Supplement of Atmos. Meas. Tech., 16, 2601–2625, 2023
https://doi.org/10.5194/amt-16-2601-2023-supplement

Atmospheric Measurement Techniques
Open Access

Supplement of

Inferring the vertical distribution of CO and $\mathrm{CO_2}$ from TCCON total column values using the TARDISS algorithm

Harrison A. Parker et al.

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

Supplemental Information:

Variable Name	Variable Description	Defining Equation
$\mathbf{A}$	TARDISS temporal averaging kernel	~
$\mathbf{A}_{\mathrm{vert}}$	TARDISS sensitivity as it relates to the vertical profile	22
$\boldsymbol{a}_{TCCON}^{\xi}$	TCCON column averaging kernel vector dotted with an integration operator	~
$\boldsymbol{a}_{\mathrm{TCCON}}$	TCCON column averaging kernel vector	~
χ	Cost of retrieval	1
$\mathbf{G}$	Gain matrix of temporal retrieval	21
DoF	Degrees of Freedom of signal	18
γ_L	Lower partial column scalar	~
γ_U	Upper partial column scalar	~
$\gamma_{a,L}$	A priori lower partial column scalar	~
$\gamma_{a,U}$	A priori upper partial column scalar	~
H	Shannon information content	19
$\mathbf{I}$	Identity matrix	~
i	Index value	~
$\boldsymbol{K}$	TARDISS Jacobian matrix	12
k_L	Lower partial column TARDISS Jacobian element for one window for one spectrum	8
k_U	Upper partial column TARDISS Jacobian element for one window for one spectrum	9
n_l	Number of levels in a vertical profile	~
n_s	Number of TCCON spectral measurements in a day	~
n_w	Number of TCCON windows used in the TARDISS retrieval	~
q	Index of the top of the lower partial column part of the profile	~
$\boldsymbol{S}_a$	A priori covariance matrix	~
$\boldsymbol{S}_\epsilon$	Model covariance matrix	~

$\boldsymbol{S}_r$	Retrieval noise matrix	26
$\boldsymbol{S}_s$	Smoothing error matrix	25
$\boldsymbol{\sigma}$	TARDISS retrieval errors	~
VEM	Validation Error Multiplier	27
$\boldsymbol{x}_{\mathrm{a},\gamma}$	Vector of a priori partial column scalar values	14
$\mathrm{x}_{\mathrm{a,TCCON}}$	TCCON a priori profile times median of TCCON VSFs for one measurement	~
$\hat{\boldsymbol{x}}_\gamma$	Retrieved state vector of partial column scalar values	16
$\hat{x}_{\gamma L}$	Lower partial column element of the retrieved state vector	~
$\hat{x}_{\gamma U}$	Upper partial column element of the retrieved state vector	~
$\boldsymbol{x}_\gamma$	Theoretical state vector of partial column scalar values	13
$\boldsymbol{x}_{L2}$	Vector of partial column scalar values calculated via the least squares method	15
$\mathbf{x}_{\mathrm{part}}$	A posteriori profile	2
$\boldsymbol{x}_{true}$	In situ measured profile	~
$\boldsymbol{y}$	Measurement vector with elements defined by $z_{TCCON} - z_{a,TCCON}$	10, 11
$\boldsymbol{\Xi}_{TCCON}$	Matrix of TCCON column averaging kernels for each window and each spectrum within a day	23
z_{TCCON}	TCCON column average mole fraction value for one window and one specturm	3
$z_{a,TCCON}$	TCCON a priori column average mole fraction times the median VSF of the windows used	~
$z_{a,L,TCCON}$	Lower partial column of TCCON a priori column average mole fraction times the median VSF of the windows used	~
$z_{a,U,TCCON}$	Upper partial column of TCCON a priori column average mole fraction times the median VSF of the windows used	~
$\hat{\boldsymbol{z}}_{comp}$	Vector of partial column mole fractions used for comparison to the smoothed column averaged in situ mole fraction	~
$\boldsymbol{z}_{PC}$	Vector of reconstructed partial column mole fraction values for a day of measurement	17
$z_{PC,L}$	Lower partial column element of the reconstructed partial column mole fraction vector	~
$z_{PC,U}$	Upper partial column element of the reconstructed partial column mole fraction vector	~
$\hat{z}_s$	Smoothed column or partial column averaged in situ mole fraction	20, 24

Table S1. Table of variable names, descriptions, and defining equation for all the variables used in the work.

S1 Temporal Assimilation

To test the influence of the number of observations included in each TARDISS retrieval, we compare the retrieved error value for each individual retrieval and with an increasing number of observations until we use the full day of observations. In this test, 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 a priori 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. S1 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. S1. We repeat this method, expanding the number of observations included until we use the entire day of observations.

The left-hand plot of Fig. S1 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. In contrast, the right-hand plot of Fig. S1 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 eight 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 a priori 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 a priori covariance matrix. Shown in Fig. S2, 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 constraints and external information to improve and inform the retrieval.

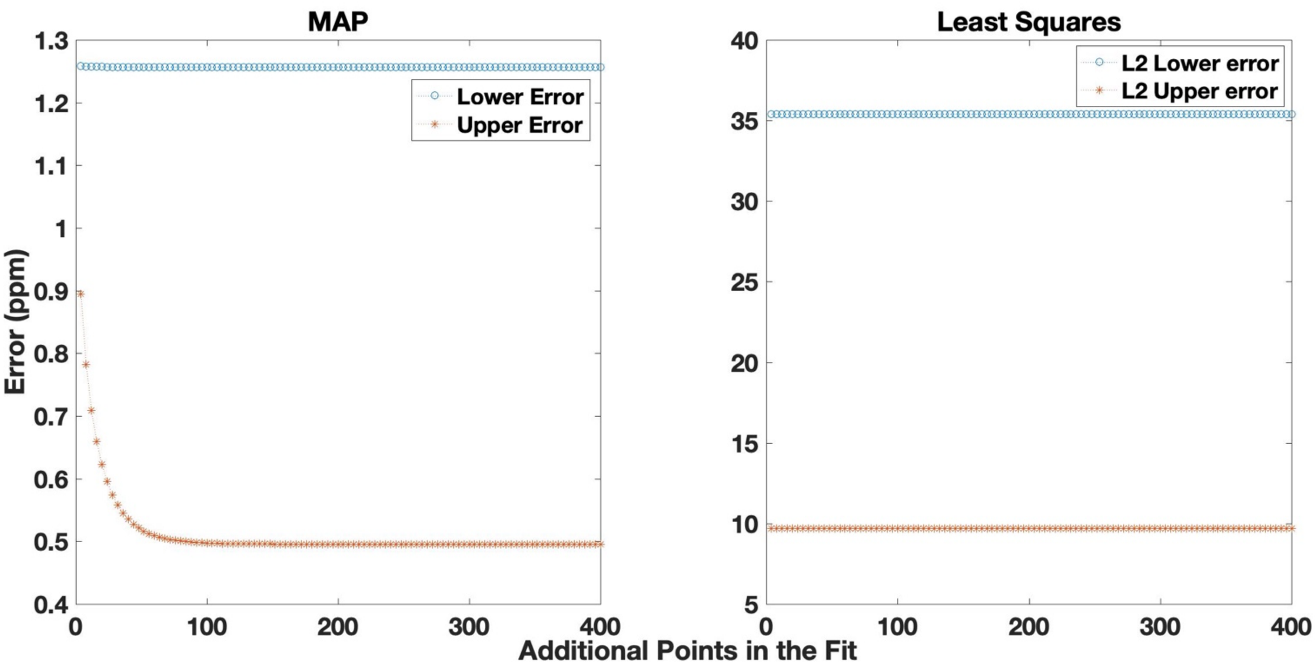

Figure S1. Errors in the retrieval of CO_2 from the midday total column measurement at the Park Falls site on July 18, 2018 using the MAP method outlined by Equation 13 and the least squares method outlined by Equation 12. The blue circles represent the error in the lower partial column and the orange 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 spectrum.

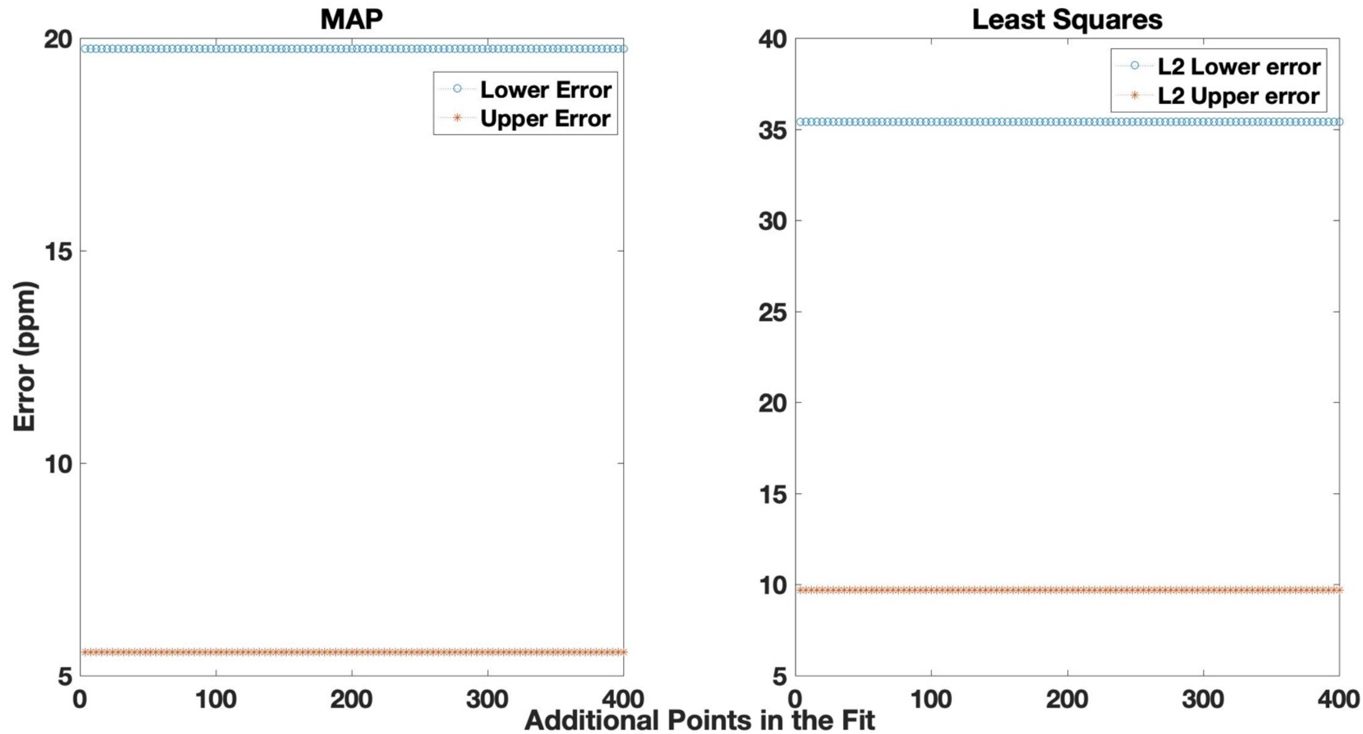

Figure S2. Same as Fig. S1, except the a priori covariance is removed from the MAP retrieval.

Site	Type	Species	Campaign/ Program	Data Availability	Dates	Altitudes
Park Falls	Aircraft	CO_2	ATom	https://doi.org/10.3334/ORNLDAAC/1925	20160822	0.79 - 12 km
	AirCore	CO_2	NOAA AirCore	https://doi.org/10.15138/6AV0-MY81	20180730 20180731	Surf. – 21km
Armstrong	Aircraft	CO_2	SEAC4RS	https://doi.org/10.3334/ORNLDAAC/1925	20130923	1.5 - 19 km
Armstrong	Aircraft	CO_2	ATom	https://doi.org/10.3334/ORNLDAAC/1925	20140820	0.79 - 12 km

Armstrong	Aircraft	CO_2	GSFC	https://doi.org/10.25925/20190319	20140820 20140822 20151002 20160210	0.6 - 13 km
Armstrong	Aircraft	CO_2	KORUS-AQ	https://doi.org/10.1525/elementa.2020.00163	20160618	0.68 - 12 km
Armstrong	AirCore	CO_2	NOAA AirCore	https://doi.org/10.15138/6AV0-MY81	20180716 20180717 20180718	Surf. – 21 km
Lamont	AirCore	CO_2, CO	NOAA AirCore	https://doi.org/10.15138/6AV0-MY81	20180723 20180725 20180727	Surf. – 21 km Surf. – 17 km
Lamont	Aircraft	CO_2, CO	NOAA GGGRN aircraft program	https://doi.org/10.1002/2014JD022591, 2015.	2008 - 2018	0.17 - 6 km
East Trout Lake	Aircraft	CO_2, CO	NOAA GGGRN aircraft program	https://doi.org/10.1002/2014JD022591, 2015.	2017 - 2020	0.17 - 7 km

Table S2. Site, measurement type, species, campaign or program, citation, and dates of the in situ profile data used in this work.

TCCON Window (cm^{-1})	Total Column Validation Slope	Total Column Validation Slope Error	Total Column Mean Ratio Deviation
CO_2			
6220	1.007	0.001	0.001
6339	1.004	0.001	0.002
6073	1.006	0.001	0.001
4852	1.006	0.001	0.003

TARDISS CO_2	1.000	0.0004	0.001
CO			
4290	1.009	0.055	0.056
2160	1.033	0.020	0.041
2111	1.053	0.020	0.052
TARDISS CO	1.047	0.019	0.052

Table S3. Comparisons of the TARDISS total column retrieval to the total column comparisons of the fits of the TCCON spectral windows 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.

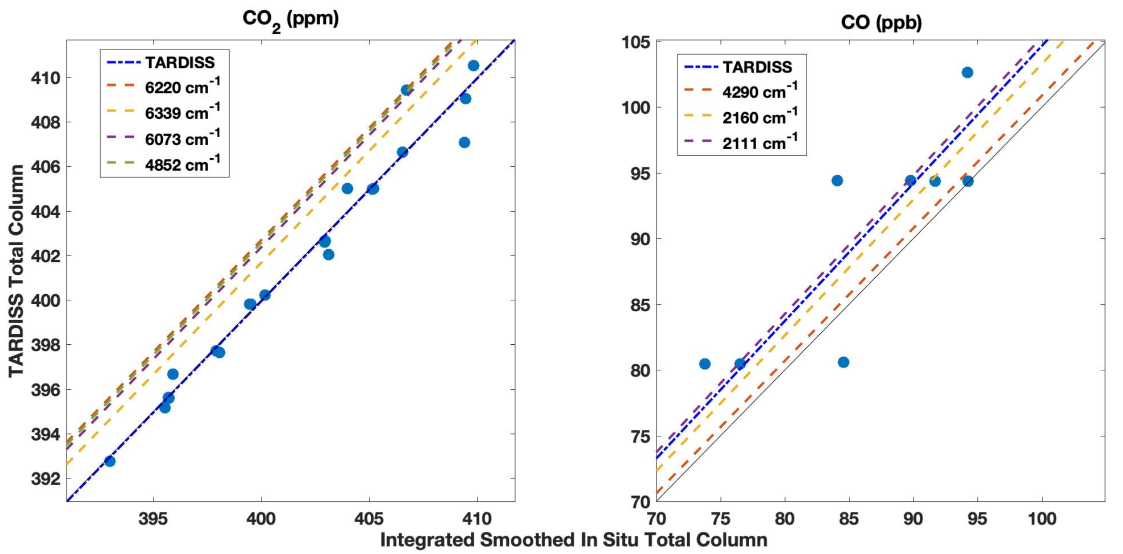

Figure S3. The direct comparisons between the total column DMF values retrieved from the TARDISS fit and the integrated, smoothed in situ partial columns for CO_2 (left) and the CO (right). The black solid line is the 1-1 line and the blue dot-dash line is the linear fit of the data with the y-intercept forced through zero. The slopes of the partial column validation of the TCCON spectral windows used in the retrieval are represented by dashed lines.

Site	Long Term Total Degrees of Freedom per Measurement	Long Term Comparison Slope	Lower partial column VEM from long-term data	Long term VEM total lower partial column error (ppm for CO_2; ppb for CO)
CO_2				

Lamont	0.0473	1.002	1.00	1.23
East Trout Lake	0.0543	1.001	1.30	1.64
CO				
Lamont	0.144	1.000	1.00	1.18
East Trout Lake	0.155	0.945	6.97	8.14

Table S4. DoF, comparison slopes, VEM, and total errors in the CO and CO_2 lower partial column retrievals for the long term comparisons performed at the Lamont and East Trout Lake sites. The values for total retrieval error and total error represent one standard deviation.

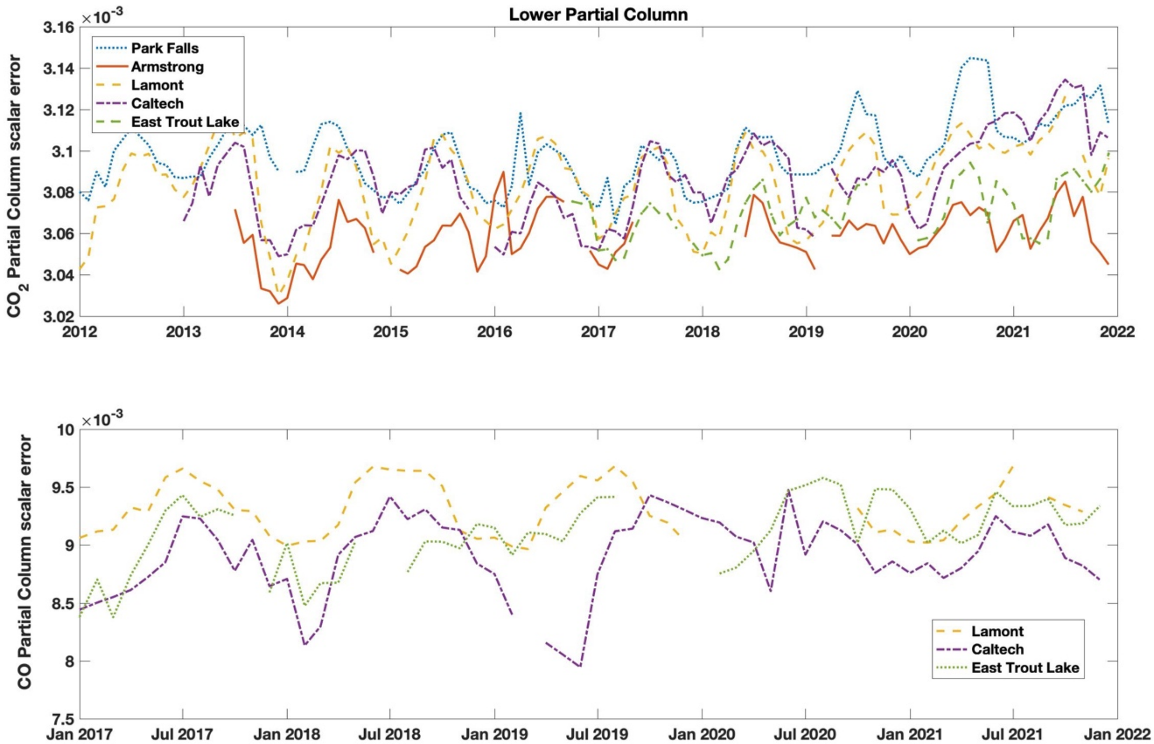

Figure S4. Monthly mean lower partial column scalar errors plotted for CO_2 (top) and CO (bottom).

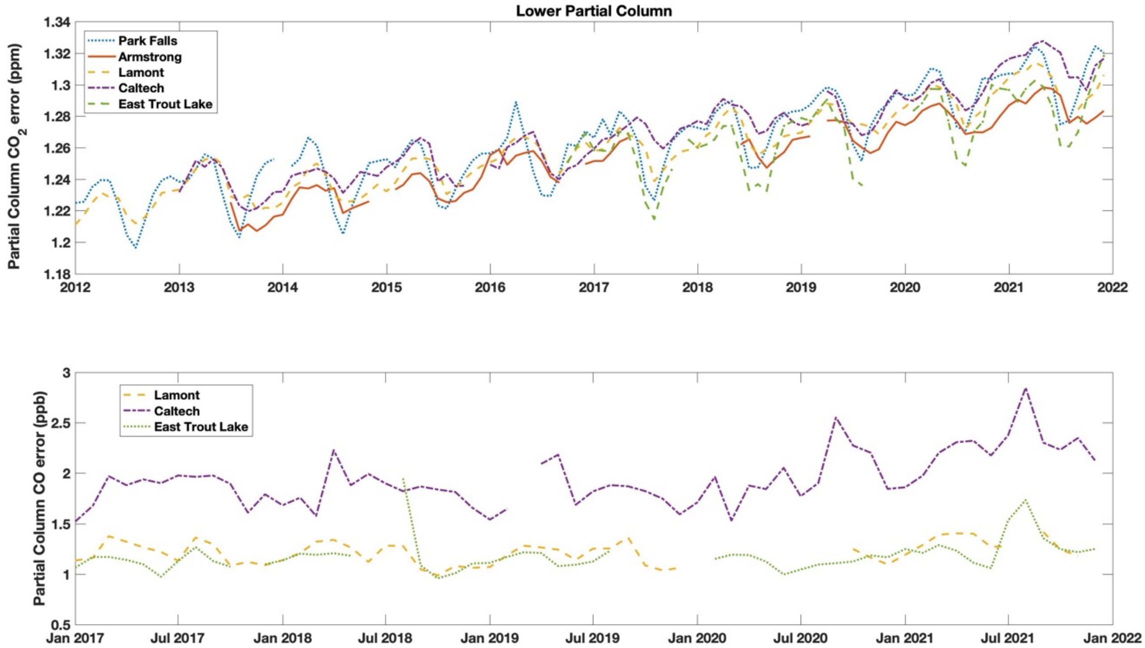

Figure S5. Monthly mean lower partial column errors plotted for CO_2 in ppm (top) and CO in ppb (bottom).

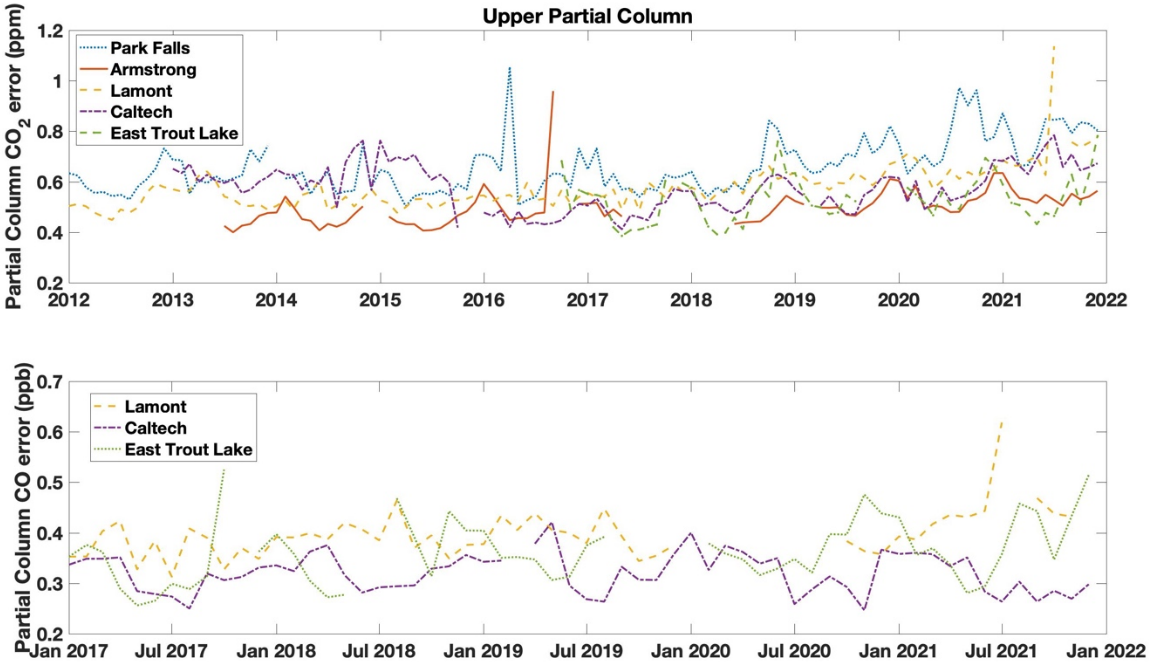

Figure S6. Monthly mean upper partial column errors plotted for CO_2 in ppm (top) and CO in ppb (bottom).

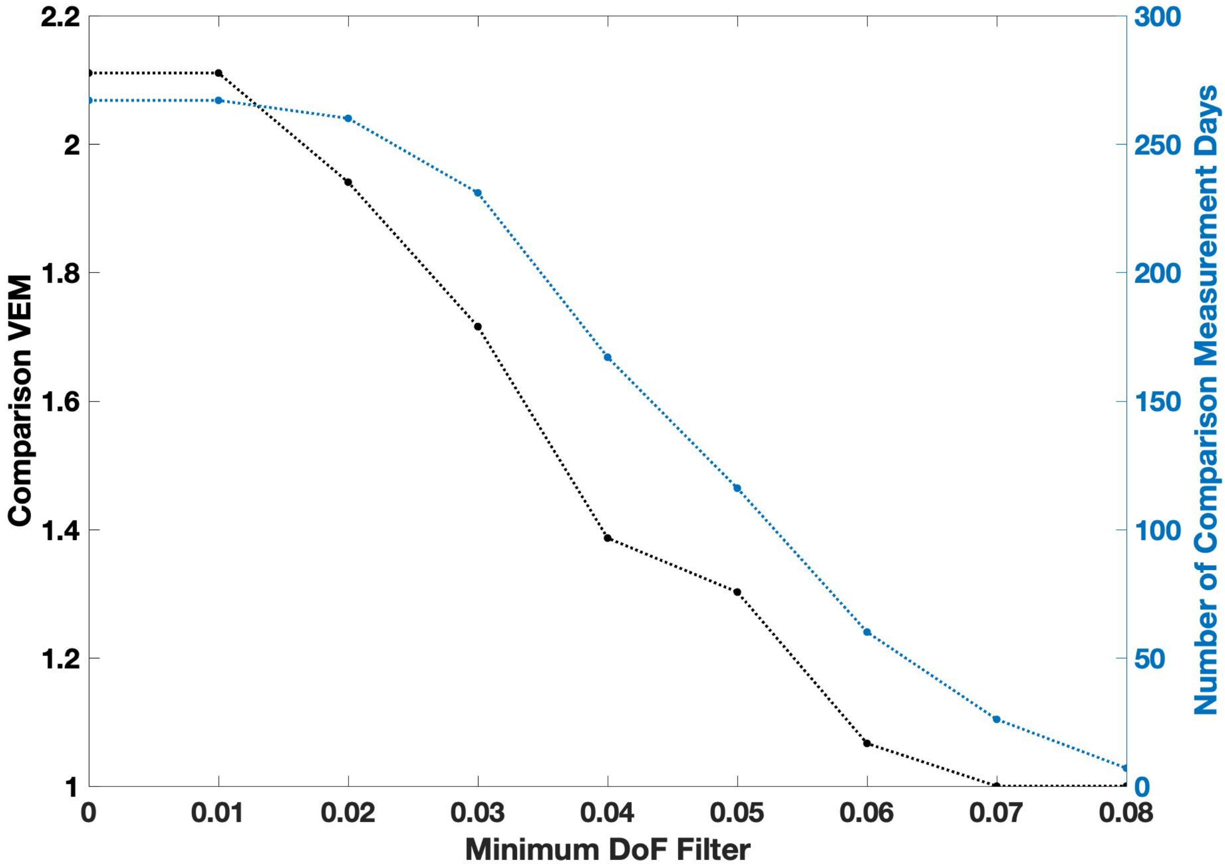

Figure S7. Comparison Validation Error Multiplier (VEM) and number of comparison days plotted by the minimum DoF per measurement filter applied to the comparison data. The retrieved data is in comparison with in situ data measured as a part of the NOAA GGGRN (Global Greenhouse Gas Reference Network) Aircraft sites from 2008 -2018 at the Lamont measurement site.

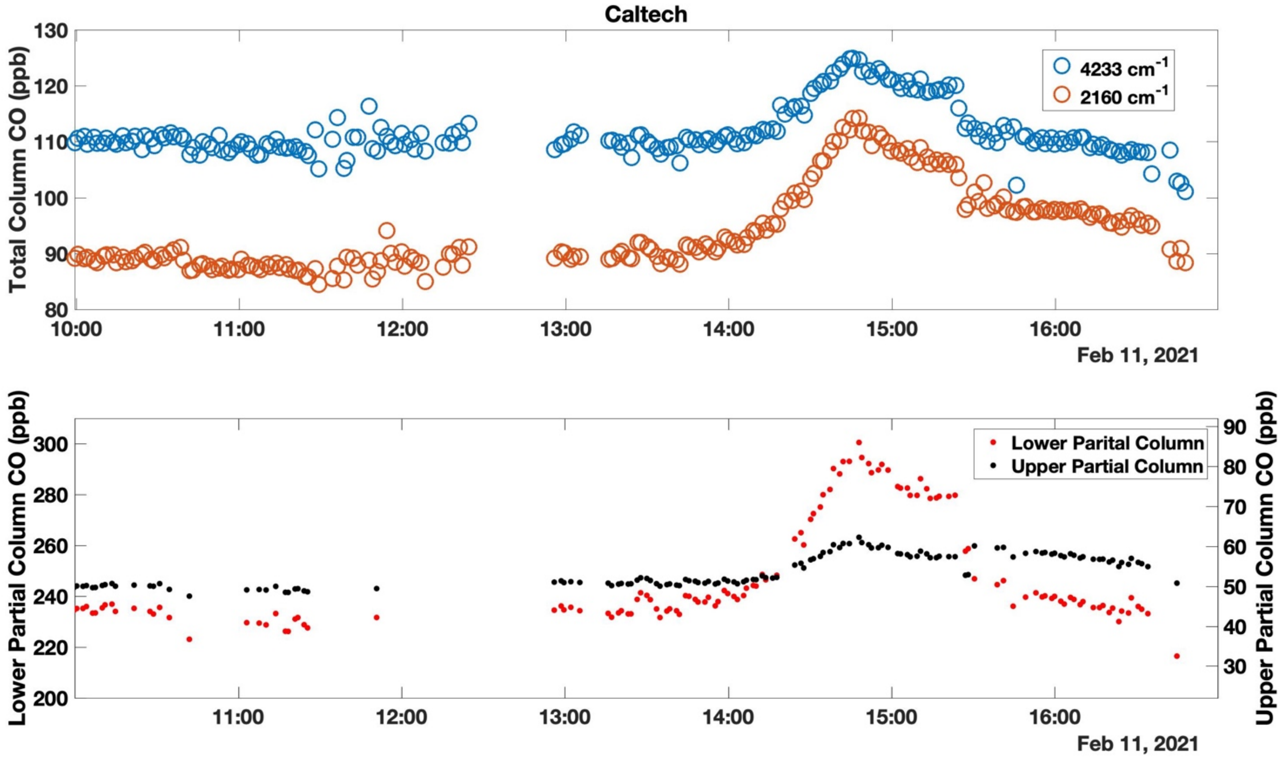

Figure S8. Example of one day of TCCON retrievals of total column CO (top) from two different spectral windows (4233 in blue and 2160 in orange) above the TARDISS partial column retrievals for CO (lower partial column in red dots, upper partial column in black dots) for the same day (bottom).