Article

Validation of Carbon Monoxide Total Column Retrievals from SCIAMACHY Observations with NDACC/TCCON Ground-Based Measurements

Philipp Hochstaffl [1],* , Franz Schreier [1] , Günter Lichtenberg [1] and Sebastian Gimeno García [1,2]

[1] DLR—German Aerospace Center, Remote Sensing Technology Institute, 82234 Oberpfaffenhofen, Germany; franz.schreier@dlr.de (F.S.); guenter.lichtenberg@dlr.de (G.L.); sebastian.gimenogarcia@eumetsat.int (S.G.G.)

[2] EUMETSAT—European Organisation for the Exploitation of Meteorological Satellites, 64283 Darmstadt, Germany

* Correspondence: philipp.hochstaffl@dlr.de; Tel.: +49-8153-28-3056

Abstract: The objective was to validate the carbon monoxide (CO) total column product inferred from Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) full-mission (2003–2011) short-wave infrared (SWIR) nadir observations using the Beer InfraRed Retrieval Algorithm (BIRRA). Globally distributed Network for the Detection of Atmospheric Composition Change (NDACC) and Total Carbon Column Observing Network (TCCON) ground-based (g-b) measurements were used as a true reference. Weighted averages of SCIAMACHY CO observations within a circle around the g-b observing system were utilized to minimize effects due to spatial mismatch of space-based (s-b) and g-b observations, i.e., disagreements due to representation errors rather than instrument and/or algorithm deficiencies. In addition, temporal weighted averages were examined and then the unweighted (classical) approach was compared to the weighted (non-classical) method. The delivered distance-based filtered SCIAMACHY data were in better agreement with respect to CO averages as compared to square-shaped sampling areas throughout the year. Errors in individual SCIAMACHY retrievals have increased substantially since 2005. The global bias was determined to be in the order of -10 parts per billion in volume (ppbv) depending on the reference network and validation strategy used. The largest negative bias was found to occur in the northern mid-latitudes in Europe and North America, and was partly caused by insufficient a priori estimates of CO and cloud shielding. Furthermore, no significant trend was identified in the global bias throughout the mission. The global analysis of the CO columns retrieved by the BIRRA shows results that are largely consistent with similar investigations in previous works.

Keywords: SCIAMACHY; BIRRA; NDACC; TCCON; validation; retrieval; carbon monoxide; mixing ratios; weighted averages

1. Introduction

Obtaining space-based (s-b) measurements of the state of Earth's atmosphere is costly compared to many other atmospheric research activities. Once the instruments needed to take these measurements are in orbit, thorough verification and validation of the delivered data is required.

Therefore, considerable efforts have been invested in establishing a ground-based (g-b) validation infrastructure for atmospheric composition, temperature, cloud, and aerosol data acquired by satellite-based remote sensing instruments. Additionally, the stringent requirements of upcoming missions such as Sentinel-5P/4/5 require comprehensive validation campaigns and solid strategies.

1.1. The Environmental Satellite (ENVISAT) and Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) Instrument

In March 2002 the ENVISAT was launched into a low-Earth orbit. The routine operation comprising the nominal measurement program began in August 2002 and lasted until April 2012 when contact was lost [1]. The satellite comprised 10 instruments with the SCIAMACHY, the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), and Global Ozone Monitoring by Occultation of Stars (GOMOS) dedicated to studying the Earth's atmosphere. The SCIAMACHY was an ultra-violet (UV) to short-wave infrared (SWIR, 1.4–3 μm) absorption spectrometer that observed the scattered and reflected solar spectral radiance transmitted through the atmosphere [2].

The SCIAMACHY's major research objective was the acquisition of information from Earth radiance spectra of various trace gases in the troposphere and the stratosphere. Among these gases were O_3, CO, CH_4, H_2O, SO_2, CO_2, and NO_2. Thus, the SCIAMACHY comprised eight detectors in the wavelength range between 214 and 2386 nm. Channel 8 observed radiance in the SWIR spectral range from 2259.38 nm to 2386.07 nm (4426–4191 cm^{-1}) at a resolution of 0.26 nm, which is equivalent to a resolving power of 8689 to 9177.

## 1.2. Carbon Monoxide (CO) from Channel 8 of SCIAMACHY

In atmospheric remote sensing, trace gas concentrations are retrieved from measured radiance or transmission spectra. The accurate retrieval of CO from Channel 8 observations of the SCIAMACHY is demanding due to the low optical depth of CO compared to the total depth in this spectral region. To be more specific, for a vertical path through the atmosphere, only about one percent of absorption is due to CO molecules. In addition, the precision of the cloud-free SCIAMACHY measurements is strongly influenced by the reflectance (albedo) of the observed ground-pixel, because it determines to a large extent the signal-to-noise ratio (SNR) of the corresponding observed spectra. The retrieval of CO can therefore only deliver total column amounts without any information on the vertical distribution. In addition, the Channel 8 detector shows temporal degradation due to an accumulation of ice, which significantly affects the throughput. This makes gas retrieval an even more challenging task [3]. Moreover, pixel degradation due to, e.g., solar radiation, is an ongoing process. According to Lichtenberg et al. [4] Channel 8 contains a substantial number of dead and bad pixels, with 40% deemed unusable by June 2009. For the CO fitting window from 4280 to 4305 cm^{-1} the number of usable pixels was reduced to only about 50 in the period between 2003 and 2005, and to even fewer afterwards.

1.3. Retrieval Codes

The retrieval of trace gas concentrations from radiance or transmission spectra poses an inverse problem that is typically solved by least-squares algorithms. Several codes have been developed for SCIAMACHY nadir SWIR spectra at different European institutes, e.g., the Weighted Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS) algorithm [5,6], the Iterative Maximum A Posteriori (IMAP)-DOAS [7] method, the Iterative Maximum Likelihood Method (IMLM) [3], and the Beer InfraRed Retrieval Algorithm (BIRRA) [8]. Recently, SCIAMACHY spectra have also been processed with the Shortwave Infrared CO Retrieval (SICOR) algorithm developed for the operational data processing of the Tropospheric Monitoring Instrument (TROPOMI) that was launched on the European Space Agency's (ESA) Sentinel-5 Precursor (S5P) mission [9,10]. Concurrent validation is a crucial task in processor/algorithm development in order to confirm applicability for scientific tasks and ultimately qualify operational algorithms or scientific tools. Based on such studies, processors can be applied on new data sets from other platforms with reasonable confidence. However, new fields of application require further validation or verification efforts.

1.4. The Network for the Detection of Atmospheric Composition Change (NDACC) and Total Carbon Column Observing Network (TCCON) Ground Truthing Networks

Fourier transform spectroscopy (FTS) is a well-established technique in passive remote sensing and observes thermal emissions (e.g., using the Infrared Atmospheric Sounding Interferometer (IASI) [11] or MIPAS [12]) or absorption (e.g., using the Atmospheric Chemistry Experiment—Fourier Transform Spectrometer (ACE-FTS) [13] or the Greenhouse Gases Observing Satellite (GOSAT) [14], which also observes in the thermal infrared range (TIR range, 8–15 μm)). High-resolution FTS is also used by numerous g-b observatories of the NDACC and TCCON [15]. The instruments associated with these networks routinely record thermal emission spectra in the mid-infrared range (MIR range, 3–8 μm) (NDACC) and solar absorption spectra in the SWIR range (TCCON) at a number of stations worldwide (see Figure 1, Table 1) and utilize these to infer information about atmospheric constituent columns and concentration profiles (NDACC only), including CO.
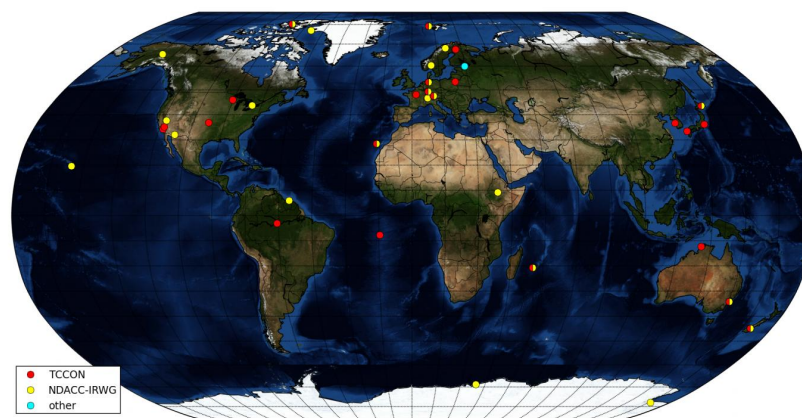
Figure 1. World map created by Feist [16] showing stations affiliated to ground-based (g-b) observing networks routinely measuring trace gases such as CO in the mid-infrared (MIR) and short-wave infrared (SWIR) ranges. The background color scheme also provides some information on the ground reflectance in the SWIR range (e.g., the variations of the reflectance on the continents and also the difference between land and the oceans).

Table 1. The Network for the Detection of Atmospheric Composition Change (NDACC) and Total Carbon Column Observing Network (TCCON) g-b Fourier transform infrared (FTIR) stations used in this validation. The last two columns indicate the time span of the g-b data used for the comparison to the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY). Further details see Appendix A and B.

Stations	Lat N [°]	Lon E [°]	Altitude [m]	NDACC (years)	TCCON (years)
Bialystock	53.23	23.03	160	-	2009–2011
Bremen	53.10	8.85	30	2003–2011	2005–2011
Darwin	−12.42	130.89	30	-	2005–2011
Garmisch	47.48	11.06	743	-	2007–2011
Izana	28.30	−16.48	2367	2003–2011	2007–2011
Jungfraujoch	46.55	7.98	3580	2003–2011	-
Kiruna	67.84	20.40	420	2003–2011	-
Lamont	36.60	−97.49	320	-	2008–2011
Lauder	−45.04	169.68	370	-	2004–2010
Ny Alesund	78.92	11.93	15	2003–2011	2005–2011
Parkfalls	45.95	−90.27	442	-	2004–2011
Thule	76.52	−68.77	220	2003–2011	-
Toronto	43.66	−79.40	174	2003–2011	-
Wollongong	−34.41	150.88	30	-	2008–2011
Zugspitze	47.42	10.98	2964	2003–2011	-

The TCCON is the reference network for the validation of greenhouse gas satellite retrievals and enables the linking of satellite retrievals to the World Meteorological Organization (WMO) reference scale [17]. In this study, the GGG2014 release of TCCON data is used. GGG stands for the whole software package. According to the TCCON website the error budget for CO is around 8.7% in total (see TCCON [18]). NDACC Fourier transform infrared (FTIR) data were obtained via the publicly available website (see http://www.ndacc.org).

1.5. Validation

In general terms, validation is defined as the process of evaluating the performance of a system against some equivalent information that is regarded as a 'true' reference [19]. In the field of remote sensing from space, compliance of observations from satellite platforms with the actual state of the atmosphere is usually assessed through validation studies. In that sense, validation means comparing measurements acquired by s-b instruments to other measurements utilizing different measuring methods, e.g., in-situ measurements or g-b soundings.

1.5.1. General Aspects

With respect to a direct comparison it stands to reason that the closer in time and space the reference measurements are acquired, the better they quantify differences due to errors in the acquisition or retrieval process, i.e., the instrument or algorithm performance. However, space-borne and reference measurements do not exactly match in time and space, nor do they address the same volumes of air. Hence, direct comparison of observations from different observing systems is affected by representation errors.

In fact, the column measured by the SCIAMACHY is an average column above the area covered by a SCIAMACHY pixel (32×120 km^2 for observations in Channel 8) which greatly extends beyond the location of the point-like g-b station. According to Verhoelst et al. [20], non-perfect co-location in space is therefore a consequence of both a difference in sampling (i.e., a satellite pixel center generally does not coincide exactly with a ground station), and a difference in the way each instrument has a smoothed perception of the real, non-homogeneous atmosphere. The air mass of sensitivity for the s-b nadir measurements contains the ground pixel footprint, an extension towards the satellite, and, especially in the SWIR range, an extension in the direction of the Sun. In the SWIR range, these

extensions correspond to a good approximation to the optical light path between the Sun, the surface reflection point, and the sensor.

Moreover, with respect to the SCIAMACHY, Borsdorff et al. [9] mention another issue, whereby for low radiance scenes, the retrieval noise error for SCIAMACHY measurements can exceed 100% of the retrieved column. That is, the comparison of individual CO columns using the SCIAMACHY to a reference is not sufficient at all. According to Gimeno García et al. [8], De Laat et al. [21], and de Laat et al. [22] averaging over an ensemble of pixels is necessary to reduce the instrument-noise error.

Finally, all of these issues need to be considered in an appropriate validation strategy of the CO product inferred from a retrieval algorithm. For mutually exclusive aspects a compromise must be made between, on the one hand, the abundance of comparison pairs, and on the other hand, non-instrumental comparison errors due to non-perfect co-location in space and time of satellite and ground-based measurements.

1.5.2. The SCIAMACHY CO Product

Early validation studies of the CO product retrieved from SCIAMACHY SWIR measurements by g-b data have been conducted by Sussmann and Buchwitz [23], Sussmann et al. [24], and Dils et al. [25]. In recent years, additional validation efforts have been carried out by Borsdorff et al. [9,10], De Laat et al. [21], and Schneising et al. [26]. Borsdorff et al. [9] additionally included Tracer Model version 5 (TM5) data for large-scale (northern and southern hemispheric Africa) intercomparison and Measurement of Ozone on Airbus In-service Aircraft (MOZAIC)/In-service Aircraft for Global Observing System (IAGOS) measurements.

In all of these studies, important aspects of the comparison of g-b and s-b data were presented. With respect to methodology, many considerations described in those articles are not only true for SCIAMACHY but are valid for all s-b to g-b comparisons. The former studies particularly include validation of CO data against measurements from g-b FTIR spectrometers from either the NDACC, the TCCON, or both. As SCIAMACHY measurements do not provide any information about the vertical distribution of CO, it is reasonable to use TCCON columns for SCIAMACHY validation. In addition, integrals of vertical CO profiles provided by the NDACC are utilized for validation of the s-b total column retrievals. Usually, if total columns calculated from vertical profiles of different remote sounding instruments are to be compared, averaging kernels have to be taken into account. Nonetheless, due to the small differences in vertical sensitivity between g-b FTIR and SCIAMACHY, and because the representation error dominates, a direct comparison (i.e., a comparison without applying total column averaging kernels) of the columns was considered possible without introducing significant intercomparison errors [23].

So far, only Borsdorff et al. [9,10] have presented full-mission data set validation of SCIAMACHY CO vertical column densities employing the SICOR algorithm. Dils et al. [25] considered SCIAMACHY data that covered the January to December 2003 time period that were produced using three different retrieval algorithms, namely the WFM-DOAS algorithm (version 0.5 for CO), the IMAP-DOAS method (version 0.9 for CO), and the IMLM (version 6.3). Sussmann and Buchwitz [23] used CO vertical profile retrievals from g-b solar FTIR measurements at Zugspitze, Germany (January–October 2003) to validate columnar CO measurements retrieved from SCIAMACHY spectra with the WFM-DOAS algorithm version 0.4.

1.6. Goals and Structure of this Study

The objective of this study is to validate the BIRRA level 2 prototype processor using the full-mission level 1b SCIAMACHY data set. Both land and ocean pixels are used in this intercomparison. In addition, the study places emphasisis on mitigating errors induced by representation deficiencies due to averaging and analyses the effect of different comparison methodologies—a crucial step in thoroughly quantifying the performance (accuracy and precision) of a retrieval algorithm or instrument.

The paper is organized in five sections. Our methodology is described in Section 2 and the results in Section 3. Finally, the discussion and conclusion are provided in Sections 4 and 5.

2. Methodology

The analysis of a solar absorption spectrum essentially reveals the slant columns of the various absorbing gases. However, the dry air mole fractions are actually advantageous for comparison. These quantities are independent of surface pressure and humidity and are therefore much more useful for satellite validation. Therefore, before comparing the observations for validation, CO needs to be harmonized across the observing systems.

2.1. Ground-Based Product Definition

The vertical column density (VCD) in [mol cm^{-2}] is the number of molecules above the surface defined as

$$N_\mu = \int_{z_{srf}}^{z_{TOA}} n_\mu(z)\,dz\,. \tag{1}$$

with n_μ the number density [mol cm^{-3}] of molecule μ. The surface pressure p is given by

$$p = \int_{z_{srf}}^{z_{TOA}} m\,n(z)\,g(z)\,dz$$

$$\approx m\,g \int_{z_{srf}}^{z_{TOA}} n(z)\,dz\,, \tag{2}$$

where m is the molar mass and g the gravitational acceleration of the Earth. In Equation (2), $m\,g$ is the column-averaged value. Within the homosphere, roughly including the troposphere, stratosphere, and mesosphere, the molar mass of air and gravity is a good approximation constant.

The column mixing ratio (CMR) f_μ of molecule μ is related to the total column according to

$$N_\mu = f_\mu\,N_{air}^{wet}\,, \tag{3}$$

with $N_{air}^{wet} = \sum N_\mu + N_{H_2O} = N_{air} + N_{H_2O}$. As indicated above, it is useful to discuss dry CMR d_μ, i.e., $\mu \neq H_2O$ rather than the true (wet) mixing ratios because dry CMR is a better tracer, not being subject to strongly varying H_2O variations. As

$$\sum_\mu f_\mu + f_{H_2O} = 1 \tag{4}$$

$$\sum_\mu d_\mu = 1 \tag{5}$$

the mole fraction and dry air mole fraction are connected according to

$$f_\mu = d_\mu\,(1 - f_{H_2O})\,. \tag{6}$$

Note that the molar masses of $m_{H_2O} = 18.02$ g mol^{-1} and $m_{air} = 28.96$ g mol^{-1} are different. $\overline{m} = m_{air}\,(1 - f_{H_2O}) + f_{H_2O}\,m_{H_2O}$ accounts for this by correcting the mass for the dry air mixing ratio with the fraction of the H_2O mixing ratio, as H_2O contributes less to the surface pressure p on a per molecule basis than air. All this together leads to Equation (3), reading as

$$N_\mu = d_\mu\,(1 - f_{H_2O})\,\frac{p}{m\,g}\,, \tag{7}$$

and therefore

$$d_\mu = \frac{N_\mu}{N_{\text{air}}^{\text{wet}} - N_{\text{H}_2\text{O}}} = \frac{N_\mu}{N_{\text{air}}} \quad . \tag{8}$$

Given that $f_{\text{H}_2\text{O}} \ll 1$, Equation (7) can be approximated according to

$$d_\mu = N_\mu \left(\frac{p}{m_{\text{air}} g} - N_{\text{H}_2\text{O}} \frac{m_{\text{H}_2\text{O}}}{m_{\text{air}}} \right)^{-1} , \tag{9}$$

which constitutes the dry air volume mixing ratio for gas μ. The dry air CMR d_μ is frequently designated as xCO and given in parts per billion in volume (ppbv).

2.1.1. NDACC

NDACC sites retrieve vertical profiles for a variety of molecules (see Table 2). Since the actual measurement obtains the slant column and not the vertical column, $\zeta = \cos(\theta)$ accounts for the local zenith angle $0 \leq \theta < \frac{\pi}{2}$. $\frac{1}{\zeta}$ is commonly known as the geometrical airmass factor. Therefore, when calculating CO dry air column mixing ratio (xCO) for the respective site, Equation (9) has to be taken into account according to

$$\text{xCO} = N_{\text{CO}} \left(\frac{1}{\zeta} \frac{p}{m_{\text{air}} g} - N_{\text{H}_2\text{O}} \frac{m_{\text{H}_2\text{O}}}{m_{\text{air}}} \right)^{-1} . \tag{10}$$

Table 2. Center wavelengths (for the NDACC) and spectral windows (for the TCCON and Beer InfraRed Retrieval Algorithm, BIRRA) used for the retrieval of CO. The NDACC relies on narrow spectral fitting regions [27], while the retrieval strategy of the TCCON prefers much wider spectral regions [28]. The BIRRA uses one spectral fitting window in Channel 8. The NDACC uses the Scale Invariant Feature Transform (SFIT) algorithm, and the TCCON applies GFIT to acquire the GGG2014 data set.

Algorithm	Wavenumber [cm^{-1}]	Interfering Species
SFIT	2057.8575	NO, COF$_2$ ...
(NDACC)	2069.6559	NO, COF$_2$ ...
	2111.5430	OCS, N$_2$ ...
	2158.2997	OCS, N$_2$ ...
GFIT	4208.70–4257.30	H$_2$O, CH$_4$
(TCCON)	4262.00–4318.80	H$_2$O, CH$_4$
BIRRA	4280.00–4305.00	H$_2$O, CH$_4$
(SCIAMACHY)		

2.1.2. The TCCON

According to Wunch et al. [15], the TCCON scales a priori profiles of retrieved molecules

$$n_\mu(z) = \alpha_\mu \, n_\mu^{\text{ref}}(z) \tag{11}$$

similarly to the BIRRA. Hence, the shape of the vertical profile is invariant with respect to variations in molecular atmospheric densities. For TCCON sites, Equation (9) is used in addition to CO with respect to the co-retrieved oxygen O_2. Assuming a dry air mole fraction of $d_{O_2} = 0.2095$, this is equivalent to the dry air column mixing ratio (details see Wunch et al. [29]). Rationing the CO and O_2 equations eliminates the denominator in Equation (9) to make d_{CO} independent of surface pressure and gravity, leading to the equation

$$\text{xCO} = 0.2095 \, \frac{N_{\text{CO}}}{N_{\text{O}_2}} . \tag{12}$$

2.2. The BIRRA

The BIRRA was developed at the Deutsches Zentrum für Luft- und Raumfahrt (DLR) and serves the operational SCIAMACHY processor for CO and CH$_4$ retrievals using Channel 8 and Channel 6 spectra, respectively. The CO total column amounts are inferred simultaneously with methane (CH$_4$) and water vapour columns and a Lambertian surface albedo from individual SCIAMACHY measurements assuming a non-scattering atmosphere. The validation study reported here is based on the BIRRA prototype version.

2.2.1. Algorithm

The BIRRA comprises the line-by-line forward model Generic Atmospheric Radiation Line-by-line InfraRed Code (GARLIC) [30,31] coupled to a least-squares [32] inversion algorithm for trace gas retrieval in the SWIR spectral region(see Gimeno García et al. [8]). In case of SWIR nadir observations, the radiative transfer equation through Earth's atmosphere with molecules μ, including H$_2$O, reduces to Beer's law, describing the radiance

$$
\begin{aligned}
I(\mathbf{x}, \nu) \quad &\equiv \quad \frac{r(\nu)}{\pi} \mu_\odot I_{\mathrm{sun}}(\nu) \times \exp\left(-\sum_\mu \alpha_\mu \, \tau_\mu^{\mathrm{ref}}(\nu) \right) \\
&\otimes \; S(\nu, \gamma, \delta) \, + \, b(\nu) \, .
\end{aligned}
\tag{13}
$$

$\mu_\odot = \cos\theta_\odot$ and I_{sun} describe the solar zenith angle and incoming radiation from the sun, respectively. The 'true' (= to be estimated by the fit) optical depth $\tau_\mu(\nu)$ is hence given by

$$
\tau_\mu(\nu) \; = \; \alpha_\mu \, \tau_\mu^{\mathrm{ref}} \, .
\tag{14}
$$

For the BIRRA, the state vector $\mathbf{x} = (\boldsymbol{\eta}, \boldsymbol{\beta})$ includes nonlinear ($\boldsymbol{\eta}$) and linear ($\boldsymbol{\beta}$) parameters to be estimated. This separation of parameters enables the algorithm to utilize a separable least-squares fit—also known as variable projection [33]—in order to estimate the unknown quantities. The molecular scaling factor(s) α_μ, the half width γ and the wavelength shift δ of the instrumental slit function S are estimated using a nonlinear least-squares fit; furthermore, the coefficients of the surface albedo r, modeled by a second-order polynomial depending on wavenumber, and the optional baseline correction b (again a polynomial but not used in this study) show up linearly within the model and are estimated using a linear least-squares fit. The model for the least-squares problem is therefore

$$
y_i \; = \; \sum_{j=1}^{n} \phi_{i,j}(\boldsymbol{\eta}) \, \beta_j \, , \qquad i = 1, 2, \ldots, m
\tag{15}
$$

with n representing the numbers of linear parameters. m is the number of observations (spectral pixels) in the microwindow chosen for the retrieval. Notice that in the separable least-squares approach it is assumed that the model functions $\phi_{i,j}(\boldsymbol{\eta})$ depend on the nonlinear parameters $\boldsymbol{\eta}$, but not on the linear parameters $\boldsymbol{\beta}$. Thus, for any given $\boldsymbol{\eta}$ a matrix $\boldsymbol{\Phi}(\boldsymbol{\eta})$ comprising the model functions, $\phi_{i,j}(\boldsymbol{\eta})$ is defined according to

$$
\boldsymbol{\Phi}(\boldsymbol{\eta}) \; = \; \begin{pmatrix}
\phi_{1,1}(\boldsymbol{\eta}) & \phi_{1,2}(\boldsymbol{\eta}) & \cdots & \phi_{1,n}(\boldsymbol{\eta}) \\
\phi_{2,1}(\boldsymbol{\eta}) & \phi_{2,1}(\boldsymbol{\eta}) & \cdots & \phi_{2,n}(\boldsymbol{\eta}) \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{m,1}(\boldsymbol{\eta}) & \phi_{m,2}(\boldsymbol{\eta}) & \cdots & \phi_{m,n}(\boldsymbol{\eta})
\end{pmatrix} \, .
\tag{16}
$$

Now the objective function to be minimized with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$ is given by

$$
\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\eta}}) \; = \; \min_{\boldsymbol{\eta}} \left\{ \min_{\boldsymbol{\beta}} \left\{ \| \mathbf{y} - \boldsymbol{\Phi}(\boldsymbol{\eta}) \boldsymbol{\beta} \|^2 \right\} \right\} \, ,
\tag{17}
$$

where the inner minimization problem forms a linear least-squares problem. The overall minimization problem can therefore be represented only in terms of the nonlinear parameters $\boldsymbol{\eta}$ according to

$$\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\eta}}) \;=\; \min_{\boldsymbol{\eta}} \left\{ \left\| \mathbf{y} - \boldsymbol{\Phi}(\boldsymbol{\eta}) \boldsymbol{\Phi}^{\dagger}(\boldsymbol{\eta}) \mathbf{y} \right\|^{2} \right\}, \tag{18}$$

with the generalized inverse $\boldsymbol{\Phi}^{\dagger}(\boldsymbol{\eta}) \;=\; \left[\boldsymbol{\Phi}^{T}(\boldsymbol{\eta}) \boldsymbol{\Phi}(\boldsymbol{\eta}) \right]^{-1} \boldsymbol{\Phi}^{T}(\boldsymbol{\eta})$. The advantages of the separable least-squares approach are that for the linear parameters $\boldsymbol{\beta}$ no initial estimate is necessary and the size of the Jacobian matrix is reduced which improves the condition and minimizes computing times. Only for nonlinear parameters such as the scale factors of the optical depth of molecule μ is a priori information on the number density (see Equation (1)) required. For CO and CH_4 this information is taken from the Air Force Geophysical Laboratory (AFGL) atmospheric constituent profiles [34], while H_2O and the auxiliary parameter temperature and pressure are taken from reanalysis data provided by the National Center for Environmental Prediction (NCEP).

2.2.2. Product Definition

Recall that in order to obtain atmospheric CO abundances, the BIRRA scales a representative CO reference profile. In general, to calculate the actual wet CMR of CO the inferred scaling factors are used according to

$$f_{CO} = \frac{\alpha_{CO} \, N_{CO}^{ref}}{\alpha_{air} \, N_{air}^{ref} + \alpha_{H_2O} \, N_{H_2O}^{ref}} \quad . \tag{19}$$

However, there is actually no parameter α_{air} showing up in the retrieval to account for the dry air CMR. Instead, differences between the retrieved and a priori CH_4 columns are used as a proxy to account for variations in the dry air column density according to

$$N_{CO} = \frac{\alpha_{CO}}{\alpha_{CH_4}} \, N_{CO}^{ref}. \tag{20}$$

This approach assumes that CH_4 is a well-mixed gas and that natural CH_4 variations in the atmosphere are small compared to variations of CO, with CH_4 changes produced solely by light path modifications (mainly by cloud shielding, for details see Gloudemans et al. [35]). Since CH_4 has strong absorption lines across the CO spectral fitting window in Channel 8, which allow is to determine the amount of CH_4 with good accuracy, this method is also suitable to detect optically thick clouds in the SCIAMACHY observations. In general, the error of the CH_4 scaling factors is 1–2 orders of magnitude smaller than the error of the CO scaling factors, depending on the scene. Intending to eliminate most of the uncertainties arising in the level-2 processing, postprocessing included the composition of quality criteria based on multiple parameters from the BIRRA output. For example, non-converging fits were filtered out, and errors of CO, H_2O and CH_4 were used to eliminate data with extremely low signal-to-noise ratios. This was done particularly for measurements on the ocean benefit from the presence of clouds due to the low reflectance of water in the SWIR range (see Figure 1, Gloudemans et al. [35]). This combination of filter criteria allows for the selection of optically thick clouds over the oceans and both cloudless scenes with acceptable small errors and cloudy scenes over land. If the light path was enhanced by,for example, aerosols, and exceeded a certain threshold (a 10% the enhancement of the CH_4 scaling factor), observations were rejected. Further details on the quality criteria chosen for the CO retrievals in this study are described in Gimeno García et al. [8] Section 2.3. The CO dry air CMR defined is therefore actually given by

$$xCO = d_{CO} = \frac{\alpha_{CO} \, N_{CO}^{ref}}{\alpha_{CH_4} \, N_{air}^{ref}} = \frac{\alpha_{CO}}{\alpha_{CH_4}} \, d_{CO}^{ref}. \tag{21}$$

2.3. Weighted Averages

It stands to reason that averaging over lots of measurements requires accounting for incorporating representation deficiencies. For example, if an average of 10 measurements is already deemed trustworthy, representation errors might not be an issue. However, in case of the SCIAMACHY's CO product, the number of measurements to be averaged for a representative mixing ratio (with respect to error of the mean) is at least in the order of 10^2.

A representative average value $\langle \mathrm{xCO} \rangle$ for both s-b and g-b data with respect to space ρ and time τ is the weighted mean according to

$$\langle \mathrm{xCO} \rangle = \frac{\sum\limits^{M} \omega \, \mathrm{xCO}}{\sum\limits^{M} \omega} , \tag{22}$$

with M representing the number of observations, and $\omega = \rho^\kappa \, \tau$ the respective weight. The weights include a spatial $\rho = r_{\mathrm{gb}} - r_{\mathrm{sb}}$ and temporal $\tau = t_{\mathrm{gb}} - t_{\mathrm{sb}}$ component, with r and t designating location and time of the observation, respectively. The CO spatial variation of column density is much larger than that of CO_2 and CH_4 [36,37]. Many NDACC and TCCON sites are located in remote regions while some sites are located in polluted areas (Table 1). Since the validation aims to compare s-b observations for both background and enhanced CO levels, the exponent κ is introduced to account for the representation deficiencies introduced by the great variability of CO in space. Typically $\langle \mathrm{xCO} \rangle$ is the monthly, seasonal, or annual mean. The corresponding standard deviation $\langle \mathrm{xCO} \rangle_{\mathrm{e}}$ is defined according to

$$\langle \mathrm{xCO} \rangle_{\mathrm{e}} = \sqrt{\frac{\sum\limits^{M} \omega \, (\mathrm{xCO} - \langle \mathrm{xCO} \rangle)^2}{\sum\limits^{M} \omega}} . \tag{23}$$

In general, if weighting is applied it needs to be applied on both datasets. Note that Equation (22) is valid with respect to both the spatial and the temporal domain.

2.3.1. Space

In the spatial domain the calculation of the averages for the SCIAMACHY data accounts for the position of acquisition relative to the location of the g-b data set. Therefore, $\tau = 1$ and $\omega = \rho^\kappa$. Note that since the location of the g-b reference site r_{gb} is fixed in time, $r_{\mathrm{gb}} = \mathrm{const}$.

2.3.2. Time

In order to calculate a representative average with respect to the temporal domain, the context was changed with $\omega = \tau$. In addition, a running average of t_{gb} needs to be introduced since t_{gb} is not constant. The interval selected for the running average is centered at the respective t_{sb} and chosen to extend the same time span as the corresponding weighted average. For example, a 30-day weighted average of the s-b CMR corresponds a running average of t_{gb} that incorporates g-b observations within $t_{\mathrm{sb}} \pm 15$ days.

2.4. Bias

In order to quantify the retrieval accuracy per reference site, the bias was calculated as an error weighted offset according to

$$b = \sum\limits^{M} \langle \mathrm{xCO}_{\mathrm{sb}} \rangle_{\mathrm{e}}^{-1} \left(\langle \mathrm{xCO}_{\mathrm{sb}} \rangle - \langle \mathrm{xCO}_{\mathrm{gb}} \rangle \right) / \sum\limits^{M} \langle \mathrm{xCO}_{\mathrm{sb}} \rangle_{\mathrm{e}} , \tag{24}$$

with $\langle \mathrm{xCO}_{\mathrm{sb}} \rangle_{\mathrm{e}}$ representing the standard deviation of the SCIAMACHY xCO in a certain time interval and M being the number of s-b and g-b averages (see Section 2.3), respectively.

In addition, the average of the s-b standard deviations at a specific site was used to characterize the accuracy of the bias b according to the standard error of the mean $s_e = \langle xCO_{sb} \rangle_e / \sqrt{M}$, with M the number of measurements incorporated in $\langle xCO_{sb} \rangle_e$. To estimate the overall performance, the global mean bias $\bar{b}$ was determined as the average of all station biases weighted by the standard deviation σ_b of the respective biases. $\bar{s}_e$ is the global (for all sites) standard error of the means. A similar approach was applied by Borsdorff et al. [9].

2.5. Averaging Multiple Years of CO

In order to exploit more observations within a given sampling area, multiannual averages for CO mixing ratios were calculated and refered to a common reference. Within the time interval 2003–2011, parameters for intercept a_n and slope b_n of xCO were estimated for each year n for both measurement systems and referred to a common reference. For this purpose, linear least-squares was applied. Subsequently, the annual estimates for intercept and slope were averaged yielding the parameters $\bar{a}$ and $\bar{b}$. The detrended data points $\hat{x}CO$ were then calculated according to

$$\hat{x}CO \rightarrow xCO_n - (a_n - \bar{a}) - (b_n - \bar{b})\, t_n. \tag{25}$$

The detrended observations $\hat{x}CO$ are largely independent with respect to the year of acquisition. Therefore more measurements belong to the common interval (e.g., a month, a year) compared to the single year analysis, and thus stricter thresholds can be imposed on filter criteria during post processing.

3. Results

The reference sites were selected with respect to the temporal data coverage and continuity of measurements as well as the reliability and completeness of auxiliary parameters such as surface pressure. The former two criteria were met if there were at least three years of observations covering the time period 2003–2011 with at least 10 months of observations. In case of NDACC stations all sites fulfilled this criteria, however, some TCCON sites such as Jet Propulsion Laboratory (JPL), Reunion Island, etc. had to be omitted. The latter criteria regarding auxiliary parameters posed an issue for some NDACC sites such as Arrival Heights, Lauder, or Wollongong. However, no TCCON sites were skipped due to this filter criterion.

3.1. Averaging of Measurements

The single most limiting factor in the global SCIAMACHY CO column product is its significant variability due to noise in the recorded spectrum. This clearly shows up in the CO values of individual observations in Figure 2. Furthermore, because Channel 8's pixel has a footprint on the Earth's surface of around 32×120 km^2 for the $180°$ nadir, a direct comparison of BIRRA-retrieved CO columns was not reasonable. It was found that even comprehensive filtering of the CO product based on the norm of the residuum ([8] Section 4) was not able to deliver the sufficient quality required for the comparison of a single or few (i.e., tens) of CO observations.

Figure 3 reveals that errors of individual SCIAMACHY retrievals have increased substantially since 2006. This seems to be primarily caused by omitting regular decontamination procedures from 2006 onward to dispose of the ice accumulated on the detector in Channel 8 (see [9] Table 1 and [2]). Furthermore, measurements conducted during this decontamination phase were commonly not suited for the adequate retrieval of parameters and most of them were filtered out during postprocessing. This effect is clearly visible in Figure 2 displaying the lack of proper CO retrievals at the turn of the years from 2003 through 2005. Also, the decreasing number of good spectral points (pixels) on the sensor plays a role (see [8] Figure 3).

Monthly-mean averages of CO mixing ratios constitute an adequate trade-off since the basic features of the CO annual cycle are preserved and at the same time they include enough measurements

to get reasonable statistics for both SCIAMACHY and NDACC/TCCON observations (see Figure 4).
Note that the validation studies cited in Section 1.5.2 also use a temporal interval of one month. Figure 4
also demonstrates that the adequate size of the temporal interval varies with the size of the sampling
area. Moreover, the results indicate that the adequate selection depends on the quality of the s-b data
and the problem investigated, hence also seasonal and annual averages were examined accordingly.

(a)　　　　　　　　　　　　　　　　　　　　　　(b)

(c)　　　　　　　　　　　　　　　　　　　　　　(d)
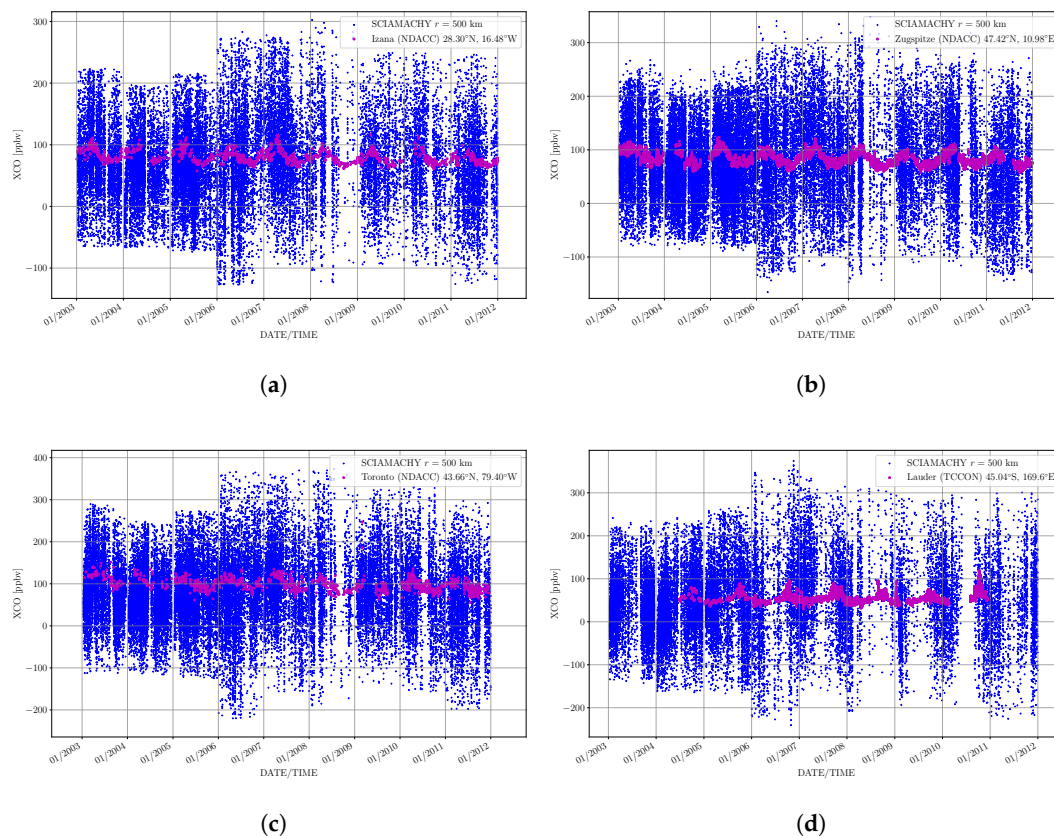
Figure 2. The SCIAMACHY and g-b reference site dry air column mixing ratios of carbon monoxide
(xCO). The plot includes the column mixing ratios (CMRs) after postprocessing within 500 km of the
reference sites (**a**) Izana, (**b**) Zugspitze, (**c**) Toronto and (**d**) Lauder from 2003 to 2011. The large scatter
of the individual SCIAMACHY CO columns is mainly caused by measurement noise, while averaging
of observations within 500 km of the reference site has only a minor contribution.

Since local enhancements shall be preserved, a related issue that also demands an adequate trade-off
is spatial averaging of the s-b observations within a given radius from the g-b reference site. In order to
include equidistant s-b measurements with respect to a reference site, latitude and longitude coordinates
of the measurements were converted to the great circle distance on the Earth's surface. In a first step, the
chosen distances for spatial averaging were 500 km, 1000 km, and 2000 km, since previous validation
studies of SCIAMACHY CO used collocation criteria within this range (e.g., Borsdorff et al. [9], De Laat
et al. [21], and Dils et al. [25]). The results in Figure 4 suggest that the smaller the sampling area, the
larger the interval for temporal averaging that should be chosen. So far, only Sussmann et al. [24]
sampled SCIAMACHY observations within 2000 km of Zugspitze. Other preceding validation studies
validated s-b data within a square- or rectangle-shaped area around the g-b site. The effect of the
different methodologies are displayed in Figure 5. Spatial filtering with respect to distance delivers a
smoother course of s-b CO averages throughout the year.

(**a**)   (**b**)
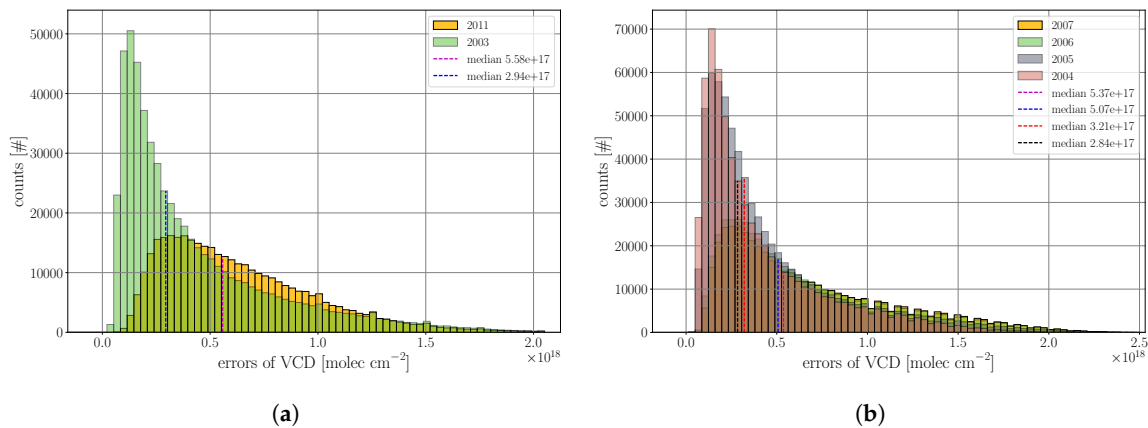
Figure 3. Histogram of the errors of the CO columns (defined in Equation (20)), represented as the unnormalized error probability function (also see Section 4.2.1 in Gimeno García et al. [8]). The data content is the global fraction of CO columns that survived after postprocessing of the BIRRA retrieval output. The dashed lines indicate the respective median. Note the different range in the x-axis. (**a**) The first and last year of the nominal SCIAMACHY operation covering the complete time span; (**b**) The occurrence frequency of total column errors for the years 2004 through 2007 indicate a substantial degradation of the quality of the CO product from 2006 onward (blue and magenta dashed lines). The years 2003 through 2005, however, show similar and significantly smaller errors.


(**a**)   (**b**)
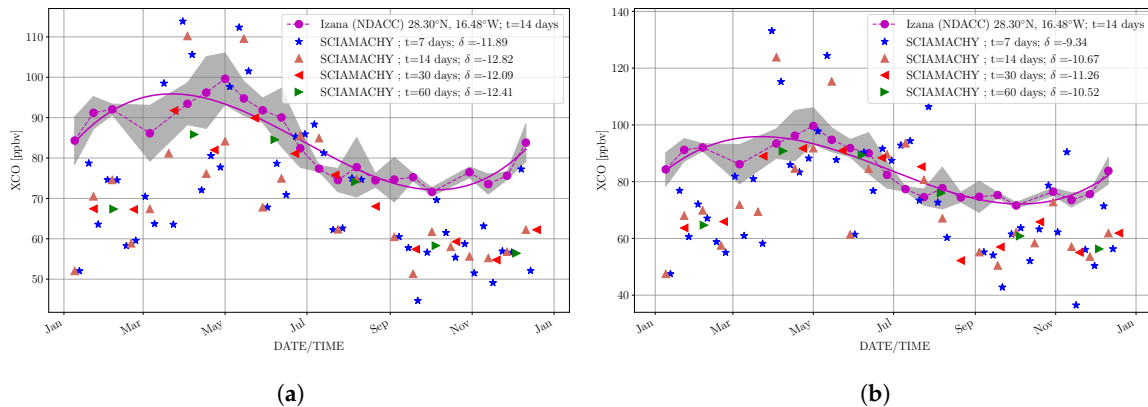
Figure 4. xCO averages for various temporal intervals within (**a**) 1000 km and (**b**) 500 km of Izana in 2003. δ designates the offset of the s-b averages to the 14-days mean of Izana. The gray shaded area and magenta solid line represent the standard deviation of the g-b means and the third-order polynomial fit on the individual observations, respectively. The SCIAMACHY xCO averages become smoother in larger temporal intervals and in the share of averages. This is in agreement within the standard deviation of the g-b site increases. Note the different range of the y-axis.

The SCIAMACHY observations within various-sized areas around two the NDACC and TCCON g-b sites, namely Bremen and Izana (see Figure 1), are compared in Figure 6 and the results demonstrate that the adequate spatial and temporal range of averaging depends on the precision of the s-b observations at a specific site. The former station is located in the plains of northern Germany, while the latter is situated on an island in the sub-tropical region in the Atlantic ocean at about 2370 m above mean sea level. With respect to Izana, a station largely surrounded by water, observations of optically thick clouds were preferred in order to get an acceptable SNR for the retrieval. Note that the bias *b* designated in the legend is the deviation of the s-b to the g-b CO averages weighted by the respective s-b standard deviation (details on the bias analysis in the subsequent section).
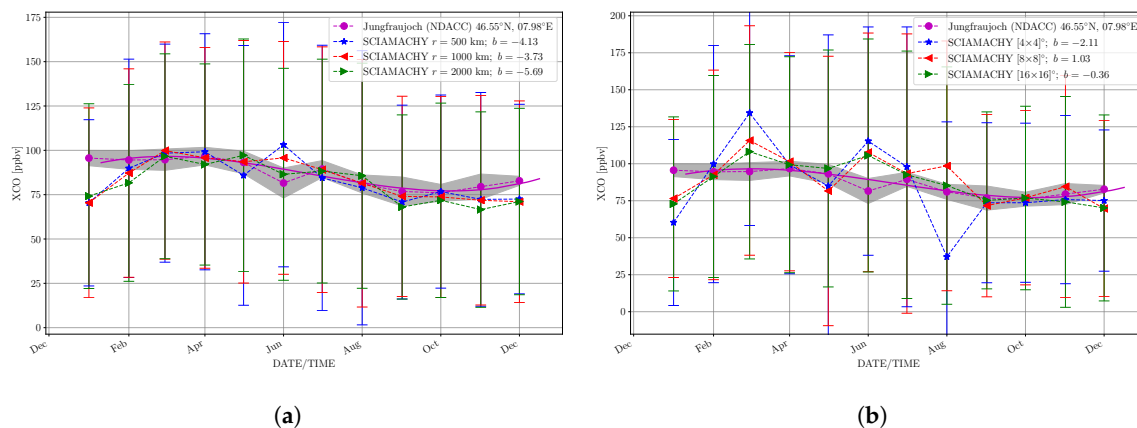
Figure 5. xCO averages for circle- and square-shaped sampling areas. (**a**) Averages of SCIAMACHY xCO values calculated for various distances from the reference site Jungfraujoch in 2003; (**b**) xCO averages within given latitudes and longitudes of the g-b site. Here, $[4 \times 4]°$ designates a square-shaped object with an extent of $\pm 2°$ in latitude and longitude from the reference site. SCIAMACHY observations within the circle-shaped sampling areas show better agreement for most cases in 2003. Note the different range of the y-axis.

Another important aspect that needs to be considered is that for spatially weighted averages (used throughout the subsequent sections), the effective radius becomes seemingly smaller since close measurements account most while observations far away from the g-b measurement site have much weaker impact on the representative average. This basically means that in the case of weighted averages, the area significantly contributing to the average of s-b measurements is reduced. Therefore, we consider averaged CO values of 1 month and within 500 km of the g-b site as an appropriate trade-off for SCIAMACHY CO validation.
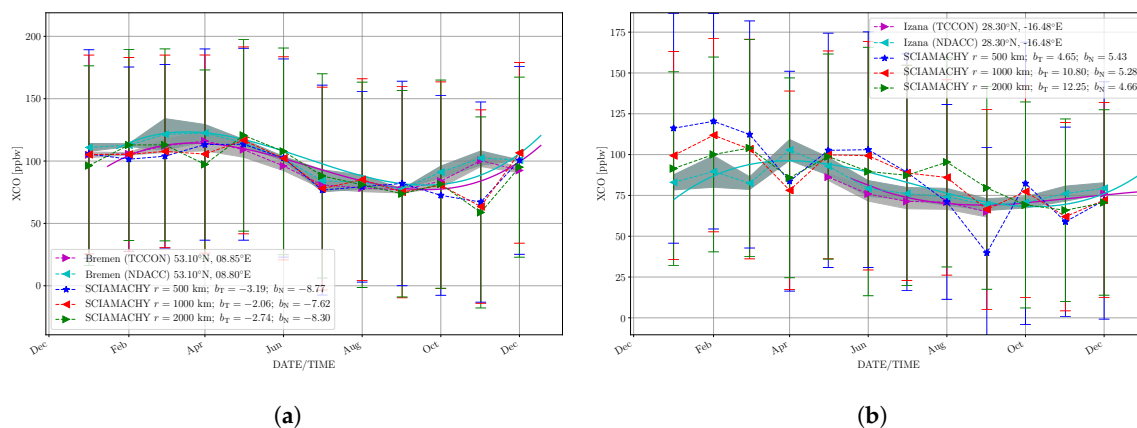
Figure 6. xCO monthly averages around the collocated NDACC and TCCON sites in 2007. (**a**) SCIAMACHY xCO values calculated for various distances from the reference site in Bremen. Measurements of the SCIAMACHY within 1000 km show the smallest bias b; (**b**) xCO averages with respect to Izana in 2003. On this site the SCIAMACHY observations within 500 km of the reference site show the smallest bias in total for the TCCON.

3.2. Bias and Weighting

The bias of SCIAMCHY observations to the g-b reference was analyzed using two different approaches. Initially, in Section 3.2.1, the classical unweighted averages for SCIAMACHY and g-b observations are used, while distance weighted averaging is treated in Section 3.2.2.

3.2.1. Unweighted Bias

Overall, Figure 7 shows that the bias of CO mixing ratios is ranging from almost 0 ppbv to -27 ppbv. In over 80% the biases range from around -5 to -15%. The global bias $\bar{b} = -12.1$ ppbv represents the mean bias of all stations weighted by the respective standard deviation. The global standard error of the mean $\bar{s}_e = 6.77$ ppbv is the average of all the stations standard errors of the mean.

In general, the results in Figure 7 agree with the findings by Borsdorff et al. [9,10] within around -5 to -10% for most sites. However, the CO mixing ratios calculated from the BIRRA retrieval have consistently low bias with respect to all sites in both networks, while Borsdorff et al. [9,10] found slightly positive biases at some sites. The data show different magnitudes of biases for different sites. Notice that some negative bias in the northern mid-latitudes in Europe and North America are significant with respect to standard deviation. These findings suggest that the comparison crucially depends on the CO a priori information which is not well characterized in polluted areas. Sites located in remote areas such as Ny Alesund, Izana, Wollongong, and Lauder do therefore show much better agreement (Figure 7).

(a)   (b)
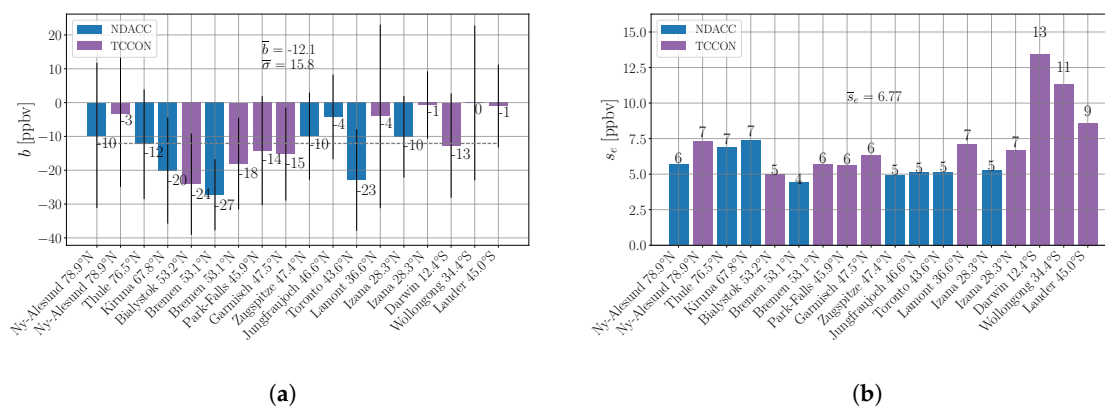
Figure 7. Mean bias of NDACC (blue) and TCCON (purple) stations with co-located SCIAMACHY CO retrievals from 2003 to 2011. Bias is the average of the monthly-mean differences weighted by the standard deviations. The total (2003–2011) bias per site was subsequently calculated using the annual weighted biases. (**a**) Global bias $\bar{b}$ and respective global standard error $\bar{\sigma}$. The dashed gray line indicates $\bar{b}$; (**b**) The standard error of the means s_e.

3.2.2. Distance Weighted Bias

The impact on xCO averages was studied using linear (r^{-1}) and quadratic (r^{-2}) inverse distance weighting of spatially close measurements (not to be confused with the error weighting of the bias also used in Section 3.2.1). In general, at most sites, the biases are decreased in the case of inverse distance weighting. This effect can be observed for both the linear inverse and quadratic inverse weighting (Figures 8 and 9). In the case of linear weighting only a few sites exhibit slightly larger biases, such as Kiruna (NDACC), Bialystok, Wollongong, and Lauder (TCCON). However, the overall bias was reduced to -11.2 ppbv while the average standard deviation was increased to 16.6 ppbv. Furthermore, most sites show that quadratic weighting (with respect to inverse distance of SCIAMACHY measurements) increases the agreement of the compared averaged values. The majority of biases observed in Figure 9 are reduced even more compared to the linear approach and the distribution throughout the sites from north to south is significantly smoothed. This could indicate that the results in the linear and quadratic weighted cases are better suited for the estimation of the instrument and retrieval deficiencies compared to the unweighted results which might include larger fractions of representation errors. In other words, it is likely that a fraction of the revealed offsets are attributed to mismatch artifacts and not only incorporate instrument or retrieval issues.

Favorable examples to study this effect more closely are g-b sites located in polluted areas, especially large cities embedded in rural areas such as Toronto. While the SCIAMACHY seems to significantly underestimate CO at Toronto in the unweighted example (Figure 7), the quadratic weighting reveals that the offset is not significant but rather suggests to be introduced by less representative observations. However, discrepancies at Kiruna, Bialystok, and Bremen in the northern mid-latitudes still seem to be significant. In the case of Bremen, the negative offset is significant with respect to both NDACC and TCCON observations, which underlines the finding of a negative bias of the SCIAMACHY SWIR observations. The global bias was reduced to 11.9 ppbv compared to the unweighted approach, but $\overline{\sigma}$ increased to 20.8 ppbv.

Table 3 reveals a discrepancy of the global bias between the NDACC and TCCON of 6.2 ppbv for the unweighted case and 5.3 ppbv in case of distance weighting (r^{-2}). In general the findings with respect to both networks are in accordance with the conclusions by Borsdorff et al. [9] in many aspects. The global $\overline{b}$ bias is larger for sites affiliated to the NDACC than those affilliated with the TCCON. With respect to the NDACC and TCCON, we found $\overline{b} \approx -14$ ppbv and $\overline{b} = -7.4$ up to -9.1, respectively, depending on the weighting while Borsdorff et al. [9] revealed values of -9.2 ppb and -1.2 ppb for the clear-sky retrievals for the NDACC and TCCON, respectively. In case of cloudy-sky retrievals, Borsdorff et al. [10] found a negative bias $\overline{b} = -6.0$ ppb which is similar to our results for the TCCON (see Table 3).

In order to underline the previous findings, it was important to confirm that the bias of the SCIAMACHY data with respect to the g-b reference did not exhibit a linear trend throughout the mission. Therefore, a bias trend t in ppbv/year was calculated for the annual biases b at each site using linear least-squares regression. The results are depicted in Figure 10. The global trend $\overline{t}$ is the average of the bias trends weighted by the standard error of the fit. In order to get a complete picture, initially all sites for all available years were investigated. It was found that the number of sites that reveal a negative bias trend exceeds the number of sites with a positive bias trend. Note that the trend is within ± 5 ppbv around most reference sites. In addition, for most sites no significant trend with respect to the accompanying standard error was identified. Most sites show a consistent standard error of about 5 ppbv/year in both the weighted and unweighted cases. Moreover, the magnitude of the standard error is in the order of (or even larger than) the calculated average trend at most sites (see Figure 10). The numbers in the figure indicate that the global trend of the bias is not significant for at least two cases. Hence, we decided that a linear trend correction of the results presented above was not required.

(a)   (b)
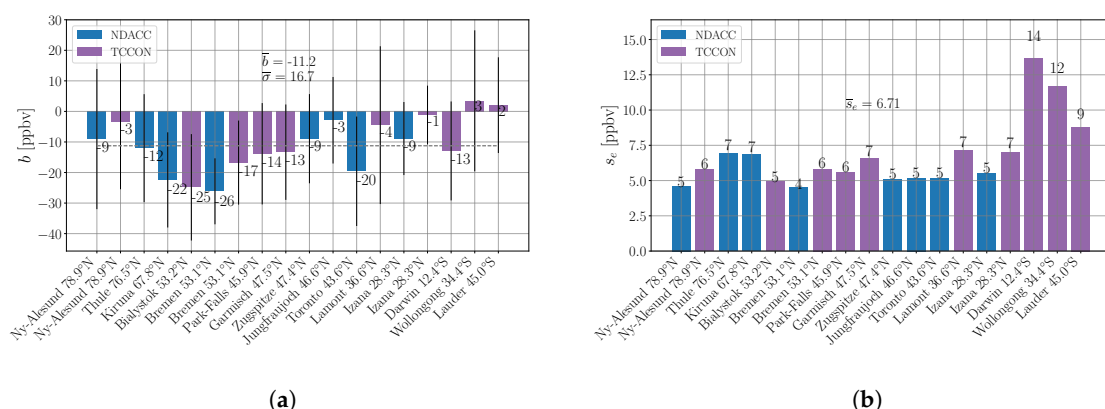
Figure 8. Mean bias, standard deviation, and standard error as in Figure 7 but with monthly-mean averages of SCIAMACHY CO weighted according to the inverse distance from the reference site. (**a**) Global monthly-mean bias $\overline{b}$ and respective standard error $\overline{\sigma}$; (**b**) The global standard error of the mean biases $\overline{s}_e$ is the average of all standard errors of the means.

(a)　　　　　　　　　　　　　　　　　　　　　　　　　　(b)
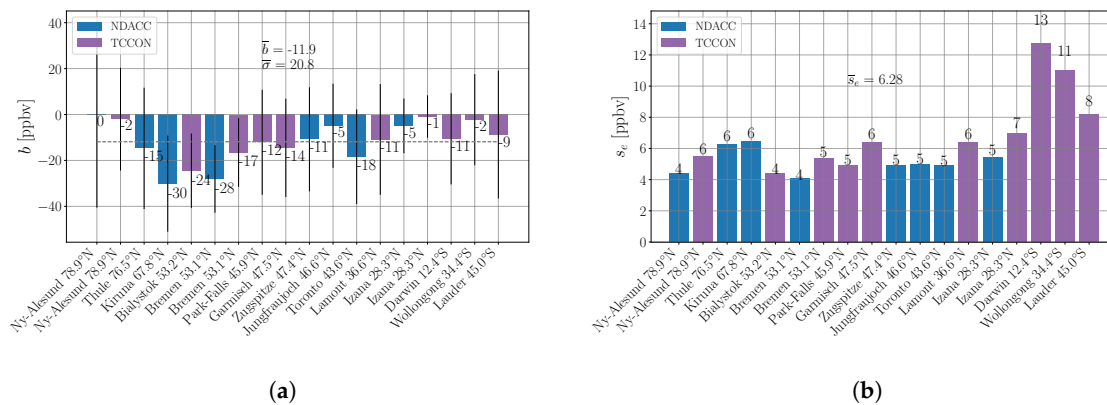
Figure 9. Same illustration as in Figure 7 and 8 but weighted inverse to the distance squared. (**a**) Global monthly-mean bias $\bar{b}$ and respective standard error $\bar{\sigma}$; (**b**) The global standard error of the mean biases $\bar{s}_e$ is the average of all standard errors of the means.

Table 3. The respective biases $\bar{b}$ (bold) and standard errors $\bar{s}_e$ separated for NDACC and TCCON. The NDACC reveals a significant bias and standard error of the means. This is also true if linear (r^{-1}) and quadratic (r^{-2}) inverse distance weighting is applied. With respect to the TCCON, the bias is less significant and closer to the magnitude of the standard error. ppbv: parts per billion in volume.

| ppbv	1		r^{-1}		r^{-2}	
NDACC	**−14.3**	5.5	**−13.8**	5.4	**−14.4**	5.1
TCCON	**−8.1**	7.7	**−7.4**	7.9	**−9.1**	7.1
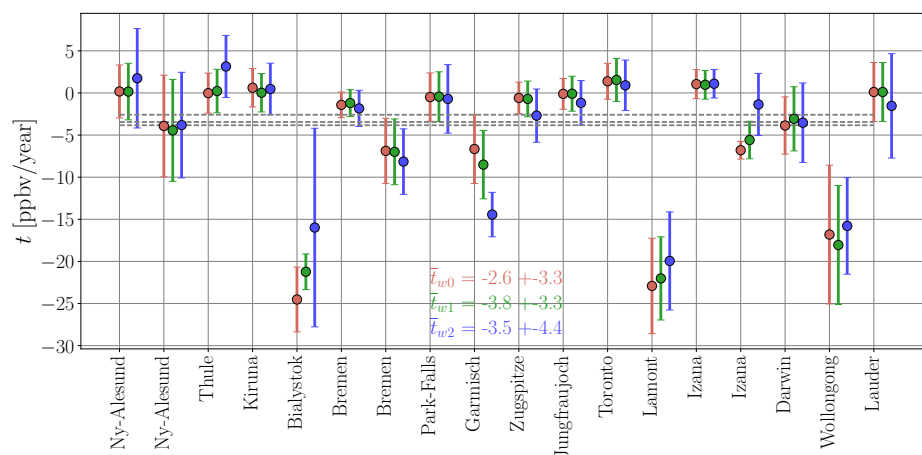
Figure 10. Linear bias trend t of the mean biases b for the classical (unweighted) and weighted approaches. No significant bias trend is identified on the global scale, however, a few sites do show a significant negative bias trend.

3.2.3. Spatio-Temporal Weighting

In general, the findings suggest that accounting for the spatial mismatch by inverse distance weighting does not necessarily require the introduction of a temporal penalty term (accounting for temporal mismatch) in order to obtain representative results for comparison. Figure 11 shows the effects of the spatio-temporal weighting on the 90-day averages for Toronto in 2003. Overall, the outcome demonstrates that temporal weighting has only minor effects on the accuracy and should in general

only be regarded as an option in areas where data quality is degraded, and therefore averaging over a larger number of observations in the space and/or time domain becomes necessary.


(a)   (b)
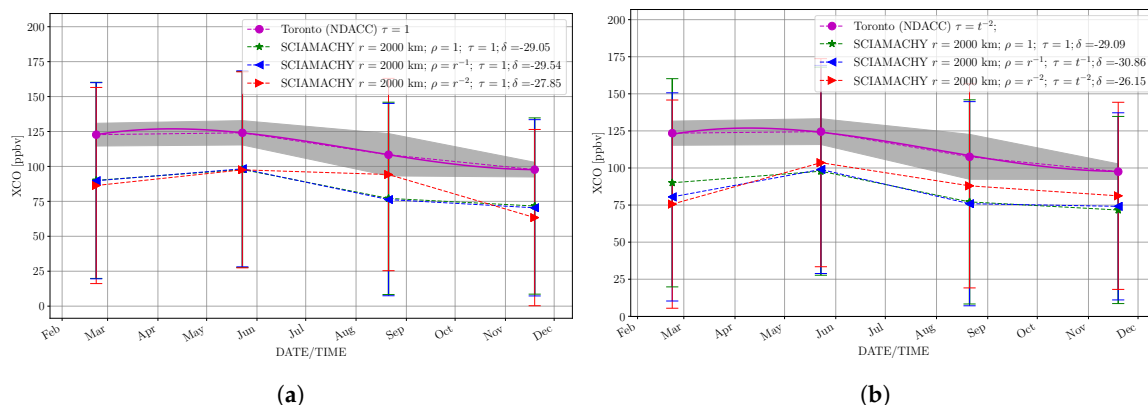
Figure 11. The 90-day averages within 2000 km from the Toronto truthing site in 2003. In (**a**) only spatial weighting was applied in the calculation of averages while in (**b**) both spatial and temporal weighting was used. δ designates the offset of the averages between xCO from SCIAMACHY and Toronto.

3.2.4. Multiannual Averages

Measurements in the SWIR channel suffered from decontamination procedures (see Section 3.1) often executed at the turn of the year. Fitting the intercept and slope based on least-squares is a process which is sensitive to outliers, and therefore, the months December and January were excluded in the calculation. In about 40% of the cases the intercept of the s-b retrievals matched within the standard deviation of the respective g-b intercept and in around 60% a positive slope of the g-b measurements matched with a positive slope of the s-b observations. Figure 12 shows two sites in the southern hemisphere where both intercept and slope coincide within the standard deviation of the g-b reference.

In Figure 13, s-b CO columns are compared to the Zugspitze (NDACC) and Garmisch (TCCON) reference sites. These g-b sites are distinct due to their close horizontal but great vertical distance (see Table 1). The permanently higher CO values at Garmisch might be an effect of the planetary boundary layer, since observations at Zugspitze are not influenced by this throughout most of the year. Notice that the 500-km circle (from where the SCIAMACHY measurements were taken) was centered at Zugspitze for both cases due to the proximity of both sites ($\approx$8 km) and the large sampling area.

Figure 13a shows the comparison for the unweighted and weighted xCO averages. The results demonstrate that some fraction of the SCIAMACHY xCO bias seen at Garmisch and Zugspitze is likely to be an effect of spatial mismatch, i.e., representation error, and should therefore not solely be considered as a retrieval or instrument flaw. Figure 13b examines how the squared inverse distance weighting is affected by variations in the size of the sampling area. It clearly reveals that the offset becomes smaller the closer the measurements are taken from the g-b site. In this case, the weighted average of SCIAMACHY observations from within 500 km shows the smallest offset with respect to the g-b sites. However, the CO annual cycle unveils in the 1000-km and 2000-km cases but is not visible in the 500-km circle.

(a)　　　　　　　　　　　　　　　　　　　　　　(b)
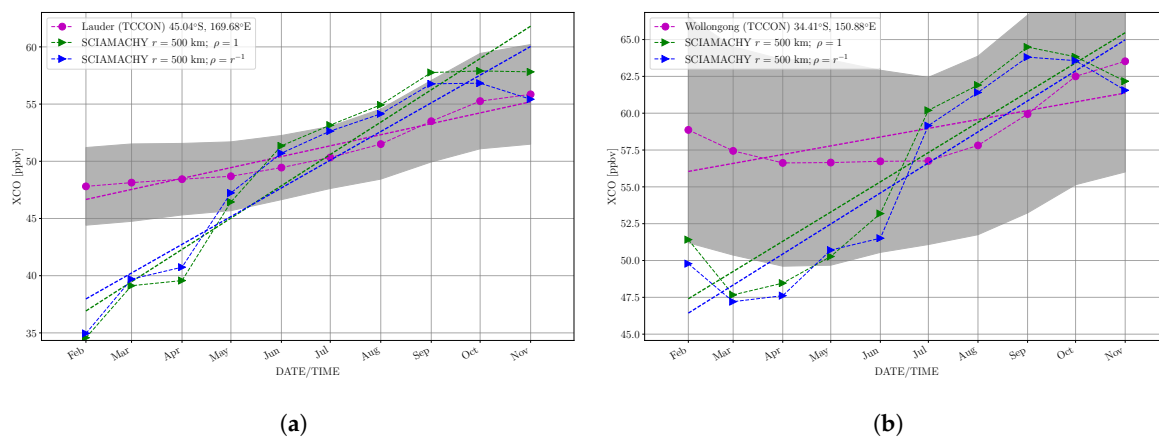
Figure 12. Comparison of intercept and slope of s-b and g-b xCO. The years 2004 through 2011 were included in the analysis for Lauder depicted in (**a**). Results for Wollongong that include data from 2008–2011 are shown in (**b**).

(a)　　　　　　　　　　　　　　　　　　　　　　(b)
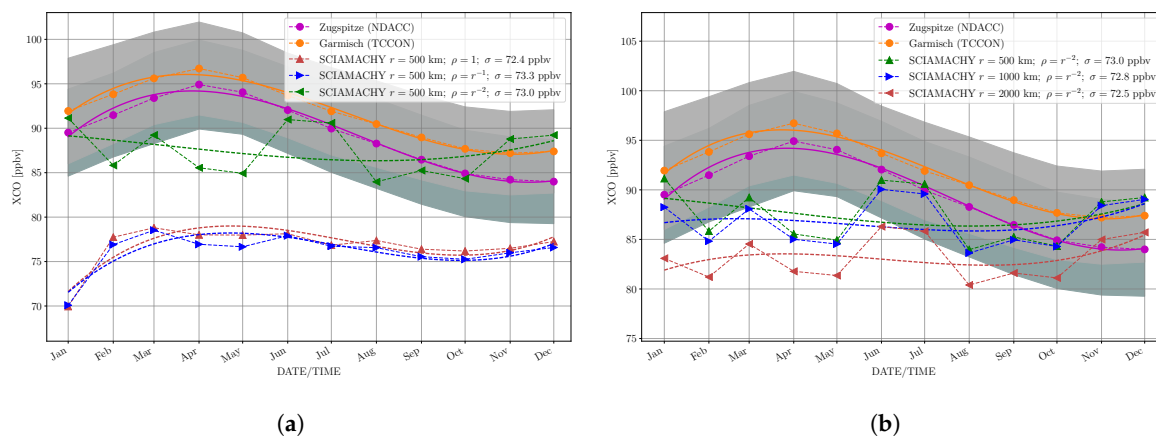
Figure 13. The effect of distance weighting on SCIAMACHY x̂CO with respect to Garmisch (2007–2011) and Zugspitze (2003–2011) reference sites. The solid and dashed lines show the three-degree polynomial fit on the individual observations and $\sigma \equiv \langle \text{xCO}_{sb} \rangle_e$. (**a**) Weighting of measurements within a given distance leads to better agreement with g-b references; (**b**) The effect of weighting partly reveals the annual cycle at the reference sites. However, the amplitude is low and a slightly negative offset remains. Note that the standard deviation of the s-b measurements is included in the legend.
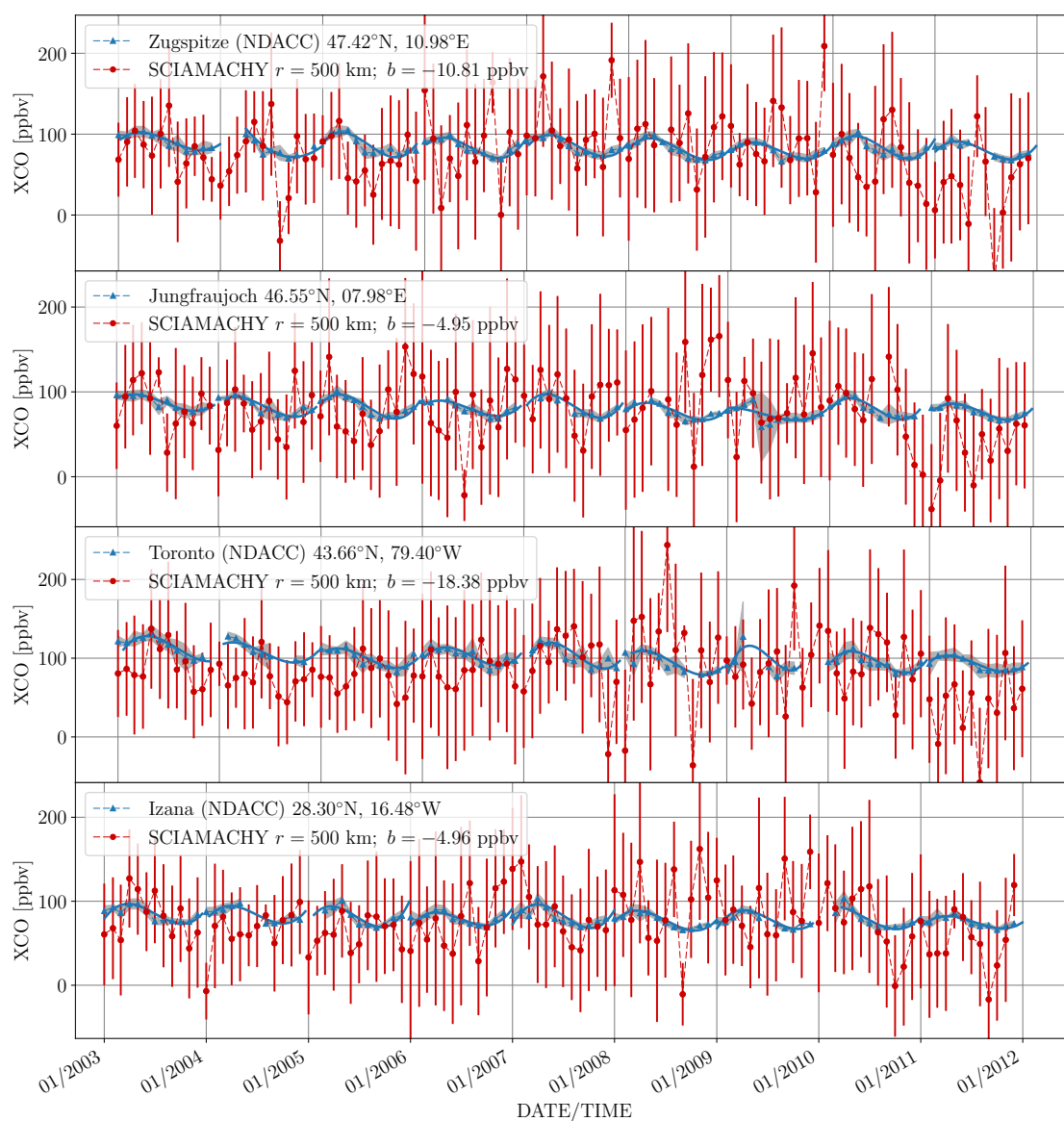
3.2.5. Time Series

For the NDACC and TCCON stations listed in Table 1, Figures 14 and 15 show the time series of dry air quadratic inverse distance-weighted monthly-mean CO CMRs. Both figures demonstrate that the large scatter of the BIRRA retrievals is mainly caused by the noise in the SCIAMACHY observations. The standard deviation of individual months sometimes even exceeds a typical mean CO concentration. The results underline that solar radiation reflected at the Earth surface or clouds reduces scattering due to a higher SNR.

3.3. Mission Averaged Global CO

The 2003–2011 averaged CO product is shown in Figure 16 with the corresponding errors in Figure 17. In order to acquire representative CO mixing ratios on a regular grid globally, averaging over the complete validation period was required. Beside increased CO concentrations over China, parts of India, and Indonesia due to pollution and wildfires, Figure 16 also reveals valuable information

of the global transport of pollution in the Earth's atmosphere. For example, the outflow of CO over the Atlantic Ocean due to biomass burning in central Africa becomes clearly visible. However, the gradients in the CO columns appear to be less pronounced due to several reasons. One is that the averaging smoothes peaks of CO that occur during the wildfire seasons over central Africa. This shows that this effect is less pronounced in regions where high concentrations of CO prevail throughout the year (see eastern parts of China). Another aspect is that the filter criteria applied also accept scenes where optically thick clouds extend into the mid-troposphere, therefore providing good SNRs but obscuring the high CO concentrations in the lower troposphere and boundary layer. This reveals that the effect is most pronounced in the region of the Inter Tropic Convergence (ITC) zone.

Figure 14. Time series of r^{-2} distance weighted monthly-mean xCO dry air column mixing ratios measured by the SCIAMACHY and NDACC stations. The CO mixing ratios were calculated using the surface pressure information per station according to the description in Section 2.1.1. Sites where no such information was available were skipped.
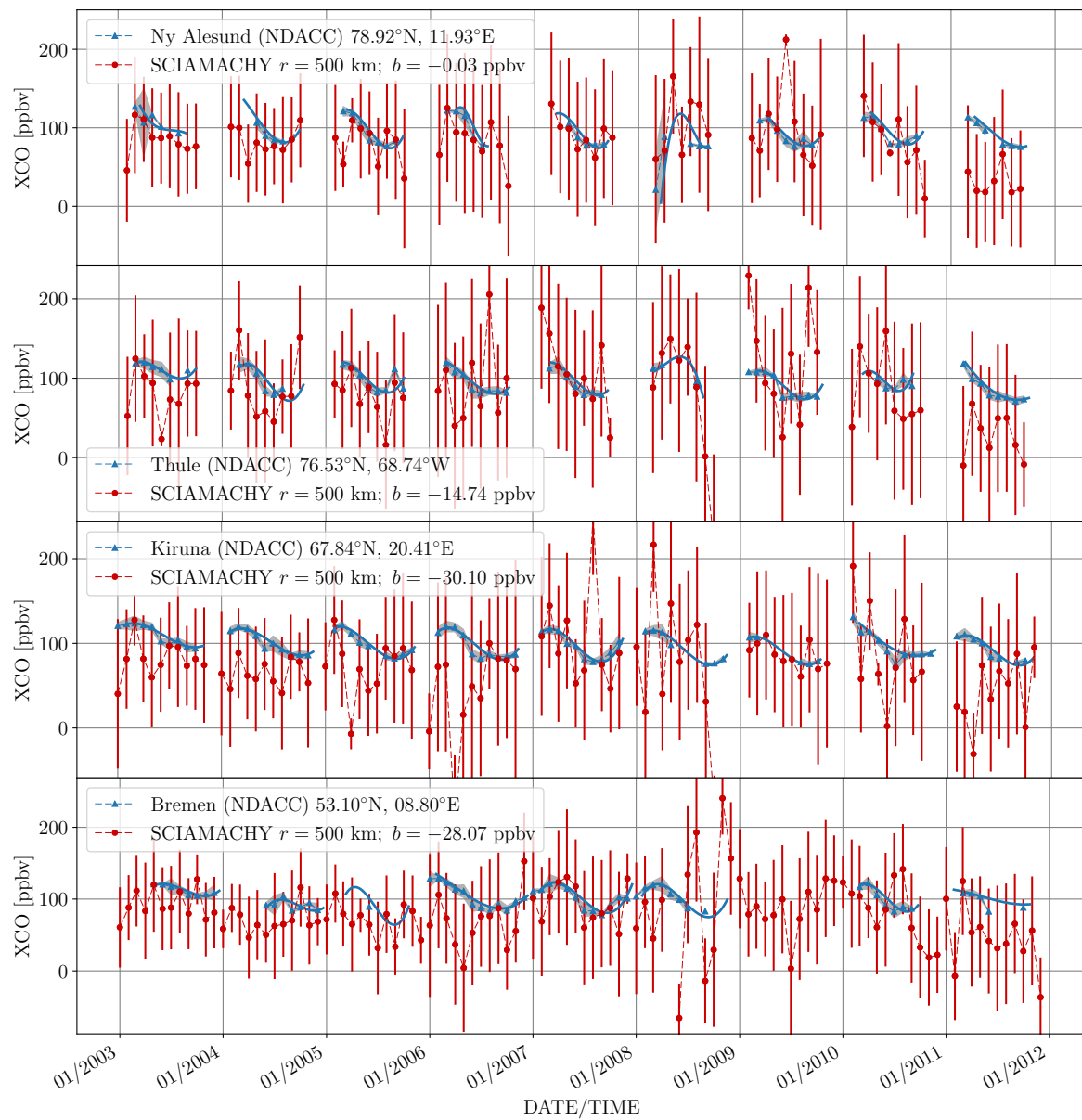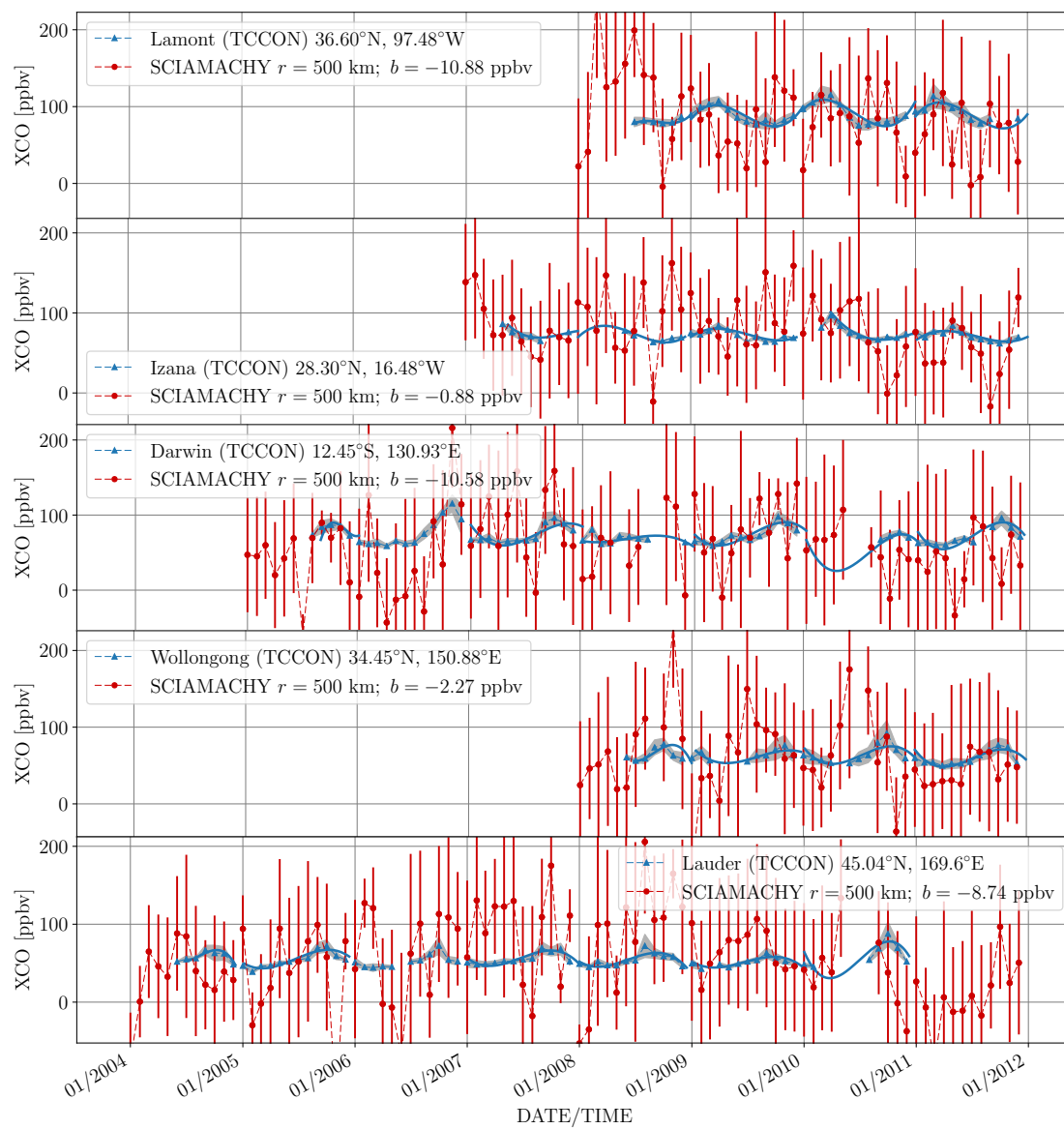
Figure 14. *Cont.*

Figure 15. Time series as in Figure 14 but for the TCCON stations. Sites that provided retrievals according to the GGG2014 data release of TCCON were used for the comparison to the BIRRA columns.
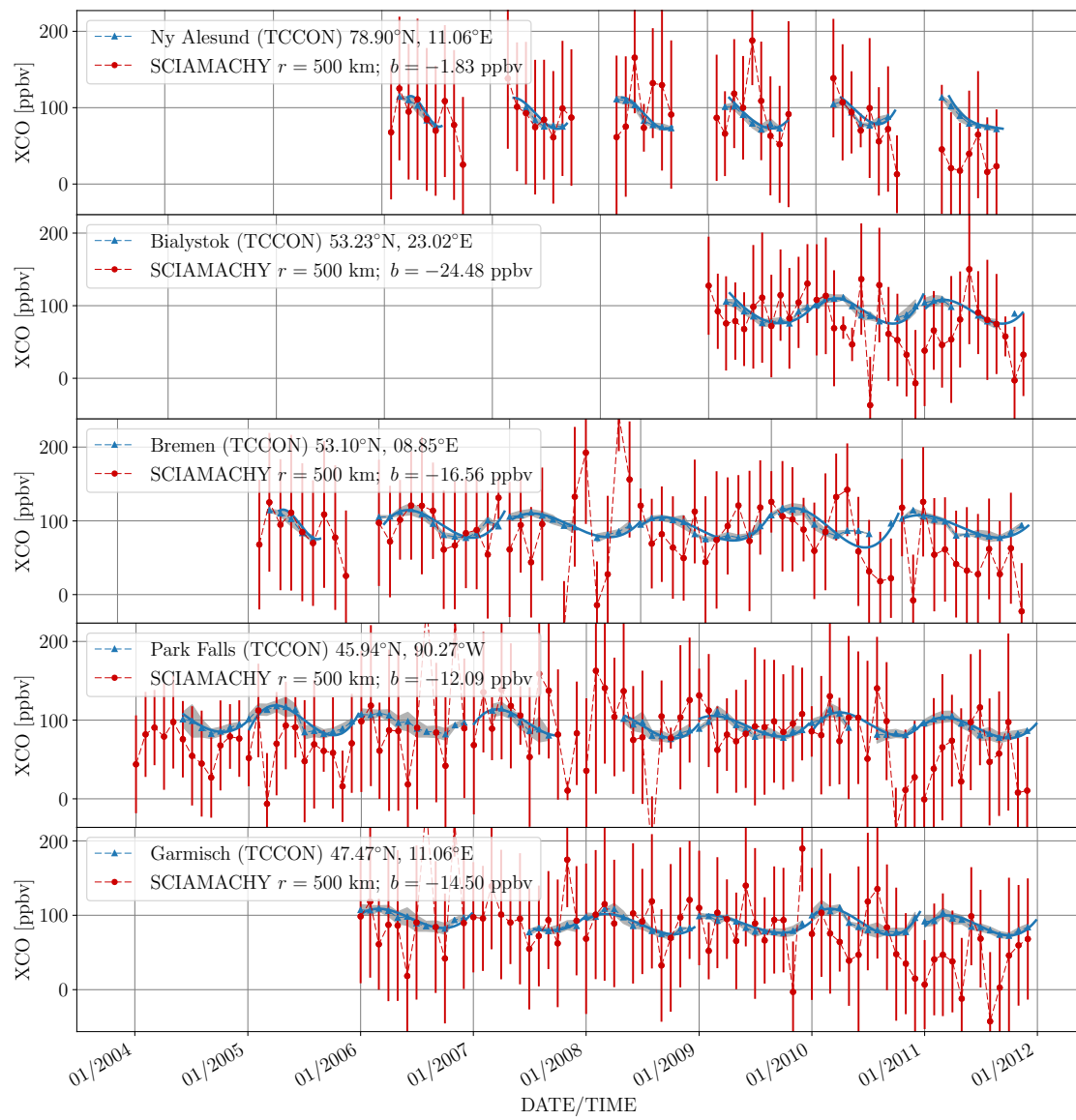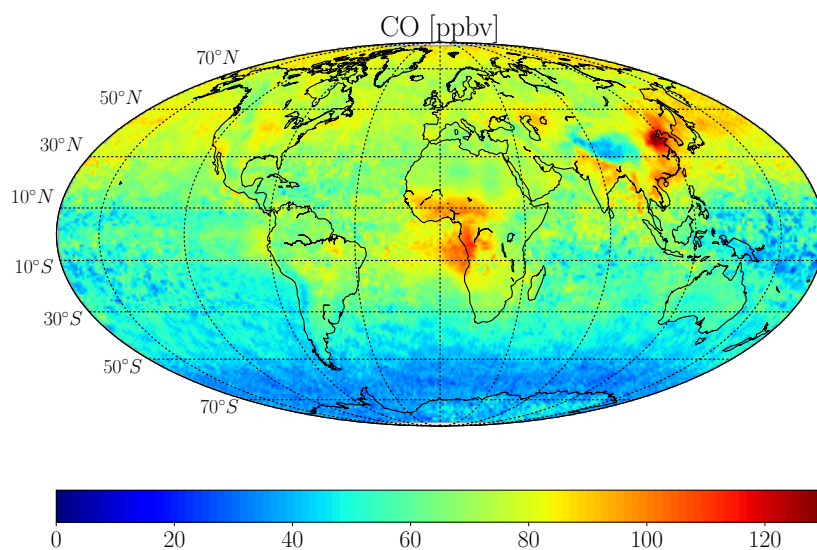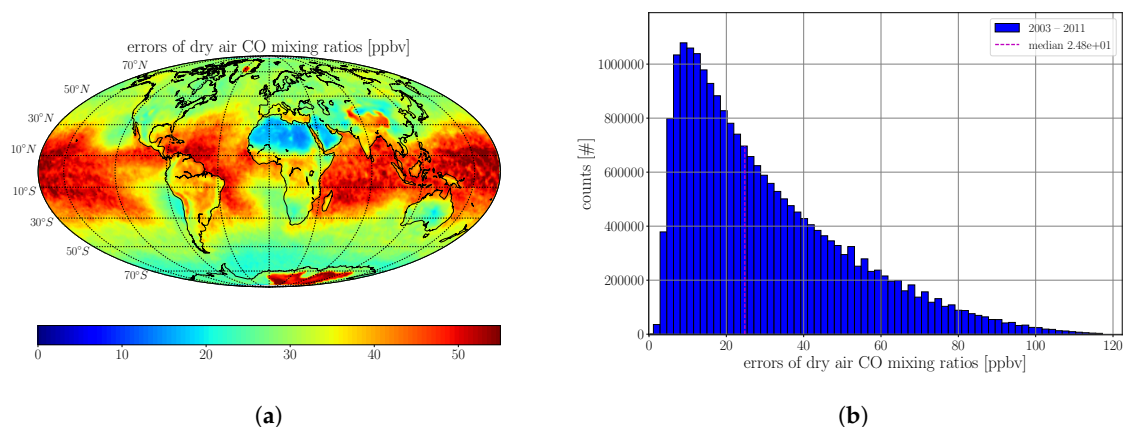
Figure 15. *Cont.*

Figure 16. Dry air CO column-averaged mixing ratios in ppbv over land and the oceans. The values are averaged from January 2003 to December 2011 on a [1×1]° latitude/longitude grid. Measurements of clear-sky scenes and optically thick cloud conditions are included.

(**a**)   (**b**)

Figure 17. (**a**) Errors of the CO total columns for the observations presented in Figure 16. It clearly shows that errors are smallest in dry regions with low cloud coverage and high surface reflectivity (see the Sahara Desert, the Arabian Peninsula, Figure 1); (**b**) Frequency distribution of the errors including years from 2003–2011 (some individual years are shown in Figure 3).

4. Discussion

The BIRRA CO total column product of the SCIAMACHY Channel 8 spectra was validated for the entire mission. In general, the results suggest that weighting is capable of mitigating some incorporated representation errors that come into play when comparing satellite-averaged CO data to averages of g-b reference sites.

A global negative bias of about −10 ppbv and a standard deviation of 15–20 ppbv (depending on the approach used for averaging and the g-b reference network used, see Table 3) was found for the SCIAMACHY xCO. Since both networks perform direct sunlight observations using FTS, the fact that SCIAMACHY measurements acquired in remote areas within the subtropical bands tend to exhibit smaller negative biases compared to observations in the mid-latitudes suggests that part of the s-b bias is caused by cloud shielding of the lower troposphere and an insufficient representation of the CO a priori profile in those regions (see Figures 7–9). In the case of Izana, a site located at about 2.4 km above mean sea level, the effect of cloud shielding is small, hence the cloud induced bias is reduced.

The full-mission analysis from Borsdorff et al. [10] also included cloudy-sky retrievals and found similar results for the bias which was estimated to be $\bar{b} = -6$ ppb. Another common feature our results is the larger negative bias for the NDACC as compared to TCCON (see Table 3). The differences might arise from the fact that both networks observe in different spectral regions and are using different retrieval approaches.

4.1. Weighted Averages

In the validation of the s-b to g-b reference, it is necessary to find representative averages of both data sets with respect to one another. Early validation studies (e.g., Sussmann and Buchwitz [23]) pointed out that even without correction, the statistical effect of reducing scatter by increasing the ensemble of averaged pixels surpasses the effect of including less representative pixels. However, a thorough validation effort should also take into account the incorporating representation issue. In particular, the variability of CO induces gradients on a regional basis which the validation methodology should account for when comparing two independent systems.

The results demonstrate that assuming xCO to be less representative with increasing distance from the reference is reasonable and that distance-weighted averages reduce the negative bias for the 2003–2011 period at most reference sites (see Figures 7 and 9). The different averaging techniques agree upon the fact that the global bias is not significant with respect to the standard deviation at most g-b locations for the 2003–2011 period (see Figures 7–9). Moreover, the methods agree upon the larger standard error of the mean bias in the Southern Hemisphere. It also became evident that the selection of larger sampling areas requires smaller temporal averages to be chosen for a given noise level (see Figure 4). The findings give confidence that the BIRRA retrievals are consistent on the global scale and for the time period of the SCIAMACHY mission.

4.2. Temporal Averages

Firstly, it is important to note that with respect to time, close temporal measurements of two observation sites do not necessarily mean similar concentrations since spatial gradients are much more pronounced compared to temporal ones (i.e., show smoother gradients). The concentration of CO in rural as well as urban areas will exhibit only weak variations throughout a monthly or seasonal period. Due to this weak fluctuation, the temporal domain, including temporal weighting in the analysis of the averages, requires utilization of the spatial distance of the measurements from the independent observing systems as a penalty term. This penalty term accounts for the decreasing representativeness of the measurements due to the spatial mismatch—although they may have been observed at the same time. Therefore, including a stronger penalty term for a spatial mismatch as compared to a temporal mismatches is regarded as a reasonable approach. It is also important to note that since spatial weighting only considers measurements with respect to their distance from the reference without taking into account directional gradients, spatial weighting is expected to deliver the best results when the CO concentration is isotropically distributed around a g-b site. However, even in cases of non-perfect isotropic distribution from the reference, the method of weighted averages seems to deliver more representative results than simple averaging without weighting (see Figure 13).

4.3. Multiannual Averages

Figures 2 and 3 clearly demonstrates that later years (2006 onward) in the SCIAMACHY dataset suffered from low SNRs, and may therefore require a modified approach for comparison. The multiannual averages, presented in Section 3.2.4, are demonstated as useful and capable in further reducing random errors. Although the multiannual averages cannot eliminate the negative bias completely (see Figure 13) and reveal disagreements in slope at most sites, the results improve for the majority of sites compared to the single-year analysis.

4.4. Spatio-Temporal Averages:

　　Section 3.2.3 shows that incorporating temporal weighting, in addition to distance weighting, is not crucial since CO columns are quite constant within a month. Hence, closer, mutually temporal observations do not necessarily mean improved representativeness (in contrast to the spatially persistent CO gradients). However, if stronger weighting on spatial mismatch is deemed useful (e.g., in areas of strong spatial gradients) one might consider including a (less pronounced) penalty term that accounts for the temporal mismatch. Quadratic spatial weighting, applied in conjunction with temporal averages beyond 30 days, should perhaps be considered (see Figure 11).

4.5. CO from TIR vs SWIR Observations

　　As an alternative to the SWIR range, atmospheric remote sensing of CO can exploit the rotational band in the microwave [38–40] or the fundamental band in the TIR range. Several s-b nadir viewing instruments observe CO in the TIR: the Atmospheric Infrared Sounder (AIRS) [41], the Cross-track Infrared Sounder (CrIS) [42,43], the Infrared Atmospheric Sounding Interferometer (IASI) [44–46], the Measurement of Pollution in the Troposphere (MOPITT) instrument [47], and the Tropospheric Emission Spectrometer (TES) [48]. Initially MOPITT retrievals utilized the TIR range only, but Worden et al. [49] demonstrated that multispectral retrievals exploiting both the near infrared (NIR, 0.8–3 μm) and TIR channels increase the sensitivity to CO.

　　A distinct advantage of the SWIR band over the TIR band is its almost uniform sensitivity down to Earth's surface. On the other hand, the CO absorption in the TIR is significantly stronger, i.e., the vertical optical depth τ_{CO} of the CO fundamental band around 4.6 μm is about a factor of 100 larger than the optical depth of the first overtone band around 2.3 μm (see, e.g., the Atmospheric Infrared Spectrum Atlas, http://eodg.atm.ox.ac.uk/ATLAS/). Moreover, the methane optical depth is larger than the carbon monoxide optical depth in the SWIR range, whereas $\tau_{CH_4} < \tau_{CO}$ in the TIR range, at around 2100 cm^{-1}.

　　The accuracy of CO columns estimated from TIR observations is typically 10% (see e.g, [41,50]), i.e., much better than for SCIAMACHY: de Laat et al. [51], and Kopacz et al. [52] found that the random instrument noise related error of a single CO total column measurement is large, typically 10–100% or larger [9], and Tangborn et al. [53] stated that total error values generally range from 20 to 100% for the observations tagged as 'good'. The smaller noise related error of TIR retrievals has been confirmed by the analysis of AIRS observations using Column EstimatoR Vertical Infrared Sounding of the Atmosphere (CERVISA), a variant of the BIRRA based on Schwarzschild's equation appropriate for the thermal infrared range [54].

5. Conclusions

　　Dry air CO total column estimates from 2.3 μm reflectance measurements using SCIAMACHY from 2003–2011 have been validated against 18 stations of the g-b NDACC and TCCON. While observations from the NDACC covered the entire SCIAMACHY mission period, the TCCON observations could only be used to validate the CO product in the later phase of the mission. It was found that after postprocessing of the BIRRA retrieval output, the actual xCO noise still varied significantly between sites, ranging from around 100 ppbv up to 200 ppbv. This was similar to findings from other authors cited in Section 1.5.2 and therefore, averaging individual s-b CO observations was an essential task in order to validate the product. Here, a distance-weighted approach was chosen, to our knowledge for the first time, to compare the CO mixing ratios from spatially distributed satellite observations to point-like g-b measurements. The approach was demonstrated reasonable in validating CO from s-b sensors. The results also suggest that part of the discrepancy between the BIRRA-retrieved columns and g-b observations arises from representation errors, inadequate description of the a priori CO profile, and cloud shielding. In an upcoming study we aim to use model data as a reference in

addition to g-b data and examine the approach of weighted averages for other s-b platforms such as GOSAT-2 and the Sentinel-5 Precursor (S5P) [55].

Abbreviations

The following abbreviations are used in this manuscript:

AIRS	Atmospheric Infrared Sounder
BIRRA	Beer InfraRed Retrieval Algorithm
CH4	Methane
CMR	Column mixing ratio
CO	Carbon monoxide
CrIS	Cross-track Infrared Sounder
DOAS	Differential Optical Absorption Spectroscopy
ENVISAT	Environmental Satellite
FTIR	Fourier transform infrared
FTS	Fourier transform spectroscopy
G-B	Ground-based
GOMOS	Global Ozone Monitoring by Occultation of Stars
IASI	Infrared Atmospheric Sounding Interferometer
IMAP	Iterative Maximum A Posteriori
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MOPITT	Measurement of Pollution in the Troposphere
NDACC	Network for the Detection of Atmospheric Composition Change
O2	Oxygen
PPBV	Parts per billion in volume
S-B	Satellite-based
S5P	Sentinel-5 Precursor
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
SICOR	Shortwave infrared CO retrieval
SWIR	Short-wave infrared
TCCON	Total Carbon Column Observing Network
TES	Tropospheric emission spectrometer
TIR	Thermal infrared
TROPOMI	Tropospheric Monitoring Instrument
VMR	Volume mixing ratio
WFM	Weighting function modified
XCO	Dry-air average column mixing ratio

Appendix A. NDACC Data Providers

The NDACC data in this publication were obtained from sites listed in Table A1. The data are publicly available (see http://www.ndacc.org).

Table A1. NDACC sites used in this publication and the coorperating institutions.

NDACC Site	Coorperating Institutions
Bremen	Institute of Environmental Physics, University of Bremen, Germany
Garmisch	Institute for Meteorology and Climate Research, Karlsruhe Institute of Technology, Gemany
Izana	Izana Atmospheric Research Center, AEMET- Meteorological State Agency, Spain
Jungfraujoch	Universtiy of Liége, Belgium
Kiruna	Institute of Space Physics, Sweden
Ny-Alesund	Institute of Environmental Physics, University of Bremen, Germany
	Alfred Wegener Institute for Polar and Marine Research, Germany
Thule	Danish Climate Center, Danish Meteorological Institute Copenhagen, Denmark
Toronto	Department of Physics, University of Toronto, Canada

Appendix B. TCCON Data Providers

The TCCON data were obtained from sites listed in Table A2. The TCCON Data Archive is hosted by CaltechDATA, California Institute of Technology, CA (US), doi:https://doi.org/10.14291/tccon.archive/1348407.

Table A2. TCCON sites with references to the data used in this publication.

TCCON Site	Reference
Bialystok	Deutscher *et al.* [56]
Bremen	Notholt *et al.* [57]
Darwin	Griffith *et al.* [58]
Garmisch	Sussmann and Rettinger [59]
Izana	Blumenstock *et al.* [60]
Lamont	Wennberg *et al.* [61]
Lauder	Sherlock *et al.* [62]
Ny-Alesund	Notholt *et al.* [63]
Park-Falls	Wennberg *et al.* [64]
Wollongong	Griffith *et al.* [65]

References

1. ESA. ESA Declares End of Mission for ENVISAT. 2012. Available online: https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/envisat (accessed on 15 January 2018).
2. Gottwald, M.; Bovensmann, H. (Eds.) *SCIAMACHY—Exploring the Changing Earth's Atmosphere*; Springer: Dordrecht, The Netherlands, 2011.
3. Gloudemans, A.; Schrijver, H.; Kleipool, Q.; van den Broek, M.; Straume, A.; Lichtenberg, G.; van Hees, R.; Aben, I.; Meirink, J. The impact of SCIAMACHY near-infrared instrument calibration on CH_4 and CO total columns. *Atmos. Chem. Phys.* **2005**, *5*, 2369–2383.
4. Lichtenberg, G.; Gimeno Garcia, S.; Schreier, F.; Slijkhuis, S.; Snel, R.; Bovensmann, H. Impact of Level 1 Quality on SCIAMACHY Level 2 Retrieval. In Proceedings of the 38th COSPAR Scientific Assembly, Bremen, Germany, 18–25 July 2010.
5. Buchwitz, M.; de Beek, R.; Bramstedt, K.; Noël, S.; Bovensmann, H.; Burrows, J.P. Global carbon monoxide as retrieved from SCIAMACHY by WFM-DOAS. *Atmos. Chem. Phys.* **2004**, *4*, 1945–1960.
6. Buchwitz, M.; de Beek, R.; Noël, S.; Burrows, J.P.; Bovensmann, H.; Bremer, H.; Bergamaschi, P.; Körner, S.; Heimann, M. Carbon monoxide, methane and carbon dioxide columns retrieved from SCIAMACHY by WFM-DOAS: year 2003 initial data set. *Atmos. Chem. Phys.* **2005**, *5*, 3313–3329.
7. Frankenberg, C.; Platt, U.; Wagner, T. Retrieval of CO from SCIAMACHY onboard ENVISAT: Detection of strongly polluted areas and seasonal patterns in global CO abundances. *Atmos. Chem. Phys.* **2005**, *5*, 1639–1644.
8. Gimeno García, S.; Schreier, F.; Lichtenberg, G.; Slijkhuis, S. Near infrared nadir retrieval of vertical column densities: methodology and application to SCIAMACHY. *Atmos. Meas. Tech.* **2011**, *4*, 2633–2657.

9. Borsdorff, T.; Tol, P.; Williams, J.E.; de Laat, J.; aan de Brugh, J.; Nédélec, P.; Aben, I.; Landgraf, J. Carbon monoxide total columns from SCIAMACHY 2.3 μm atmospheric reflectance measurements: Towards a full-mission data product (2003–2012). *Atmos. Meas. Tech.* **2016**, *9*, 227–248.

10. Borsdorff, T.; aan de Brugh, J.; Hu, H.; Nédélec, P.; Aben, I.; Landgraf, J. Carbon monoxide column retrieval for clear-sky and cloudy atmospheres: a full-mission data set from SCIAMACHY 2.3 μm reflectance measurements. *Atmos. Meas. Tech.* **2017**, *10*, 1769–1782.

11. Clerbaux, C.; Boynard, A.; Clarisse, L.; George, M.; Hadji-Lazaro, J.; Herbin, H.; Hurtmans, D.; Pommier, M.; Razavi, A.; Turquety, S.; et al. Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmos. Chem. Phys.* **2009**, *9*, 6041–6054.

12. Fischer, H.; Birk, M.; Blom, C.; Carli, B.; Carlotti, M.; von Clarmann, T.; Delbouille, L.; Dudhia, A.; Ehhalt, D.; Endemann, M.; et al. MIPAS: an instrument for atmospheric and climate research. *Atmos. Chem. Phys.* **2008**, *8*, 2151–2188.

13. Bernath, P.; McElroy, C.T.; Abrams, M.C.; Boone, C.D.; Butler, M.; Camy-Peyret, C.; Carleer, M.; Clerbaux, C.; Coheur, P.F.; Colin, R.; et al. Atmospheric Chemistry Experiment (ACE): Mission overview. *Geophys. Res. Lett.* **2005**, *32*, L15S01, doi:10.1029/2005GL022386.

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

15. Wunch, D.; Toon, G.C.; Blavier, J.F.L.; Washenfelder, R.A.; Notholt, J.; Connor, B.J.; Griffith, D.W.T.; Sherlock, V.; Wennberg, P.O. The Total Carbon Column Observing Network. *Philos. Trans. R. Soc. A* **2011**, *369*, 2087–2112.

16. Feist, D. Global FTS-Network. 2015. Available online: https://oc.bgc-jena.mpg.de (accessed on 20 June 2016).

17. Warnek, T.; Notholt, J.; Chen, H. *Validation and Calibration of Greenhouse-Gas Satellite Observations by Ground-Based Remote Sensing Measurements*; University of Bremen: Bibliothekstraße, Bremen, 2015.

18. TCCON. Total Carbon Column Observing Network—Wiki. 2018. Available online: https://tccon-wiki.caltech.edu/Network_Policy/Data_Use_Policy/Data_Description#Measurement_precision_(repeatability) (accessed on 11 January 2018).

19. Einarsson, B. (Ed.) *Accuracy and Reliability in Scientific Computing*; SIAM: Philadelphia, PA, USA, 2005.

20. Verhoelst, T.; Granville, J.; Hendrick, F.; Köhler, U.; Lerot, C.; Pommereau, J.P.; Redondas, A.; Van Roozendael, M.; Lambert, J.C. Metrology of ground-based satellite validation: co-location mismatch and smoothing issues of total ozone comparisons. *Atmos. Meas. Tech.* **2015**, *8*, 5039–5062.

21. De Laat, A.T.J.; Gloudemans, A.M.S.; Schrijver, H.; Aben, I.; Nagahama, Y.; Suzuki, K.; Mahieu, E.; Jones, N.B.; Paton-Walsh, C.; Deutscher, N.M.; et al. Validation of five years (2003–2007) of SCIAMACHY CO total column measurements using ground-based spectrometer observations. *Atmos. Meas. Tech.* **2010**, *3*, 1457–1471.

22. De Laat, A.; Gloudemans, A.; Schrijver, H.; van den Broek, M.; Meirink, J.; Aben, I.; Krol, M. Quantitative analysis of SCIAMACHY carbon monoxide total column measurements. *Geophys. Res. Lett.* **2006**, *33*, L07807.

23. Sussmann, R.; Buchwitz, M. Initial validation of ENVISAT/SCIAMACHY columnar CO by FTIR profile retrievals at the Ground-Truthing Station Zugspitze. *Atmos. Chem. Phys.* **2005**, *5*, 1497–1503.

24. Sussmann, R.; Stremme, W.; Buchwitz, M.; de Beek, R. Validation of ENVISAT/SCIAMACHY columnar methane by solar FTIR spectrometry at the Ground-Truthing Station Zugspitze. *Atmos. Chem. Phys.* **2005**, *5*, 2419–2429.

25. Dils, B.; De Maziere, M.; Muller, J.; Blumenstock, T.; Buchwitz, M.; de Beek, R.; Demoulin, P.; Duchatelet, P.; Fast, H.; Frankenberg, C.; et al. Comparisons between SCIAMACHY and ground-based FTIR data for total columns of CO, CH_4, CO_2 and N_2O. *Atmos. Chem. Phys.* **2006**, *6*, 1953–1976.

26. Schneising, O.; Bergamaschi, P.; Bovensmann, H.; Buchwitz, M.; Burrows, J.P.; Deutscher, N.M.; Griffith, D.W.T.; Heymann, J.; Macatangay, R.; Messerschmidt, J.; et al. Atmospheric greenhouse gases retrieved from SCIAMACHY: Comparison to ground-based FTS measurements and model results. *Atmos. Chem. Phys.* **2012**, *12*, 1527–1540.

27. Meier, A. *Spectroscopic Atlas of Atmospheric Microwindows in the Middle Infrared*; IRF Technical Report; IRF Institute för Rymdfysik: Kiruna, Sweden, 2004.

28. Kiel, M.; Wunch, D.; Wennberg, P.O.; Toon, G.C.; Hase, F.; Blumenstock, T. Improved retrieval of gas abundances from near-infrared solar FTIR spectra measured at the Karlsruhe TCCON station. *Atmos. Meas. Tech.* **2016**, *9*, 669–682.

29. Wunch, D.; Toon, G.C.; Sherlock, V.; Deutscher, N.M.; Liu, C.; Feist, D.G.; Wennberg, P.O. *The Total Carbon Column Observing Network's GGG2014 Data Version*; Technical Report; Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2015.

30. Schreier, F.; Gimeno García, S.; Hedelt, P.; Hess, M.; Mendrok, J.; Vasquez, M.; Xu, J. GARLIC—A General Purpose Atmospheric Radiative Transfer Line-by-Line Infrared-Microwave Code: Implementation and Evaluation. *J. Quant. Spectrosc. Radiat. Transf.* **2014**, *137*, 29–50.

31. Schreier, F.; Gimeno García, S.; Vasquez, M.; Xu, J. Algorithmic vs. finite difference Jacobians for infrared atmospheric radiative transfer. *J. Quant. Spectrosc. Radiat. Transf.* **2015**, *164*, 147–160.

32. Dennis,, J., Jr.; Gay, D.; Welsch, R. Algorithm 573: NL2SOL—An Adaptive Nonlinear Least–Squares Algorithm. *ACM Trans. Math. Softw.* **1981**, *7*, 369–383.

33. Golub, G.; Pereyra, V. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Probl.* **2003**, *19*, R1–R26.

34. Anderson, G.; Clough, S.; Kneizys, F.; Chetwynd, J.; Shettle, E. *AFGL Atmospheric Constituent Profiles (0–120 km)*; Technical Report TR-86-0110; AFGL: Hanscom AFB, MA, USA, 1986.

35. Gloudemans, A.; de Laat, A.; Schrijver, H.; Aben, I.; Meirink, J.; van der Werf, G. SCIAMACHY CO over land and oceans: 2003–2007 interannual variability. *Atmos. Chem. Phys.* **2009**, *9*, 3799–3813.

36. Buschmann, M.; Deutscher, N.M.; Sherlock, V.; Palm, M.; Warneke, T.; Notholt, J. Retrieval of xCO$_2$ from ground-based mid-infrared (NDACC) solar absorption spectra and comparison to TCCON. *Atmos. Meas. Tech.* **2016**, *9*, 577–585.

37. Ostler, A.; Sussmann, R.; Rettinger, M.; Deutscher, N.M.; Dohe, S.; Hase, F.; Jones, N.; Palm, M.; Sinnhuber, B.M. Multistation intercomparison of column-averaged methane from NDACC and TCCON: Impact of dynamical variability. *Atmos. Meas. Tech.* **2014**, *7*, 4081–4101.

38. Filipiak, M.; Harwood, R.; Jiang, J.; Li, Q.; Livesey, N.; Manney, G.; Read, W.; Schwartz, M.; Waters, J.; Wu, D. Carbon Monoxide Measured by the EOS Microwave Limb Sounder on Aura: First Results. *Geophys. Res. Lett.* **2005**, *32*, L14825.

39. Pumphrey, H.; Filipiak, M.; Livesey, N.; Schwartz, M.; Boone, C.; Walker, K.; Bernath, P.; Ricaud, P.; Barret, B.; Clerbaux, C.; et al. Validation of middle-atmosphere carbon monoxide retrievals from the Microwave Limb Sounder on Aura. *J. Geophys. Res.* **2007**, *112*, D24S38.

40. Dupuy, E.; Urban, J.; Ricaud, P.; Le Flochmoën, E.; Lautié, N.; Murtagh, D.; De La Noë, J.; El Amraoui, L.; Eriksson, P.; Forkman, P.; et al. Strato-mesospheric measurements of carbon monoxide with the Odin Sub-Millimetre Radiometer: Retrieval and first results. *Geophys. Res. Lett.* **2004**, *31*, L20101.

41. McMillan, W.W.; Barnet, C.; Strow, L.; Chahine, M.T.; McCourt, M.L.; Warner, J.X.; Novelli, P.C.; Korontzi, S.; Maddy, E.S.; Datta, S. Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. *Geophys. Res. Lett.* **2005**, *32*, L11801.

42. Gambacorta, A.; Barnet, C.; Wolf, W.; King, T.; Maddy, E.; Strow, L.; Xiong, X.; Nalli, N.; Goldberg, M. An Experiment Using High Spectral Resolution CrIS Measurements for Atmospheric Trace Gases: Carbon Monoxide Retrieval Impact Study. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1639–1643.

43. Fu, D.; Bowman, K.W.; Worden, H.M.; Natraj, V.; Worden, J.R.; Yu, S.; Veefkind, P.; Aben, I.; Landgraf, J.; Strow, L.; et al. High-resolution tropospheric carbon monoxide profiles retrieved from CrIS and TROPOMI. *Atmos. Meas. Tech.* **2016**, *9*, 2567–2579.

44. George, M.; Clerbaux, C.; Hurtmans, D.; Turquety, S.; Coheur, P.F.; Pommier, M.; Hadji-Lazaro, J.; Edwards, D.P.; Worden, H.; Luo, M.; et al. Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors. *Atmos. Chem. Phys.* **2009**, *9*, 8317–8330.

45. Fortems-Cheiney, A.; Chevallier, F.; Pison, I.; Bousquet, P.; Carouge, C.; Clerbaux, C.; Coheur, P.F.; George, M.; Hurtmans, D.; Szopa, S. On the capability of IASI measurements to inform about CO surface emissions. *Atmos. Chem. Phys.* **2009**, *9*, 8735–8743.

46. Illingworth, S.; Remedios, J.; Boesch, H.; Moore, D.; Sembhi, H.; Dudhia, A.; Walker, J. ULIRS, an optimal estimation retrieval scheme for carbon monoxide using IASI spectral radiances: sensitivity analysis, error budget and simulations. *Atmos. Meas. Tech.* **2011**, *4*, 269–288.

47. Deeter, M.N.; Emmons, L.K.; Francis, G.L.; Edwards, D.P.; Gille, J.C.; Warner, J.X.; Khattatov, B.; Ziskin, D.; Lamarque, J.F.; Ho, S.P.; et al. Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument. *J. Geophys. Res.* **2003**, *108*, doi:10.1029/2002JD003186.

48. Rinsland, C.; Luo, M.; Logan, J.; Beer, R.; Worden, H.; Worden, J.; Bowman, K.; Kulawik, S.; Rider, D.; Osterman, G.; et al. Nadir Measurements of Carbon Monoxide Distributions by the Tropospheric Emission Spectrometer onboard the Aura Spacecraft: Overview of Analysis Approach and Examples of Initial Results. *Geophys. Res. Lett.* **2006**, *33*, L22806.

49. Worden, H.; Deeter, M.; Edwards, D.; Gille, J.; Drummond, J.; Nédélec, P. Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals. *J. Geophys. Res.* **2010**, *115*, D16314.

50. Emmons, L.; Pfister, G.; Edwards, D.; Gille, J.; Sachse, G.; Blake, D.; Wofsy, S.; Gerbig, C.; Matross, D.; Nedelec, P. Measurements of Pollution in the Troposphere (MOPITT) validation exercises during summer 2004 field campaigns over North America. *J. Geophys. Res.* **2007**, *112*, D12S02.

51. De Laat, A.; Gloudemans, A.; Aben, I.; Krol, M.; Meirink, J.; van der Werf, G.; Schrijver, H. Scanning Imaging Absorption Spectrometer for Atmospheric Chartography carbon monoxide total columns: Statistical evaluation and comparison with chemistry transport model results. *J. Geophys. Res.* **2007**, *112*, D12310.

52. Kopacz, M.; Jacob, D.J.; Fisher, J.A.; Logan, J.A.; Zhang, L.; Megretskaia, I.A.; Yantosca, R.M.; Singh, K.; Henze, D.K.; Burrows, J.P.; et al. Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES). *Atmos. Chem. Phys.* **2010**, *10*, 855–876.

53. Tangborn, A.; Stajner, I.; Buchwitz, M.; Khlystova, I.; Pawson, S.; Burrows, J.; Hudman, R.; Nedelec, P. Assimilation of SCIAMACHY total column CO observations: Global and regional analysis of data impact. *J. Geophys. Res.* **2009**, *114*, D07307.

54. Schreier, F.; Gimeno García, S.; Lichtenberg, G.; Hoffmann, P. Intercomparison of Carbon Monoxide Retrievals from SCIAMACHY and AIRS Nadir Observations. In Proceedings of the Atmospheric Science Conference, Barcelona, Spain, 7–11 September 2009; Volume SP-676.

55. Landgraf, J.; aan de Brugh, J.; Scheepmaker, R.; Borsdorff, T.; Hu, H.; Houweling, S.; Butz, A.; Aben, I.; Hasekamp, O. Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements. *Atmos. Meas. Tech.* **2016**, *9*, 4955–4975.

56. Deutscher, N.M.; Notholt, J.; Messerschmidt, J.; Weinzierl, C.; Warneke, T.; Petri, C.; Grupe, P.; Katrynski, K. TCCON data from Bialystok (PL), Release GGG2014.R1. TCCON Data Archive, hosted by CaltechDATA, 2015. doi:https://doi.org/10.14291/tccon.ggg2014.bialystok01.R1/1183984.

57. Notholt, J.; Petri, C.; Warneke, T.; Deutscher, N.M.; Buschmann, M.; Weinzierl, C.; Macatangay, R.C.; Grupe, P. TCCON data from Bremen (DE), Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA, 2014. doi:https://doi.org/10.14291/tccon.ggg2014.bremen01.R0/1149275.

58. Griffith, D.W.T.; Deutscher, N.M.; Velazco, V.A.; Wennberg, P.O.; Yavin, Y.; Keppel-Aleks, G.; Washenfelder, R.; Toon, G.C.; Blavier, J.F.; Paton-Walsh, C.; Jones, N.B.; Kettlewell, G.C.; Connor, B.; Macatangay, R.C.; Roehl, C.; Ryczek, M.; Glowacki, J.; Culgan, T.; Bryant, G. TCCON data from Darwin (AU), Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA, 2014. doi:https://doi.org/10.14291/tccon.ggg2014.darwin01.R0/1149290.

59. Sussmann, R.; Rettinger, M. TCCON data from Garmisch (DE), Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA, 2014. doi:https://doi.org/10.14291/tccon.ggg2014.garmisch01.R0/1149299.

60. Blumenstock, T.; Hase, F.; Schneider, M.; Garcia, O.E.; Sepulveda, E. TCCON data from Izana (ES), Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA, 2014. doi:https://doi.org/10.14291/tccon.ggg2014.izana01.R0/1149295.

61. Wennberg, P.O.; Wunch, D.; Roehl, C.; Blavier, J.F.; Toon, G.C.; Allen, N. TCCON data from Lamont (US), Release GGG2014.R1. TCCON Data Archive, hosted by CaltechDATA, 2016. doi:https://doi.org/10.14291/tccon.ggg2014.lamont01.R1/1255070.

62. Sherlock, V.; Connor, B.; Robinson, J.; Shiona, H.; Smale, D.; Pollard, D. TCCON data from Lauder (NZ), 120HR, Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA, 2014. doi:https://doi.org/10.14291/tccon.ggg2014.lauder01.R0/1149293.

63. Notholt, J.; Warneke, T.; Petri, C.; Deutscher, N.M.; Weinzierl, C.; Palm, M.; Buschmann, M. TCCON data from Ny Ålesund, Spitsbergen (NO), Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA, 2017. doi:https://doi.org/10.14291/tccon.ggg2014.nyalesund01.R0.

64. Wennberg, P.O.; Roehl, C.; Wunch, D.; Toon, G.C.; Blavier, J.F.; Washenfelder, R.; Keppel-Aleks, G.; Allen, N.; Ayers, J. TCCON data from Park Falls (US), Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA, 2014. doi:https://doi.org/10.14291/tccon.ggg2014.parkfalls01.R0/1149161.

65. Griffith, D.W.T.; Velazco, V.A.; Deutscher, N.M.; Paton-Walsh, C.; Jones, N.B.; Wilson, S.R.; Macatangay, R.C.; Kettlewell, G.C.; Buchholz, R.R.; Riggenbach, M. TCCON data from Wollongong (AU), Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA, 2014. doi:https://doi.org/10.14291/tccon.ggg2014.wollongong01.R0/1149291.