https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

A non-linear data driven approach to bias correction of XCO2 for OCO-2 NASA ACOS version 10

William Keely[1], Steffen Mauceri[2], Sean Crowell[3], Christopher O'Dell[4]

[1]Data Science and Analytics Institute, University of Oklahoma, Norman, OK, USA
[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[3]School of Meteorology, University of Oklahoma, Norman, OK, USA
[4]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA

Correspondence to: William Keely (william.r.keely@ou.edu)

Abstract.

Measurements of column averaged, dry air mole fraction of CO_2 (termed XCO_2) from the Orbiting Carbon Obersvatory-2 (OCO-2) contain systematic errors and regional scale biases; often induced by forward model error or nonlinearity in the retrieval. Operationally, these biases are corrected for by a multiple linear regression model fit to co-retrieved variables that are highly correlate with XCO_2 error. The operational bias correction is fit in tandem with a hand-tuned quality filter which limits error variance and reduces the regime of interaction between state variables and error to one that is largely linear. While the operational correction and filter are successful in reducing biases in retrievals, they do not allow for throughput or correction of data in which biases become nonlinear in predictors or features. In this paper, we demonstrate a clear improvement in the reduction of error variance over the operational method using a robust data driven, non-linear method. We further illustrate how the operational quality filter can be relaxed when used in conjunction with a non-linear bias correction, which allows for an increase of sounding throughput by 16% while maintaining the residual error of the operational correction. The method can readily be applied to future ACOS algorithm updates, OCO-2's companion instrument OCO-3, and to other retrieved atmospheric state variables of interest.

1 Introduction

Carbon dioxide (CO_2) is a key contributor to radiative forcing and hence, rising levels in the atmosphere are of concern due to their influence on future climate. Following a long history of critical in situ measurements of CO_2 at key sites around the world that allowed us to better understand the carbon cycle on continental scales, the era of space-based remote sensing began with the Greenhouse gases Observing SATellite (GOSAT) mission in 2009 (Kuze et al., 2009) and was followed by the Orbital Carbon Observatory 2 (OCO-2) in 2014. These data have yielded substantial scientific insights, such as a much more dynamic

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

tropical carbon cycle compared with previous understanding (e.g., Liu et al, 2017; Palmer et al, 2019; Crowell et al, 2019; Peiro et al, 2021) as well as studies into power plant emissions and plumes (Nassar et al., 2017).

OCO-2 measures reflected solar radiances, from which column averaged CO_2 dry air mole fractions (XCO_2) are retrieved with the NASA Atmospheric CO_2 Observations from Space (ACOS) algorithm (Crisp et al., 2012; O'Dell et al., 2012; Connor et al., 2008). Radiances are measured in the near-infrared Oxygen A band near 0.76 μm; the shortwave infrared weak CO_2 band near 1.6 μm; and the shortwave infrared strong CO_2 band near 2.05 μm). ACOS is based on Bayesian optimal estimation (Rogers, 2000) that adjusts input parameters (e.g., XCO_2, aerosols, surface characteristics, surface pressure) in order to maximize agreement between a modelled spectrum (derived by a radiative transfer model) and OCO-2 measurements. The parameters that best explain the measured radiances are labeled as the "retrieved" parameters. ACOS has undergone continuous improvement since the initial version.

Since the radiances contain uncorrected calibration artifacts and the modeled representation of the atmospheric radiative transfer is not perfect, retrieved parameters contain systematic biases. The inverse problem is underconstrained and leads to posterior errors in retrieved parameters that are correlated. In order to correct for errors in XCO2 arising from these types of dependencies, a multiple linear regression (MLR) bias correction with co-retrieved state variables or features used as predictors is fit to the difference (ΔXCO_2) between the ACOS retrieved XCO_2 and a truth proxy estimate of XCO_2. This method was first introduced for ACOS retrievals applied to the GOSAT instrument (Wunch et al., 2011), and later extended to OCO-2 and OCO-3 (O'Dell et al., 2018; Taylor et al., 2020, 2023). The multiple linear regression (MLR) bias correction is fit in tandem with a quality filter of epmerically defined thresholds on a set of features. The bias correction and quality filter are derived iteratively, with filter thresholds chosen restricting features to a range in which the relationship between ΔXCO_2 and the parameters are mostly linear, improving the goodness of fit for the multilinear regression, which is then used inturn to retune the quality filter thresholds. The combined bias correction and quality filtering process is derived manually, so that the final product must be hand-tuned for each algorithm update. After the feature-based correction, a footprint correction and global TCCON offset are applied. The combined bias correction and quality filter is robust across a set of ground truth proxy metrics and greatly reduce both mean bias and error scatter of XCO_2 retrieved from OCO-2. Full details of the operational bias correction and filtering can be found in O'Dell et al. (2018).

A drawback of applying quality filter is the exclusion of data due to the linear assumption of the bias correction. to which the quality filter limits the regime of interaction between state vector variables and ΔXCO_2. Due to loss of data, the bias correction and quality filter are often disregarded for local studies (Nassar et al., 2017; Mendonca et al., 2021) or too limiting for certain regions (Jacobs et. al., 2020). Applying non-linear data driven techniques is shown to improve the precision of the bias correction relative to the operational linear correction for GOSAT/GOSAT-2 (Noël et al., 2021) and TROPOMI (Schneising

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

et al., 2019). Specific correction of 3D cloud biases for OCO-2 retrieved XCO_2 (Massie et al. 2016) using a non-linear method in addition to the operational correction, is demonstrated in Mauceri et al. (2022).

This research demonstrates a general non-linear bias correction approach for OCO-2 build 10 (B10, Taylor et al., 2023) via an interpretable machine learning method. Our non-linear bias correction is shown to reduce systematic errors and increases the percentage of good quality soundings by allowing for the relaxation of the hand-tuned thresholds employed with the standard quality threshold method. Pursuing such a reproduceable data driven approach will be useful for future OCO-2,3 ACOS algorithm updates and for upcoming missions such as GeoCarb (Moore et al. 2018).

2 Data

In order to develop a bias correction, we define three truth proxy data sets for the true atmospheric column mode fraction. ΔXCO_2 is then set as the difference between the raw ACOS retrieval of XCO_2 and the truth proxy estimate of XCO_2:

$$\Delta XCO_2 = XCO_{2,ACOS} - XCO_{2,Proxy} \quad (1)$$

We use the same proxy data sets used in the development of the operational bias correction (Osterman et al. 2020): co-located OCO-2 soundings with Total Carbon Column Observing Network (TCCON), a collection of small-area clusters of soundings for which XCO_2 is not expected to vary above the instrument noise, and a set of modelled mole fractions whose underlying surface flux is constrained by the NOAA global in situ network (Masarie, K., et al. 2014). Data sets include soundings from November 2014 through to Febuary 2019. Each truth proxy captures a different scale of retrieval error and as such give complementary information as described in O'Dell et al. (2018). All data sets were sampled in conjunction with corresponding locations and times in the OCO-2 B10 L2 lite files which can be found here: (**https://disc.gsfc.nasa.gov/datasets/**). Spatial coverage and sounding count is shown in Figure 1. The newest version available Lever 2 product is build 11 (B11), however at the time of this writing is still under going re-processing.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

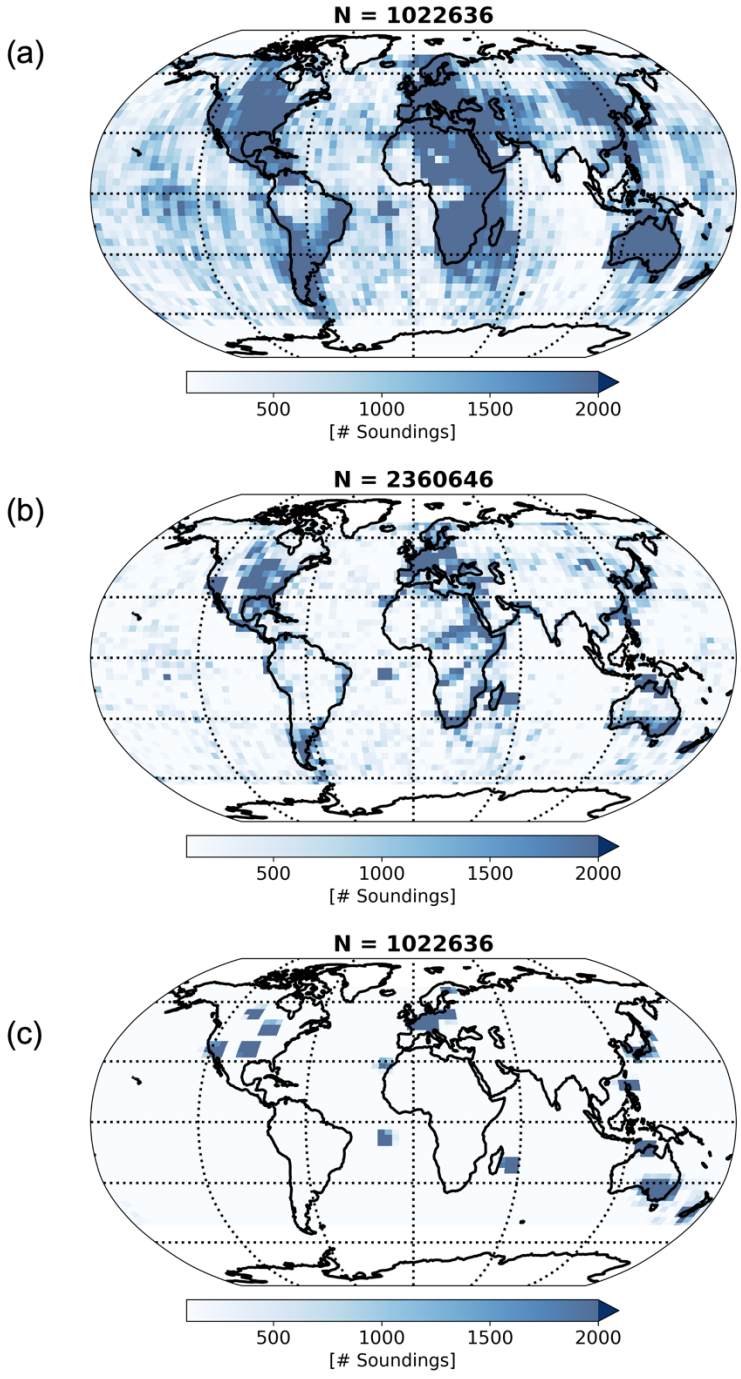


Figure 1: Spatial coverage for each truth proxy. The mean of a set of flux models is shown in (a), small area approximation is shown in (b), and TCCON is shown in (c).

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

2.1 TCCON truth proxy

TCCON is a system of ground-based sun-looking Fourier Transform Spectrometers with growing global coverage, that retrieve dry air mole column averaged measurements of the trace greenhouse gases from radiances in similar spectral bands to OCO-2. Since each site has been extensively validated against WMO-traceable in situ observations aboard aircraft, TCCON offers the most accurate comparison for XCO_2 (Wunch et al., 2010). While TCCON is well calibrated, site coverage is limited outside of North America, Europe, and Oceania. The TCCON data set therefore is spatially the sparsest of the three truth proxies and offering non-uniform point comparisons. Only OCO-2 soundings located near each site in space (2.5º lat, 5º lon) and time (2 h) are compared with TCCON retrievals (Wunch et al., 2017; Wunch et al., 2011).

Table 1. TCCON sites used for bias correction and filtering proxy the TCCON truth proxy for B10 ACOS.

TCCON (station name* = island)	Continent	Latitude	Altitude (m)	Operational date range (YYYYMM-YYYYMM)	Data citation
Saga*	Asia	33.2º N	7	201106-present	Kawakami et al. (2014)
Orléans	Europe	48.0º N	130	200908-present	Warneke et al. (2019)
Garmisch	Europe	47.5º N	740	200707-present	Sussman and Rettinger (2018)
Tsukuba*	Asia	36.1º N	188	200812-present	Morino et al. (2018a)
Sodankylä	Europe	67.4º N	30	200901-present	Kivi et al. (2014)
Rikubetsu	Asia	43.5º N	380	201311-present	Morino et al. (2016)
Izaña*	Africa	28.3º N	237	200705-present	Blumenstock et al. (2017)
JPL	N. America	34.2º N	390	201103-201307 201706-201805	Wennberg et al. (2016a)
Bialystok	Europe	53.2º N	180	200903-201810	Deutscher et al. (2019)
Bremen	Europe	53.1º N	27	200407-present	Notholt et al. (2019)
Wollongong	Australia	34.4º N	30	200805-present	Griffith et al. (2014b)
Park Falls	N. America	45.9º N	440	200405-present	Wennberg et al. (2017)
Réunion*	Africa	20.9º N	87	201109-present	De Maziére et al. (2017)
Anmyeondo	Asia	36.5º N	30	201408-present	Goo et al. (2014)
Darwin	Australia	12.4º S	30	200508-present	Griffith et al. (2014a)
Lauder*	Australia	45.0º S	370	200406-present	Pollard et al. (2019)
Lamont	N. America	33.2º N	320	200807-present	Wennberg et al. (2016b)
Karlsruhe	Europe	33.2º N	116	200909-present	Hase et al. (2015)
Manaus	S. America	3.2º S	49.2	201408-201506	Dubey et al. (2014)

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Paris	Europe	48.8° N	60	201409-present	Te et al. (2014)
Burgos*	Asia	18.5° N	35	201703-present	Morino et al. (2018b)

100

2.2 Small area approximation truth proxy

The small area approximation described in O'Dell et al. (2018) offers insight into small scale drivers of bias and retrieval variability. The small area approximation truth proxy assumes that XCO2 within a 100km neighborhood is largely uniform for
105 a given overpass by OCO-2. This assumption is evaluated in Worden et al. (2017), where it was found by using a fairly high-resolution atmospheric model (GEOS-5), variance of XCO_2, is around 0.1 ppm per 100km. The proxy overs improved spatial coverage compared to TCCON, but struggles to capture biases with low variability over the small area.

2.3 Flux models truth proxy

A set of flux inversion models form the largest of the truth proxy data sets, both in number of soundings and in spatial coverage.
110 The models included in this truth proxy set are found in Table 1. The posterior XCO_2 fields produced by the models are sampled along OCO-2 tracks, the proxy is then computed as the average of the models at every sounding where there is good agreement (within 1.5 ppm) among models (O'Dell et al., 2018; Osterman et al., 2020).

115 **Table 2. Flux models used for the model mean truth proxy. TM5 – Transport model 5, TM3 – Transport model 3, LMDZ – Laboratoire de Meteorology, EnKF – Ensemble Kalman Filter, 4D-Var – 4 Dimensional Variation.**

Model name	Institute	Transport model	Inverse method	Citation
CarbonTracker	NOAA Global Monitoring Laboratory	TM5	EnKF	Peters et al. (2007) CarbonTracker (2021)
CarboScope	Max Planck Institute for Biogeocehmistry	TM3	4D-Var	Rödenbeck (2005); Rödenbeck et al. (2018) CarboScope (2021)
CAMS	Copernicus Amosphere Monitoring Service	LMDZ	4D-Var	Chevallier et al. (2010) CAMS (2021)

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

3 Methods

3.1 Gradient boosting

To model systematic error from co-retrieved state vector elements, we employee a machine learning method known as extreme gradient boosting or XGBoost (Chen et al. 2016). XGBoost can fit both linear and non-linear relationships. This method is highly interpretable compared to other machine learning algorithms such as neural networks, which often require post-hoc methods to understand sensitivity to variables of importance. While methods like polynomial or spline fits are arguably more interpretable non-linear models, they do not capture multiplicative interactions among the input features that gradient boosters can (Friedman et al. 2001).

XGBoost is an ensemble model where a set of simple models known as regression trees (Breiman 1984) are fit and the average is then taken across the ensemble. XGBoost and gradient boosting in general are similar to other regression tree ensembling methods such as random forest, however, they provide improved predictive performance by fitting trees iteratively where each new tree is fit on the residuals of the prior trees in the ensemble. By limiting individual tree complexity and including L_1 and L_2 norm regularization strategies (the strength of each norm is regulated by λ, γ in Eq. 2) into the ensembling process, XGBoost is also highly robust to overfitting outliers present in the data. The regularization parameters must be tuned during model training. To do this we hold out small subset from the training data to evaluate different parameter selections. Final XGBoost models use $\lambda = 1.05$, and $\gamma = 7.65$.

Model interpretability is provided by the splitting criteria for developing the individual trees in the XGBoost ensemble. Trees are grown or fit by selecting features that provide high information gain (Eq. 2). Information gain is calculated by evaluating the sum of the gradients G and hessians H of the loss function at left and right leaf nodes when selecting features during tree fitting (for our experiements we use Mean Squared Error as the loss function shown in Eq. 3). Features that are informative for reducing residual error during tree development yield high gain values. These values can be summed across trees in the ensemble to produce a ranking of feature importance.

$$Information\ Gain = \frac{1}{2}\left[\frac{G_{Left}^2}{H_{Left}+\lambda} + \frac{G_{Right}^2}{H_{Right}+\lambda} - \frac{(G_{Left}+G_{Right})^2}{H_{Left}+H_{Right}+\lambda}\right] - \gamma\,, \tag{2}$$

$$Mean\ Squared\ Error\ loss = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2, \tag{3}$$

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

3.2 Quality filtering

Soundings of the lowest quality are typically caught by the O_2 A-band pre-processor (Taylor et al. 2012) and IMAP-DOAS (Frankenberg et al. 2005) algorithms due to clouds and low SNR in the continuum, and are then screened out before being run through the L2 retrieval algorithm (Taylor et al., 2016). After retrieval, an additional number of soundings are flagged and removed for which the ACOS algorithm failed to converge or for which the chi-squared difference between modelled and measured spectra is too large. Additionally, large unphysical outliers present in the tails of the conditional distributions of several atmospheric state variables are also removed by hand using domain expert selected thresholds. Finally, users can select for high-quality retrievals using the binary XCO_2 quality flag (QF) with "good" data having a QF = 0, and "poor" data having QF = 1. The operational XCO_2 quality flag is derived using a set of filters applied to the state vector variables found in conjunction with linear parametric bias correction. An initial linear correction is fit on soundings that have passed the pre-processing filtering steps. Each filter is then hand-selected, QF = 1 data is removed, and the bias correction is re-fit until a final set of filters and linear model weights are derived that sufficiently reduce mean bias and scatter (O'Dell et al 2018).

To assess the ability of the non-linear method to correct QF = 1 data and the potential for increased throughput of well corrected data, we derive a new quality flag (QFNew). Our flag is developed in a similar fashion to the B10 quality flag for use with the non-linear correction. The first step is to start with the same set of state vector variables and associated thresholds. Next, thresholds are relaxed for a selection of state vector variables that allow for higher sounding throughput, while maintaining or reducing corrected ΔXCO_2 across truth proxies. Thresholds are never set to be more constraining than the B10 values in order to not remove soundings that are already considered to be of passing quality.

3.3 Training and test split

For training and evaluating the non-linear correction, we subset each of our truth proxy datasets into a training and testing datasets. First, datasets are split by surface type and operation mode into ocean glint (Ocean G) and land nadir/glint (Land NG). In the B10 correction, both land modes are combined in the bias correction due to low variance in feature importance between nadir and glint (O'Dell et al. 2018). To compare to the operational correction, we also combine both modes over land. Once each truth proxy set is split into surface type and operation mode; we must further split the subsets into a training set for fitting the bias correction models, and a test set that is heldout for evaluation of model performance. To ensure that model performance is indicative of how well the models generalize to unseen data, we split the data by time. Models are trainined on data from 2014, 2015, 2016, and 2017, then evaluated on data from 2018. For model parameter tuning, we use approximately 10% of the training dataset.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

3.4 Experiment Design

Two methods are used in our bias correction: a non-linear machine learning model called XGBoost; and as a baseline, we also train a MLR model similar to the hand-tuned model used in the operational correction. XGBoost and MLR models are trained separately for Land nadir, Land glint, and Ocean glint observations. The predictor variables, or features are the same for both model types. This, allows for comparison between the non-linear model and baseline linear method to properly assess that improved fit is coming from the captured non-linearity and not just the inclusion of the additional predictors. We also compare our approach to the operational correction for B10.

Three experiments were conducted. First, we train models independently on each truth proxy to identify a subset of features to be used as predictors in our non-linear bias correction. Next we fit a non-linear model on QF = 0 + 1 data and a model on QF = 0 data then, evaluate each model on soundings that were assigned a "good" quality flag of 0 with the operational quality filter (QF = 0). This is to ensure the ability of the nonlinear method to reproduce the linear model, which is the currently accepted community standard. Secondly, we evaluate the model trained on QF = 0 + 1 data on exluded regime of data labelled QF = 1 where non-linear relationships between ΔXCO_2 and predictors become more pronounced. Finally, we derive a new quality flag (QFNew) used in conjunction with the non-linear correction to evaluate the ability of the method to increase sounding throughput by correcting outside of the regime of the operational flag.

4 Results

4.1 Feature Selection

We select informative features for our bias correction following an iterative procedure. In the first step, we train XGBoost models for each proxy by surface type and operation mode (6 models in total). These initial models are trained using a large subset of 27 co-retrievied state vector variables, which are potentially informative for correcting ΔXCO_2 from the B10 L2Lite files. The resulting models are used to rank features according to their information gain which is defined in Eq. 2. Features that are less informative are removed from the set and new models are trained with the reduced feature set. Afterwards, feature importance is once again evaluated. To ensure robustness to correlation among features we (which information gain does not account for) we calculate Pearson's correlation values between features and remove less informative features that are highly correlated with others. This process is iteratively repeated until reaching a relatively small subset of maximally informative features. These features are combined to train the final bias correction models: seven features for the Land NG correction and five features for the Ocean G. The minimum subset of maximally informative features varied between truth proxies as seen in Figure 2. For Land NG, 4 features are sufficient for the small area proxy, while the model mean and TCCON proxies require a larger subset of features. Similarly, for Ocean G, information gain drops off faster for small area and model mean proxies, while a larger set of features is required for TCCON. The resulting features used in the final models and a brief description is shown in Table 1.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Features used for the land correction include *co2_grad_del* which is the change in the retrieved vertical profile of CO_2. The quantity $dp\_frac = X_{CO2,raw}(1 - {P_{ap,SCO2}}/{P_{ret}})$ is used in both the operational and for the non-linear models first derived by Kiel et al. (2019). This difference ratio is a transformation of the previously used surface pressure changes from the prior that considers the smaller dry air column over higher elevations and is defined as where $X_{CO2,raw}$ is the uncorrected retrieval of the column average and $P_{ap,SCO2}$ and P_{ret} are the apriori surface pressure for the strong band and retrieved pressure respectively. For the operational correction, dp_frac replaces former surface pressure difference terms, and is also more informative than other surface pressure differences features for the non-linear method.

The *h2o_ratio* is the ratio of XH_2O estimated by single band retrivals from the strong and weak CO_2 bands separately using the IMAP-DOAS algorithm, which can differ from unity in the presence of atmospheric scattering (Taylor et al. 2016). In addition to these variables, we use three aerosol features for our bias correction over land scenes. The first being the sum of dust, water, and sea salt optical thickness termed *DWS***.** We include retrieved ice particle optical depth (*aod_ice*) and the finer stratospheric aerosol optical depth (*aod_stratear*). The last feature used for land, as well as for ocean glint, is the albedo slope for the strong CO_2 band termed *albedo_slope_sco2*. This variable represents the slope of the reflectance across the strong CO_2 spectral band for land soundings and the slope of the Lambertian component of the combined Cox-Munk and Lambertian Bidirectional Reflectance Distribution Function (BRDF) for ocean soundings (Cox and Munk, 1954).

In addition to the albedo_slope_sco2, four additional variables are used for the correction of Ocean G scenes. These are *dp_sco2*, the difference between the prior surface pressure and retrieved surface pressure from the strong CO_2 band, *snr_wco2*, the estimated signal to noise ratio derived during optimal estimation; and finally, *rms_rel_wco2* which is the percent residual error from the modeled radiance for the weak CO_2 to the measured radiance. *co2_grad_del* is used for ocean glint correction models.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

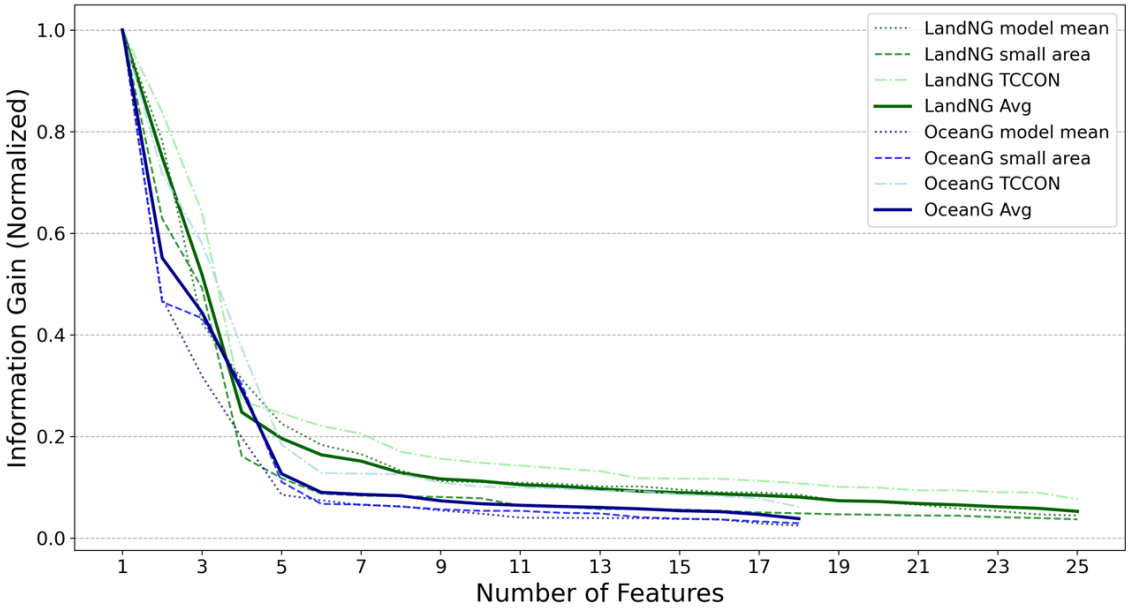


Figure 2. Number of features plotted against their information gain for Land NG and Ocean G truth proxies. Average gain for Land NG and Ocean G are plotted in the solid dark green and solid dark blue lines respectively. Features with gain below ~0.1 do not significantly increase the model's predictive skill and therefore, are not included.

Table 3. Selected features for in our bias correction models. The first column shows state vector variable names as defined in the B10 L2 files, second provides a brief description, and the last column shows which region and viewing mode correction the variable is used for.

State Variable	Description	Surface Type & Operation Mode
dpfrac	Surface pressure difference that considers smaller dry air columns over higher elevations (Kiel et al. 2019).	Land NG
h2o_ratio	Ratio of retrieved H_2O column in weak and strong CO_2 bands by IMAP-DOAS.	Land NG
DWS	Additive combination of retrieved dust, water, and sea salt aerosol optical depth.	Land NG
aod_stratear	Retrieved upper tropo+stratospheric aerosol optical depth at 0.755 microns.	Land NG
aod_ice	Retrieved ice cloud optical depth at 0.755 microns.	Land NG
co2_grad_del	Large unphysical variation between the retrieved vertical CO_2 profile and prior.	Land NG/Ocean G

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

albedo_slope_sco2	Retrieved strong band reflectance slope(land) or slope of Lambertian albedo component of BRDF (ocean).	Land NG/Ocean G
dp_sco2	Surface pressure difference between the retrieved and prior, evaluated for the strong CO_2 band location on the ground.	Ocean G
rms_rel_wco2	RMSE of the L2 fit residuals in the weak CO_2 band relative to the signal.	Ocean G
snr_wco2	The estimated signal-to-noise ratio in the continuum of the weak CO_2 band.	Ocean G

4.1 Model evaluation for QF = 0

To ensure that the non-linear method generalizes the linear relationships largely observed for QF = 0, we evaluate two XGBoost models: one which is fit on QF = 0 + 1, and one fit on QF = 0, to a MLR fit on the same feature set as the non-linear models. As the operational quality flag is hand-tuned by re-fitting a MLR, the regime between the variables selected for correction and systematic error are reduced to mostly linear relationships. The non-linear method has only a marginal improvement over the MLR and B10 correction on soundings that are passed by the operational quality filter over land (**0.04-0.06 ppm**), and a slightly more substantial improvement over ocean (**0.06-0.10 ppm**), on the evaluation data. We found that retraining the XGBoost models on QF = 0 data does not offer a substantial reduction in error despite initial XGBoost models being trained on un-filtered data. Due to XGBoost's robustness to overfitting QF = 1 outliers, we forgo the iterative refitting approach that is required for the MLR and operational correction by training once on QF = 0 + 1 data. Table 2 shows the QF = 0 RMSE results for XGBoost models trained on both QF = 0 + 1 data and QF = 0 data, alongside the MLR model fit to the filtered regime for 2018 and B10 operational correction.

Table 4: RMSE scores for 2018 on QF=0 data. Results are shown for Land and Ocean data by truth proxy and model. Two XGBoost models are shown: one trained on QF = 0 + 1 ($XGBoost_{QF0+1}$) data and the evaluated on QF = 0, and another ($XGBoost_{QF=0}$) trained and evaluated on only QF = 0 data. A multiple linear regression ($MLR_{QF=0}$) is also fit for QF = 0 using the same feature set. In the last column, RMSE for operationally corrected XCO_2 (B10) is shown.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Land QF=0				
Truth Proxy	**XGBoost$_{QF=0+1}$ RMSE**	**XGBoost$_{QF=0}$ RMSE**	**MLR$_{QF=0}$ RMSE**	**B10 RMSE**
Small area	0.81 ppm	0.81 ppm	0.89 ppm	0.84 ppm
TCCON	1.11 ppm	1.10 ppm	1.20 ppm	1.15 ppm
Model mean	1.06 ppm	1.05 ppm	1.14 ppm	1.11 ppm
Ocean QF=0				
Truth Proxy	**XGBoost$_{QF=0+1}$ RMSE**	**XGBoost$_{QF=0}$ RMSE**	**MLR$_{QF=0}$ RMSE**	**B10 RMSE**
Small area	0.45 ppm	0.44 ppm	0.58 ppm	0.51 ppm
TCCON	0.80 ppm	0.79 ppm	0.89 ppm	0.90 ppm
Model mean	0.66 ppm	0.66 ppm	0.81 ppm	0.75 ppm

4.2 Correcting Outside of the Filtered Regime

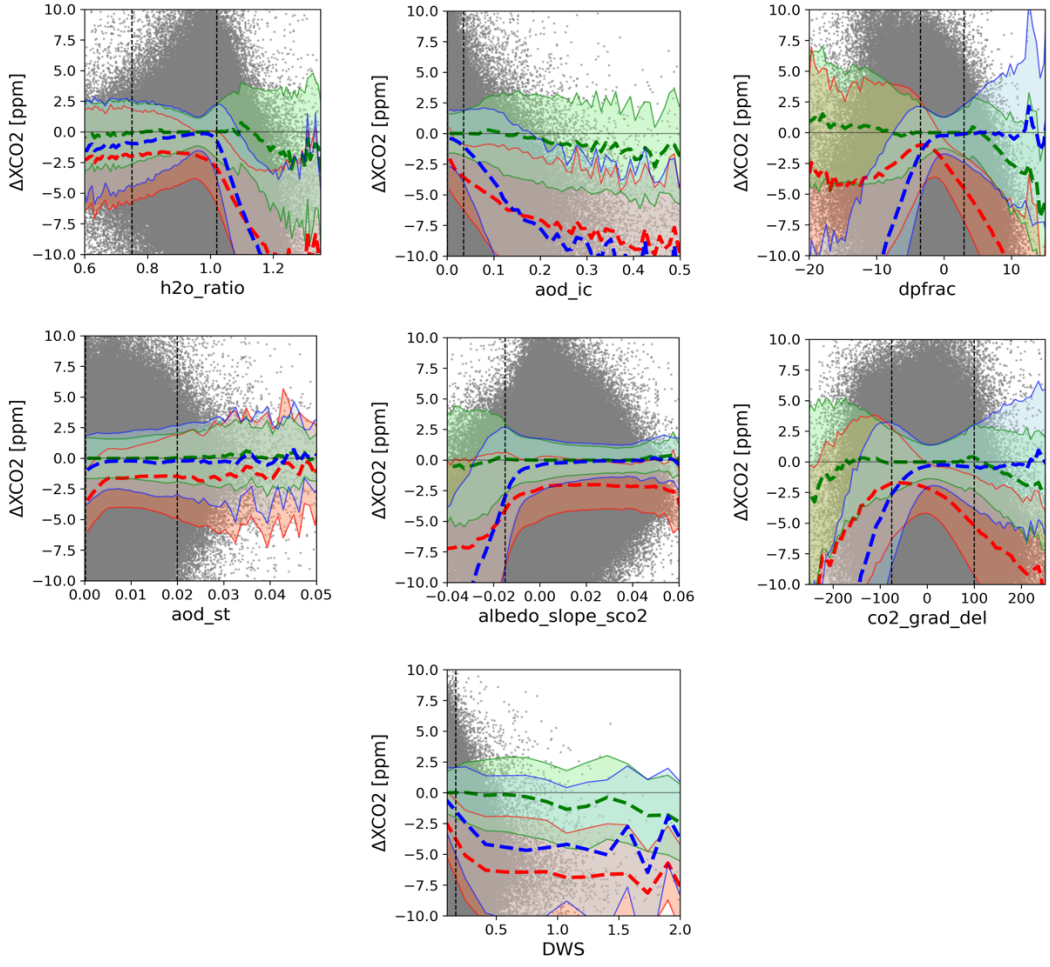
Correction of systematic error outside of the quality filtered regime (QF = 1) is difficult to fit with a linear model. Strong non-linearities are observed for many of the co-retrieved state vector variables and ΔXCO_2. For many variables this behaviour is observed over un-physical values in a few spurious soundings and is easily filtered out. However, for variables such as h2o_ratio which are responsible for the bulk of the quality filtering (variables that correlate with greater than 0.5 ppm bias or 2 ppm scatter) exhibit such characteristics over their marginal distributions (O'Dell et al. 2018). The dependent linear correction and quality filter is prohibitive for correcting and passing data in these regions of the domain. Figure 3 and 4 illustrate the interaction between state variables chosen for correction and ΔXCO_2. The non-linear model (green) improves both mean and variance of ΔXCO_2 over both the raw ΔXCO_2 (red) before correction, and B10 correction (blue). Table 3 displays the RMSE scores of the XGBoost corrected XCO_2 and operational corrected XCO_2 for QF = 1 data. The non-linear correction provides a large improvement in reducing the residual error for QF = 1 data over the operational correction with a **1.41-2.63 ppm** for land data and **1.11-1.43 ppm** for ocean. These errors are still significantly larger than the corresponding QF = 0 errors.

Table 5: RMSE scores for 2018 on QF = 1 data. XGBoost corrected XCO_2 and operationally corrected XCO_2 (B10) for Land and Ocean data.

Land QF = 1		
Truth Proxy	**XGBoost RMSE**	**B10 RMSE**
Small area	1.77 ppm	3.18 ppm
TCCON	2.42 ppm	5.05 ppm

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Model mean	2.35 ppm	4.14 ppm
	Ocean QF = 1	
Truth Proxy	**XGBoost RMSE**	**B10 RMSE**
Small area	1.08 ppm	2.19 ppm
TCCON	1.18 ppm	2.45 ppm
Model mean	1.39 ppm	2.82 ppm



290

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Figure 3. ΔXCO_2 vs land features. Mean interaction and 2σ Stddev for uncorrected ΔXCO_2 plotted in red, XGBoost corrected in green and B10 corrected in blue. The vertical black dotted lines indicate B10 QF filters and individual soundings are shown with grey scatter.

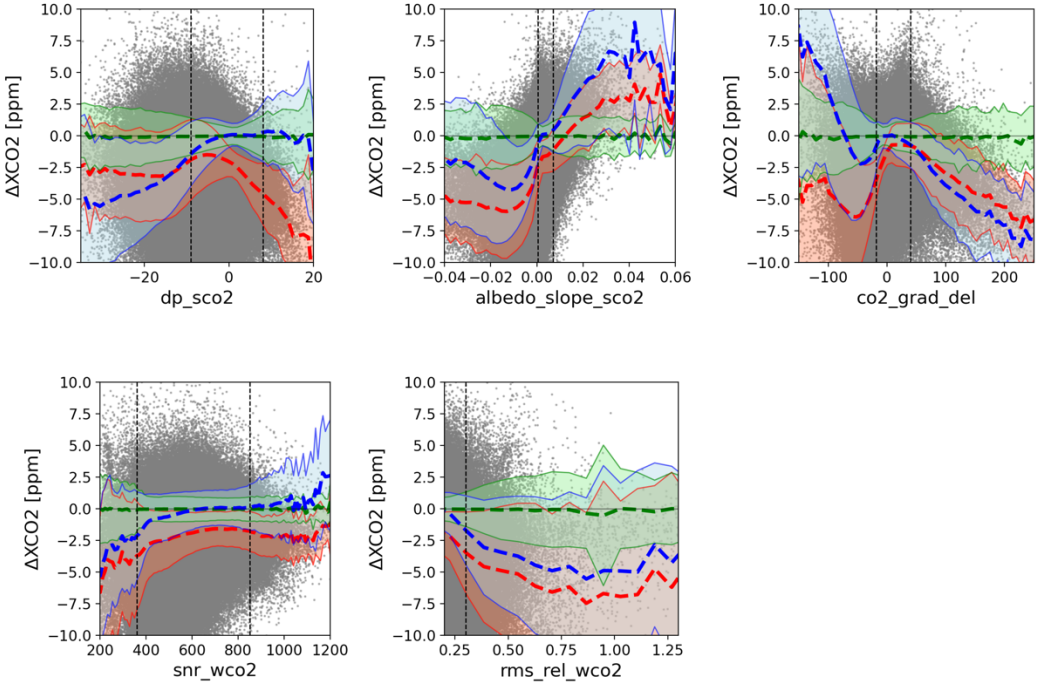


Figure 4. ΔXCO_2 vs ocean features. Mean interaction and 2σ Stddev for uncorrected ΔXCO_2 plotted in red, XGBoost corrected in green and B10 corrected in blue. The vertical black dotted lines indicate B10 QF filters and individual soundings are shown with grey scatter.

4.4 Comparison to B10

For the operational correction, regression weights for the linear model are hand-selected that have good agreement in their correction across truth proxies. The full operational correction also includes a fixed correction for each of OCO-2's eight footprints as described in Osterman et al. 2020. To provide a fair comparison between the full correction models, we also apply the footprint correction after applying the non-linear feature correction. Table 4 shows the mean and 1σ standard deviation for each bias correction and QF regime. The largest improvement in the non-linear method over B10 comes when correcting QF=1 data. Achieving a 66% improvement in reduction of error variance for land, and a 72% improvement for ocean data respectively. The improvement in correction over B10 is less significant for QF=0 with improvement of 8% for land and 4% for ocean.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Regionally, the non-linear correction shows up to a 0.5 ppm improvement over northern Africa, where the B10 correction appears to underestimate ΔXCO_2 in comparison. A reduction in biases is also observed in large parts of South America's tropical and sub-tropical regions as well as parts of tropical Asia shown in Figure 5a. These regions also contain the largest difference in Land NG correction between the methods with an average difference (B10-XGBoost) of -0.5 ppm. There is a slight positive difference between methods over the Amazon Basin and Congo Rainforest (Figure 5e). Figures 5c, and 5d illustrate the improvement of the non-linear method to correct QF=1 data over the operational approach. For QF = 1, where the interaction between features and error is non-linear, large biases in XCO_2 remain after operational correction. The XGBoost model reduces these remaining biases in many regions, indicating that there may still be usable data that is filtered out by the operational QF when paired with the non-linear correction.

Table 6. Comparison of mean and stddev XGBoost corrected XCO_2, XCO_2 after the operational correction (B10) and un-corrected XCO_2 (Raw) for 2018 and all QF filter regimes for both Land and Ocean data.

QF = 0		
Surface/Mode	**XGBoost**	**B10**
Land NG	0.00±0.98 ppm	-0.07±1.03 ppm
Ocean G	0.01±0.61 ppm	0.17±0.67 ppm
QF = 1		
Surface/Mode	**XGBoost**	**B10**
Land NG	0.02±2.24 ppm	-1.26±3.84 ppm
Ocean G	-0.03±1.32 ppm	-1.01±2.47 ppm
QF = 0 + 1		
Surface/Mode	**XGBoost**	**B10**
Land NG	0.00±1.59 ppm	-0.53±2.58 ppm
Ocean G	0.00±0.94 ppm	-0.26±1.70 ppm

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

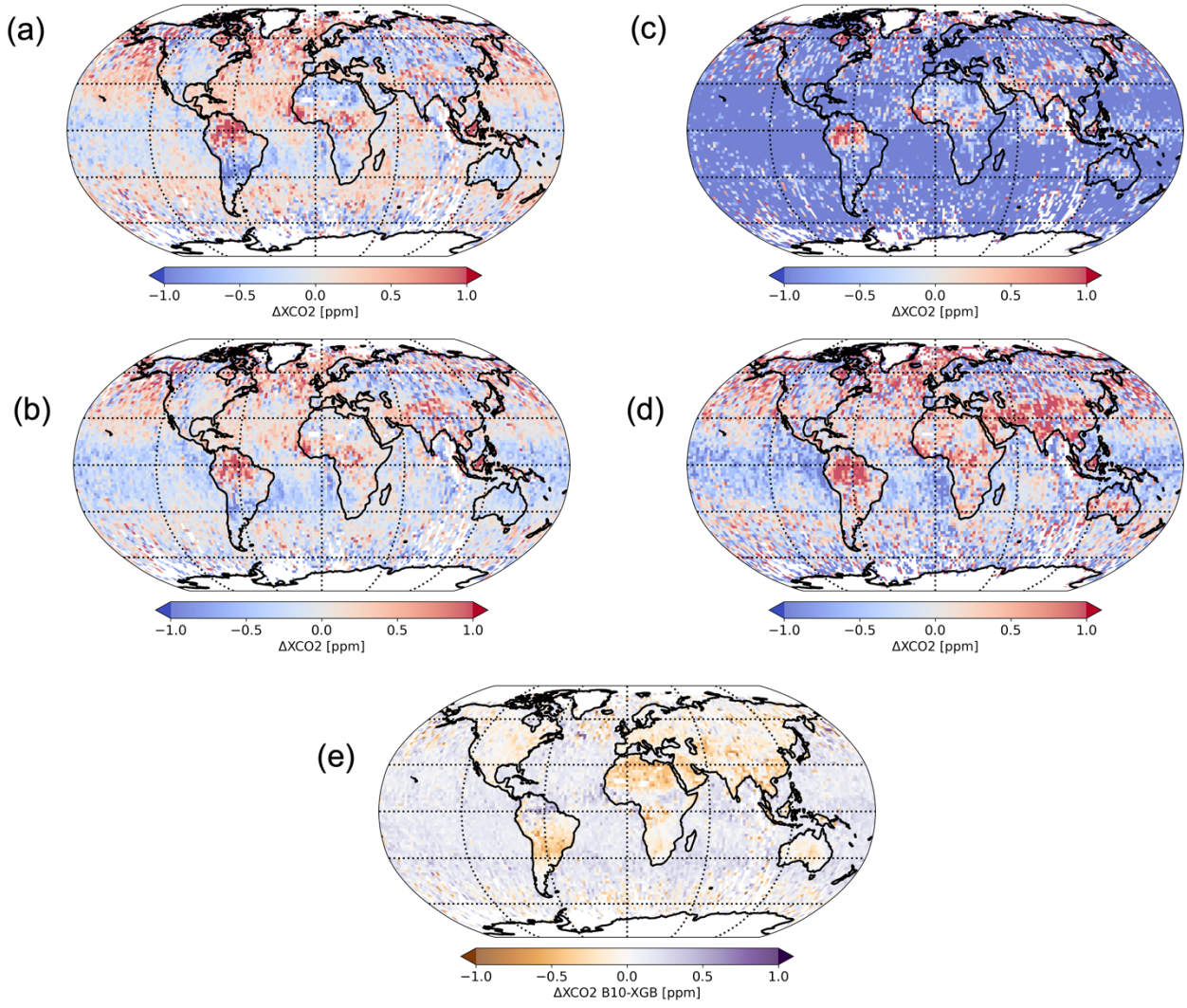


Figure 5. Remaining XCO_2 biases (ΔXCO_2) after correction for 2016-2018, binned to a 2°x2° resolution. ΔXCO_2 after the B10 correction for QF=0 is shown in (a), ΔXCO_2 after the XGBoost correction for QF=0 is shown in (b), ΔXCO_2 after the B10 correction for QF=1 is shown in (c), ΔXCO_2 after the XGBoost correction for QF=1 is shown in (d), and difference (B10 – XGB) for QF=0 is shown in (e).

4.5 Increased sounding throughput

One of the benefits of the non-linear bias correction is the potential for increased throughput of well-corrected QF = 1 data. Improved throughput of well corrected data would be of benefit to point analysis studies where data is limited by the operational QF, and potentially of benefit to flux models as well. To provide an empirical example of this, we create a modified

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

version of the operational XCO_2 quality flag combined with the gradient booster trained on QF = 0 + 1 data. We take a conservative approach where initial filter values are set equal to those of the operational quality filtering. Then, we select a few variables for which the filters are relaxed to increase sounding throughput while maintaining the RMSE of the combined operational correction and quality filter. With our new quality flag (QFNew), we are able to increase sounding throughput by approximately **16%** over the B10 QF while matching the RMSE of the B10 correction as shown in Table 5.

For many features, the quality filters were not changed from the operational filters, as relaxing filters on variables that are already passing most of their conditional distributions would allow for only marginal improvements in throughput at the cost of large systematic errors. Therefore, we select only features for which large portions of the marginal distributions are removed by the operational flag and where the non-linear correction improves both mean and variance of ΔXCO_2. The relaxed filters for these variables are shown in Figure 8 and 9 by the vertical red dashed lines. The light green and blue curves (land and ocean) show the improved mean ΔXCO_2 after the XGBoost correction and the black curve shows the ΔXCO_2 prior to any bias correction. The operational filter also minimizes the unit-less metric of the binned standard deviation of ΔXCO_2 divided by the posterior XCO_2 uncertainty below a value of 3 ppm/ppm (Osterman et al. 2020). When tuning QFNew, we also aim to minimize this metric. Higher throughput of well-corrected data is observed in northern and central Africa, the Amazon basin, and in latitiudes above 60º north as seen in Figure 6. While selection of these variables and the relaxation of their filter values is subjective, this empirical result illustrates the benefit of quality flag derived in conjunction with the non-linear bias correction. Future work will focus on the automation of defining the the quality flag thesholds using a data driven approach.

Table 7. Improvements in sounding through put and RMSE for XGBoost correction and QFNew = 0.

Region/Truth Proxy	XGBoost RMSE	B10 % Passing	QFNew % Passing
LandNG	1.03 ppm	60%	70%
OceanG	0.68 ppm	63%	76%

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

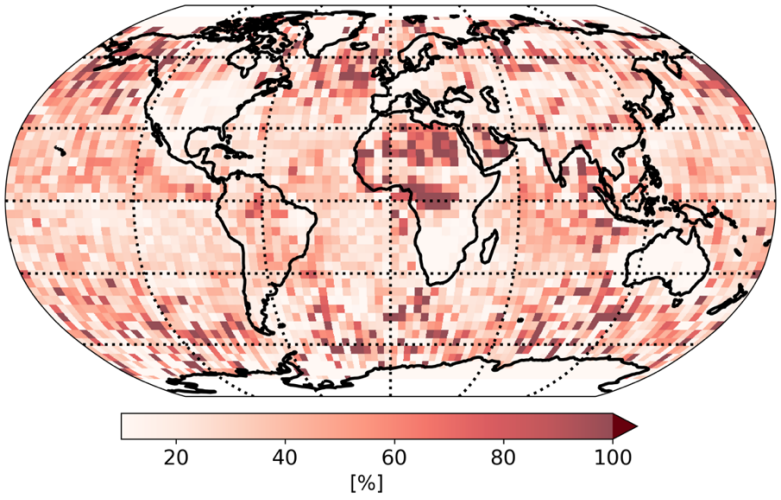


Figure 6. Relative increase in percent passing QFNew over B10 QF for 2016-2018 aggregated by 4º×4º bins.

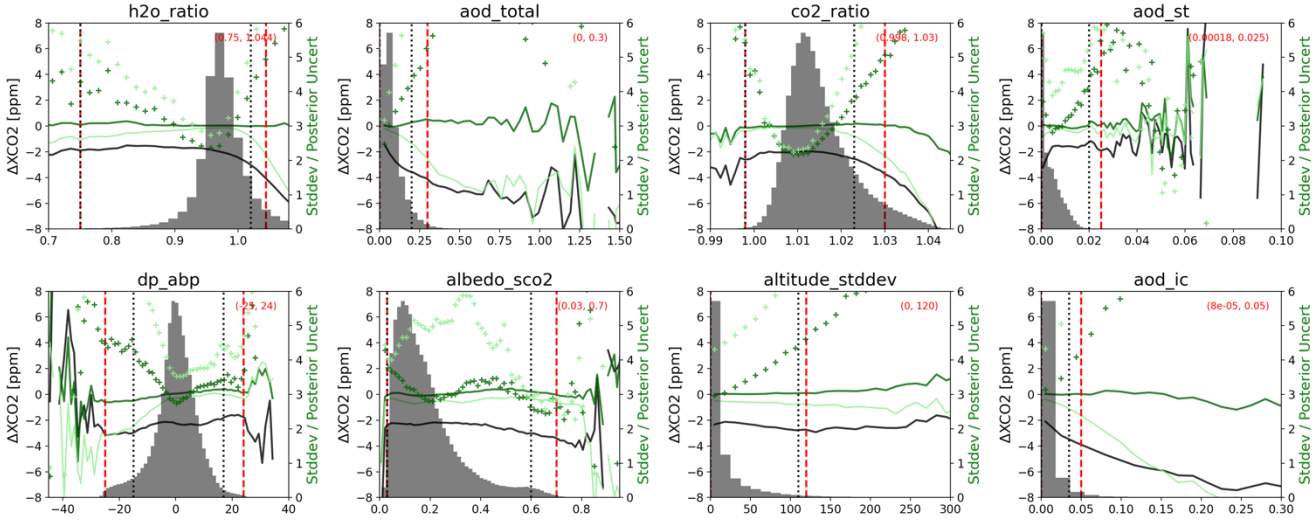


Figure 7: Variables selected for land QFNew: the difference between the uncorrected retrieval and the model mean truth proxy is shown with the black curve. The difference after the non-linear correction is show by the light green curve. The binned Std error divided by the posterior uncertainty of XCO2 is show by the green pluses and right y-axis. B10 QF filters are indicated by the black vertical dashed lines and QFNew is shown by the red dashed lines.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

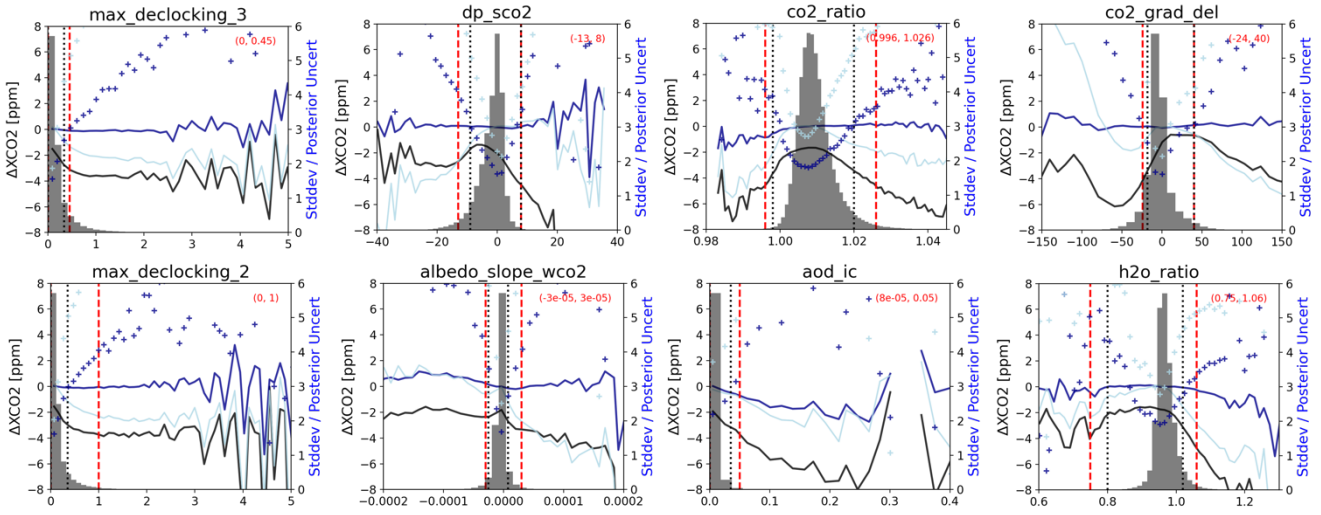


Figure 8: Variables selected for ocean QFNew: the difference between the raw retrieval uncorrected retrieval and the model mean truth proxy is shown with the black curve. The difference after the non-linear correction is show by the light blue curve. The binned Std error divided by the posterior uncertainty of XCO2 is show by the blue diamonds and right y-axis. B10 filters are indicated by the black vertical dashed lines and a potential filter is shown by the red dashed lines.

5 Discussion and Future Work

5.1 Evaluating feature importance between filter regimes

To further evaluate feature importance between QF = 0 and QF = 1, we train a set of XGBoost models individually one each truth proxy set and filtering regime. The differences in feature importance between QF = 0 and Qf = 1, are shown in Figure 10. Feaures with positive differences are more informative for correcting QF = 0 data, while features with negative differences are more informative for QF = 1. The two features with positive difference (co2_grad_del for Land NG and dp_sco2 for Ocean G) exhibit high correlation with ΔXCO_2 (O'Dell et al. 2018) for QF filtered data and are used in the operational bias correction. These features are shown to be more informative for correction of QF = 0 data. Other variables used by the operational correction become equally or more informative for correcting ΔXCO_2 in the regime of QF = 1 data.

For Land NG, there is a large negative difference in importance for h2o_ratio indicating the feature is highly informative for correcting QF = 1 data. To explain the high importance for the h2o_ratio for QF = 1 correction, we look to the non-linear interaction outside of the bound imposed by the operational filter which filters a large portion of the conditional distribution (see Figure 4). The operational filter removes soundings with a h2o_ratio of less than 1.023 reducing the regime of interaction to one that is not highly correlated with systematic error in XCO_2. When modelling bias outside of the filter bound h2o_ratio

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

increases in importance and is informative to correct for bias. These larger values of h2o_ratio observed outside of the filter bound are explained in Taylor et al. (2016) where it was shown that surface albedo for the strong CO_2 band is generally lower than the weak CO_2 band. When larger quantities of aerosols are present, there is a greater effect of multiple scattering between the surface and atmosphere causing a decrease in depth in the absorption features for CO_2 and H_2O. The bias correction model also favours the additional aerosols and dpfrac parameter for correction of QF = 1 soundings. There exists some disagreement in overall importance among truth proxies, noteably dpfrac, h2o_ratio, and aod_ice are more informative for the TCCON proxy in QF = 1 data.

For Ocean G, there are significant increases in feature importance for dp_sco2 on QF = 0 data and albedo_slope_sco2 on QF = 1 data. These features account for ~7.5 ppm and ~8.0 ppm in absolute variation of ΔXCO_2 over the majority of their domains. For dp_sco2, there is difference in agreement on information gain between the truth proxies, where the feature is more informative for the TCCON proxy for QF = 0 correction but likely equally important for both regimes in the small area proxy. The features snr_wco2 and co2_grad_del are preferential for correcting across filtering regimes, and have strong agreement among proxy sets.

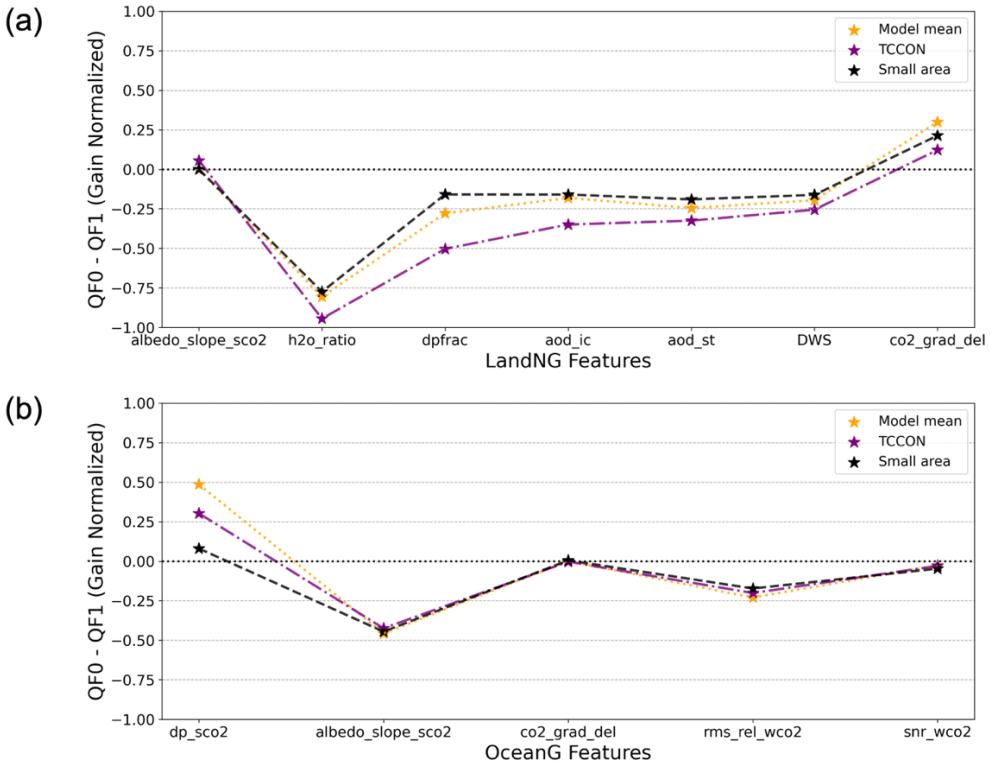
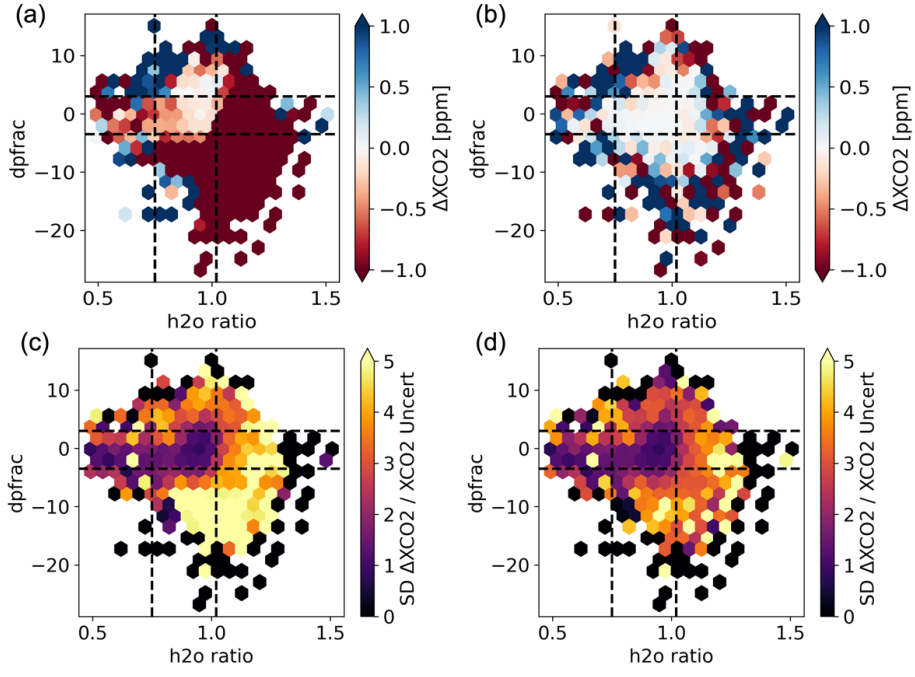


Figure 9: Difference in feature importance between QF=0 and QF=1. Differences shown by truth proxy for Land (a) and Ocean (b). Y-axis shows change in normalized information gain.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

5.2 Potential for further improving data throughput

Figure 11 further illustrates how the shape of the filtering or decision surface can effect data throughput. Soundings are binned by two state vector features: h2o_ratio and dpfrac. Figure 11b, and Figure 11d show the improvement in reduction of mean ΔXCO_2 and in the error divided by the posterior uncertainty, from the non-linear correction. The QF filters for each feature are indicated by the black dashed lines and the interior of the intersection of these filters inidicates the region of state space that is labeled as QF = 0 (Note: the additional filters of the QF further reduce the data that is passed in this region). Significant portions of the distribution, where the non-linear method is able to accurately correct, lay outside of this filtered region and are labeled QF = 1. A data driven filter can be constructed using similar interpretable machine learning techniques and produce a unified correction/filtering product. Furthermore, moving away from the binary quality flag to a ternary ("very good", "good", "bad") will likely provide an improved data product for end users. Data driven methods for quality filtering have already proven to be useful in the northern high latitudes (Mendonca et al. 2021) and a genetic algorithm was previously used to derive the Warn Levels which complement the operational quality flag found in early OCO-2 data versions (Mandrake et al. 2015). An important task for such future work will be to ensure that the machine learning method learns a physically consistent filter that can increase data throughput while still limiting variance of error and ΔXCO_2.



https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Figure 10. Hex bin plots show conditional distributions of ΔXCO_2 vs. dp_frac and h2o_ratio. Remaining ΔXCO_2 after the operational correction for B10 is shown in (a). Remaining ΔXCO_2 after the non-linear correction is shown in (b). Binned sttdev of ΔXCO_2 divided by the posteriori uncertainty from the retrieved X_{CO2} is shown in (c) for the operational correction for B10 and (d) for the non-linear correction. B10 QF filter thresholds for both features are shown with black dashed lines for reference.

6 Conclusion

We demonstrate a reproducible data driven approach for selecting co-retrieved state vector variables and other features that can be used to predict biases in the ACOS estimate of XCO_2. The use of the non-linear method allows for decoupling of the dependent bias correction and filter used in operation, as the filter no longer needs to limit the correction function to a linear fit. By doing so, this method achieves a 66% and 72% improvement in reduction of the error variance over the operational correction, for land and ocean data respectively; that is removed by the quality filter. To utilize this improvement in correction, we derive a new quality flag (QFNew), by relaxing select filter thresholds from the operational quality flag. Using the proposed QFNew flag, we are able to increase data throughput by 16% while maintaining a comparable residual error to the operational B10 correction.

The data driven approach to bias correction helps to reduce the amount of hand tuning of regression weights and iterative filter selection by a domain expert. The workflow outlined in this research is easily reproducable for future ACOS algorithm updates, and for OCO-2's companion instrument, OCO-3, aboard the ISS. The method of correction may also be adapted to other variables of interest such as surface pressure, and other trace gases.

Author contribution

WK conducted the experiments, and formulated the manuscript and figures. CO prepared and provided the truth proxy data sets used. SM, SC, and CO provided significant conceptual input for experiment design and analysis of results. All authors provided thorough review and comment on the final paper.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgment

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

The authors would like to thank the institutions that provide data from the TCCON instruments, and to the OCO-2 algorithm and science teams at CSU/CIRA and JPL. CarbonTracker results were provided by NOAA ESRL, Boulder, Colorado, USA, from the website at http:// carbontracker.noaa.gov (last access: 10 January 2022).

Financial support

WK and SC were supported by the NASA GeoCarb Mission (80LARC17C0001). SM was supported by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Adminitration (80NM0018D0004). CO was supported by a subcontract with the Jet Propulsion Laboratory.

References

Blumenstock, T., Hase, F., Schneider, M., Garcia, O. E., and Sepulveda, E.: TCCON data from Izaña (ES), Release GGG2014R1, TCCON data archive, hosted by CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.IZANA01.R1, 2017.

Breiman, L. (1984). Classification and Regression Trees (1st ed.). Routledge. https://doi.org/10.1201/9781315139470

CAMS: Copernicus Atmosphere Monitoring Service, available at: https://atmosphere.copernicus.eu/sites/default/files/2019-08/CAMS73_2018SC1_D73.1.4.1-2018-v1_201907_v1.pdf (last access: 10 January 2022), 2021.

CarbonTracker: CarbonTracker, available at: https://carbontracker.noaa.gov (last access: 10 January 2022), 2021.

CarboScope: CarboScope, available at: https://www.bgc-jena.mpg.de/CarboScope (last access: 10 January 2022), 2021.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining – KDD"16, 785–794, ACM Press, New York, New York, USA, 2016.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

from a global 21 year reanalysis of atmospheric measurements, J. Geophys. Res., 115, D21307, https://doi.org/10.1029/2010JD013887, 2010.

Connor, B. J., Bösch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: Inverse method and prospective error analysis, J. Geophys. Res., 113, A05305, https://doi.org/10.1029/2006JD008336, 2008.

Cox, C. and Munk, W. H.: The measurement of the roughness of the sea surface from photographs of the sun's glitter, J. Opt. Soc. Am., 44, 838–850, 1954.

Chevallier, F., Viovy, N., Reichstein, M., and Ciais, P.: On the assignment of prior errors in Bayesian inversions of CO_2 surface fluxes, Geophys. Res. Lett., 33, 1–5, https://doi.org/10.1029/2006GL026496, 2006.

Chevallier, F., Broquet, G., Pierangelo, C., and Crisp, D.: Probabilistic global maps of the CO2 column at daily and monthly scales from sparse satellite measurements, J. Geophys. Res., 122, 7614–7629, 2017.

Crisp, D., Atlas, R. M., Breon, F. M., Brown, L. R., Burrows, J. P., Ciais, P., . . . Schroll, S. The Orbiting Carbon Observatory (OCO) mission. Advances in Space Research, 34(4), 700-709. doi: https://doi.org/10.1016/j.asr.2003.08.062 2004.

Crisp, D., Fisher, B. M., O'Dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Mandrake, L., McDuffie, J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O'Brien, D. M., Oyafuso, F., Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T. E., Thompson, D. R., Wennberg, P. O., Wunch, D., and Yung, Y. L.: The ACOS CO_2 retrieval algorithm – Part II: Global XCO2 data characterization, Atmos. Meas. Tech., 5, 687–707, https://doi.org/10.5194/amt-5-687-2012, 2012.

Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A. M., Oyafuso, F. A., Frankenberg, C., O'Dell, C. W., Bruegge, C. J., Doran, G. B., Eldering, A., Fisher, B. M., Fu, D., Gunson, M. R., Mandrake, L., Osterman, G. B., Schwandner, F. M., Sun, K., Taylor, T. E., Wennberg, P. O., and Wunch, D.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, Atmos. Meas. Tech., 10, 59–81, https://doi.org/10.5194/amt-10-59-2017, 2017.

Crowell, S., Baker, D., Schuh, A., Basu, S., Eldering, A., Feng, L., Crisp, D., O'Dell, C. W., Oda, T., Sweeney, C., Palmer, P. I., and Jones, D. B. A.: The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network, Atmos. Chem. Phys., 19, 9797–9831, https://doi.org/10.5194/acp-19-9797-2019, 2019.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., Metzger, J.-M., Duflot, V., and Cammas, J.-P.: TCCONdata from Réunion Island (RE), Release GGG2014.R1, TCCON data archive, hosted by CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.REUNION01.R1, 2017.

Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., and Grupe, P.: TCCON data from Bialystok (PL), Release GGG2014.R2, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.BIALYSTOK01.R2, 2019.

Dubey, M.K., Lindenmaier, R., Henderson, B. G., Green, D., Allen, N. T., Roehl, C. M., Blavier, J.-F., Butterfield, Z. T., Love, S., Hamelmann, J. D., and Wunch, D.: TCCON data from Four Corners (US), Release GGG2014.R0,TCCONdataarchive,CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.FOURCORNERS01.R0/1149272, 2014.

Frankenberg, C., Platt, U., and Wagner, T.: Iterative maximum a posteriori (IMAP)-DOAS for retrieval of strongly absorbing trace gases: Model studies for CH4 and CO2 retrieval from near infrared spectra of SCIAMACHY onboard ENVISAT, Atmos. Chem. Phys., 5, 9–22, https://doi.org/10.5194/acp-5-92005, 2005.

Friedman, J. H.: Greedy function approximation: a gradient boosting machine, Ann. Stat., 29, 1189–1232, https://doi.org/10.1214/aos/1013203451, 2001.

Goo, T.-Y., Oh, Y.-S., and Velazco, V. A.: TCCON data from Anmeyondo (KR), Release GGG2014.R0, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/ TCCON.GGG2014.ANMEYONDO01.R0/1149284, 2014.

Griffith, D. W., Deutscher, N. M., Velazco, V. A., Wennberg, P. O., Yavin, Y., Aleks, G. K., Washenfelder, R. a., Toon, G. C., Blavier, J.-F., Murphy, C., Jones, N., Kettlewell, G., Connor, B. J., Macatangay, R., Roehl, C., Ryczek, M., Glowacki, J., Culgan, T., and Bryant, G.: TCCONdatafromDarwin(AU),Release GGG2014R0, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/tccon.ggg2014.darwin01.R0/1149290, 2014a.

Griffith, D. W., Velazco, V. A., Deutscher, N. M., Murphy, C., Jones, N., Wilson, S., Macatangay, R., Kettlewell, G., Buchholz, R. R., and Riggenbach, M.: TCCON data from Wollongong (AU), Release GGG2014R0, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/tccon.ggg2014.wollongong01.R0/ 1149291, 2014b.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Hase, F., Blumenstock, T., Dohe, S., Gross, J., and Kiel, M.: TCCON data from Karlsruhe (DE), Release GGG2014R1, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/tccon.ggg2014.karlsruhe01.R1/1182416, 2015.

Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Taura, C., Fukamachi, T., and Sakashita, M.: TCCON data from Saga (JP), Release GGG2014R0, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/tccon.ggg2014.saga01.R0/1149283, 2014.

Kivi, data R., Heikkinen, P., and Kyrö, E.: TCCON from Sodankylä (FI), Release GGG2014.R0, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/tccon.ggg2014.sodankyla01.R0/1149280, 2014.

Kulawik, S.S.; O'Dell, C.; Nelson, R.R.; Taylor, T.E. Validation of OCO-2 error analysis using simulated retrievals. *Atmos. Meas. Tech. 12*, 5317–5334 doi: https://doi.org/10.5194/amt-12-5317-2019 2019

Liu, J., Bowman, K. W., Schimel, D. S., Parazoo, N. C., Jiang, Z., Lee, M., Bloom, A. A., Wunch, D., Frankenberg, C., Sun, Y., O'Dell, C. W., Gurney, K. R., Menemenlis, D., Gierach, M., Crisp, D., and Eldering, A.: Contrasting carbon cycle responses of the tropical continents to the 2015–2016 El Niño, Science, 358, eaam5690, https://doi.org/10.1126/science.aam5690, 2017.

Mandrake, L., Frankenberg, C., O'Dell, C. W., Osterman, G., Wennberg, P., and Wunch, D.: Semi-autonomous sounding selection for OCO-2, Atmos. Meas. Tech., 6, 2851–2864, https://doi.org/10.5194/amt-6-2851-2013, 2013.

Mandrake, L., O'Dell, C. W., Wunch, D., Wennberg, P. O., Fisher, B., Osterman, G. B., and Eldering, A.: Orbiting Carbon Observatory-2 (OCO-2) Warn Level, Bias Correction, and Lite File Product Description, Tech. rep., Tech. rep., Jet Propulsion Laboratory, California Institute of Technology, Pasasdena, available at: https://disc.sci.gsfc.nasa.gov/OCO-2/documentation/oco-2-v7/OCO2_XCO2_Lite_Files_and_Bias_Correction_0915_sm.pdf (last access: 16 October 2015), Pasadena, CA, 2015.

Morino, I., Matsuzaki, T., and Horikawa, M.: TCCON data from Tsukuba (JP), 125HR, Release GGG2014R2, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.TSUKUBA02.R2, 2018a.

Morino, I., Velazco, V. A., Hori, A., Uchino, O., and Griffith, D. W. T.: TCCON data from Burgos, Ilocos Norte (PH), Release GGG2014.R0, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.BURGOS01.R0, 2018b.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

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.

Massie, S. T., Schmidt, S. K., Eldering, A., and Crisp, D.: Observational evidence of 3-D cloud effects in OCO-2 CO_2 retrievals, J. Geophys. Res., 122, 7064–7085, https://doi.org/10.1002/2016JD026111, 2016.

Mauceri, S., Massie, S., Schmidt, S.: Correcting 3D cloud effects in XCO2 retrievals from OCO-2, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-202 , in-review, 2022

Moore, B., Crowell, S. M. R., Rayner, P. J., Kumer, J., O'Dell, C. W., O'Brien, D., Utembe, S., Polonsky, I., Schimel, D., and Lemen, J.: The Potential of the Geostationary Carbon Cycle Ob- servatory (GeoCarb) to Provide Multi-scale Constraints on the Carbon Cycle in the Americas, Front. Environ. Sci., 6, https://doi.org/10.3389/fenvs.2018.00109, 2018.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D.: Quantifying CO_2 Emissions From Individual Power Plants From Space, Geophys. Res. Lett., 44, 10045–10053, https://doi.org/10.1002/2017GL074702, 2017.

Noël, S., Reuter, M., Buchwitz, M., Borchardt, J., Hilker, M., Schneising, O., Bovensmann, H., Burrows, J. P., Di Noia, A., Parker, R. J., Suto, H., Yoshida, Y., Buschmann, M., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Liu, C., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Petri, C., Pollard, D. F., Rettinger, M., Roehl, C., Rousogenous, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Vrekoussis, M., and Warneke, T.: Retrieval of greenhouse gases from GOSAT and GOSAT-2 using the FOCAL algorithm, Atmos. Meas. Tech., 15, 3401–3437, https://doi.org/10.5194/amt-15-3401-2022, 2022.

Notholt, J., Petri, C., Warneke, T., Deutscher, N. M., Palm, M., Buschmann, M., Weinzierl, C., Macatangay, R. C., and Grupe, P.: TCCON data from Bremen (DE), Release GGG2014.R1, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.BREMEN01.R1, 2019.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO_2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99–121, https://doi.org/10.5194/amt-5-99-2012, 2012.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, P. Natl. Acad. Sci. USA, 104, 18925–18930, https://doi.org/10.1073/pnas.0708986104, 2007.

Pollard, D. F., Robinson, J., and Shiona, H.: TCCON data from Lauder (NZ), Release GGG2014.R0, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.LAUDER03.R0, 2019.

Rödenbeck, C.: Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Tech. rep., Max Planck Institute for Biogeochemistry, Jena, Germany, 2005.

Rödenbeck, C., Zaehle, S., Keeling, R., and Heimann, M.: How does the terrestrial carbon exchange respond to interannual climatic variations? A quantification based on atmospheric CO2 data, Biogeosciences, 15, 2481–2498, https://doi.org/10.5194/bg-15-2481-2018, 2018.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific, Singapore, 2000.

Schneising O, Buchwitz M, Reuter M, Bovensmann H, Burrows JP, Borsdorff T, Deutscher NM, Feist DG, Griffith DW, Hase F, Hermans C.: A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor. Atmospheric Measurement Techniques. https://doi.org/10.5194/amt-12-6771-2019 2019.

Sussmann, from R. Garmisch and Rettinger, (DE), Release M.: TCCON data GGG2014R2, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.GARMISCH01.R2, 2018.

Taylor, T. E., O'Dell, C. W., Frankenberg, C., Partain, P. T., Cronk, H. Q., Savtchenko, A., Nelson, R. R., Rosenthal, E. J., Chang, A. Y., Fisher, B., Osterman, G. B., Pollock, R. H., Crisp, D., Eldering, A., and Gunson, M. R.: Orbiting Carbon

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

Observatory-2 (OCO-2) cloud screening algorithms: validation against collocated MODIS and CALIOP data, Atmos. Meas. Tech., 9, 973–989, https://doi.org/10.5194/amt-9-973-2016, 2016.

Taylor, T.E., Eldering, A., Merrelli, A., Kiel, M., Somkuti, P., Cheng, C., Rosenberg, R., Fisher, B., Crisp, D., Basilio, R. and Bennett, M.: OCO-3 early mission operations and initial (vEarly) XCO2 and SIF retrievals. *Remote Sensing of Environment*, *251*, p.112032, https://doi.org/10.1016/j.rse.2020.112032, 2020.

Taylor, T.E., O'Dell, C.W., Baker, D., Bruegge, C., Chang, A., Chapsky, L., Chatterjee, A., Cheng, C., Chevallier, F., Crisp, D., Dang, L., Drouin, B., Eldering, A., Feng, L., Fisher, B., Fu, D., Gunson, M., Hammerle, V., Keller, G.R., Kiel, M., Kurosu, T., Lambert, A., Laughner, J., Lee, R., Liu, J., Mandrake, L., Marchetti, Y., McGarragh, G., Merrelli, A., Nelson, R.R., Osterman, G., Oyafuso, F., Palmer, P.I., Payne, V.H., Rosenberg, R., Somkuti, P., Spiers, G., To, C., Wennberg, P.O., Yu, S., and Zong, J.: Evaluating the consistency between OCO-2 and OCO-3 XCO2 estimates derived from the NASA ACOS version 10 retrieval algorithm, *Atmos. Meas. Tech. Discuss*, in review, 2023.

Te, Y., Jeseck, P., and Janssen, C.: TCCON data from (FR), Release GGG2014R0, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/tccon.ggg2014.paris01.R0/1149279, 2014.

Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N. M., Petri, C., and Grupe, P.: TCCON data from Orléans (FR), Release GGG2014R1, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.ORLEANS01.R1, 2019.

Wennberg, P. O., Roehl, C. M., Blavier, J.-F., Wunch, D., and Allen, N. T.: TCCON data from Jet Propulsion Laboratory (US), 2011, Release GGG2014.R1, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.JPL02.R1/1330096, 2016a.

Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., Allen, N., Dowell, P., Teske, K., Martin, C., and Martin., J.: TCCON data from Lamont (US), Release GGG2014R1, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/tccon.ggg2014.lamont01.R1/1255070, 2016b.

Wennberg, P. O., Roehl, C., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R. a., Keppel-Aleks, G., Allen, N., and Ayers, J.: TCCON data from Park Falls (US), Release GGG2014R1, TCCON data archive, CaltechDATA [data set], https://doi.org/10.14291/TCCON.GGG2014.PARKFALLS01.R1, 2017.

https://doi.org/10.5194/egusphere-2023-362
Preprint. Discussion started: 12 April 2023

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, Atmos. Meas. Tech., 3, 1351–1362, https://doi.org/10.5194/amt-3-1351-2010, 2010.

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of CO_2 total columns from space, Atmos. Chem. Phys., 11, 12317–12337, https://doi.org/10.5194/acp-11-12317-2011, 2011.

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Notholt, J., Warneke, T., Petri, C., De Maziere, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO_2 measurements with TCCON, Atmos. Meas. Tech., 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017.