XCO2 estimates from the OCO-2 measurements using a neural network approach

Leslie David, Francois-Marie Breon, Frederic Chevallier

Atmos. Meas. Tech., 14, 117–132, 2021
https://doi.org/10.5194/amt-14-117-2021

Atmospheric Measurement Techniques
Open Access
EGU

XCO_2 estimates from the OCO-2 measurements using a neural network approach

Leslie David, François-Marie Bréon, and Frédéric Chevallier

Laboratoire des Sciences du Climat et de l'Environnement/IPSL, CEA-CNRS-UVSQ,
Université Paris-Saclay, 91198 Gif-sur-Yvette, France

Correspondence: François-Marie Bréon (fmbreon@cea.fr)

Received: 5 May 2020 – Discussion started: 2 June 2020
Revised: 14 November 2020 – Accepted: 16 November 2020 – Published: 7 January 2021

Abstract. The Orbiting Carbon Observatory (OCO-2) instrument measures high-resolution spectra of the sun's radiance reflected at the earth's surface or scattered in the atmosphere. These spectra are used to estimate the column-averaged dry air mole fraction of CO_2 (XCO_2) and the surface pressure. The official retrieval algorithm (NASA's Atmospheric CO_2 Observations from Space retrievals, ACOS) is a *full-physics algorithm* and has been extensively evaluated. Here we propose an alternative approach based on an artificial neural network (NN) technique. For training and evaluation, we use as reference estimates (i) the surface pressures from a numerical weather model and (ii) the XCO_2 derived from an atmospheric transport simulation constrained by surface air-sample measurements of CO_2. The NN is trained here using real measurements acquired in nadir mode on cloud-free scenes during even-numbered months and is then evaluated against similar observations during odd-numbered months. The evaluation indicates that the NN retrieves the surface pressure with a root-mean-square error better than 3 hPa and XCO_2 with a 1σ precision of 0.8 ppm. The statistics indicate that the NN trained with a representative set of data allows excellent accuracy that is slightly better than that of the full-physics algorithm. An evaluation against reference spectrophotometer XCO_2 retrievals indicates similar accuracy for the NN and ACOS estimates, with a skill that varies among the various stations. The NN–model differences show spatiotemporal structures that indicate a potential for improving our knowledge of CO_2 fluxes. We finally discuss the pros and cons of using this NN approach for the processing of the data from OCO-2 or other space missions.

1 Introduction

During the past decades, natural fluxes have absorbed about half of the anthropogenic emissions of CO_2 (Knorr, 2009), but there is large uncertainty about the spatial distribution of this sink over time and therefore on the processes that control it. A growing network of high-precision atmospheric CO_2 measurements has been used together with meteorological information to constrain the sources and sinks of CO_2 using a technique known as *atmospheric inversion* (e.g., Peylin et al., 2013), but the lack of data in large regions of the globe, such as the tropics, does not allow the monitoring of these fluxes with enough space–time resolution. Early attempts to complement this network with satellite retrievals from sensors that were not specifically designed for this purpose were not successful (Chevallier et al., 2005), but a series of dedicated instruments were put into orbit when the Greenhouse Gases Observing Satellite (GOSAT, Yokota et al., 2009) and the second Orbiting Carbon Observatory (OCO-2 Eldering et al., 2017) were launched in 2009 and 2014, respectively. These were still operational at the time of writing. This new and evolving constellation is directly supported by Japanese, US, Chinese, and European space agencies (CEOS Atmospheric Composition Virtual Constellation Greenhouse Gas Team, 2018). All missions have adopted the same CO_2 observation principle that consists of measuring the solar irradiance reflected at the earth's surface in selected spectral bands. Along the double atmospheric path (down-going and up-going), the sunlight is absorbed by atmospheric molecules at specific wavelengths. The resulting absorption lines on the measured spectra make it possible to estimate the amount of gas between the surface and the top of the atmosphere. CO_2 shows many such absorption lines around 1.61 and 2.06 µm

that are used to estimate the CO_2 column. Similarly, the oxygen lines around 0.76 µm are used to estimate the surface pressure and can also be used to infer the sunlight atmospheric path, leading to the column-averaged dry air mole fraction of CO_2, referred to as XCO_2 (O'Brien and Rayner, 2002; Crisp et al., 2004).

One main difficulty in the retrieval of XCO_2 from the measured spectra results from the presence of atmospheric particles that scatter light and change its atmospheric path. Accounting for aerosols, in particular, is challenging because aerosols are highly variable in amount and in vertical distribution. Another major difficulty results from modeling errors. The radiative transfer models that are used for the retrieval leave significant residuals between the measured and modeled spectra, even after the XCO_2 and aerosol amount have been inverted for a best fit (Crisp et al., 2012; O'Dell et al., 2018). As a consequence of the various uncertainties in the retrieval process, raw XCO_2 retrievals show significant biases against reference ground-based retrievals (Wunch et al., 2011b, 2017). These biases, together with the comparison against modeling results, led to the development of empirical corrections to the retrieved XCO_2. In the case of the OCO-2 V8r retrievals generated by NASA's Atmospheric CO_2 Observations from Space (ACOS), these corrections amount to roughly half that of the "signal", i.e., of the difference between the prior and the retrieved XCO_2 (O'Dell et al., 2018).

The limitations in the full-physics retrieval method, despite considerable effort and progress (e.g., O'Dell et al., 2018; Reuter et al., 2017; and Wu et al., 2018 in the case of OCO-2), encourage developing alternative approaches. Here, we want to re-evaluate the potential of an artificial neural network technique (NN) to estimate XCO_2 from the measured spectra. A NN-based technique was already used by Chédin et al. (2003) for a fast retrieval of midtropospheric mean CO_2 concentrations from some meteorological satellite radiometers. These authors trained their NNs on a large ensemble of radiance simulations made using a reference radiation model and assuming diverse atmospheric and surface conditions. NN-based approaches are also commonly used for the retrieval of other species from various high-spectral-resolution satellite radiance measurements because of their computational efficiency (e.g., Hadji-Lazaro et al., 1999).

A NN approach requires a large and representative training dataset. A standard method for problems similar to that discussed here is to use a radiative transfer model and to generate a large ensemble of pseudo-observations based on assumed atmospheric and surface parameters. However, as mentioned above, the radiative transfer models have deficiencies that are rather small, but nevertheless significant with respect to the high-precision objective of the CO_2 measurements. In addition, there may be some wrong assumptions and unknown instrumental defects that are not accounted for in the forward modeling. We thus prefer to avoid using such radiative transfer models and rather base the training on a fully empirical approach (see, e.g., Aires et al., 2005). We use real OCO-2 observations together with collocated estimates of the surface pressure and XCO_2. The retrievals from the NN approach are evaluated against model estimates of surface pressure and XCO_2, as well as observations from the Total Carbon Column Observing Network (TCCON, Wunch et al., 2011a). In the following, Sect. 2 presents the approach while Sect. 3 describes the results. Section 4 discusses the results and the way forward.

2 Data and method

Our NN estimates XCO_2 and the surface pressure from nadir spectra measured by the OCO-2 satellite over land. OCO-2 has eight cross-track footprints (e.g., Eldering et al., 2015), but we only use footprint #4 in the following for simplicity. If successful, the same approach can be applied to all footprints. The focus on nadir measurements here is motivated by the complication introduced by the Doppler effect in glint mode, which is the other pointing mode for OCO-2 routine science operations: the absorption lines affect pixel elements that vary among the spectra. These variations of the position of the absorption line may cause additional difficulty to the NN training. The solar lines in the nadir spectra are also affected by Doppler shifts due to the motion of the earth and satellite relative to the sun, but this concerns a limited set of spectral elements that are affected by the solar (Fraunhofer) lines. The development of a glint-mode NN is therefore left for a future study.

We use spectral samples in the three bands of the instrument (around 0.76, 1.61 and 2.06 µm). They have footprints of $\sim 3\,km^2$ on the ground. In principle, each band is described by 1016 pixel elements, but some are marked as bad either because some of the corresponding detectors have died or because of known temporary or permanent issues. We systematically remove 15 pixel elements that are flagged in about 80 % of the spectra and 478 pixels in the band edges. Conversely, we do not remove the spectra that are affected by the deep solar lines and we let the NN handle these specific features. Because the information in the spectrum is mostly in the relative depth of the absorption lines, and not in their overall amplitude, we normalize each spectrum by a radiance that is representative of the offline values (i.e., the mean of the 90 %–95 % range for each spectrum). This essentially removes the impact of the variations in the surface albedo and in the sun irradiance linked to the solar zenith angle. Other input choices may be attempted in the future.

As input to the NN, we add the observation geometry (sun zenith angle and relative azimuth). The sun zenith angle drives the atmospheric pathlength and is then required for the interpretation of the absorption line depth in terms of atmospheric optical depth. The azimuth was not included in our first attempts, but when later included it led to a significant improvement in the results. Although the NN technique does not allow for a clear physical interpretation, we assume

that the information brought by the relative azimuth is linked to the polarization of the molecular scattering contribution to the measurements that varies with the azimuth.

The NN exploits these 2557 input variables to compute two variables only: XCO_2 and the surface pressure. It is structured as a multilayer perceptron (Rumelhart et al., 1988) with one hidden layer of 500 neurons that use a sigmoid activation function. The number of hidden layers is somewhat arbitrary and based on a limited sample of trials. Lower quality estimates were obtained with 50 neurons whereas the training time increased markedly for 1000 neurons and more. The weights of the input variables to the hidden neurons and the weights of the hidden variables to the output variables are adjusted iteratively with the standard Keras library (Keras Team, 2015). Figure A1 in the Appendix illustrates the convergence process. The NN cost function (a.k.a. loss) becomes fairly constant for a test dataset after about 100 iterations, whereas it continues to decrease for the training dataset, indicating an overfitting of the data. The iteration is stopped when there is no decrease of the test loss for 50 iterations. There is a factor of 3 to 4 between the loss of the training dataset and that of the test, which confirms the overfitting of the former.

Note that the NN estimate does not use any a priori information on surface pressure or the CO_2 profile after the training is done. Also, no explicit information is provided on the altitude, location, or time period of the observation. The NN estimates are therefore only driven by the OCO-2 spectrum measurements, together with the observation geometry (sun zenith and relative azimuth). The observation geometry varies with the latitude and the season so that the NN may infer some location information from this input. Conversely, it is the same from one year to the next and, at a given date, for all longitudes. Thus, there is no information on the longitude or the year of observation in the geometry parameters that are provided to the network.

The NN training is based on OCO-2 radiance measurements (V8r) acquired during even-numbered months between January 2015 and August 2018. The 4-year period allows varying the global background CO_2 dry air mole fraction by $\sim 2\,\%$, as much as typical XCO_2 seasonal variations in the northern extratropics (see, e.g., Fig. 1 of Agustí-Panareda et al., 2019). Our evaluation dataset is based on observations during the odd-numbered months of the same period. In both cases, we make use of XCO_2 estimates and the quality control filters of the ACOS L2Lite V9r products: only observations with *xco2_quality_flag* $= 0$ are used. We also consider the *warn level*, *outcome flag* and *cloud_ flag_idp* that are provided in the V8r L2lite and L2Std files. For NN training, we only use the best quality observations, i.e., those with a *warn level* lower or equal to 2, a *cloud_flag* of 3 (very clear) and an *outcome flag* of 1. This choice is based on the evaluation of the surface pressure estimates described below (Fig. 3). This distinction leads to about 131 000 observations for the training. For the evaluation of the NN estimates, we use less restrictive criteria and accept observations with outcome_flag of either 1 or 2, and cloud_flag of 2 or 3. These choices are justified below. The spatial distribution of the observations that are used for the training is shown in Fig. A2 of the Appendix. The training dataset covers most regions of the globe with the exception of South America. The underrepresentation of this subcontinent stems from both the high cloudiness and impact of cosmic rays that leads to missing pixel elements (see below).

For the reference surface pressure (training and evaluation), an obvious choice is the use of numerical weather analyses corrected for the sounding altitude. Indeed, the typical accuracy for surface pressure data is on the order of 1 hPa (Salstein et al., 2008). For convenience, we use the surface pressure that is provided together with the OCO-2 data and is derived from the Goddard Earth Observing System, Version 5, Forward Processing for Instrument Teams (GEOS5-FP-IT) created at the Goddard Space Flight Center Global Modeling and Assimilation Office (Suarez et al., 2008; and Lucchesi et al., 2013). There is no such obvious choice for XCO_2 as there is no global-scale highly accurate dataset of XCO_2 and we thus rely here on best estimates from a modeling approach. We use the CO_2 atmospheric inversion of the Copernicus Atmosphere Monitoring Service (CAMS, http://atmosphere.copernicus.eu, last access: 28 January 2020, Chevallier et al., 2010; version 18r2). This product was released in July 2019 and contributed, e.g., to the Global Carbon Budget 2019 of Friedlingstein et al. (2019). It results from the assimilation of CO_2 surface air-sample measurements in a global atmospheric transport model run at spatial resolution 1.90° in latitude and 3.75° in longitude over the period from 1979–2018 and using the adjoint of this transport model. Neither satellite retrievals nor TCCON observations were used for this modeling. For each OCO-2 observation, XCO_2 is computed from the collocated concentration vertical profile, through a simple integration weighted by the pressure width of the model layers. Note that the model layers use "dry" pressure coordinates so that there is no need for a water vapor correction in the vertical integration. The GEOS5-FP-IT surface pressure and the XCO_2 from CAMS are used both for the training and the evaluation, although using independent datasets (odd- and even-numbered months).

Many measured spectra lack one or several spectral pixels. This is particularly the case over South America, as a consequence of the South Atlantic cosmic ray flux anomaly that impacts the OCO-2 detector in this region. We therefore devised a method to interpolate the spectra and to fill the missing pixels. Our method first sorts all spectral pixels as a function of the measured radiance in a large number of complete measured spectra. The pixel ranks are averaged to generate a rank representative of the full dataset. Then, when a pixel element is missing in a spectrum, we look for its typical rank and we average the radiances of the two pixel elements that have the ranks just above and below. The procedure is applied even when several pixel elements are missing in a spectrum,

except when these are successive in the typical ranking. The procedure described here fills the missing elements and the NN can then be applied to the corrected spectrum to estimate the surface pressure and XCO_2.

3 Results

Figure 1 shows a density histogram of the GEOS5 FP-IT surface pressure analysis and of the NN estimate for the evaluation dataset (odd-numbered months). Clearly, there is an excellent agreement between the two over a very wide range of surface pressures. There is no significant bias and the standard deviation is 2.9 hPa. The equivalent ACOS V8r retrieval shows a bias of 1.5 hPa and a standard deviation of 3.4 hPa, slightly larger than that of the NN approach. Note that the ACOS statistics are those of the ACOS retrieval–minus–prior statistics (see Sect. 2). Interpreting them in terms of error is counterintuitive because the Bayesian retrieval is supposed to be better than the prior, but in practice radiation modeling errors lead to a different interpretation (see, e.g., the discussion in Sect. 4.3.4 of O'Dell et al., 2018).

Both NN and ACOS correlations with GEOS5 FP-IT are very high (0.997 and 0.996) although the best fit shows a very small deviation from the 1 : 1 line. Interestingly, the best-fit deviations from the 1 : 1 line are of opposite sign (slopes 0.99 and 1.01). The results of the NN are surprisingly good given the simplicity of the approach and given that the NN estimate does not use any a priori information or ancillary information such as the surface altitude or temperature profile, contrarily to the ACOS estimate. The quality of the NN results for the estimate of the surface pressure is a first demonstration of the potential of the approach. Note that the retrieval accuracy holds over a very large range of surface pressures (the relative variations of XCO_2 are much smaller), although there is some indication of biases for the lowest pressures that are underrepresented in the training dataset. These biases of ≈ 5 hPa affect the observations over high-elevation surfaces such as the Tibetan Plateau or the US Rocky Mountains.

Figure 2 is similar to Fig. 1 but for XCO_2. There is no significant bias between the NN estimate and the CAMS model, while the standard deviation is 0.84 ppm. The bias-corrected ACOS retrievals show a slight bias against the CAMS model and the standard deviation (1.14 ppm) is larger than that of the NN approach. Note that the statistics given here are affected by CAMS modeling errors that may eventually be corrected with the help of the satellite information. The best fit slope deviations from the 1 : 1 line are larger than for the surface pressure; the slopes are 0.93 for the NN and 0.87 for ACOS.

Figures 1 and 2, together with the quantitative assessment of the precision, are given for the observations that are clear according to ACOS (cloud flag = 2 or 3), that have a warn level of 2 or less, that may include missing pixel elements, and that have an outcome flag of 1 or 2. This choice is based on a prior performance analysis. We have analyzed how the performance of the NN approach varies with the quality indicators. For this objective, we have compared the retrieved surface pressure against the value derived from the numerical weather data, as in Fig. 1, and we have evaluated the statistics of their difference as a function of the quality flags. First (figure not shown), there is no significant difference between the cases when the measured spectra are complete and those when one or several missing pixel elements have been interpolated with the method described above. Conversely, the statistics vary with the cloud flag and the warn level, as shown in Fig. 3. We only use the spectra for which an ACOS retrieval is available. Among those, and according to the flag *cloud_flag_idp*, about 53 % are labeled as "very clear" while 43 % are "probably clear". The statistics are slightly better for the former than they are for the latter. Conversely, the rather rare "definitely cloudy" and "probably cloudy" show deviations that are significantly larger. This result was highly expected since our NN did not learn how to handle clouds in the spectra. Therefore, all "definitely cloudy" and "probably cloudy" soundings are outside the domain covered by the training dataset. Note also that the observations used here have all been classified as "clear" by ACOS preprocessing. Thus, most OCO-2 observations are not used here and Fig. 3 should not be interpreted as the ability to retrieve the surface pressure in cloudy conditions. Most (78 %) of the observations have a warn level of 0. The deviation statistics increase with the warn level, both in terms of bias and standard deviation. In comparison, the difference in the statistics for an outcome flag of 1 and 2 are small. Besides, more than half of the ACOS retrievals have an outcome flag of 2, which encourages us not to reject those for further use. Based on this analysis, we retain all spectra that are very clear (cloud flag of 2 or 3) and that have a warn level of 2 or less.

We made a figure similar to Fig. 3 but based on the XCO_2 estimates (not shown). Although the results are similar in terms of sign (i.e., increase of the deviations with the warn levels), the signal is not as obvious (there is less relative difference between one warn level and another, or for the various cloud flags). Our interpretation is that the relative accuracy of the surface pressure used as a reference estimate is much better than that of the NN retrieval, whereas the accuracy of the XCO_2 from CAMS is not much better than that of the NN. As a consequence, variations in the accuracy of the NN do not show up as clearly for XCO_2 than they do for the surface pressure.

A standard method to evaluate an algorithm that estimates XCO_2 from spaceborne observation is the comparison of its products against estimates from TCCON retrievals. These estimates use ground-based solar absorption spectra recorded by Fourier transform infrared spectroscopy and have been tuned with airborne in situ profiles (Wunch et al., 2010). To take advantage of the full potential of the TCCON retrievals for the bias correction and validation of the XCO_2 estimates, the OCO-2 platform can be oriented so that the instrument

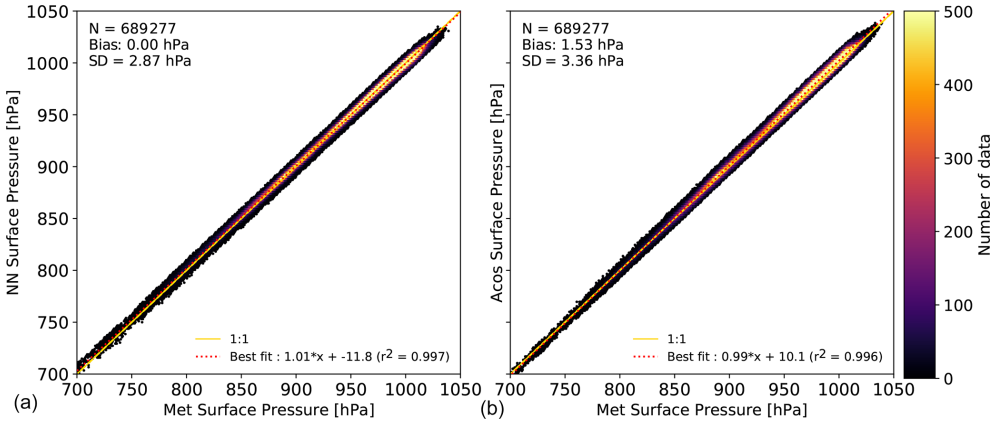

Figure 1. Density histogram of the surface pressure retrieved from the OCO-2 satellite measurements against those derived from GEOS-FP-IT. Panel **(a)** is for the NN approach while panel **(b)** is for the ACOS V9r retrieval (using the official bias correction). The figure insets provide the number of data points, the bias, the standard deviation, the equation of the best linear fit, and the correlation. The yellow line is the 1 : 1 line whereas the red dotted line is the best linear fit.

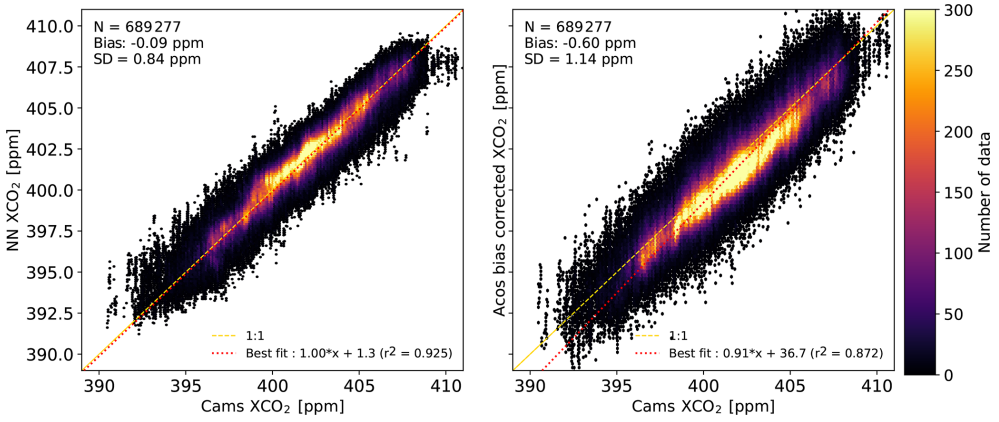

Figure 2. Same as Fig. 1 but for XCO_2. In this case, the reference data are the CAMS v18r2 simulation.

field of view is close to the surface station. The ACOS full-physics algorithm can handle these spectra that are acquired in neither nadir nor glint geometries, but the NN was trained solely on nadir spectra and cannot yet be applied to the observations acquired in target mode. We thus have to rely on nadir measurements acquired in the vicinity of TCCON sites. In the following, we use nadir measurements that are within 5° in longitude and 1.5° in latitude to the TCCON site. The XCO_2 estimates (either from ACOS, the NN, or the model sampled at the OCO-2 measurement location) are averaged for a given overpass. Similarly, we average the TCCON estimates of XCO_2 within 30 min of the satellite overpass. No attempt was made to correct for the different weighting functions of the surface and spaceborne remote sensing estimates. The comparisons are shown in Fig. 4 for each TCCON station listed in Table 1.

Overall, the biases and standard deviations of the differences to TCCON observations are -0.34 ± 1.40 ppm for the NN, -0.47 ± 1.49 ppm for ACOS and 0.04 ± 1.09 ppm for CAMS. The statistics per station are provided in Table 1. Two stations, Paris and Pasadena, show a large negative bias for both estimates, which may be interpreted as the impact of the city on the atmosphere sampled by the TCCON measurement, while the atmosphere sampled by the distant satellite may be less affected. Conversely, there is no such negative bias for other stations that are located close to large cities, such as Tsukuba, a suburb of the Tokyo Metropolitan area. Zugspitze is rather specific due to its high altitude. The comparison against TCCON indicates that the NN approach performs similarly to ACOS, if not better. The dispersion is larger for one approach versus the other for some stations, while the opposite is true for other stations. Note also that the CAMS model performs better than both satellite retrievals for most stations. This observation provides further justification to the use of this model for training the NN.

The evaluation of the algorithm performance is limited by the distance between the satellite estimate and its surface validation. This is inherent to the use of nadir-only observations

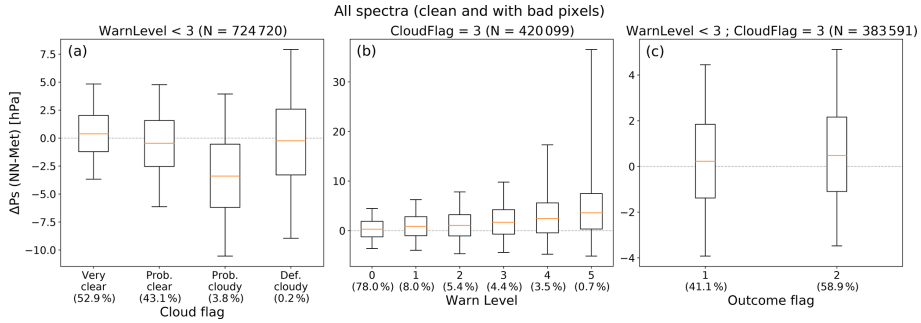

Figure 3. Statistics on the difference between the surface pressure retrieved by the NN approach and those derived from the weather analyses as a function of various quality parameters. In these figures, the red line is the median, the boxes indicate the 25 % and 75 % percentiles, and the whiskers indicate the 5 %–95 % range. Panel **(a)** shows the statistics as a function of the cloud flag, panel **(b)** is as a function of the warn level, while panel **(c)** is as a function of the outcome flag.

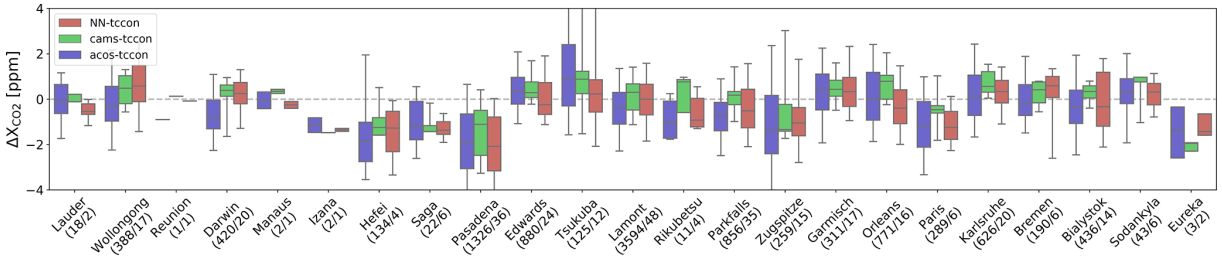

Figure 4. Statistics of the differences between the NN retrieval (red), the CAMS model (green) or the bias-corrected ACOS retrievals (blue), and the TCCON retrievals. The boxes indicate the 25 %–75 % percentiles and the median is shown by the horizontal line within the box. The whiskers indicate the 5 %–95 % percentiles. Stations are ordered by increasing latitudes. The numbers below the station name indicate the number of individual observations and coincidence days used for the statistics. The references for the various TCCON observations are provided in Table 1.

that are seldom located close to the TCCON sites. A reduction of the distance results in less coincidences, which leads to a validation dataset of poor representativeness. Note that the CAMS model was sampled at the location of the satellite observations, so that the higher performance of the model versus the satellite products cannot be caused by a higher proximity to the TCCON station.

We now investigate whether the model–minus–NN differences are purely random or contain some spatial or temporal structures. This question is important because if the differences show a random structure there is little hope to use these data to improve the surface fluxes used in the CAMS product. Conversely, if the XCO_2 differences do show some structures, they can be attributed to surface flux errors in the CAMS product that may then be corrected through inverse atmospheric modeling. There is no certainty, however, as a spatial structure in the NN–minus–CAMS difference can also be interpreted as a bias in the satellite estimate.

We first show (Fig. 5) the difference between the NN estimates of the surface pressure and the numerical weather analyses. These are monthly maps of the NN-minus-CAMS difference for 3 years of the period at a $5° \times 5°$ resolution. We only present the odd-numbered months as the others were used for the training and therefore do not show any significant differences. There are very clear spatial patterns of a few hPa that are not expected and should be interpreted as a bias in the NN approach. The biases over the high mountains and plateaus have already been mentioned. In addition, positive biases tend to occur in the high latitudes and negative biases toward the tropics. The structures show additional spatial and temporal patterns and are therefore more complex than just a latitude function. The same figure but based on the ACOS retrievals (Fig. A3) displays large-scale structures with different spatial patterns; the surface pressure bias is mostly negative over northern latitudes and positive over low latitudes. A histogram (Fig. 6) of the monthly differences, such as those shown on Fig. 5, confirms that the amplitude of the surface pressure biases is larger with ACOS than it is with the NN. The NN or ACOS surface pressure bias is -0.33 hPa or 1.39 hPa, respectively, and the standard deviation is 2.12 or 2.79 hPa, respectively.

Figure 7 is similar to Fig. 5 but for XCO_2 differences between the NN estimate and the CAMS model. As for the surface pressure, there are clear spatial patterns with amplitudes of 1 to 2 ppm. The question is whether these are mostly linked to monthly biases in the CAMS model or to the NN.

Table 1. TCCON stations used in this paper (Fig. 4). The data were obtained from the http://tccondata.org website at during the summer of 2019 (last access: 1 August 2019).

Stations	[lat; long]	Altitude [m]	Reference	Biases NN/ACOS/CAM	SD NN/ACOS/CAM
Lauder	[−45.04; 169.68]	370	Sherlock et al. (2017)	−0.48/−0.25/0.076	0.43/1.51/0.16
Wollongong	[−34.41; 150.88]	30	Griffith et al. (2017b)	0.60/−0.20/0.42	1.21/1.32/0.60
Reunion	[−20.90; 55.49]	90	De Maziere et al. (2017)	−0.08/−0.90/0.13	–/–/–
Darwin	[−12.43; 130.89]	30	Griffith et al. (2017a)	0.19/−0.69/0.23	0.80/1.09/0.72
Manaus	[−3.21; −60.6]	50	Dubey et al. (2017)	−0.25/−0.05/0.34	0.43/1.04/0.26
Izana	[28.3; −16.48]	2300	Blumenstock et al. (2017)	−1.35/−1.14/−1.48	0.18/0.92/0.0
Hefei	[31.90; 118.67]	30	Liu et al. (2018)	−1.47/−1.58/−1.01	1.11/1.76/0.63
Saga	[33.24; 130.29]	10	Shiomi et al. (2017)	−1.36/−1.03/−1.15	0.57/1.22/0.59
Pasadena	[34.14; −118.13]	240	Wennberg et al. (2017b)	−2.12/−1.87/−1.41	1.57/1.64/1.17
Edwards	[34.96; −117.88]	700	Iraci et al. (2017)	0.07/0.41/0.50	1.00/1.01/0.64
Tsukuba	[36.05; 140.12]	30	Morino et al. (2017a)	0.42/1.43/1.05	2.13/2.53/1.61
Lamont	[36.6; −97.49]	320	Wennberg et al. (2017c)	−0.03/−0.38/0.16	1.07/1.21/0.94
Rikubetsu	[43.46; 1473.77]	390	Morino et al. (2017b)	−0.57/−0.84/0.47	0.84/1.07/0.98
Parkfalls	[45.94; −90.27]	440	Wennberg et al. (2017a)	−0.41/−0.75/0.11	1.15/1.01/0.72
Zugspitze	[47.42; 11.06]	2960	Sussmann and Rettinger (2017b)	−0.85/−1.14/−0.83	1.45/1.85/1.36
Garmisch	[47.48; 11.06]	740	Sussmann and Rettinger (2017a)	0.40/0.28/0.43	0.98/1.29/0.62
Orleans	[47.97; 2.11]	130	Warneke et al. (2017)	−0.35/0.13/0.66	1.06/1.38/0.67
Paris	[48.85; 2.36]	60	Te et al. (2017)	−1.29/−1.24/−0.62	1.30/1.66/1.23
Karlsruhe	[49.1; 8.44]	110	Hase et al. (2017)	0.26/0.21/0.75	0.80/1.29/0.55
Bremen	[53.10; 8.85]	7	Notholt et al. (2017)	0.30/−0.07/0.36	1.11/1.02/0.45
Bialystok	[53.23; 23.02]	180	Deutscher et al. (2017)	−0.11/−0.32/0.33	1.31/1.30/0.42
Sodankyla	[67.37; 26.63]	190	Kivi et al. (2017)	0.26/0.24/0.61	0.79/1.36/0.80
Eureka	[80.05; −86.42]	600	Strong et al. (2017)	−1.02/−1.50/−2.16	1.01/2.25/0.41

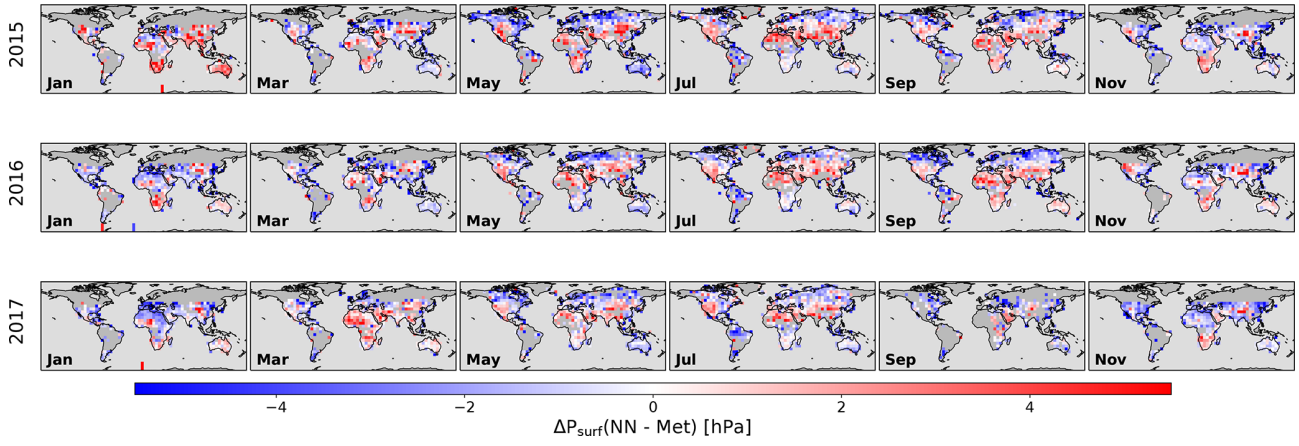

Figure 5. Difference between the NN estimates of the surface pressure and the numerical weather analyses. The differences have been averaged at monthly and spatial 5° × 5° resolutions. The results are shown for 3 years and only for the months that were not used for the training.

The first hypothesis is of course more favorable as it would indicate that the satellite data can bring new information to constrain the surface fluxes. However, the analysis of the surface pressure that shows biases of several hPa suggests that the NN XCO_2 estimate may also show biases with spatially coherent patterns. Interestingly, the patterns vary in time and are not correlated with those of the surface pressure. Further analysis, in particular atmospheric flux inversion, is necessary for a proper interpretation of the NN–CAMS differences.

The differences of ACOS estimates to the CAMS model also show patterns of similar amplitude as those in Fig. 7 (Fig. A4). However, there is no clear correspondence between these patterns and those obtained using the NN prod-

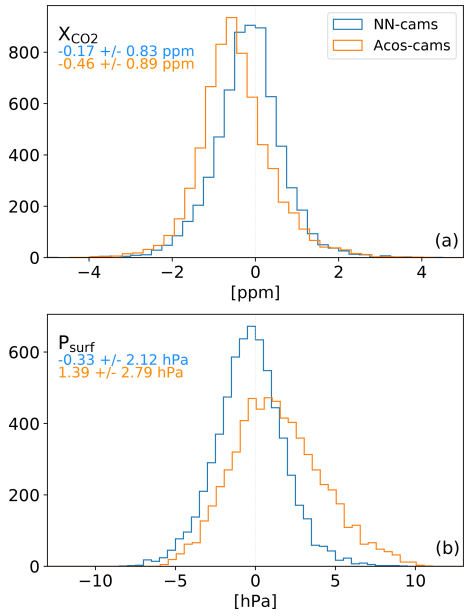

Figure 6. Histogram of the monthly mean differences at 5° resolution (such as those shown in Fig. 5) between the satellite retrievals and the CAMS model: **(a)** is for XCO_2 while **(b)** is for the surface pressure. The blue line is for the NN product while the orange line is for ACOS.

uct. The differences between the satellite products and the CAMS model are small, but these contain the information that may be used to improve our knowledge on the surface fluxes. The absence of a clear correlation between the spatiotemporal pattern from the NN and ACOS approaches indicate that their use would lead to very different corrections on the surface fluxes if used as input of an atmospheric inversion approach. Figure 6a shows the histogram of these monthly mean differences. The histograms are very similar for the two satellite products, although the standard deviation of the difference to the CAMS model is slightly larger for ACOS than it is for the NN approach (0.89 vs. 0.83 ppm).

4 Discussion and conclusion

The use of the same product for the NN training and its evaluation may be seen as a weakness of our analysis. One may argue that the NN has learned from the model and generates an estimate (either the surface pressure or XCO_2) that is not based on the spectra but rather on some prior information. Let us recall that the NN input does not contain any information on the location or date of the observation. This is a strong indication that the information is derived from the spectra as the NN does not "know" the CAMS value that corresponds to the observation location. Yet, the NN input also includes the observation geometry (sun angle and azimuth) that is somewhat correlated with the latitude and day of the year. One may then argue that the NN learns from this indirect information on the observation location and then generates an estimate based on the corresponding CAMS value. However, since the observation geometry is exactly the same from one year to the next, there is no information, direct or indirect, on the observation year in the NN input. Thus, the XCO_2 growth rate that is accurately retrieved by the NN method (see Fig. 7) is necessarily derived from the spectra. A similar argument can be made on the spatial variation across the longitudes.

To further demonstrate that the NN retrieves XCO_2 from the spectra rather than from the prior, we performed an additional experiment. The training is based only on even-numbered months. As a consequence, the prior does not include any direct information on the odd-numbered months. For these months, the best prior estimate is a linear interpolation between the two adjacent even-numbered months. We can then analyze how the NN estimate compares with the CAMS product, which accounts for the true synoptic variability, and a degraded version of CAMS based on a linear interpolation between the two adjacent months. This comparison is shown in Fig. 8. The center figure compares the true CAMS value and that derived from the temporal interpolation. As expected, both are highly correlated (the seasonal cycle and the growth rate are kept in the interpolated values) but nevertheless show a difference standard deviation of 0.89 ppm. This value can be interpreted as the synoptic variability of XCO_2 present in CAMS but not captured in the interpolated product. The comparison of the NN estimate against CAMS (right) and the interpolated CAMS (left) shows significantly better agreement to the former. Thus, the NN product reproduces some XCO_2 variability that is not contained in the training prior. It provides further demonstration that the NN estimates relies on the spectra rather than on the time/space variations of the training dataset.

The results shown above indicate that the NN approach allows an estimate of surface pressure and XCO_2 with a precision that is similar or better than that of the operational ACOS algorithm. The lack of independent "truth" data does not allow a full quantification of the product precision and accuracy. However, there are indications that the accuracy on the surface pressure is better than 3 hPa rms, while the precision (standard deviation) of XCO_2 is better than 0.9 ppm. The data used for the XCO_2 product evaluation has its own error that is difficult to disentangle from that of the estimate based on the satellite observation. It may also contain a bias that is propagated to the NN through its training.

One obvious advantage of the NN approach is the speed of the computation, which is several orders of magnitude higher than that of the full-physics algorithm. This is significant given the current reprocessing time of the OCO-2 dataset despite the considerable computing power made available for the mission. It also bears interesting prospects for future XCO_2 imaging missions that will bring even higher data volume (e.g., Pinty et al., 2017).

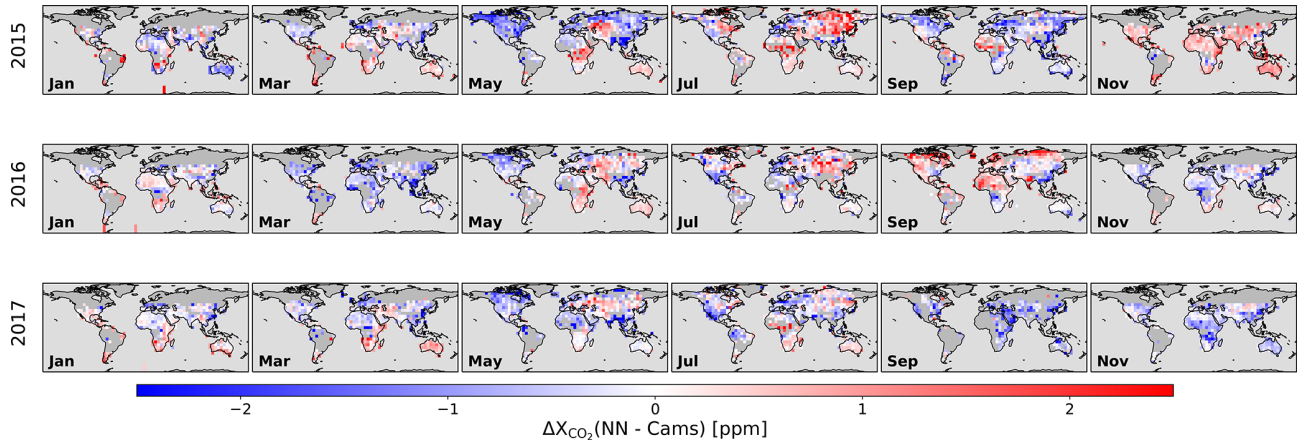

Figure 7. Same as Fig. 5 but for the difference between the XCO_2 estimated by the NN approach and that derived from the CAMS model.

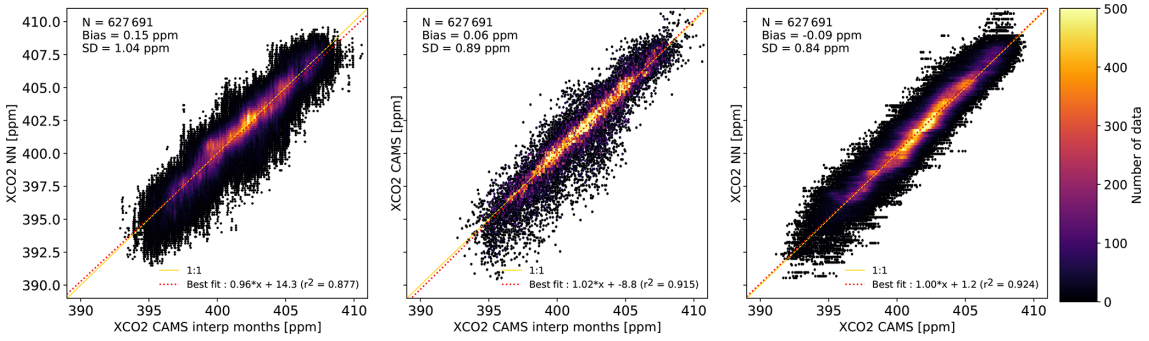

Figure 8. Scatterplots of XCO_2 estimated by the NN, the CAMS model, and the CAMS model that has been interpolated in time from adjacent months (see text for details). Note that the number of points is less than in Fig. 2 because the edge months could not always be interpolated.

Another advantage is that the NN approach described in this paper does not require the extensive debias procedure necessary for the ACOS product (O'Dell et al., 2018; Kiel et al., 2019). Per construction, there is no bias between the NN estimates and the dataset used for the NN training. The NN approach therefore requires less effort and fewer resources.

There are, however, a number of drawbacks for the NN approach described in this paper.

One obvious drawback is the use of a CO_2 model simulation in the training while the main purpose of the satellite observation is to improve our current knowledge of atmospheric CO_2 and its surface fluxes. Our argument is that although the CAMS simulation used here has high skill (as demonstrated in Fig. 4), it may have positive or negative XCO_2 biases for some months and some areas. These biases are independent from the measured spectra so that the NN training will aim at average values. As a consequence, the NN product could in principle be of higher quality than the CAMS product, even though the same model has been used as the reference estimate for the training (see, e.g., Aires et al., 2005).

Another drawback of the NN approach is that it does not directly provide its averaging kernel. The averaging kernel vector reports the sensitivity of the retrieved total column to changes in the concentration profile (Connor et al., 1994). It is a combination of physical information (about radiative transfer) and statistical information (about the prior information). It is needed for a proper comparison with 3D atmospheric models (e.g., Chevallier, 2015). When comparing with model simulations, for instance for atmospheric inversion, we may wish to neglect the NN implicit prior information. This hypothesis leads to a homogeneous pressure weighting over the vertical, as this is the product that the NN was trained to simulate. Alternatively, we could decide to neglect the difference in prior information between the NN and the full-physics algorithm and use typical averaging kernels of the latter. A third, more involved option would be to perform a detailed sensitivity study of the NN, based on radiative transfer simulations.

Similarly, the current version of our neural network does not provide a posterior uncertainty. A Monte Carlo approach using various training datasets could be use in the future for such an estimate.

Also, because of the CO_2 growth rate, the developed NN cannot be safely used to process observations that are ac-

quired later than a few weeks after the last data of the training dataset, as the observed CO_2 may then be outside of its range. Therefore, the use of the neural network approach for near real-time applications would require frequent updates of the training phase.

We acknowledge the fact that the NN product evaluated here is not fully independent from the ACOS product. Indeed, we use the cloud flag and the quality diagnostic from ACOS to select the spectra that are of sufficient quality. If we aim for some kind of operational product, there is a need to design a procedure to identify these good quality spectra. One option would be to compare the surface pressure retrieved by the NN to the numerical weather analysis estimate and to reject cases with significant deviations (e.g., differences larger than 3 hPa).

Despite these drawbacks, the results presented here show that a neural network has a large potential for the estimate of XCO_2 from satellite observations, such as those from OCO-2, the forthcoming MicroCarb (Pascal et al., 2017), or the CO_2M constellation (Sierk et al., 2018), which aims to measure anthropogenic emissions. It is rather amazing that a first attempt leads to trueness and precision numbers that are similar or better than those of the full-physics algorithm. There are several paths to improvement. One is to provide the NN with some ancillary information such as the surface altitude or a proxy of the atmospheric temperature. Another is to train the NN with model estimates (such as those of CAMS used here) but that have been better sampled for their assumed precision, for instance through a multimodel evaluation. Also, one could train the NN with observations acquired during a few days of each month, rather than the even-numbered months as done here, so that the evaluation dataset would provide a better evaluation of the seasonal cycle.

Our next objective is to attempt a similar NN approach for the measurements that were acquired in the glint mode. As explained above, the glint observations may be more difficult to reproduce by the NN than those acquired in the nadir mode. However, we were very much surprised by the performance of the NN with the nadir data and cannot preclude being surprised again. Finally, we shall analyze the spatial structure of the NN retrievals in regions that are expected to be homogeneous and in regions where structures of anthropogenic origin are expected (e.g., Nassar et al., 2017; Reuter et al., 2019).

Appendix A

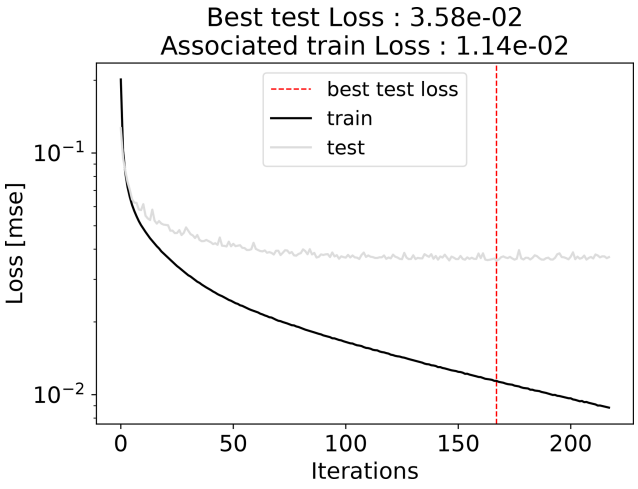

Figure A1. Illustration of the iterative convergence of the NN during its training. The loss is an indicator of the difference between the NN estimate and the dataset. One dataset is used for the best estimate of the NN weights whereas another independent one is used for the evaluation of the NN capability. The NN is stopped when there is no further reduction of the loss for the test dataset for 50 iterations. The weights for the NN are those obtained for the lowest loss of the test dataset (iteration 167 on the figure).

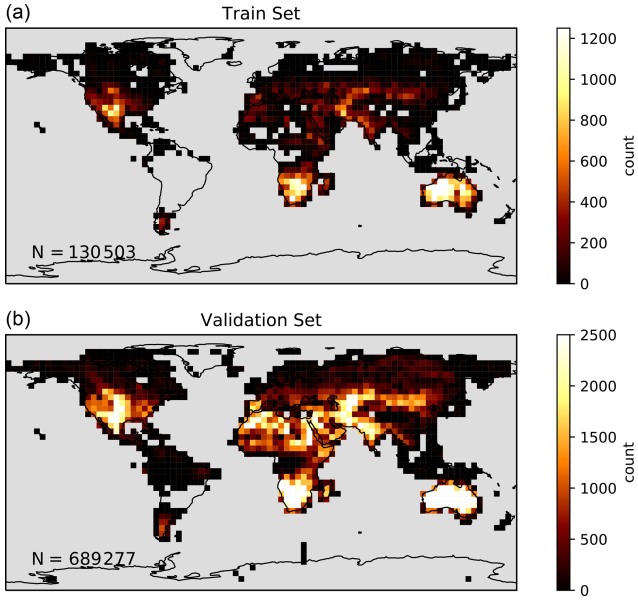

Figure A2. Spatial density of the observations that were used for the training **(a)** and validation **(b)** processes.

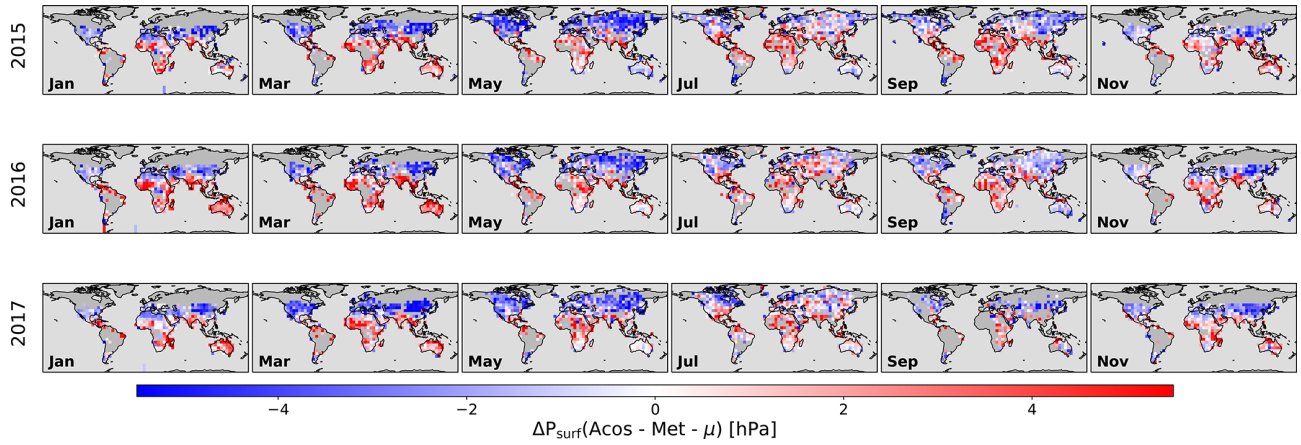

Figure A3. Same as Fig. 5 but for the surface pressure retrieved by the ACOS algorithm. The mean bias over the full period (μ) is removed so that the differences are centered on zero.

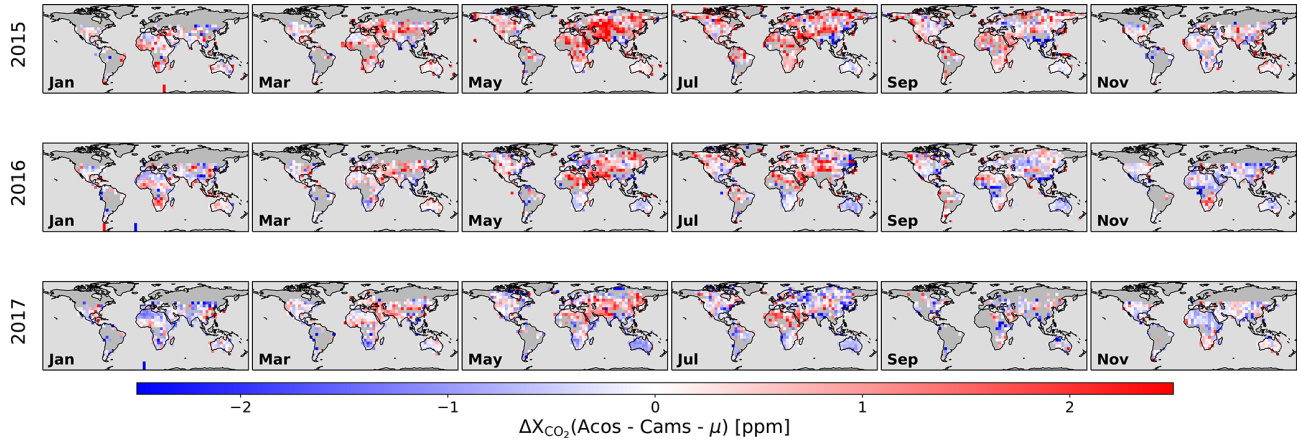

Figure A4. Same as Fig. 7 but for the XCO_2 retrieved by the ACOS algorithm. The mean bias over the full period (μ) is removed so that the differences are centered on zero.

Code and data availability. The code used in this paper and the CAMS model simulations are available, upon request, from the author. The OCO-2 and TCCON data can be downloaded from their respective websites. The TCCON data is available from the website http://tccondata.org (TCCON, 2021), whereas the OCO-2 data is available from http://ocov2.jpl.nasa.gov (NASA Jet Propulsion Laboratory, 2021).

Author contributions. FMB designed the study. LD developed the code and performed the computations. All authors shared the result analysis.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. OCO-2 L1 and L2 data were produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the ACOS/OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center. TCCON data were obtained from the TCCON Data Archive (http://tccondata.org, last access: 5 January 2021). We warmly thank those who made these data available.

Financial support. This work was in part funded by CNES, the French space agency, in the context of the preparation for the MicroCarb mission.

Review statement. This paper was edited by Piet Stammes and reviewed by Christopher O'Dell and one anonymous referee.

References

Agustí-Panareda, A., Diamantakis, M., Massart, S., Chevallier, F., Muñoz-Sabater, J., Barré, J., Curcoll, R., Engelen, R., Langerock, B., Law, R. M., Loh, Z., Morguí, J. A., Parrington, M., Peuch, V.-H., Ramonet, M., Roehl, C., Vermeulen, A. T., Warneke, T., and Wunch, D.: Modelling CO_2 weather – why horizontal resolution matters, Atmos. Chem. Phys., 19, 7347–7376, https://doi.org/10.5194/acp-19-7347-2019, 2019.

Aires, F., Prigent, C., and Rossow, W. B.: Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: 2. Global statistical relationships, J. Geophys. Res., 110, D11103, https://doi.org/10.1029/2004JD005087, 2005.

Blumenstock, T., Hase, F., Schneider, M., Garcia, O. E., and Sepulveda, E.: TCCON data from Izana (ES), Release GGG2014R1, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.izana01.R1, 2017.

CEOS: A Constellation Architecture for Monitoring Carbon Dioxide and Methane from Space, Technical Report, University of Zurich, Switzerland, Department of Informatics, available at: http://ceos.org/document_management/Meetings/Plenary/32/documents/CEOS_AC-VC_White_Paper_Version_1_20181009.pdf (last access: 18 October 2019), 2018.

Chédin, A., Serrar, S., Scott, N. A., Crévoisier, C., and Armante R.: First global measurement of midtropospheric CO_2 from NOAA polar satellites: Tropical zone, J. Geophys. Res., 108, 4581, https://doi.org/10.1029/2003JD003439, 2003.

Chevallier, F.: On the statistical optimality of CO_2 atmospheric inversions assimilating CO_2 column retrievals, Atmos. Chem. Phys., 15, 11133–11145, https://doi.org/10.5194/acp-15-11133-2015, 2015.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res., 110, D24309, https://doi.org/10.1029/2005JD006390, 2005.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Frohlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morgu, 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 from a global 21 year reanalysis of atmospheric measurements, J. Geophys. Res.-Atmos., 115, D21307, https://doi.org/10.1029/2010JD013887, 2010.

Connor, B. J., Siskind, D. E., Tsou, J. J., Parrish, A., and Remsberg, E. E.: Ground-based microwave observations of ozone in the upper stratosphere and mesosphere, J. Geophys. Res., 99, 16757–16770, 1994.

Crisp, D., Atlas, R. M., Breon, F.-M., Brown, L. R., Burrows, J. P., Ciais, P., Connor, B. J., Doney, S. C., Fung, I. Y., Jacob, D. J., Miller, C. E., O'Brien, D., Pawson, S., Randerson, J. T., Rayner, P., Salawitch, R. J., Sander, S. P., Sen, B., Stephens, G. L., Tans, P. P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission, Adv. Space Res., 34, 700–709, 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 X_{CO2} data characterization, Atmos. Meas. Tech., 5, 687–707, https://doi.org/10.5194/amt-5-687-2012, 2012.

De Maziere, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., Metzger, J.-M., Duflot, V., and Cammas, J.-P.: TCCON data from Reunion Island (RE), Release GGG2014R0, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.reunion01.R1, 2017.

Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., Grupe, P., and Katrynski, K.: TCCON data form Bialystok (PL), Release GGG2014R2, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.bialystok01.R2, 2017.

Dubey, M., Henderson, B., Green, D., Butterfield, Z., Keppel-Aleks, G., Allen, N., Blavier, J.-F., Roehl, C., Wunch, D., and Lindenmaier, R.: TCCON data from Manaus (BR), Release GGG2014R0, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.manaus01.R0/1149274, 2017.

Eldering, A., Pollock, R., Lee, R. A. M., Rosenberg, R., Oyafuso, F., Crisp, D., Chapsky, L., and Granat, R.: Orbiting Carbon Observatory (OCO) – 2 Level 1B Theoretical Basis Document, available at: http://disc.sci.gsfc.nasa.gov/OCO-2/documentation/oco-2-v7/OCO_2_L1B_ATBD.V7.pdf (last access: 16 June 2016), 2015.

Eldering, A., Wennberg, P. O., Crisp, D., Schimel, D., Gunson, M. R., Chatterjee, A., Liu, J., Schwandner, F. M., Sun, Y., O'Dell, C. W., Frankenberg, C., Taylor, T., Fisher, B., Osterman, G. B., Wunch, D., Hakkarainen, J., Tamminen, J., and Weir, B.: The Orbiting Carbon Observatory-2 early science investigations of regional carbon dioxide fluxes, Science, 358, eaam5745, https://doi.org/10.1126/science.aam5745, 2017.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2019, Earth Syst. Sci. Data, 11, 1783–1838, https://doi.org/10.5194/essd-11-1783-2019, 2019.

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.: TCCON data from Darwin (AU), Release GGG2014R0, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.darwin01.R0/1149290, 2017a.

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, CDIAC, https://doi.org/10.14291/tccon.ggg2014.wollongong01.R0/1149291, 2017b.

Hadji-Lazaro, J., Clerbaux, C., and Thiria, S.: An inversion algorithm using neural networks to retrieve atmospheric CO total columns from high-resolution nadir radiances, J. Geophys. Res., 104, 23841–23854, https://doi.org/10.1029/1999JD900431, 1999.

Hase, F., Blumenstock, T., Dohe, S., Gross, J., and Kiel, M.: TCCON data from Karlsruhe (DE), Release GGG2014R1, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.karlsruhe01.R1/1182416, 2017.

Iraci, L. T., Podolske, J., Hillyard, P. W., Roehl, C., Wennberg, P. O., Blavier, J.-F., Landeros, J., Allen, N., Wunch, D., Zavaleta, J., Quigley, E., Osterman, G., Albertson, R., Dunwoody, K., and Boyden, H.: TCCON data from Edwards (US), Release GGG2014R1, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.edwards01.R1/1255068, 2017.

Keras Team: Keras, available at: https://github.com/fchollet/keras (last access: 5 January 2021), GitHub, 2015.

Kiel, M., O'Dell, C. W., Fisher, B., Eldering, A., Nassar, R., MacDonald, C. G., and Wennberg, P. O.: How bias correction goes wrong: measurement of X_{CO2} affected by erroneous surface pressure estimates, Atmos. Meas. Tech., 12, 2241–2259, https://doi.org/10.5194/amt-12-2241-2019, 2019.

Kivi, R., Heikkinen, P., and Kyrö, E.: TCCON data from Sodankyla (FI), Release GGG2014R0, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.sodankyla01.R0/1149280, 2017.

Knorr, W.: Is the airborne fraction of anthropogenic CO_2 emissions increasing?, Geophys. Res. Lett., 36, L21710, https://doi.org/10.1029/2009GL040613, 2009.

Liu, C., Wang, W., and Sun, Y.: TCCON data from Hefei (PCR), Release GGG2014R0, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.hefei01.R0, 2018.

Lucchesi, R.: File Specification for GEOS-5 FP-IT (Forward Processing for Instrument Teams), Technical Report, NASA Goddard Spaceflight Center, Greenbelt, MD, USA, available at: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20150001438.pdf (last access: 4 December 2018), 2013.

Morino, I., Matsuzaki, T., and Shishime, A.: TCCON data from Tsukuba (JP), 125HR, Release GGG2014R2, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.tsukuba02.R2, 2017a.

Morino, I., Yokozeki, N., Matzuzaki, T., and Shishime, A.: TCCON data from Rikubetsu (JP), Release GGG2014R2, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.rikubetsu01.R2, 2017b.

NASA Jet Propulsion Laboratory: Orbiting Carbon Observatory-2, available at: http://ocov2.jpl.nasa.gov, last access: 5 January 2021.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D., 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.

Notholt, J., Petri, C., Warneke, T., Deutscher, N. M., Buschmann, M., Weinzierl, C., Macatangay, R., and Grupe, P.: TCCON data from Bremen (DE), Release GGG2014R0, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.bremen01.R0/1149275, 2017.

O'Brien, D. M. and Rayner, P. J.: Global observations of the carbon budget 2. CO_2 column from differential absorption of reflected sunlight in the 1.61 µm band of CO_2, J. Geophys. Res., 107, ACH6-1, https://doi.org/10.1029/2001JD000617, 2002.

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Man-

drake, 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.

Pascal, V., Buil, C., Loesel, J., Tauziede, L., Jouglet, D., and Buisson, F.: An improved microcarb dispersive instrumental concept for the measurement of greenhouse gases concentration in the atmosphere, Proc. SPIE, 10563, https://doi.org/10.1117/12.2304219, 2017.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699–6720, https://doi.org/10.5194/bg-10-6699-2013, 2013.

Pinty, B., Janssens-Maenhout, G., Dowell, M., Zunker, H., Brunhes, T., Ciais, P., Dee, D., Denier van der Gon, H. A. C., Dolman, H., Drinkwater, M., Engelen, R., Heimann, M., Holmlund, K., Husband, R., Kentarchos, A., Meyer, A., Palmer, P., and Scholze, M.: An operational anthropogenic CO_2 emissions monitoring and verification support capacity – Baseline requirements, Model components and functional architecture, European Commission Joint Research Centre, Reading, UK, EUR 28736 EN, https://doi.org/10.2760/08644, 2017.

Reuter, M., Buchwitz, M., Schneising, O., Noël, S., Rozanov, V., Bovensmann, H., and Burrows, J. P.: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering – Part 1: Radiative Transfer and a Potential OCO-2 XCO_2 Retrieval Setup, Remote Sens., 9, 1159, https://doi.org/10.3390/rs9111159, 2017.

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., Bovensmann, H., and Burrows, J. P.: Towards monitoring localized CO_2 emissions from space: co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites, Atmos. Chem. Phys., 19, 9371–9383, https://doi.org/10.5194/acp-19-9371-2019, 2019.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning Internal Representations by Error Propagation, in: Readings in Cognitive Science, edited by: Collins, A. and Smith, E. E., Elsevier, https://doi.org/10.1016/B978-1-4832-1446-7.50035-2, 399–421, 1988.

Salstein, D. A., Ponte, R. M., and Cady-Pereira, K.: Uncertainties in atmospheric surface pressure fields from global analyses, J. Geophys. Res., 113, D14107, https://doi.org/10.1029/2007JD009531, 2008.

Sherlock, V., Connor, B. J., Robinson, J., Shiona, H., Smale, D., and Pollard, D.: TCCON data from Lauder (NZ), 125HR, Release GGG2014R0, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.lauder02.R0/1149298, 2017.

Shiomi, K., 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, CDIAC, https://doi.org/10.14291/tccon.ggg2014.saga01.R0/1149283, 2017.

Sierk, B., Bézy, J.-L., Löscher, A., and Meijer, Y.: The European CO_2 Monitoring Mission: Observing anthropogenic greenhouse gas emissions from space, Proceedings of the International Conference on Space Optics, Chania, Greece, 9–12 October 2018, 237–250, https://doi.org/10.1117/12.2535941, 2018.

Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., Fogal, P., Drummond, J., Batchelor, R., and Lindenmaier, R.: TCCON data from Eureka (CA), Release GGG2014R3, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.eureka01.R3, 2017.

Suarez, M. J., Rienecker, M. M., Todling, R., Bacmeister, J., Takacs, L., Liu, H. C., Gu, W., Sienkiewicz, M., Koster, R. D., and Gelaro, R.: The GEOS-5 Data Assimilation System-Documentation of Versions 5.0.1, 5.1.0, and 5.2.0, Technical Report, NASA Goddard Spaceflight Center, Greenbelt, MD, USA, available at: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120011955.pdf (last access: 4 December 2018), 2008.

Sussmann, R. and Rettinger, M.: TCCON data from Garmisch (DE), Release GGG2014R2, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.garmisch01.R2, 2017a.

Sussmann, R. and Rettinger, M.: TCCON data from Zugspitze (DE), Release GGG2014R1, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.zugspitze01.R1, 2017b.

TCCON: Homepage, available at: http://tccondata.org, last access: 5 January, 2021.

Te, Y., Jeseck, P., and Janssen, C.: TCCON data from Paris (FR), Release GGG2014R0, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.paris01.R0/1149279, 2017.

Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N. M., Petri, C., Grupe, P., Vuillemin, C., Truong, F., Schmidt, M., Ramonet, M., and Parmentier, E.: TCCON data from Orléans (FR), Release GGG2014R1, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.orleans01.R1, 2017.

Wennberg, P. O., Roehl, C., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., Allen, N., and Ayers, J.: TCCON data from Park Falls (US), Release GGG2014R1, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.parkfalls01.R1, 2017a.

Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., and Allen, N.: TCCON data from Caltech (US), Release GGG2014R1, TCCON data archive, CDIAC, https://doi.org/10.14291/tccon.ggg2014.pasadena01.R1/1182415, 2017b.

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, hosted by CDIAC, https://doi.org/10.14291/tccon.ggg2014.lamont01.R1/1255070, 2017c.

Wu, L., Hasekamp, O., Hu, H., Landgraf, J., Butz, A., aan de Brugh, J., Aben, I., Pollard, D. F., Griffith, D. W. T., Feist, D. G., Koshelev, D., Hase, F., Toon, G. C., Ohyama, H., Morino, I., Notholt, J., Shiomi, K., Iraci, L., Schneider, M., de Maz-

ière, M., Sussmann, R., Kivi, R., Warneke, T., Goo, T.-Y., and Té, Y.: Carbon dioxide retrieval from OCO-2 satellite observations using the RemoTeC algorithm and validation with TCCON measurements, Atmos. Meas. Tech., 11, 3111–3130, https://doi.org/10.5194/amt-11-3111-2018, 2018.

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

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W., Sherlock, V., and Wennberg, P. O.: The total carbon column observing network, Philos. T. Roy. Soc. A, 369, 2087–2112, https://doi.org/10.1098/rsta.2010.0240, 2011a.

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, 2011b.

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) X_{CO2} measurements with TCCON, Atmos. Meas. Tech., 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global concentrations of CO_2 and CH_4 retrieved from GOSAT: First preliminary results, Scientific Online Letters on the Atmosphere, 5, 160–163, https://doi.org/10.2151/sola.2009-041, 2009.