 00163, https://doi.org/10.1525/elementa.2020.00163, 2021.

Crosson, E. R.: A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor, Appl. Phys. B, 92, 403–408, https://doi.org/10.1007/s00340-008-3135-y, 2008.

Deeter, M. N.: Vertical resolution and information content of CO profiles retrieved by MOPITT, Geophys. Res. Lett., 31, L15112, https://doi.org/10.1029/2004GL020235, 2004.

Emmons, L. K., Pfister, G. G., Edwards, D. P., Gille, J. C., Sachse, G., Blake, D., Wofsy, S., Gerbig, C., Matross, D., and Nédélec, P.: Measurements of Pollution in the Troposphere (MOPITT) validation exercises during summer 2004 field campaigns over North America, J. Geophys. Res., 112, D12S02, https://doi.org/10.1029/2006JD007833, 2007.

Hedelius, J. K., Toon, G. C., Buchholz, R. R., Iraci, L. T., Podolske, J. R., Roehl, C. M., Wennberg, P. O., Worden, H. M., and Wunch, D.: Regional and urban column CO trends and anomalies as observed by MOPITT over 16 years, Geophys Res Atmos, https://doi.org/10.1029/2020JD033967, 2021.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A

Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sensing of Environment, 66, 1–16, https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

Iraci, L. T., Podolske, J. R., Roehl, C., Wennberg, P. O., Blavier, J.-F., Allen, N., Wunch, D., and Osterman, G. B.: TCCON data from Edwards (US), Release GGG2020.R0 (R0), https://doi.org/10.14291/TCCON.GGG2020.EDWARDS01.R0, 2022.

Karion, A., Sweeney, C., Tans, P., and Newberger, T.: AirCore: An Innovative Atmospheric Sampling System, Journal of Atmospheric and Oceanic Technology, 27, 1839–1853, https://doi.org/10.1175/2010JTECHA1448.1, 2010.

Keppel-Aleks, G., Wennberg, P. O., and Schneider, T.: Sources of variations in total column carbon dioxide, Atmos. Chem. Phys., 11, 3581–3593, https://doi.org/10.5194/acp-11-3581-2011, 2011.

Keppel-Aleks, G., Wennberg, P. O., Washenfelder, R. A., Wunch, D., Schneider, T., Toon, G. C., Andres, R. J., Blavier, J.-F., Connor, B., Davis, K. J., Desai, A. R., Messerschmidt, J., Notholt, J., Roehl, C. M., Sherlock, V., Stephens, B. B., Vay, S. A., and Wofsy, S. C.: The imprint of surface fluxes and transport on variations in total column carbon dioxide, Biogeosciences, 9, 875–891, https://doi.org/10.5194/bg-9-875-2012, 2012.

Kerzenmacher, T., Dils, B., Kumps, N., Blumenstock, T., Clerbaux, C., Coheur, P.-F., Demoulin, P., García, O., George, M., Griffith, D. W. T., Hase, F., Hadji-Lazaro, J., Hurtmans, D., Jones, N., Mahieu, E., Notholt, J., Paton-Walsh, C., Raffalski, U., Ridder, T., Schneider, M., Servais, C., and De Mazière, M.: Validation of IASI FORLI carbon monoxide retrievals using FTIR data from NDACC, Atmospheric Measurement Techniques, 5, 2751–2761, https://doi.org/10.5194/amt-5-2751-2012, 2012.

Kuai, L., Wunch, D., Shia, R.-L., Connor, B., Miller, C., and Yung, Y.: Vertically constrained CO2 retrievals from TCCON measurements, Journal of Quantitative Spectroscopy and Radiative Transfer, 113, 1753–1761, https://doi.org/10.1016/j.jqsrt.2012.04.024, 2012.

Masarie, K. A., Peters, W., Jacobson, A. R., and Tans, P. P.: ObsPack: a framework for the preparation, delivery, and attribution of atmospheric greenhouse gas measurements, Earth Syst. Sci. Data, 6, 375–384, https://doi.org/10.5194/essd-6-375-2014, 2014.

Olsen, S. C.: Differences between surface and column atmospheric CO 2 and implications for carbon cycle research, J. Geophys. Res., 109, D02301, https://doi.org/10.1029/2003JD003968, 2004.

Parrish, D. D., Xu, J., Croes, B., and Shao, M.: Air quality improvement in Los Angeles—perspectives for developing cities, Frontiers of Environmental Science & Engineering, 10, 1–13, https://doi.org/10.1007/s11783-016-0859-5, 2016.

Pougatchev, N. S., Connor, B. J., and Rinsland, C. P.: Infrared measurements of the ozone vertical distribution above Kitt Peak, J. Geophys. Res., 100, 16689, https://doi.org/10.1029/95JD01296, 1995.

Roche, S., Strong, K., Wunch, D., Mendonca, J., Sweeney, C., Baier, B., Biraud, S. C., Laughner, J. L., Toon, G. C., and Connor, B. J.: Retrieval of atmospheric CO2 vertical profiles from ground-based near-infrared spectra, Atmos. Meas. Tech., 14, 3087–3118, https://doi.org/10.5194/amt-14-3087-2021, 2021.

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, Repr., World Scientific, Singapore, 240 pp., 2008.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, Journal of Geophysical Research: Atmospheres, 108, n/a-n/a, https://doi.org/10.1029/2002JD002299, 2003.

Sepúlveda, E., Schneider, M., Hase, F., Barthlott, S., Dubravica, D., García, O. E., Gomez-Pelaez, A., González, Y., Guerra, J. C., Gisi, M., Kohlhepp, R., Dohe, S., Blumenstock, T., Strong, K., Weaver, D., Palm, M., Sadeghi, A., Deutscher, N. M., Warneke, T., Notholt, J., Jones, N., Griffith, D. W. T., Smale, D., Brailsford, G. W., Robinson, J., Meinhardt, F., Steinbacher, M., Aalto, T., and Worthy, D.: Tropospheric CH4 signals as observed by NDACC FTIR at globally distributed sites and comparison to GAW surface in situ measurements, Atmospheric Measurement Techniques, 7, 2337–2360, https://doi.org/10.5194/amt-7-2337-2014, 2014.

Shan, C., Wang, W., Liu, C., Guo, Y., Xie, Y., Sun, Y., Hu, Q., Zhang, H., Yin, H., and Jones, N.: Retrieval of vertical profiles and tropospheric CO2 columns based on high-resolution FTIR over Hefei, China, Opt. Express, 29, 4958, https://doi.org/10.1364/OE.411383, 2021.

Sweeney, C., Karion, A., Wolter, S., Newberger, T., Guenther, D., Higgs, J. A., Andrews, A. E., Lang, P. M., Neff, D., Dlugokencky, E., Miller, J. B., Montzka, S. A., Miller, B. R., Masarie, K. A., Biraud, S. C., Novelli, P. C., Crotwell, M., Crotwell, A. M., Thoning, K., and Tans, P. P.: Seasonal climatology of CO2 across North America from aircraft measurements in the NOAA/ESRL Global Greenhouse Gas Reference Network, Journal of Geophysical Research, 36, https://doi.org/10.1002/2014JD022591, 2015.

Tans, P.: System and method for providing vertical profile measurements of atmospheric gases, , 15, 2009.

Tans, P.: Fill dynamics and sample mixing in the AirCore, Atmos. Meas. Tech., 15, 1903–1916, https://doi.org/10.5194/amt-15-1903-2022, 2022.

Thompson, C. R., Wofsy, S. C., Prather, M. J., Newman, P. A., Hanisco, T. F., Ryerson, T. B., Fahey, D. W., Apel, E. C., Brock, C. A., Brune, W. H., Froyd, K., Katich, J. M., Nicely, J. M., Peischl, J., Ray, E., Veres, P. R., Wang, S., Allen, H. M., Asher, E., Bian, H., Blake, D., Bourgeois, I., Budney, J., Bui, T. P., Butler, A., Campuzano-Jost, P., Chang, C., Chin, M., Commane, R., Correa, G., Crounse, J. D., Daube, B., Dibb, J. E., DiGangi, J. P., Diskin, G. S., Dollner, M., Elkins, J. W., Fiore, A. M., Flynn, C. M., Guo, H., Hall, S. R., Hannun, R. A., Hills, A., Hintsa, E. J., Hodzic, A., Hornbrook, R. S., Huey, L. G., Jimenez, J. L., Keeling, R. F., Kim, M. J., Kupc, A., Lacey, F., Lait, L. R., Lamarque, J.-F., Liu, J., McKain, K., Meinardi, S., Miller, D. O., Montzka, S. A., Moore, F. L., Morgan, E. J., Murphy, D. M., Murray, L. T., Nault, B. A., Neuman, J. A., Nguyen, L., Gonzalez, Y., Rollins, A., Rosenlof, K., Sargent, M., Schill, G., Schwarz, J. P., Clair, J. M. St., Steenrod, S. D., Stephens, B. B., Strahan, S. E., Strode, S. A., Sweeney, C., Thames, A. B., Ullmann, K., Wagner, N., Weber, R., Weinzierl, B., Wennberg, P. O., Williamson, C. J., Wolfe, G. M., and Zeng, L.: The NASA Atmospheric Tomography (ATom) Mission: Imaging the Chemistry of the Global Atmosphere, Bulletin of the American Meteorological Society, 103, E761–E790, https://doi.org/10.1175/BAMS-D-20-0315.1, 2022.

Toon, O. B., Maring, H., Dibb, J., Ferrare, R., Jacob, D. J., Jensen, E. J., Luo, Z. J., Mace, G. G., Pan, L. L., Pfister, L., Rosenlof, K. H., Redemann, J., Reid, J. S., Singh, H. B., Thompson, A. M., Yokelson, R., Minnis, P., Chen, G., Jucks, K. W., and Pszenny, A.: Planning, implementation, and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by

Regional Surveys (SEAC4RS) field mission, J. Geophys. Res. Atmos., 121, 4967–5009, https://doi.org/10.1002/2015JD024297, 2016.

Turquety, S., Clerbaux, C., Law, K., Gloudemans, A. M. S., Schrijver, H., Boone, C. D., Bernath, P. F., and Edwards, D. P.: CO emission and export from Asia: an analysis combining complementary satellite measurements (MOPITT, SCIAMACHY and ACE-FTS) with global modeling, Atmos. Chem. Phys., 18, 2008.

Van Rooy, P., Tasnia, A., Barletta, B., Buenconsejo, R., Crounse, J. D., Kenseth, C., Meinardi, S., Murphy, S., Parker, H., Schulze, B., Seinfeld, J. H., Wennberg, P. O., Blake, D. R., and Barsanti, K. C.: Observations of Volatile Organic Compounds in the Los Angeles Basin during COVID-19, ACS Earth Space Chem., acsearthspacechem.1c00248, https://doi.org/10.1021/acsearthspacechem.1c00248, 2021.

Wennberg, P. O., Roehl, C.M., Wunch, D, Blavier, J.-F., Toon, G. C., Allen, N. T., Treffers, R., and Laughner, J.: TCCON data from Caltech (US), Release GGG2020.R0 (R0), https://doi.org/10.14291/TCCON.GGG2020.PASADENA01.R0, 2022a.

Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., and Allen, N. T.: TCCON data from Lamont (US), Release GGG2020.R0 (R0), https://doi.org/10.14291/TCCON.GGG2020.LAMONT01.R0, 2022b.

Wennberg, P. O., Roehl, C. M., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., and Allen, N. T.: TCCON data from Park Falls (US), Release GGG2020.R0 (R0), https://doi.org/10.14291/TCCON.GGG2020.PARKFALLS01.R0, 2022c.

Wiacek, A., Taylor, J. R., Strong, K., Saari, R., Kerzenmacher, T. E., Jones, N. B., and Griffith, D. W. T.: Ground-Based Solar Absorption FTIR Spectroscopy: Characterization of Retrievals and First Results from a Novel Optical Design Instrument at a New NDACC Complementary Station, J. Atmos. Oceanic Technol., 24, 432–448, https://doi.org/10.1175/JTECH1962.1, 2007.

Wofsy, S. C., Afshar, S., Allen, H. M., Apel, E. C., Asher, E. C., Barletta, B., Bent, J., Bian, H., Biggs, B. C., Blake, D. R., Blake, N., Bourgeois, I., Brock, C. A., Brune, W. H., Budney, J. W., Bui, T. P., Butler, A., Campuzano-Jost, P., Chang, C. S., Chin, M., Commane, R., Correa, G., Crounse, J. D., Cullis, P. D., Daube, B. C., Day, D. A., Dean-Day, J. M., Dibb, J. E., DiGangi, J. P., Diskin, G. S., Dollner, M., Elkins, J. W., Erdesz, F., Fiore, A. M., Flynn, C. M., Froyd, K. D., Gesler, D. W., Hall, S. R., Hanisco, T. F., Hannun, R. A., Hills, A. J., Hintsa, E. J., Hoffman, A., Hornbrook, R. S., Huey, L. G., Hughes, S., Jimenez, J. L., Johnson, B. J., Katich, J. M., Keeling, R. F., Kim, M. J., Kupc, A., Lait, L. R., McKain, K., Mclaughlin, R. J., Meinardi, S., Miller, D. O., Montzka, S. A., Moore, F. L., Morgan, E. J., Murphy, D. M., Murray, L. T., Nault, B. A., Neuman, J. A., Newman, P. A., Nicely, J. M., Pan, X., Paplawsky, W., Peischl, J., Prather, M. J., Price, D. J., Ray, E. A., Reeves, J. M., Richardson, M., Rollins, A. W., Rosenlof, K. H., Ryerson, T. B., Scheuer, E., Schill, G. P., Schroder, J. C., Schwarz, J. P., St.Clair, J. M., Steenrod, S. D., Stephens, B. B., Strode, S. A., Sweeney, C., Tanner, D., Teng, A. P., Thames, A. B., Thompson, C. R., Ullmann, K., Veres, P. R., Wagner, N. L., Watt, A., Weber, R., Weinzierl, B. B., Wennberg, P. O., Williamson, C. J., Wilson, J. C., et al.: ATom: Merged Atmospheric Chemistry, Trace Gases, and Aerosols, Version 2, https://doi.org/10.3334/ORNLDAAC/1925, 2021.

Wolfe, G. M., Kawa, S. R., Hanisco, T. F., Hannun, R. A., Newman, P. A., Swanson, A., Bailey, S., Barrick, J., Thornhill, K. L., Diskin, G., DiGangi, J., Nowak, J. B., Sorenson, C., Bland, G., Yungel,

J. K., and Swenson, C. A.: The NASA Carbon Airborne Flux Experiment (CARAFE): instrumentation and methodology, Atmos. Meas. Tech., 11, 1757–1776, https://doi.org/10.5194/amt-11-1757-2018, 2018.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmospheric Measurement Techniques, 3, 1351–1362, https://doi.org/10.5194/amt-3-1351-2010, 2010.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369, 2087–2112, https://doi.org/10.1098/rsta.2010.0240, 2011.

Xu, K., Pingintha-Durden, N., Luo, H., Durden, D., Sturtevant, C., Desai, A. R., Florian, C., and Metzger, S.: The eddy-covariance storage term in air: Consistent community resources improve flux measurement reliability, Agricultural and Forest Meteorology, 279, 107734, https://doi.org/10.1016/j.agrformet.2019.107734, 2019.

Zhou, M., Langerock, B., Vigouroux, C., Sha, M. K., Ramonet, M., Delmotte, M., Mahieu, E., Bader, W., Hermans, C., Kumps, N., Metzger, J.-M., Duflot, V., Wang, Z., Palm, M., and De Mazière, M.: Atmospheric CO and CH4 time series and seasonal variations on Reunion Island from ground-based in situ and FTIR (NDACC and TCCON) measurements, Atmos. Chem. Phys., 18, 13881–13901, https://doi.org/10.5194/acp-18-13881-2018, 2018.

Zhou, M., Langerock, B., Vigouroux, C., Sha, M. K., Hermans, C., Metzger, J.-M., Chen, H., Ramonet, M., Kivi, R., Heikkinen, P., Smale, D., Pollard, D. F., Jones, N., Velazco, V. A., García, O. E., Schneider, M., Palm, M., Warneke, T., and De Mazière, M.: TCCON and NDACC XCO measurements: difference, discussion and application, Atmos. Meas. Tech., 12, 5979–5995, https://doi.org/10.5194/amt-12-5979-2019, 2019.