



On the potential of a neural-network-based approach for estimating XCO₂ from OCO-2 measurements

François-Marie Bréon, Leslie David, Pierre Chatelanaz, and Frédéric Chevallier

Laboratoire des Sciences du Climat et de l'Environnement/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay,
91191 Gif-sur-Yvette, France

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

Received: 5 October 2021 – Discussion started: 27 October 2021

Revised: 22 July 2022 – Accepted: 1 August 2022 – Published: 15 September 2022

Abstract. In David et al. (2021), we introduced a neural network (NN) approach for estimating the column-averaged dry-air mole fraction of CO₂ (XCO₂) and the surface pressure from the reflected solar spectra acquired by the OCO-2 instrument. The results indicated great potential for the technique as the comparison against both model estimates and independent TCCON measurements showed an accuracy and precision similar to or better than that of the operational ACOS (NASA's Atmospheric CO₂ Observations from Space retrievals – ACOS) algorithm. Yet, subsequent analysis showed that the neural network estimate often mimics the training dataset and is unable to retrieve small-scale features such as CO₂ plumes from industrial sites. Importantly, we found that, with the same inputs as those used to estimate XCO₂ and surface pressure, the NN technique is able to estimate latitude and date with unexpected skill, i.e., with an error whose standard deviation is only 7° and 61 d, respectively. The information about the date mainly comes from the weak CO₂ band, which is influenced by the well-mixed and increasing concentrations of CO₂ in the stratosphere. The availability of such information in the measured spectrum may therefore allow the NN to exploit it rather than the direct CO₂ imprint in the spectrum to estimate XCO₂. Thus, our first version of the NN performed well mostly because the XCO₂ fields used for the training were remarkably accurate, but it did not bring any added value.

Further to this analysis, we designed a second version of the NN, excluding the weak CO₂ band from the input. This new version has a different behavior as it does retrieve XCO₂ enhancements downwind of emission hotspots, i.e., a feature that is not in the training dataset. The comparison against the reference Total Carbon Column Observing Net-

work (TCCON) and the surface-air-sample-driven inversion of the Copernicus Atmosphere Monitoring Service (CAMS) remains very good, as in the first version of the NN. In addition, the difference with the CAMS model (also called *innovation* in a data assimilation context) for NASA Atmospheric CO₂ Observations from Space (ACOS) and the NN estimates is correlated.

These results confirm the potential of the NN approach for an operational processing of satellite observations aiming at the monitoring of CO₂ concentrations and fluxes. The true information content of the neural network product remains to be properly evaluated, in particular regarding the respective input of the measured spectrum and the training dataset.

1 Introduction

There is a growing interest for the monitoring of CO₂ from space. The aim is not so much the atmospheric concentration, which is already known with high accuracy, but rather the CO₂ fluxes. Indeed, there is a need to monitor natural fluxes of CO₂ to better understand their driving factors and to improve land and ocean models (Peylin et al., 2013). There is also a strong societal requirement to monitor the CO₂ anthropogenic emissions at national and more detailed scales. For these objectives, a series of dedicated instruments have been put in orbit since the Greenhouse Gases Observing Satellite (GOSAT, Yokota et al., 2009) and the second Orbiting Carbon Observatory (OCO-2, Eldering et al., 2017), launched in 2009 and 2014, respectively, and still operated at the time of writing. This new and evolving constellation is directly supported by Japanese, US, Chinese, and European space agen-

cies (CEOS Atmospheric Composition Virtual Constellation Greenhouse Gas Team, 2018). The OCO-3 instrument was launched in 2019 and is flying attached to the International Space Station (ISS) with a focus on the imagery of cities and industrial sites (Taylor et al., 2020). These targets are also the main focus of the CO2M mission under development at ESA.

These missions all use the same general principle to estimate the CO₂ concentration in the atmosphere. They measure the reflected solar light at high spectral resolution, which allows identification of absorption lines whose depth is related to the total amount of gas along the atmospheric path (O'Brien and Rayer, 2002). Atmospheric CO₂ shows a number of such lines close to 1.61 and 2.06 μm so that these spectral regions are targeted. Because the absorption is more intense at 2.06 μm , this measurement channel is often referred to as the *strong-CO₂* (or sCO₂) band, whereas the 1.61 μm is the *weak-CO₂* (wCO₂) band. The line depth is also affected by the surface pressure and the number of scattering particles in the atmosphere. To identify and account for their contribution, an additional measurement is made around the oxygen absorption band at 0.76 μm (O₂ band). The combination of these measurements makes it possible to estimate the column-averaged dry-air mole fraction of CO₂, referred to as XCO₂ (Crisp et al., 2004). Note that the MicroCarb instrument, to be launched by CNES in 2022, will have a fourth band at 1.27 μm . This band serves the same purpose as the O₂ band; it has the advantage of being spectrally closer to the CO₂ bands and the disadvantage of being affected by air-glow (Bertaux et al., 2020).

The interpretation of measured spectra in terms of XCO₂ is achieved through *full physics* algorithms that explicitly account for the absorption by CO₂, O₂, and water vapor; for scattering in the atmosphere; and for non-Lambertian reflection on the Earth surface. The modeling must also account for the instrument line shape function and Doppler effects. The inversion process is iterative and starts from a prior estimate of all atmospheric parameters. It is very computer-time-consuming. The processing of OCO-2 data has shown systematic differences between the measured spectra and those modeled after inversion, which led to the development of empirical corrections to the measured spectra (Crisp et al., 2012; O'Dell et al., 2018). In addition, raw XCO₂ 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 XCO₂.

The need for empirical corrections to the full-physics algorithms and the considerable computer load motivated us to develop an alternative approach described in David et al. (2021). We used an artificial network technique (NN) which is purely empirical, without the use of any radiative transfer model. Our hypothesis was that the CAMS (Copernicus Atmosphere Monitoring Service) model constrained by surface air sample measurements provides a fairly accurate

estimate of the atmospheric CO₂ concentration, including the growth rate over multiple years (Chevallier et al., 2019; see also Fig. 8). Indeed, the seasonal cycle of CO₂ together with the growth rate generates a set of XCO₂ samples with a well-known variability. The uncertainties on the modeling (≈ 1 ppm) are small with respect to the range of XCO₂ samples that is available in the multi-year dataset (20 ppm). As a consequence, although CAMS is not the truth, it may be used for supervised learning. Note that other 4D descriptions of the atmospheric composition could have been used for our work. We chose CAMS mostly for practical reasons; the same procedure may be attempted with another modeling dataset.

In practice, we used a series of OCO-2 spectra from a 5-year dataset for the NN training. We then applied the NN to the observations that were not used in the training and compared their estimates to both the same CAMS model used for the training and also the fully independent set of Total Carbon Column Observing Network (TCCON, Wunch et al., 2011a) observations. The results indicated an accuracy and precision that were similar to, if not better than, that of the ACOS algorithm.

More recent results challenged our interpretation of the NN skill. In particular, the XCO₂ estimates of the NN did not show significant enhancement downwind of large power plants, unlike the product of the NASA Atmospheric CO₂ Observations from Space (ACOS) full-physics algorithm. This is shown in the following together with our interpretation. A new version of the NN resulted from this interpretation and retains the high accuracy of the first version, while being much more independent from the training dataset.

In the following, Sect. 2 describes the main characteristic of the NN approach and the training procedure. Section 3 presents the limitation of the first version of the NN, as it shows no innovation with respect to the training dataset. Section 4 describes and justifies a new version of the NN approach. Section 5 discusses the results, suggests directions for improvements, and concludes.

2 Data and method

The NN described in this paper estimates XCO₂ from spectra measured by the OCO-2 satellite over land. Most of the analysis is made with the spectra acquired in nadir mode, but we have also developed a version for glint acquisition that is described and commented on at the end of Sect. 4. Conversely to the analysis in David et al. (2021), we now use all cross-track footprints. A single NN is used to process all footprints even though the spectral elements of different footprints correspond to different sampled wavelengths.

We use spectral samples in the three bands of the instrument (around 0.76, 1.61, and 2.06 μm). They have footprints of $\sim 3 \text{ km}^2$ on the ground. In principle, each band is described by 1016 samples, but some are marked as bad ei-

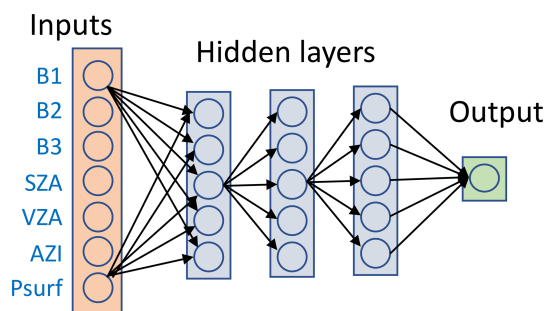


Figure 1. Graphical representation of the NN used in this paper. The outputs from all neurons feed in all neurons of the next layer. There is a variable number of hidden layers. Similarly, there is a choice of the number of neurons in each layer. Not all inputs are used for the various versions of the NN that are described in this paper.

ther because some of the corresponding detectors died at some stage or because of known temporary or permanent issues. We systematically remove 15 spectral samples that are flagged in about 80 % of the spectra and 478 pixels in the band edges. Conversely, we do not remove the samples 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 solar irradiance linked to the sun zenith angle.

Figure 1 offers a graphical representation of the NN. As input, we use the three band spectra (or a subset; see below) and the observation geometry (sun and view zenith angle: SZA and VZA, and relative azimuth: AZI). Some versions also use the surface pressure (P_{surf}) as input. No explicit information is provided to the NN regarding the location or date of the observation. The inputs feed all the neurons of a first “hidden” layer. We use a fully connected neural network, which means that all the neurons are connected to the neurons of the previous and next layer. We have attempted NN versions with a variable number of hidden layers (a single one was used in David et al., 2021). Each neuron computes a weighted sum of the inputs and derives a single output on the basis of either a sigmoid function or a “rectified linear unit”. The loss is derived from the mean absolute error. The weights of the input variables to the neurons are adjusted iteratively with the standard Keras library (Keras Team, 2015) for an optimal agreement between the NN output and a reference.

The NN training is based on OCO-2 radiance measurements (v10r) acquired between February 2015 and December 2019. We make use of XCO₂ estimates and the quality control filters of the ACOS L2Lite v9r products (Eldering et al., 2015): only observations with *xco2_quality_flag* = 0 are used. For the validation of the NN estimates, we also use ob-

servations with relaxed quality requirements. For versions of the NN that use the surface pressure as input, we use the estimate that is provided together with the OCO-2 data, and they are derived from the Goddard Earth Observing System, Version 5, Forward Processing for Instrument Teams (GEOS5-FP-IT) created at Goddard Space Flight Center Global Modeling and Assimilation Office (Suarez et al., 2008; Lucchesi, 2013). The weather model pressures have been adjusted to the sounding surface height.

Our analysis makes use of the CAMS CO₂ atmospheric inversion (Chevallier et al., 2010; version 19r1). This product was released in July 2020 and contributed, e.g., to the Global Carbon Budget 2020 (Friedlingstein et al., 2020). It results from the assimilation of CO₂ surface air sample measurements in a global atmospheric transport model run at spatial resolution of 1.90° in latitude and 3.75° in longitude over the period 1979–2019 and using the adjoint of this transport model. Neither satellite retrievals nor TCCON observations were used for this modeling. For each OCO-2 observation, XCO₂ 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 XCO₂ from CAMS is used for both the training and the evaluation, although using independent datasets: the “training” dataset is a 3 % random sample of the full dataset. The observations that are used for the training are earmarked and not used for further evaluation.

3 Initial results and interpretation

David et al. (2021) described a first version of the NN approach to estimate XCO₂. In this first version, the surface pressure was not used as input, and the training was made on observations acquired during even months, while the validation used observations of the odd months. The results were surprisingly good in that the statistical difference to both the CAMS modeling and the independent TCCON observations indicated an accuracy similar to or better than that of the CAMS product. Further analysis posterior to the publication was worrisome, however.

First, we found that well-documented local enhancements of XCO₂ in the ACOS product (e.g., Nassar et al., 2017; Reuter et al., 2019), also referred to as *plumes*, did not show up in the NN product. We analyzed in particular a case over South Africa acquired on 31 August 2016, an illustration of which is provided in Fig. 2. Over a distance of ≈ 100 km, the ACOS product shows several well-identified enhancements of ≈ 5 ppm, whereas the NN product does not show any significant pattern. The presence of large coal power plants upwind of the OCO-2 observations makes the enhancements trustworthy. We found many similar cases where the NN did not display an XCO₂ plume where ACOS did. We concluded

that the NN did reproduce the seasonal variation in XCO₂ together with the growth rate but was unable to identify small-scale features. Since all observations are processed independently, we could not interpret this apparent incoherence.

Second, we made an experiment where the training dataset is biased by 1 ppm for the observations acquired during a single month (within the full period of 50+ months). When applied to the validation dataset, the differences to CAMS show a bias of ≈ 0.5 ppm, but only for the observations that are within a few weeks of the biased period (Fig. A1). This is rather surprising as the observation date is not an input of the NN. Still, these results provide a clear indication that this version of the NN is somehow sensitive to the observation date.

To investigate the issue, we developed and trained a new NN with the same inputs, but aiming at estimating the date, latitude, and longitude. For the training, we used the true values of these parameters, and we analyzed how the NN was able to make an estimate based on the inputs (the spectra and the observation geometry). Figure 3 shows the histograms of the errors when applied to the independent dataset.

The results indicate that the NN approach is able to make a reasonable estimate of the location and date of the observation based on the spectra and the observation geometry. The standard deviation of the latitude error is on the order of 7°, and there is no significant difference with the footprint. One may expect that this information is largely derived from the observation geometry that changes with the latitude (both the SZA and the azimuth do). One argument in favor of this hypothesis is that the precision of the longitude estimate is much worse, with a standard deviation on the order of 58°. Indeed, for a given day, the observation geometry is nearly the same for all successive orbits; thus, there is no information in the observation geometry to estimate the longitude, while there is such information for the latitude. As for the date, the standard deviation is ≈ 61 d, or 2 months. Clearly then, in the input data of the NN, there is indirect information about the observation date and latitude, and this was a surprise to us. Indeed, when describing the NN approach in David et al. (2021), we argued that the NN had no information on the measurement date, as successive observations from the same day of year and location, but different years, were made with the exact same observation geometry.

The various histograms of Fig. 3 were made using a single (O₂) band, a combination of the O₂ band with either CO₂ band, and all three bands. The most striking difference between the various histograms is for the date estimate. Indeed, the accuracy strongly degrades when the wCO₂ band is not included. The combination of O₂+wCO₂ bands leads to a much better accuracy (a factor of more than 3 on the standard deviation) than that obtained with O₂+sCO₂. The other differences on the histograms are not as large.

How does the NN indirect information on the observation date, and why is this information somehow contained in the wCO₂ band? Our best interpretation is that the weak CO₂

spectrum is sensitive to the upper atmosphere CO₂ concentration that is rather well mixed while increasing regularly in time. The absorption lines in the sCO₂ band are much stronger so that their centers are saturated in the spectra. As a consequence, the CO₂ signal is more in the line wings, which are more sensitive to the higher pressure (lower altitude) levels. The wCO₂ lines are not saturated and the spectrum shape may provide the information for an estimate of the high-altitude CO₂ concentration. We investigated another hypothesis which the wCO₂ detector shows an evolution in time, that could be used by the NN to infer the observation date. However, we did not find any indication of such behavior. Thus, at this point, the stratospheric CO₂ hypothesis is physically plausible and is our best hypothesis because we have no other. Note however that we have investigated the correlation between the longitudinal anomalies of stratospheric CO₂ in the CAMS model and the error on the date estimate by the NN approach. No such correlation was found. Thus, either our hypothesis is wrong or the description of the longitudinal variations in stratospheric CO₂ in CAMS offers a poor representation of the reality. Both hypotheses are plausible.

These results clearly demonstrate that the input data to the NN provide indirect information on the date and latitude. Atmospheric simulations such as those of CAMS indicate that XCO₂ variations are mostly a function of time and latitude. Indeed, on average, the deviations of XCO₂ along the longitudes are on the order of 0.5 ppm (standard deviation). They are however larger (≈ 1 ppm) over the Northern Hemisphere where most of the observations analyzed here are acquired. We hypothesize that our first version of the NN, as published in David et al. (2021), obtains a proxy of the latitude and date and outputs the corresponding CAMS value. Based on the CAMS simulation, we found that the typical uncertainty on the position and date ($\sigma_{\text{lat}} = 7^\circ$, $\sigma_{\text{lon}} = 58^\circ$ and $\sigma_{\text{date}} = 60$ d) leads to a 1σ error of 0.91 ppm on XCO₂ (difference between the values at the true and perturbed location and date). This value appears consistent with the precision obtained with our first version of the NN. Note however that this statistical difference gets larger when considering locations consistent with the OCO-2 observations that are used here. The important point is that the error increases considerably (a factor of 2) for degraded precisions on the location and date with a different version of the NN that is discussed below.

4 A new version of the neural network

As shown above, the NN appears to use the wCO₂ band to derive a proxy of the observation date, which makes it possible, together with the proxies of the location, to estimate XCO₂ based on the statistical distribution of the CAMS XCO₂. To avoid this feature, an option is to not use the information from the wCO₂ band. We therefore developed a similar version of the NN but without this band (i.e., only the O₂ and sCO₂, together with the observation geometry).

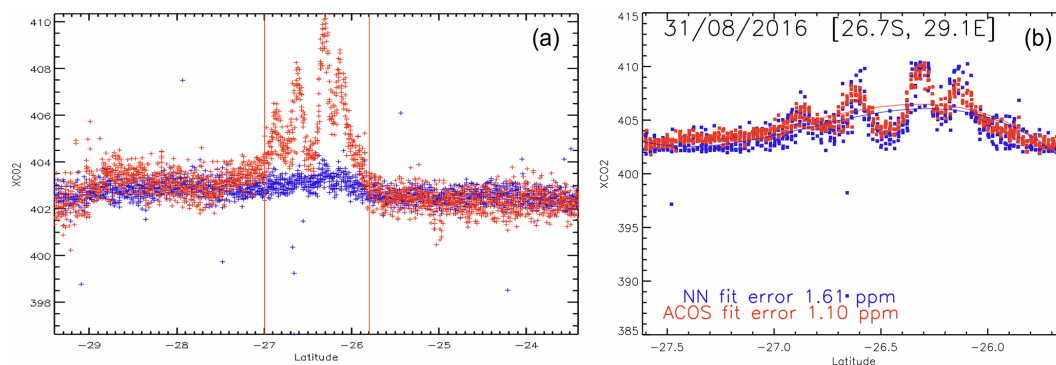


Figure 2. XCO₂ estimated by the ACOS algorithm (red) and the NN approach (blue) as a function of latitude using its initial version as published in David et al. (2021) (a) and the new version presented in this paper (b). The ACOS product showed a number of XCO₂ enhancements that are not shown by the NN estimates. The plumes are observed downwind of large coal power plants, which make these features trustworthy. The date is 31 August 2016.

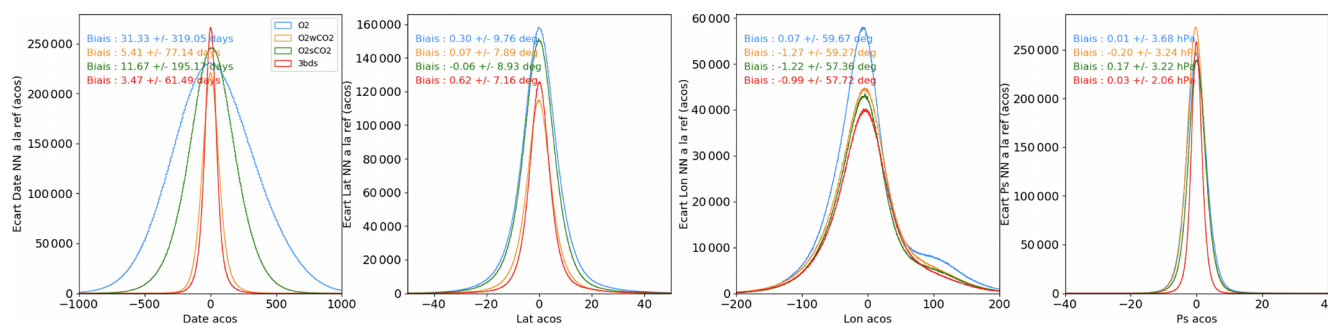


Figure 3. Analysis of the ability of the NN to estimate the date, location (latitude, longitude), and surface pressure from the input spectra and observation geometry. The graphs show the histograms of the differences between the NN estimate and the true value. Several versions of the NN were analyzed using all three bands (red), only the wCO₂ and O₂ bands (orange), sCO₂ and O₂ (green), and only the O₂ band (blue).

With this version, the behavior of the NN changes markedly. The most important feature is that the NN now reproduces the XCO₂ plumes that are shown by the output of the ACOS algorithm. Two representative examples are shown in Fig. 4. These cases demonstrate that the NN does produce XCO₂ features that are not in the training database, as we expected. The NN is trained on the variations in XCO₂ caused by the atmospheric growth rate and the surface flux seasonal cycle. It identifies signatures in the spectra that relate to the CO₂ atmospheric content. These signatures can then be used for an estimate of XCO₂, even for situations that are poorly reproduced in the training dataset.

In addition to the change in the band selection, and posterior to the result shown in Fig. 4, we made several other modifications to the NN algorithm.

1. We decided to use the surface pressure from the weather forecast model as an additional input to the NN. In David et al. (2021), the surface pressure was an output of the NN model. It was used to demonstrate the capability of the NN approach to interpret the spectral shapes in terms of atmospheric parameters. Indeed, the

estimate of the surface pressure could be compared to an independent estimate from numerical weather analyses which are known to be precise within $\approx 1\%$. However, the surface pressure may alternatively provide useful information to the NN for the interpretation of the spectra, as it does in the full-physics algorithms in the form of a prior estimate and also for the derivation of the bias-corrected product.

2. We decided to increase the number of NN hidden layers to five (instead of one in David et al., 2021). Our experience indicates that, with a larger number of layers, there is less over-fitting of the training spectra; i.e., there is a better agreement between the loss of the training and that of the test dataset. An increased number of hidden layers also leads to slightly better performance, in particular for the NN that was designed for the land-glint observations (see below).
3. We developed a similar approach for the glint cases (still over land). Our initial fear was that it would be more difficult for the NN to handle glint observations because

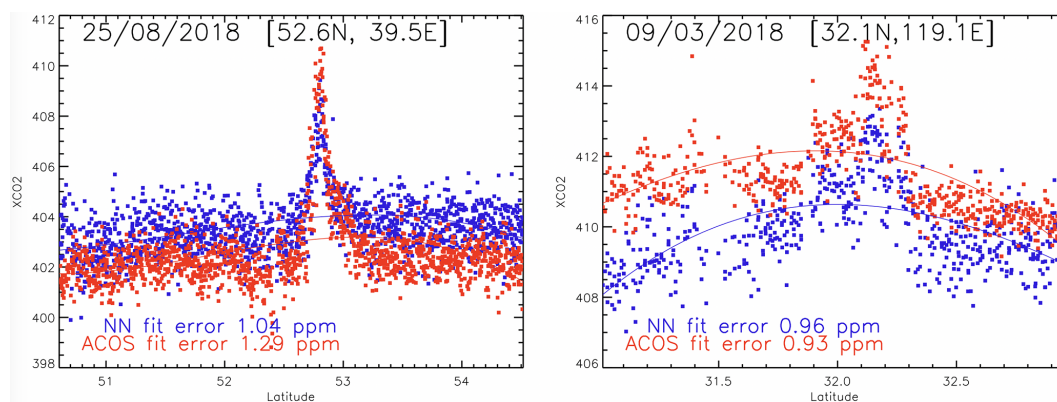


Figure 4. Two examples of XCO₂ plumes that are captured by the ACOS bias-corrected XCO₂ estimates. These were not shown by the first version of the NN algorithm (an example shown in Fig. A1) but are well captured by the second version that does not use the wCO₂ band (shown here). The NN estimates are in blue whereas the ACOS estimates are in red. The lines are simple polynomial fits on the XCO₂ estimates and do not aim at capturing the plume signature. These cases were identified and discussed in Reuter et al. (2019).

of (i) larger variations in the optical path than for the nadir mode and (ii) the Doppler effect that may affect the absorption line positions on the input spectra. This is why our first attempts focused on the nadir cases, but there is a need to also exploit the many observations acquired in glint mode.

Figure 5 shows the inter-comparison of the XCO₂ estimated from CAMS, ACOS, and NN. All three datasets are highly consistent, with a statistical difference around 1 ppm and little bias. Let us recall that there is no satellite data input to the version of CAMS that is used here, so that it is fully independent from ACOS. The 1.06 ppm standard deviation of their differences demonstrates that both product precisions are better than this number. CAMS and the NN are not as independent because the latter is trained with the former (but using different space-time locations). Let us stress that any bias in CAMS may be transferred to the NN product. Thus, a high agreement between CAMS and the NN product is not a demonstration of the latter accuracy. Still, it has been shown that the NN retrieves features that are not in CAMS, which indicates some independence between the satellite product and the model. The standard deviation of their differences is 0.85 ppm. The quadratic difference between NN and ACOS is a strong function of the sCO₂ albedo as shown in Fig. 6: it decreases from ≈ 1.5 to ≈ 0.75 ppm as the sCO₂ band albedo increases from 0.10 to 0.45. A better accuracy of the satellite product with stronger surface albedo is expected as (i) the measurement signal-to-noise ratio gets higher and (ii) the relative contribution of atmospheric scattering to the signal decreases. The precision estimate is also a function of the O₂ band albedo, but this effect is not as strong and the O₂ band albedo shows less variability than that of the sCO₂ band.

Figure 5 also shows that the slope of the best fit of the satellite products against CAMS is close to 1 but with a dif-

ference of opposite sign (0.99 and 1.02). As a consequence, there is a more significant slope deviation from 1 (0.97) between the two satellite products.

Figure 7 provides further information on the differences between the remotely sensed products and the CAMS estimate. The histograms are close to Gaussian and confirm that NN is closer to CAMS than the ACOS counterpart. An interesting feature is that both the NN-CAMS and ACOS-CAMS differences depend on the cloud flag (*cloud_flag_idp*), which indicates that this flag has some value. The difference between the cloud contamination histogram remains small however, and does not deserve to disqualify the observations with a cloud flag of 2. Here, we only use the “definitely clear” and “probably clear” cases (flags of 3 and 2). The population of the lower value cases (“definitely cloudy” and “probably cloudy”) is much smaller, and the histograms for these cases are not shown, while they show further degradation. It is difficult to elaborate further as the true nature of the cloud contamination in the cases classified as probably clear is unknown.

Figure 8 is based on the satellite product innovation, i.e., the difference to the model estimates. Indeed, one may consider that the model provides current knowledge on the XCO₂ distribution, constrained by surface air sample measurements and atmospheric transport. The satellite product has the potential to improve this knowledge, but only as much as the difference with the model estimate. Typical values are around 1 ppm. The interesting result shown by Fig. 8 is that the two satellite estimates are significantly correlated. This provides further evidence that the NN estimate is not only a reconstruction of the training dataset (CAMS) with some noise. Indeed, when NN differs from the model, ACOS, the independent satellite product tends to agree.

Finally, Fig. 9 shows a comparison of the model and remotely sensed estimates of XCO₂ against the reference retrievals of the TCCON network. Although the OCO-2 satel-

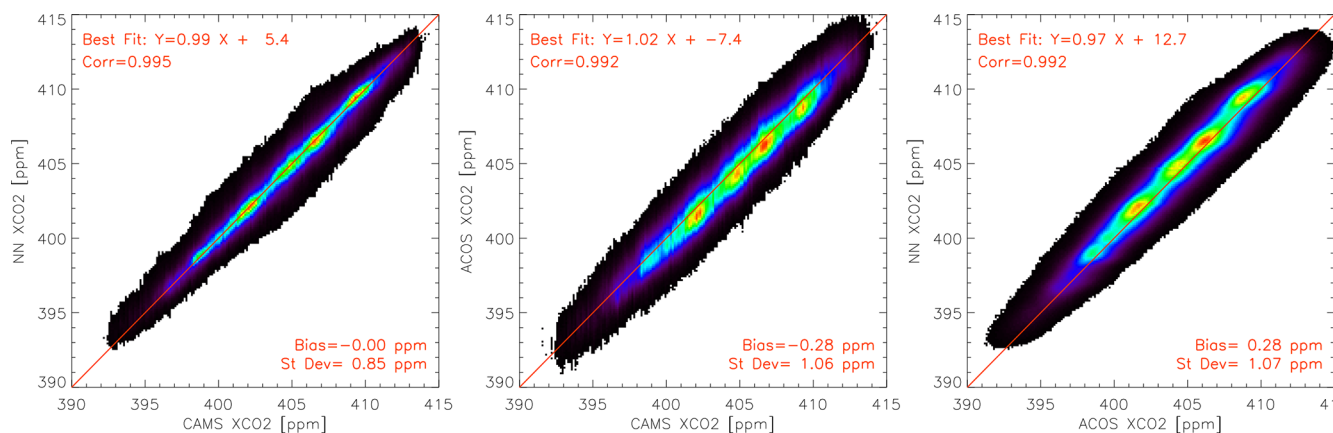


Figure 5. Inter-comparison of XCO₂ estimated from CAMS, ACOS, and the NN. The density histogram is based on nadir observations from February 2015 to December 2019. A similar figure for the glint cases is shown in Fig. A3.

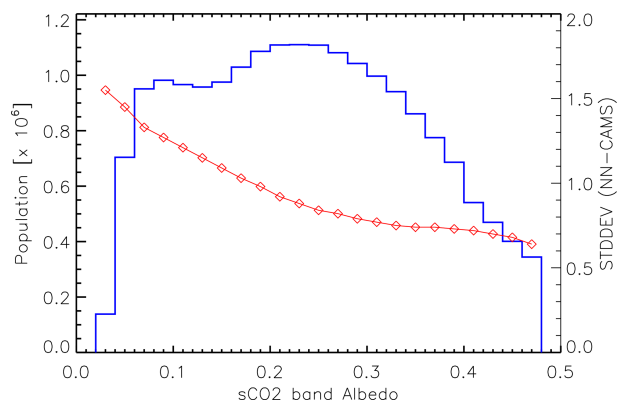


Figure 6. Standard deviation of the NN–CAMS difference as a function of the sCO₂ band albedo (red, right scale). The computation is made over 0.02 bins, the population of which is shown by the blue line (left scale).

lite platform can be oriented so that the instrument field of view is close to the surface station, we only use nadir data here. Indeed, the NN was not trained on the target data and can only be used to process measurements that have been acquired in observation configurations that are similar to those of the training. We thus have to rely on nadir or glint 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. For the reference, we average the TCCON estimates of XCO₂ 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. Statistics per station are provided in Table 1. The biases vary significantly among stations, although they are generally less than 1 ppm (in amplitude). Two stations, Pasadena and Zugspitze, show a large negative bias for both satellite estimates and the model. For Pasadena, it may be interpreted as the impact of the city on

the atmosphere sampled by the TCCON measurement, while the atmosphere at the location of the satellite observation (which may be several hundred kilometers away) is less affected. Zugspitze is a high-altitude site (2960 m), so that the atmospheric column sampled by the sun photometer does not have the same vertical representativeness as that of the satellite observation (in addition to the spatial distance that is common with other sites). A large negative bias is also found at Eureka (80.05° N). The fact that the difference with the CAMS model at this site is much larger than for other sites could hint at an issue in the sun photometer product there. Conversely, there are large positive biases at Burgos and Ny-Ålesund (78.9° N, very close to the latitude of Eureka). Since the model and satellite estimates somewhat agree, one may also question the TCCON calibration at these sites. For other stations, which form the large majority, the biases are smaller than 1 ppm, and there is a fair consistency between the satellite products in the sense that the sign of their bias is the same in most cases. The range of the difference with TCCON varies among stations. The best satellite–TCCON agreement is found at the Lamont station, which, interestingly, is also the one with the most coincidences. Excellent agreement is also seen at Darwin, Edwards, Park Falls, and Bremen. The comparison with TCCON does not allow favoring of one satellite estimate versus the other. Focusing on the stations with a large number of observation (25 overpasses or more), the NN estimates appear slightly better than ACOS at Darwin, Edwards, Garmisch, Orléans, and Białystok, while they are the opposite at Saga, Park Falls, and Sodankylä. The figure (and table) also clearly shows that the CAMS product offers a better agreement with the TCCON data than any of the satellite estimates in most cases. The high quality of the CAMS modeling used in this paper, at least over the TCCON site, provides further justification of its use as a training dataset.

We have applied a very similar procedure to the OCO-2 observations acquired in glint mode over land. An evaluation

Table 1. TCCON stations used in this paper (Figs. 8 and A4). The data have been obtained from the <https://iccondata.org/web/site/on4/February2021>.

Stations	lat, long	Alt [m]	Reference	Nadir						Glint					
				CAMS/ACOS/NN			CAMS/ACOS/NN			CAMS/ACOS/NN			CAMS/ACOS/NN		
				North/Nobs	Biases	SD	North/Nobs	Biases	SD	North/Nobs	Biases	SD	North/Nobs	Biases	SD
Lauder	−45.04, 169.68	370	Sherlock et al. (2017)	11/855	0.4	0.5	1.1	0.2	1.2	1.0	53/5508	0.4	0.5	0.8	0.3
Wollongong	−34.41, 150.88	30	Griffith et al. (2017b)	38/8687	0.1	−0.2	0.2	0.7	1.3	1.3	23/4059	−0.6	−0.9	−0.7	1.2
Réunion Island	−20.90, 55.49	90	De Mazière et al. (2017)	6/126	0.2	1.0	0.6	0.3	1.4	1.2	6/123	−0.8	−0.5	−1.2	0.6
Darwin	−12.43, 130.89	30	Griffith et al. (2017a)	26/6940	0.4	−0.1	0.4	0.7	1.1	1.0	28/7432	0.4	−0.2	0.4	0.6
Burgos	18.53, 120.65	35	Velazco et al. (2017)	4/104	1.3	2.0	1.5	0.7	1.3	1.0	9/789	0.3	0.6	0.2	0.4
Izana	28.3, −16.48	2300	Blumenstock et al. (2017)	15/211	−0.5	0.6	0.6	0.8	1.3	1.1	6/183	−0.1	1.0	0.5	0.6
Hefei	31.90, 118.67	30	Liu et al. (2018)	6/545	−0.2	0.3	0.3	1.0	2.1	2.0	9/1581	0.3	−0.5	−0.1	1.1
Saga	33.24, 130.29	10	Kawakami et al. (2014)	25/1968	−0.4	−0.4	−0.3	0.8	1.5	1.6	33/2934	−0.3	−0.5	−0.8	0.6
Pasadena	34.14, −118.13	240	Wennerg et al. (2017b)	65/19041	−1.8	−1.7	−1.6	1.1	1.5	1.5	51/13589	−1.6	−1.2	−1.5	1.1
Edwards	34.96, −117.88	700	Iraci et al. (2017)	64/23494	0.3	0.5	0.5	0.6	1.2	1.0	70/25890	0.3	0.6	0.1	0.7
Tsukuba	36.05, 140.12	30	Morino et al. (2017a)	30/3602	0.6	0.5	0.9	1.1	1.9	1.7	24/1535	−0.1	−0.1	−0.2	0.9
Ameyondo	36.54, 126.33	30	Guo et al. (2014)	9/1177	−0.2	0.1	−0.6	0.7	1.4	1.5	1/396	0.2	−0.2	−0.3	0.2
Lamont	36.6, −97.49	320	Wennerg et al. (2017c)	94/44404	−0.0	−0.1	0.3	0.9	1.0	1.0	96/41423	0.1	−0.0	0.0	0.6
Rikubetsu	43.46, 1473.77	390	Morino et al. (2017b)	20/1988	0.1	−0.2	0.2	0.9	1.6	1.3	17/2192	0.5	0.3	0.3	0.5
Park Falls	45.94, −90.27	440	Wennerg et al. (2017a)	59/7560	−0.0	−0.2	0.3	0.8	1.0	1.2	45/5150	−0.3	−0.3	−0.4	1.0
Zugspitze	47.42, 11.06	2960	Sussmann and Rettinger (2017b)	18/846	−2.0	−2.3	−1.4	1.1	1.8	1.6	9/1404	−1.0	−0.2	−0.2	1.4
Garmisch	47.48, 11.06	740	Sussmann and Rettinger (2017a)	41/7745	0.3	0.4	0.8	0.9	1.5	1.5	38/7248	0.3	0.7	0.2	0.9
Orléans	47.97, 2.11	130	Warneke et al. (2017)	49/17239	0.8	0.3	0.6	0.5	1.2	1.1	34/10418	0.8	0.6	0.5	0.7
Paris	48.85, 2.36	60	Té et al. (2017)	12/3614	−0.0	−0.5	−0.0	0.9	1.3	1.2	14/3421	−0.6	−0.4	−0.9	1.6
Karlsruhe	49.1, 8.44	110	Hase et al. (2017)	38/11363	0.6	0.6	0.8	0.7	1.4	1.3	45/8883	0.4	0.6	0.1	0.6
Bremen	53.10, 8.85	7	Notholt et al. (2017)	14/2866	0.7	0.3	0.9	0.4	0.9	1.1	9/1415	1.1	0.9	0.6	1.0
Bratsk	53.23, 23.02	180	Deutscher et al. (2017)	34/4575	0.3	−0.0	0.8	0.7	1.3	1.1	31/8326	0.4	0.2	0.4	0.3
Sodankylä	67.37, 26.63	190	Kivi et al. (2017)	29/3508	0.4	0.2	0.3	0.8	1.3	1.5	21/4698	0.8	0.8	0.3	1.2
Ny-Ålesund	78.9, 11.9	20	Notholt et al. (2019)	10/65	1.2	2.1	2.2	0.6	2.5	2.3	7/18	1.0	3.1	1.4	0.7
Eureka	80.05, −86.42	600	Strong et al. (2017)	3/42	−1.4	−1.3	−0.5	1.0	2.0	1.9	9/429	0.6	−0.1	−0.8	1.5

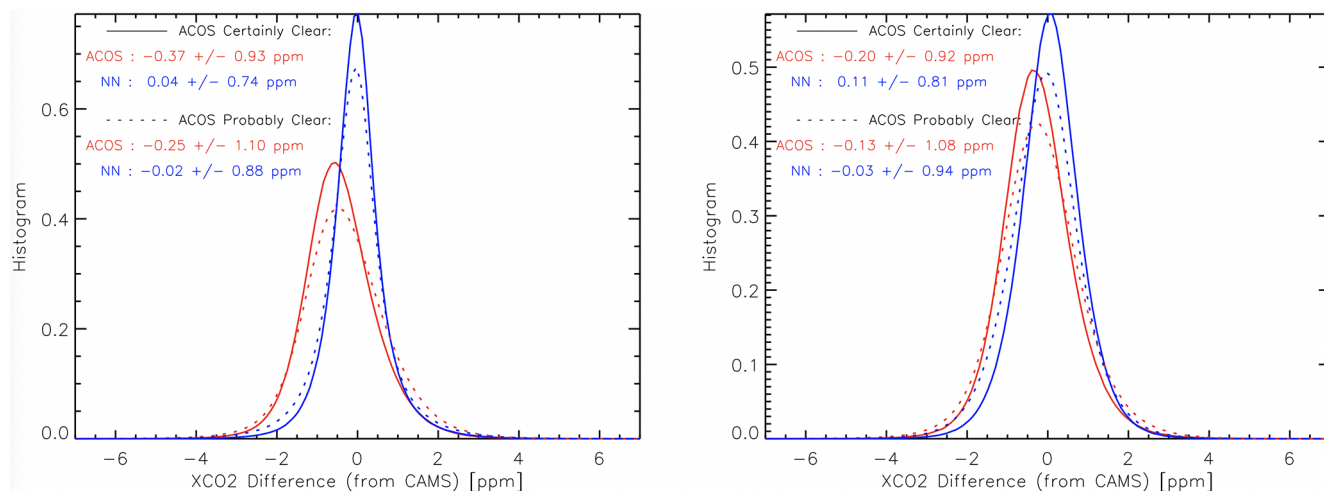


Figure 7. Histogram of the differences between either one of the two satellite datasets and the CAMS model. We distinguish cases when the flag *cloud_flag_idp* is “certainly clear” and “probably clear”. The left figure is for the nadir dataset, whereas the right figure is for the glint.

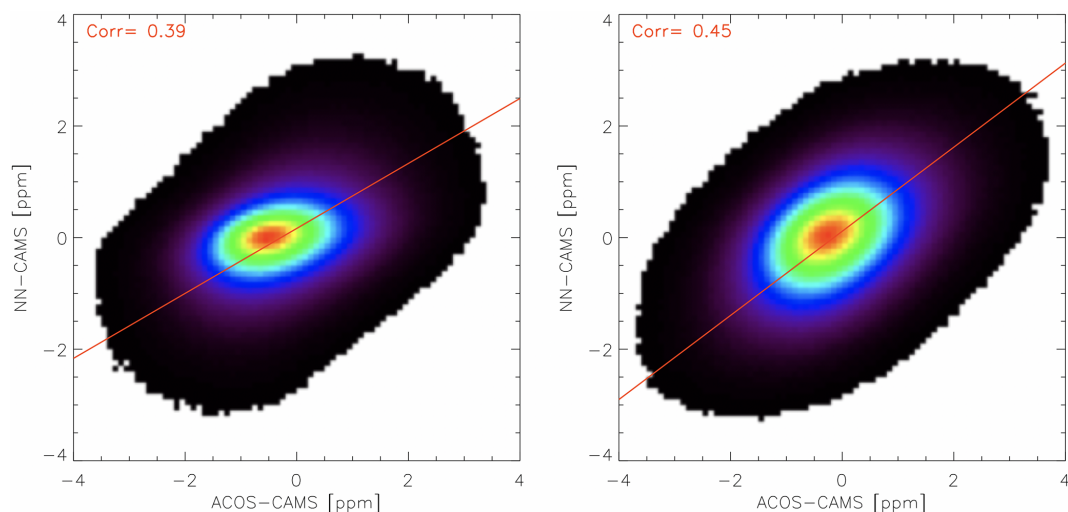


Figure 8. Density histogram of the innovation, i.e., the difference between the satellite product and the model estimates, differences between either one of the two satellite datasets and the CAMS model. The red line shows the result of a linear fit through the data points aiming at a minimization of the distance to the best line. The left figure is for the nadir dataset, whereas the right figure is for the glint.

of the estimate performance is shown in Figs. 7, 8, A2, and A3. The conclusions are very similar to those obtained for nadir. The agreement with CAMS is slightly degraded with respect to the nadir cases (0.92 vs. 0.85 ppm for the certainly clear observations) but somewhat closer than that of ACOS (Figs. 7 and A2). The deviations from the model of the two satellite estimates are significantly correlated, and the correlation coefficient is even larger than that derived for nadir observations (0.45 vs. 0.39, Fig. 8). The comparison with the TCCON estimates leads to the same conclusions as those described above for the nadir cases.

5 Discussion and conclusion

This paper follows on from David et al. (2021), in which we described a neural-network-based technique to estimate XCO₂ and the surface pressure from the OCO-2 spectral measurements. An important message is that our interpretation of the results in that earlier study was incorrect. The NN developed in that paper reproduced the statistical variations of the training dataset (CAMS) and was unable to generate features, such as plume from emission hot-spots. Thus, contrary to our claims, the NN method, as presented in that paper, could not be used to process OCO-2 and generate XCO₂ estimates with any real value. We have shown here that a NN-based procedure is able to estimate the latitude and date

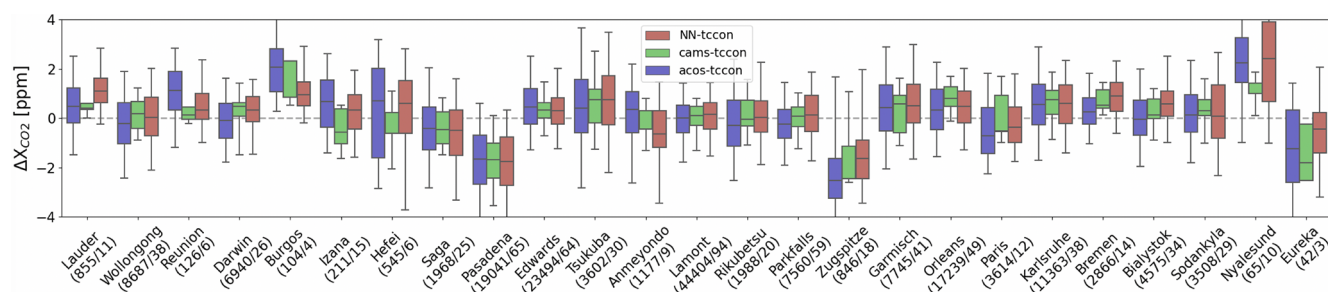


Figure 9. Statistics of the differences between the NN retrieval (red), the CAMS model (green), or the bias-corrected ACOS retrievals (blue) and the TCCON retrievals. The boxes indicate the 25 %–75 % percentiles, and the median is shown by the horizontal line within the box. The whiskers indicate the 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 of the various TCCON observations are provided in Table 1. Figure A4 provides similar results for the glint case.

of the observation with a reasonable accuracy. This was unexpected as we wrote in David et al. (2021)

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.

Our interpretation was wrong. In fact, the NN input can somehow be used by the NN for a fairly accurate estimate of the latitude and date. Because most XCO₂ variations are a function of latitude and date, this information could be used by the NN to generate a reasonable estimate, i.e., one that mimics the main variations in the training dataset.

A question remains on the indirect information that is used by the NN to estimate the observation date. The fact that the precision on the date estimate is much better when using a combination of the O₂+wCO₂ rather than of the O₂+sCO₂ suggests that the information lies in the wCO₂ band. Our best hypothesis is that the wCO₂ spectrum contains some information on the stratospheric CO₂ whose concentration is well-mixed while increasing regularly with time and implicitly contains, therefore, information on the observation date. Further testing this hypothesis would require, for instance, the identification of some anomaly in the stratospheric CO₂ (linked to a specific atmospheric circulation) that would show up as a significant error on the date estimate made by the NN. We have not been able to identify such a feature.

Despite this initial setback, we have continued our analysis of the potential of the NN to process the OCO-2 spectra. A strong motivation relied on the results obtained for the estimate of the surface pressure. Indeed, David et al. (2021) showed that the NN could estimate the surface pressure with an accuracy on the order of 3 hPa. The spatial and temporal variations in the surface pressure, at the scale of the potential accuracies on the date and location, are much larger than this number, so that the NN estimate cannot rely on this kind of

indirect information. This provided a strong indication that the NN method has the potential to extract meaningful information from the spectrum itself.

We have therefore developed a new version of the NN excluding the wCO₂ band from the inputs. In this version, the behavior of the NN is much different from the earlier version as it generates features that are not in the training dataset. This clearly shows that the NN uses the signature of XCO₂ contained in the sCO₂ spectra to make an XCO₂ estimate. The accuracy of this estimate is similar to the one obtained with the first version of the NN and similar to that of the ACOS products. This is confirmed by the comparison of the XCO₂ estimates against the TCCON retrievals. Another strong argument that the NN XCO₂ estimate contains true information and is not only a noisy copy of the training dataset is that the innovations of the two satellite estimates, i.e., the differences to the model data, are significantly correlated (Fig. 8).

Note that we use here a single neural network for the eight footprints of the OCO-2 instrument. We analyzed whether the result performance, assessed as the standard deviation of the differences with CAMS, is a function of the footprint. The statistics are very similar for all, except for footprint 2, which shows slightly higher deviation for both the ACOS and the NN satellite products (a difference of ≈ 0.1 ppm to the mean of ≈ 1 ppm).

These results confirm that the NN technique has a strong potential to process the OCO-2 observations, as well as those from forthcoming missions aiming at the observation of CO₂ from space such as the forthcoming MicroCarb (Pascal et al., 2017) or the CO2M constellation (Sierk et al., 2019). As discussed above, the current version does not use the wCO₂ band at all, and this may be seen as a loss of useful information. There is therefore a need to select appropriate spectral samples in the wCO₂ band rather than discarding them all. It requires improved understanding of the indirect information that is used by the NN to estimate the observation date and location.

The NN technique has two obvious advantages compared to the physical methods that are used to process the OCO-2 observations as well as other instruments with similar objectives: (i) a much smaller computational burden and (ii) no need for a de-bias procedure (O'Dell et al., 2018; Kiel et al., 2019). Our implementation still faces remaining challenges, which we discussed in David et al. (2021).

The first challenge is the cloud detection. All the analysis described in this paper relies on the ACOS cloud detection, and only the observations identified as clear are processed. Our analysis demonstrates the potential of the NN approach but is currently not independent from ACOS. We are currently evaluating independent approaches for the cloud detection. Although the NN described here aims at an estimate of XCO₂, we have shown earlier that the same tool can be used for an estimate of the surface pressure with a 1σ precision on the order of 3 hPa for clear-sky cases. Numerical weather analyses are actually better than that (Salstein et al., 2008). Thus, one may use the comparison of the surface pressure estimate from the NN to the numerical weather data for an easy identification of perturbations to the spectra that are linked to cloud or large aerosol contamination. This would allow an easy and rapid quality indicator for the selection of observations that may be used for XCO₂ estimates, either using a physics-based algorithm or a NN approach. This idea remains to be evaluated.

The second challenge concerns the absence of a quantitative indication of the amount of information that the NN takes from its prior information (contained in the training database) vs. the amount of information that the NN takes from the measured spectra. For Bayesian full-physics retrievals, these weights are represented by the *averaging kernel* (Rodgers, 1990), which allows a clean comparison of each retrieval with 3D atmospheric models, at least in theory (see the discussion about the practical difficulties in Chevallier, 2015). The NN training targets the CO₂ column with a homogeneous weighting along the vertical, but this can hardly be achieved without some contribution from the prior information. This challenge may be evaluated in the future on the basis of radiative transfer simulations.

The third challenge concerns the absence of a quality indicator with the XCO₂ estimate. With the physical methods, the spectrum residuals provide an efficient means to identify cases when no satisfactory agreement can be found between the measured and modeled spectra. With the NN approach, there is no uncertainty associated with each retrieval. Our analysis has shown that the apparent precision (evaluated against CAMS) is a strong function of the surface albedo. There may be other geophysical variables that pilot the uncertainty. To provide precision estimates for each NN-based XCO₂ estimate, ensembles of randomized trainings, where uncertain parameters or input/output variables are varied adequately (e.g., Chau et al., 2022), or analytical estimates (Aires et al., 2004) should be explored.

The last challenge concerns the need for a high-quality training dataset, in the context of increasing XCO₂. The comparison against the TCCON observations (Fig. 8) demonstrates that the CAMS inversion product meets this requirement. In fact, there are strong indications that CAMS remains better than the satellite products, at the very least in terms of global precision. However, because of the atmospheric growth rate of CO₂, the training must be regularly updated. Indeed, with a frozen training dataset, the true real-time XCO₂ progressively leaves the training range. The NN approach requires a training dataset that is representative of the observation and would then lead to underestimates. For quasi-near-real-time data assimilation (e.g., Massart et al., 2016), the training dataset must therefore gradually integrate recent high-quality XCO₂ data, but without sacrificing robustness.

As a final remark, we call for caution. We have been tricked by the NN ability to generate a consistent description of the atmospheric XCO₂ in our first analysis. It is difficult to ensure that we are not tricked again. The source of the information that leads to a fairly accurate estimate of the date, when using the weak CO₂ band, remains unclear. As a consequence, although it is demonstrated that the new version of the NN generates structures that are not in the training dataset, there may be biases in the CAMS modeling that have a significant influence on the NN product.

Appendix A

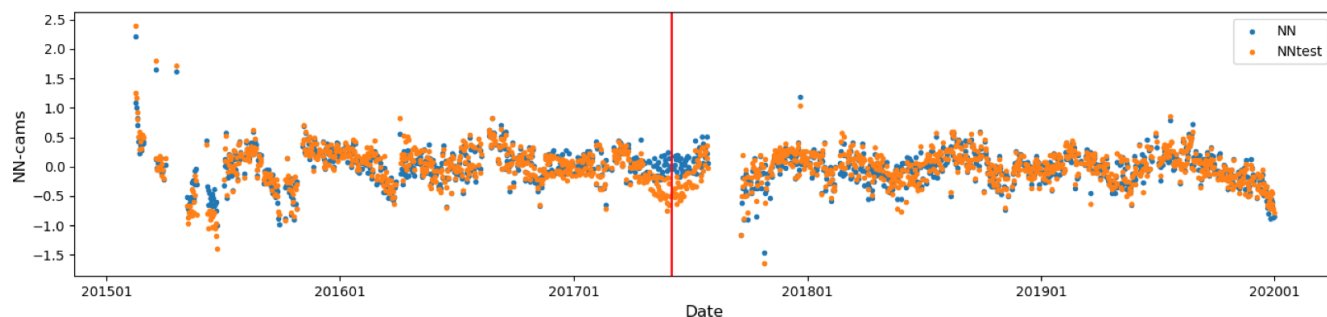


Figure A1. Mean difference, at daily scale, between the NN XCO₂ estimate and the CAMS model. The blue dots show the results for the nominal training. The orange dots show the results when the training was made with a dataset biased by 1 ppm but only the observations of June 2017, the center of which is indicated by the red vertical line.

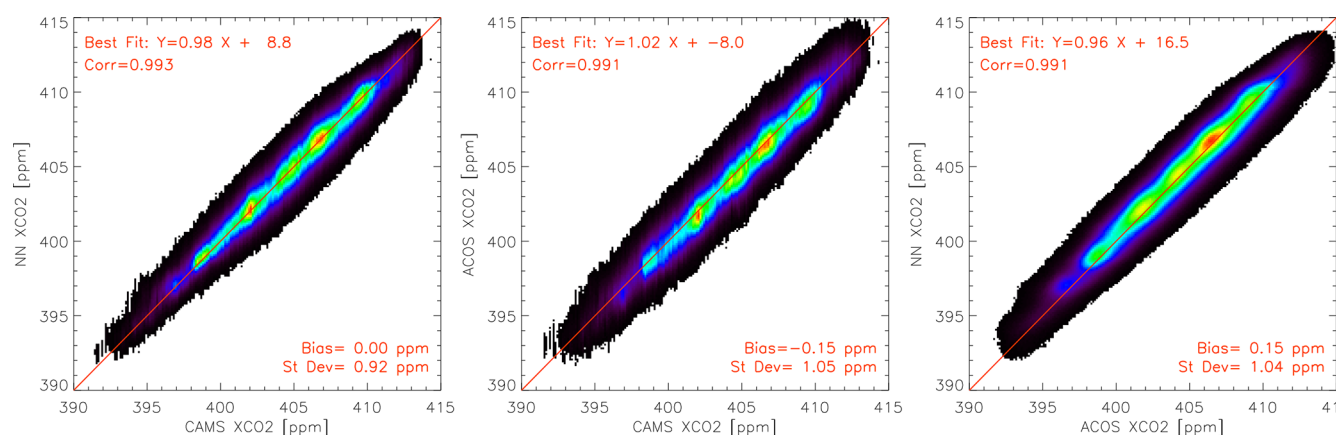


Figure A2. Same as Fig. 5, but for the glint observations.

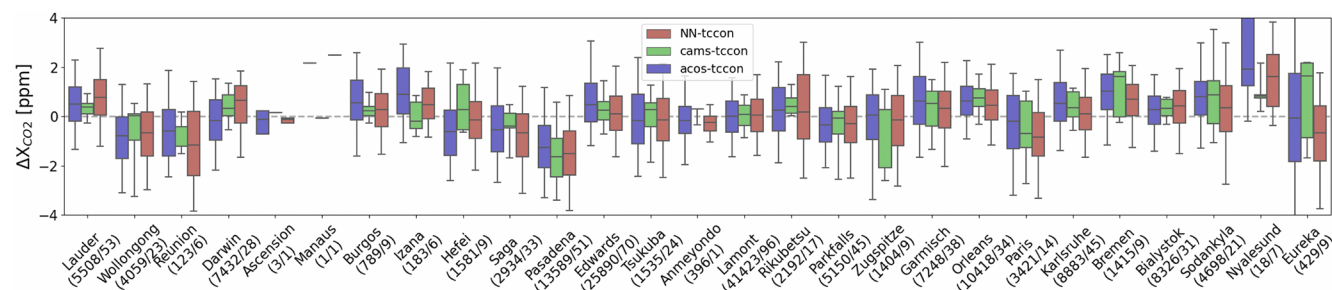


Figure A3. Same as Fig. 9, but for the glint observations.

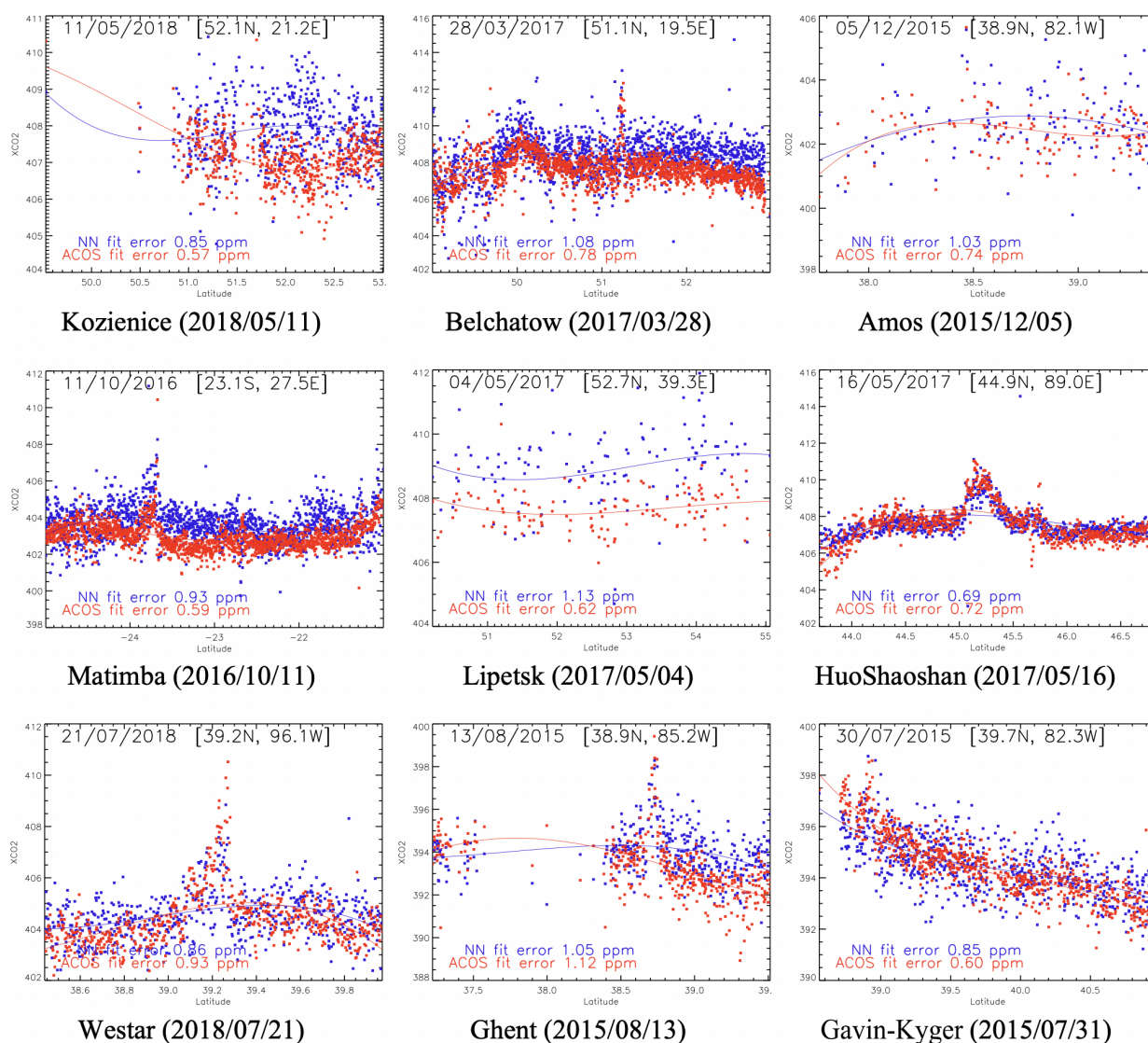


Figure A4. Examples of anthropogenic CO₂ plumes as seen by the OCO₂ instrument processed with the ACOS algorithm (red) and the neural network described in this paper (blue). The cases have been identified and described in Reuter et al. (2019) and Nassar et al. (2021).

Code and data availability. The codes used in this paper and the CAMS model simulations are available, upon request, from the author. The OCO-2 can be downloaded from the NASA OCO-2 archive depository (<https://disc.gsfc.nasa.gov/datasets?keywords=oco-2>, last access: 1 February 2022), and TCCON data can be downloaded from the TCCON Data Archive (<https://tccondata.org/>, last access: 4 February 2021).

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

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. This work was in part funded by CNES, the French space agency, in the context of the preparation for the MicroCarb mission, and, to a smaller extent, by the Copernicus Atmosphere Monitoring Service, implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Commission.

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

– <https://tccondata.org/> (last access: 4 February 2021). We warmly thank those who made these data available.

Review statement. This paper was edited by Folkert Boersma and reviewed by Christopher O'Dell and Sihe Chen.

References

- Aires, F., Prigent, C., and Rossow, W. B.: Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 2. Output errors, *J. Geophys. Res.*, 109, D10304, <https://doi.org/10.1029/2003JD004174>, 2004.
- Bertaux, J.-L., Hauchecorne, A., Lefèvre, F., Bréon, F.-M., Blanot, L., Jouglet, D., Lafrique, P., and Akaev, P.: The use of the 1.27 μm O₂ absorption band for greenhouse gas monitoring from space and application to MicroCarb, *Atmos. Meas. Tech.*, 13, 3329–3374, <https://doi.org/10.5194/amt-13-3329-2020>, 2020.
- Blumenstock, T., Hase, F., Schneider, M., García, O. E., and Sepúlveda, E.: TCCON data from Izana (ES), Release GGG2014R1, CaltechDATA [data set], <https://doi.org/10.14291/tccon.ggg2014.izana01.R1>, 2017.
- CEOS Atmospheric Composition Virtual Constellation Greenhouse Gas Team: A Constellation Architecture for Monitoring Carbon Dioxide and Methane from Space, Tech. Rep., University of Zurich, Department of Informatics, http://ceos.org/document_management/Meetings/Plenary/32/documents/CEOS_AC_VC_White_Paper_Version_1_20181009.pdf (last access: 18 October 2019), 2018.
- Chau, T. T. T., Gehlen, M., and Chevallier, F.: A seamless ensemble-based reconstruction of surface ocean $p\text{CO}_2$ and air–sea CO₂ fluxes over the global coastal and open oceans, *Biogeosciences*, 19, 1087–1109, <https://doi.org/10.5194/bg-19-1087-2022>, 2022.
- Chevallier, F.: On the statistical optimality of CO₂ atmospheric inversions assimilating CO₂ column retrievals, *Atmos. Chem. Phys.*, 15, 11133–11145, <https://doi.org/10.5194/acp-15-11133-2015>, 2015.
- 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.: CO₂ surface fluxes at grid point scale estimated from a global 21 year re-analysis of atmospheric measurements, *J. Geophys. Res.-Atmos.*, 115, D21307, <https://doi.org/10.1029/2010JD013887>, 2010.
- Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., and Cozic, A.: Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions, *Atmos. Chem. Phys.*, 19, 14233–14251, <https://doi.org/10.5194/acp-19-14233-2019>, 2019.
- 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 CO₂ retrieval algorithm – Part II: Global XCO₂ data characterization, *Atmos. Meas. Tech.*, 5, 687–707, <https://doi.org/10.5194/amt-5-687-2012>, 2012.
- David, L., Bréon, F.-M., and Chevallier, F.: XCO₂ estimates from the OCO-2 measurements using a neural network approach, *Atmos. Meas. Tech.*, 14, 117–132, <https://doi.org/10.5194/amt-14-117-2021>, 2021.
- De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scollas, F., Kumps, N., Metzger, J.-M., Duflo, V., and Cammas, J.-P.: TCCON data from Reunion Island (RE), Release GGG2014.R0, CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.REUNION01.R0/1149288>, 2017.
- Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., Grupe, P., and Katrynski, K.: TCCON data from Bialystok (PL), Release GGG2014R2, CaltechDATA [data set], <https://doi.org/10.14291/tccon.ggg2014.bialystok01.R2>, 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, http://disc.sci.gsfc.nasa.gov/OCO-2/documentation/oco-2-v7/OCO2_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., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneeth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metz, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.:

- Global Carbon Budget 2020, *Earth Syst. Sci. Data*, 12, 3269–3340, <https://doi.org/10.5194/essd-12-3269-2020>, 2020.
- Goo, T.-Y., Oh, Y.-S., and Velazco, V. A.: TCCON data from Anmeyondo (KR), Release GGG2014.R0, CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.ANMEYONDO01.R0/1149284>, 2014.
- Griffith, D. W., Deutscher, N. M., Velazco, V. A., Wennberg, P. O., Yavin, Y., Aleks, G. K., Washenfelder, R. a., Toon, G. C., Blavier, J.-F., Murphy, C., Jones, N., Kettlewell, G., Connor, B. J., Macatangay, R., Roehl, C., Ryzek, M., Glowacki, J., Culf, T., and Bryant, G.: TCCON data from Darwin (AU), Release GGG2014R0, CaltechDATA [data set], <https://doi.org/10.14291/tcon.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 Riggensbach, M.: TCCON data from Wollongong (AU), Release GGG2014R0, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.wollongong01.R0/1149291>, 2017b.
- Hase, F., Blumenstock, T., Dohe, S., Gross, J., and Kiel, M.: TCCON data from Karlsruhe (DE), Release GGG2014R1, CaltechDATA [data set], <https://doi.org/10.14291/tcon.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, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.edwards01.R1/1255068>, 2017.
- Keras Team: Keras, GitHub, <https://github.com/fchollet/keras> (last access: 10 January 2021), 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 XCO₂ 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 Sodankylä (FI), Release GGG2014R0, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.sodankyla01.R0/1149280>, 2017.
- Liu, C., Wang, W., and Sun, Y.: TCCON data from Hefei (PCR), Release GGG2014R0, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.hefei01.R0>, 2018.
- Lucchesi, R.: File Specification for GEOS-5 FP-IT, GMAO Office Note No. 2 (Version 1.2), 60 pp., http://gmao.gsfc.nasa.gov/pubs/office_notes (last access: 4 December 2018), 2013.
- Massart, S., Agustí-Panareda, A., Heymann, J., Buchwitz, M., Chevallier, F., Reuter, M., Hilker, M., Burrows, J. P., Deutscher, N. M., Feist, D. G., Hase, F., Sussmann, R., Desmet, F., Dubey, M. K., Griffith, D. W. T., Kivi, R., Petri, C., Schneider, M., and Velazco, V. A.: Ability of the 4-D-Var analysis of the GOSAT BESD XCO₂ retrievals to characterize atmospheric CO₂ at large and synoptic scales, *Atmos. Chem. Phys.*, 16, 1653–1671, <https://doi.org/10.5194/acp-16-1653-2016>, 2016.
- Morino, I., Matsuzaki, T., and Shishime, A.: TCCON data from Tsukuba (JP), 125HR, Release GGG2014R2, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.tsukuba02.R2>, 2017a.
- Morino, I., Yokozeki, N., Matzuzaki, T., and Shishime, A.: TCCON data from Rikubetsu (JP), Release GGG2014R2, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.rikubetsu01.R2>, 2017b.
- NASA OCO-2 archive depository: <https://disc.gsfc.nasa.gov/datasets?keywords=oco-2>, last access: 1 February 2022.
- Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D., and Crisp, D.: Quantifying CO₂ emissions from individual power plants from space, *Geophys. Res. Lett.*, 44, 10045–10053, <https://doi.org/10.1002/2017GL074702>, 2017.
- Nassar, R., Mastrogiacomio, J.-P., Bateman-Hemphill, W., McCracken, C., MacDonald, C. G., Hill, T., O'Dell, C. W., Kiel, M., and Crisp, D.: Advances in quantifying power plant CO₂ emissions with OCO-2, *Remote Sens. Environ.*, 264, 112579, <https://doi.org/10.1016/j.rse.2021.112579>, 2021.
- 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, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.bremen01.R0/1149275>, 2017.
- Notholt, J., Warneke, T., Petri, C., Deutscher, N. M., Weinzierl, C., Palm, M., and Buschmann, M.: TCCON data from Ny Ålesund, Spitsbergen (NO), Release GGG2014.R1, Version R1, CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.NYALESUND01.R1>, 2019.
- O'Brien, D. M. and Rayner, P. J.: Global observations of the carbon budget 2. CO₂ column from differential absorption of reflected sunlight in the 1.61 µm band of CO₂, *J. Geophys. Res.*, 107, 4354, <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., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Té, 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, *P. SPIE*, 10563, 10563K-1–10563K-9, <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-Luijckx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO₂ inversions, *Biogeosciences*, 10, 6699–6720, <https://doi.org/10.5194/bg-10-6699-2013>, 2013.
- Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., Bovensmann, H., and Burrows, J. P.: Towards monitoring localized CO₂ emissions from space: co-located regional CO₂ and NO₂ enhancements observed by the OCO-2 and SSP satellites, *Atmos. Chem. Phys.*, 19, 9371–9383, <https://doi.org/10.5194/acp-19-9371-2019>, 2019.

- Rodgers, C. D.: Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, 95, 5587–5595, 1990.
- 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, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.lauder02.R0/1149298>, 2017.
- Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Taura, C., Fukamachi, T., and Sakashita, M.: TCCON data from Saga (JP), Release GGG2014R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.SAGA01.R0/1149283>, 2014.
- Sierk, B., Bézy, J.-L., Löscher, A., and Meijer, Y.: The European CO₂ Monitoring Mission: observing anthropogenic greenhouse gas emissions from space, *Proc. SPIE*, 11180, <https://doi.org/10.1117/12.2535941>, 2019.
- 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, CaltechDATA [data set], <https://doi.org/10.14291/tcon.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, Tech. rep. 20120011955, NASA Goddard Spaceflight Center, Greenbelt, MD, USA, <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, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.garmisch01.R2>, 2017a.
- Sussmann, R. and Rettinger, M.: TCCON data from Zugspitze (DE), Release GGG2014R1, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.zugspitze01.R1>, 2017b.
- Té, Y., Jeseck, P., and Janssen, C.: TCCON data from Paris (FR), Release GGG2014R0, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.paris01.R0/1149279>, 2017.
- Taylor, T. E., Eldering, A., Merrelli, A., Kiel, M., Somkuti, P., Cheng, C., Rosenberg, R., Fisher, B., Crisp, D., Basilio, R., Bennett, M., Cervantes, D., Chang, A., Dang, L., Frankenberg, C., Haemmerle, V. R., Keller, G. R., Kurosu, T., Laughner, J. L., Lee, R., Marchetti, Y., Nelson, R. R., O'Dell, C. W., Osterman, G., Pavlick, R., Roehl, C., Schneider, R., Spiers, G., To, C., Wells, C., Wennberg, P. O., Yelamanchili, A., and Yu, S.: OCO-3 early mission operations and initial (vEarly) XCO₂ and SIF retrievals, *Remote Sens. Environ.*, 251, 112032, <https://doi.org/10.1016/j.rse.2020.112032>, 2020.
- TCCON Data Archive: Total Carbon Column Observing Network (TCCON), hosted by CaltechDATA, <https://tccondata.org/>, last access: 4 February 2021.
- Velazco, V., Morino, I., Uchino, O., Deutscher, N., Bukosa, B., Belikov, D., Oishi, Y., Nakajima, T. Y., Macatangay, R. C., Nakatsuru, T., Maksyutov, S., Schwandner, F. M., and Griffith, D.: Total Carbon Column Observing Network Philippines: Toward Quantifying Atmospheric Carbon in Southeast Asia, *Climate, Disaster and Development Journal*, 2, 1–12, <https://doi.org/10.18783/cddj.v002.i02.a01>, 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 Orleans (FR), Release GGG2014R1, CaltechDATA [data set], <https://doi.org/10.14291/tcon.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, CaltechDATA [data set], <https://doi.org/10.14291/tcon.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, CaltechDATA [data set], <https://doi.org/10.14291/tcon.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, CaltechDATA [data set], <https://doi.org/10.14291/tcon.ggg2014.lamont01.R1/1255070>, 2017c.
- 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 CO₂ 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 Mazière, 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., Té, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO₂ 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 CO₂ and CH₄ retrieved from GOSAT: First preliminary results, *Scientific Online Letters on the Atmosphere (SOLA)*, 5, 160–163, <https://doi.org/10.2151/sola.2009-041>, 2009.