ESTIMATE OF XCO2 FROM OCO-2 OBSERVATIONS USING A NEURAL NETWORK APPROACH

FM Bréon, L David, P Chatelanaz, F. Chevallier

Laboratoire des Sciences du Climat et de l'Environnement

Context

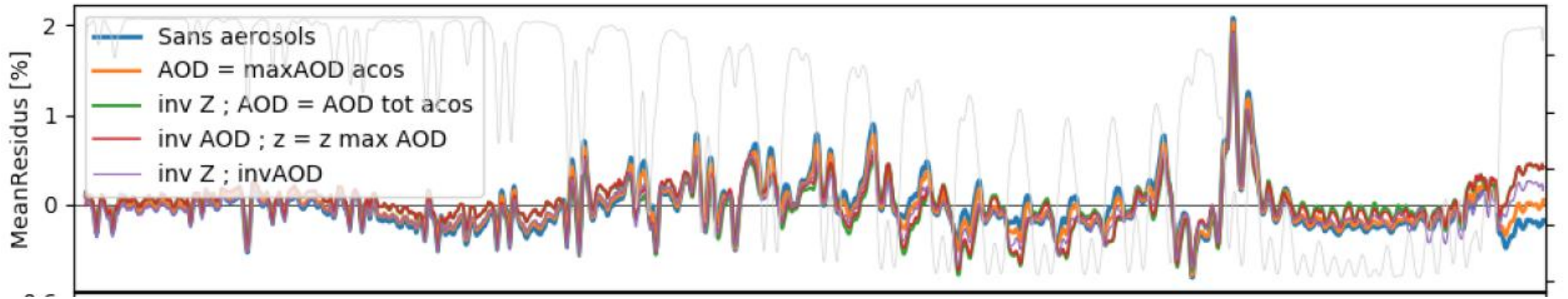
Several satellite missions have been launched (OCO-2, OCO-3, GOSAT, TANSAT) or are under development (MicroCarb, CO2M…) for the **monitoring of atmospheric CO2 from space**

These missions use the reflected solar spectra at high spectral resolution, and use the depth of absorption lines to estimate the amount of CO2 along the path

Atmospheric scattering due to molecule and aerosols must be accounted for (Very computer intensive)

Comparison of **measured and modelled spectra** (after inversion) show **systematic differences** that require empirical corrections

Estimates of **Column-averaged concentrations** (XCO2) show **biases** and empirical de-bias techniques have been developed

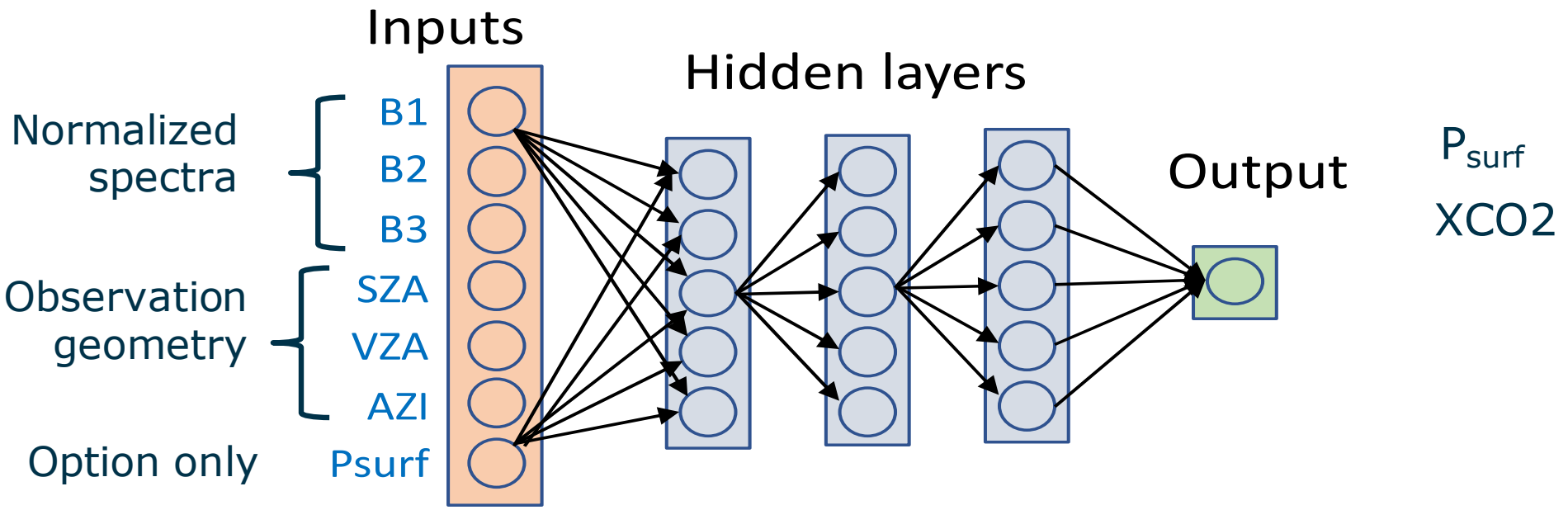


Justification and method

Since the estimates require some empirical correction, why not try a full empirical approach ?

NO radiative transfer model

Use a Neural Network approach



Spectra are normalized by their maximum radiance (reduces surface reflectance variations)

NN Training

For the training, we use only high quality observations

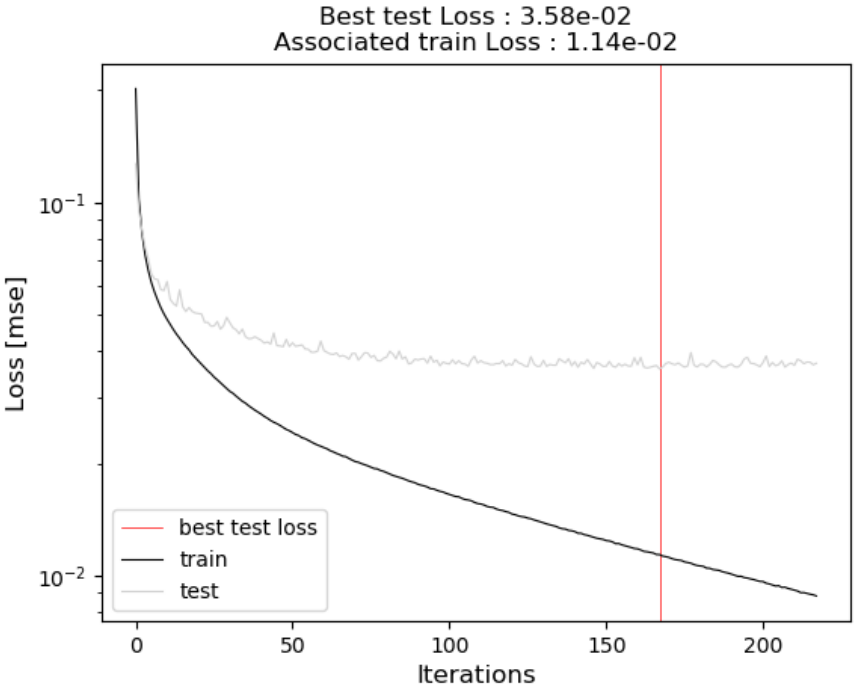
P_{surf} "truth" is numeric weather data corrected for altitude (ACOS prior). "Met" hereafter

CO_2 "truth" is derived from a global CO_2 simulation (CAMS) where the fluxes have been optimized against the near-surface measurements only

The vertical profile is averaged with an homogeneous vertical weighting function

Training uses observations acquired during first week of each month 2015-2019

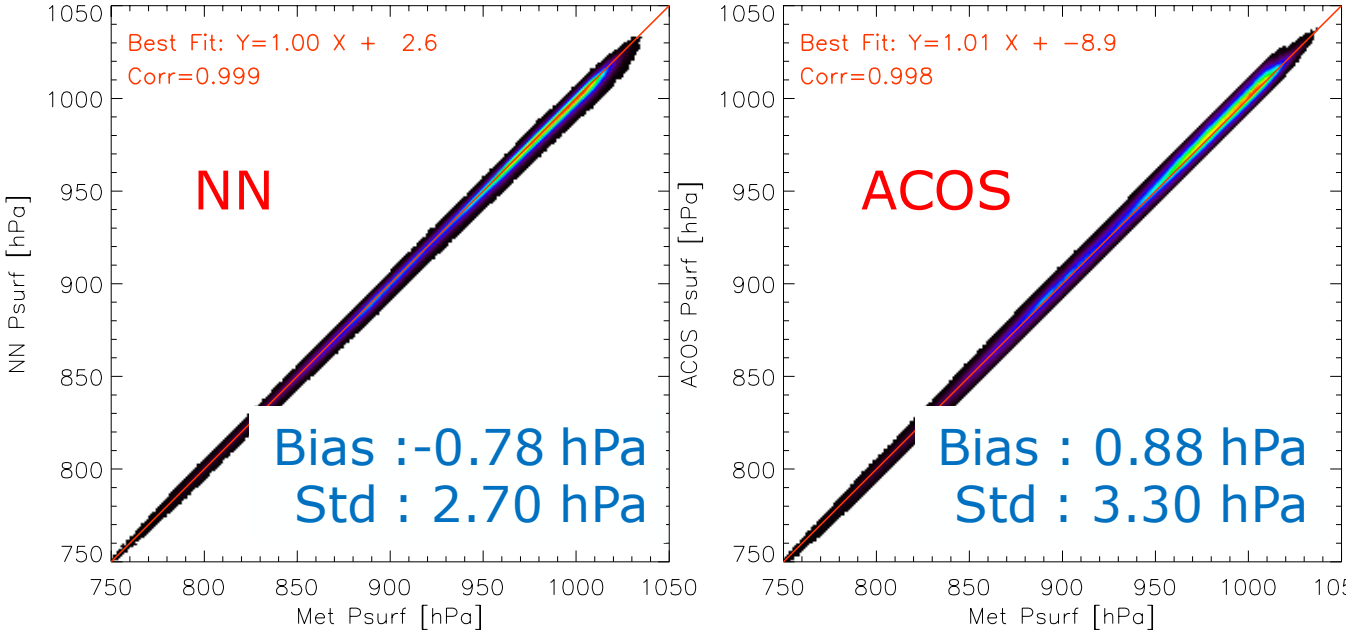
In the following, we use ACOS product v10. XCO2 from ACOS is bias-corrected (used for comparison)



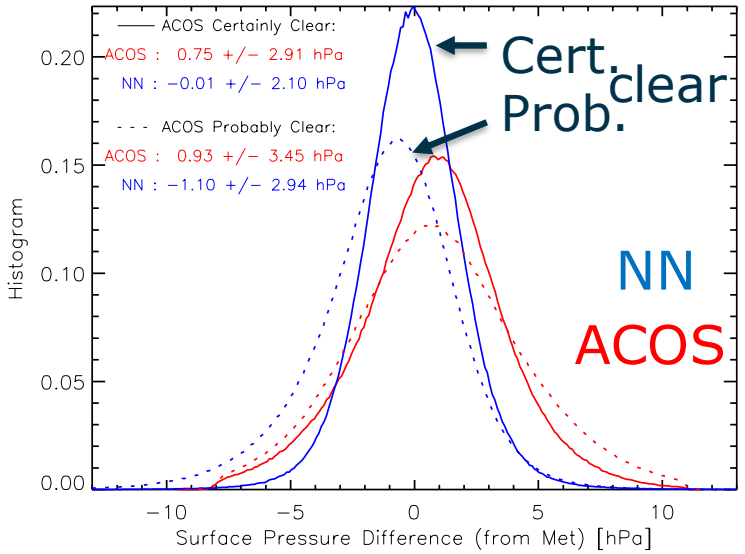
Validation : Psurf

Validation based on observations acquired during rest of the month (first week excluded)

With respect to Met



Histogram of differences



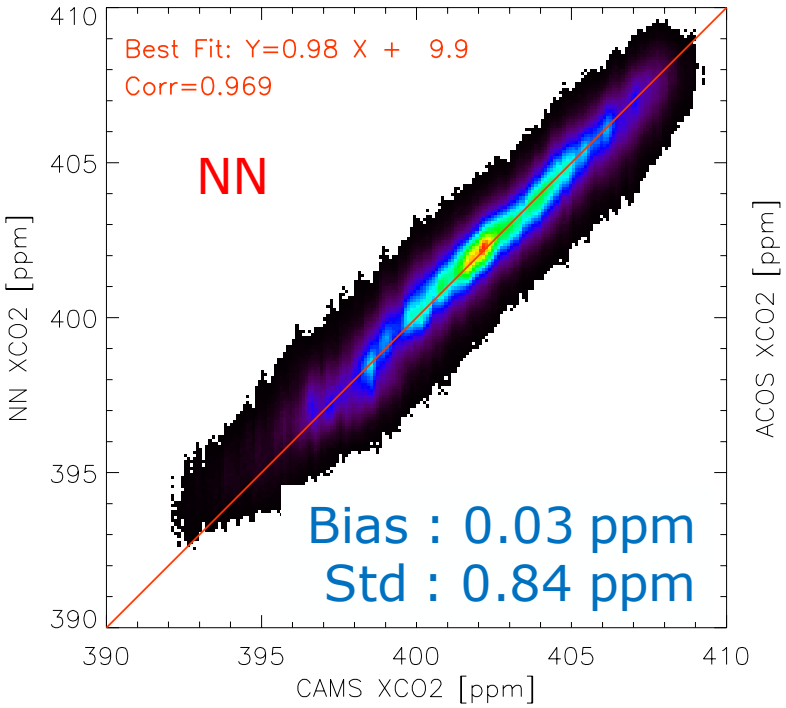
NN approach retrieves the surface pressure with better precision and accuracy than the ACOS algorithm does (although the later uses validation data as prior)

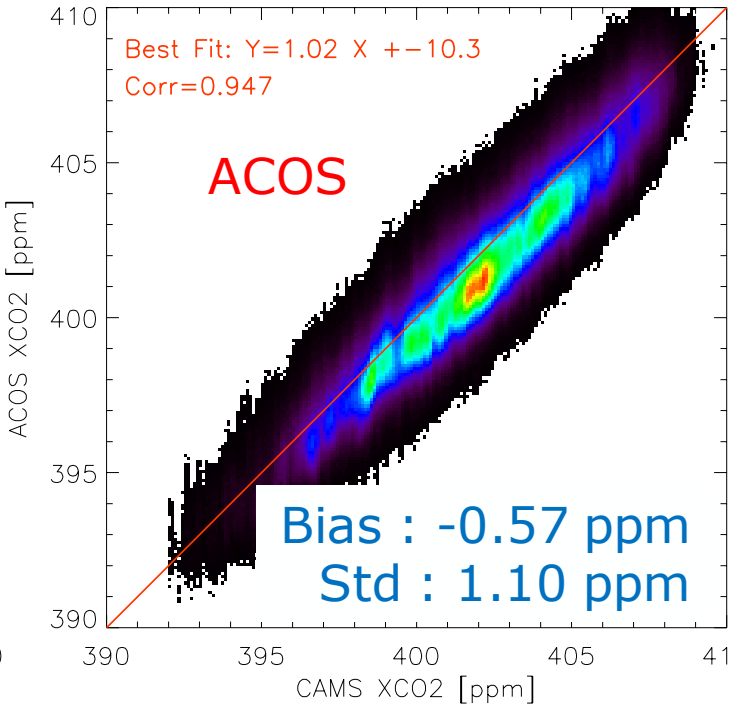
This is with no prior information, and no information on the temperature profile

Validation : XCO2

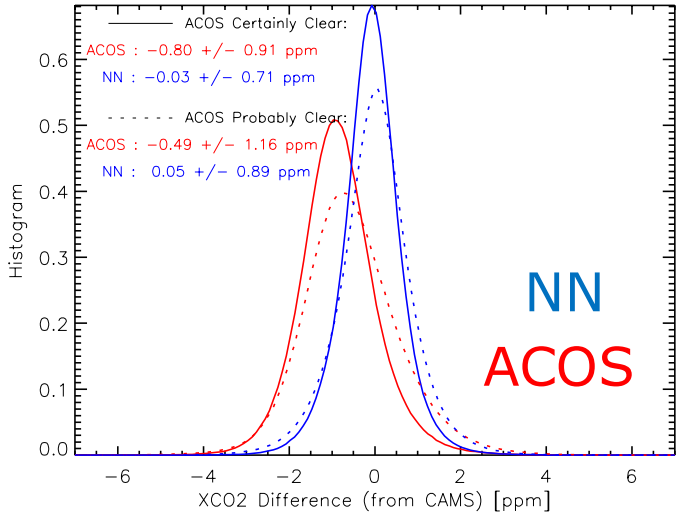
Same dataset as Psurf

With respect to CAMS





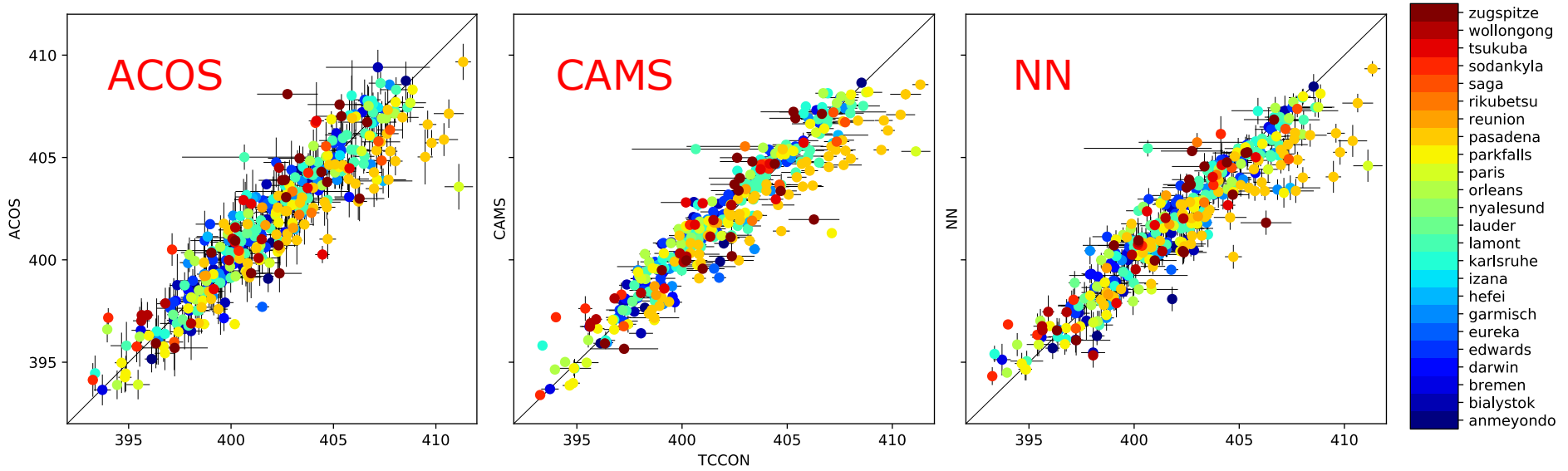
Histogram of differences



NN approach retrieves XCO2 with smaller bias and better precision than the ACOS algorithm does

Again, this is without any prior for the NN estimates

Validation against TCCON



TCCON (network of spectro-photometers) provides a dataset for independent validation

The CAMS model seems to be closer to TCCON than both satellite estimates

Slightly less dispersion for the Neural Network (NN) product than ACOS

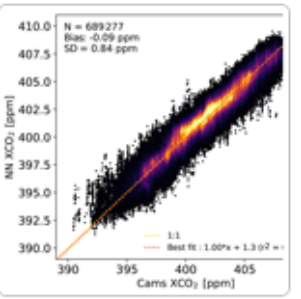
Published !

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

Research article

07 Jan 2021

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



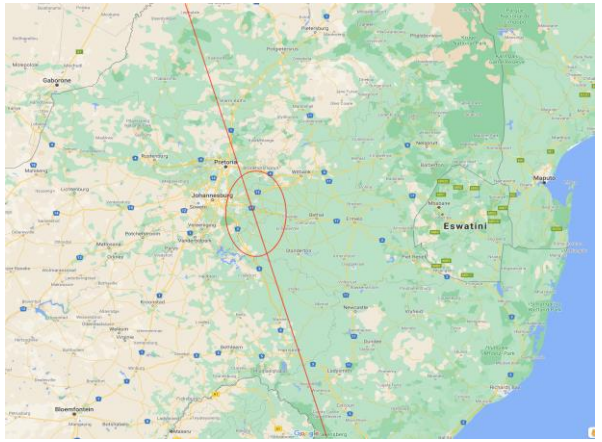
Leslie David, François-Marie Bréon, and Frédéric Chevallier
Laboratoire des Sciences du Climat et de l'Environnement/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, 91198 Gif-sur-Yvette, France

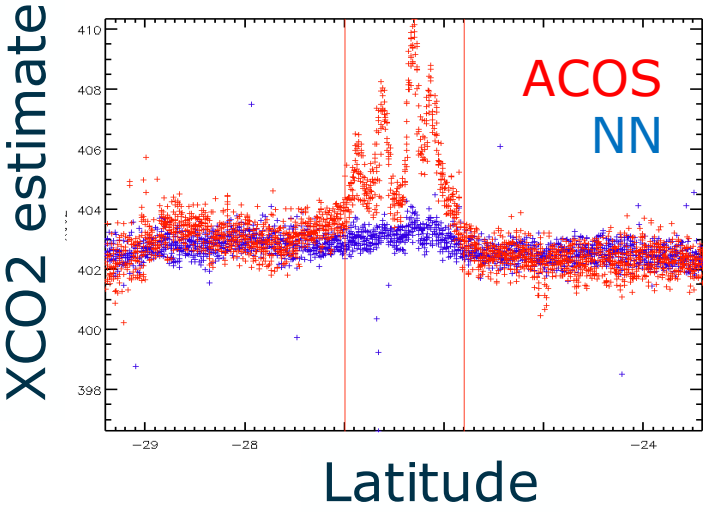
Correspondence: François-Marie Bréon (fmbreon@cea.fr)

Received: 05 May 2020 – Discussion started: 02 Jun 2020 – Revised: 14 Nov 2020 – Accepted: 16 Nov 2020 – Published: 07 Jan 2021

Paper describing these results were published in AMT early 2021

Analysis of a small scale signal





Over South Africa. The OCO-2 subtrack passes to the East of Johanesburg, in a region with several coal power plants. ACOS sees a XCO2 signal with a large amplitude. **The NN hardly sees anything**.

Yet, the ACOS observation is very credible given its shape and its location.

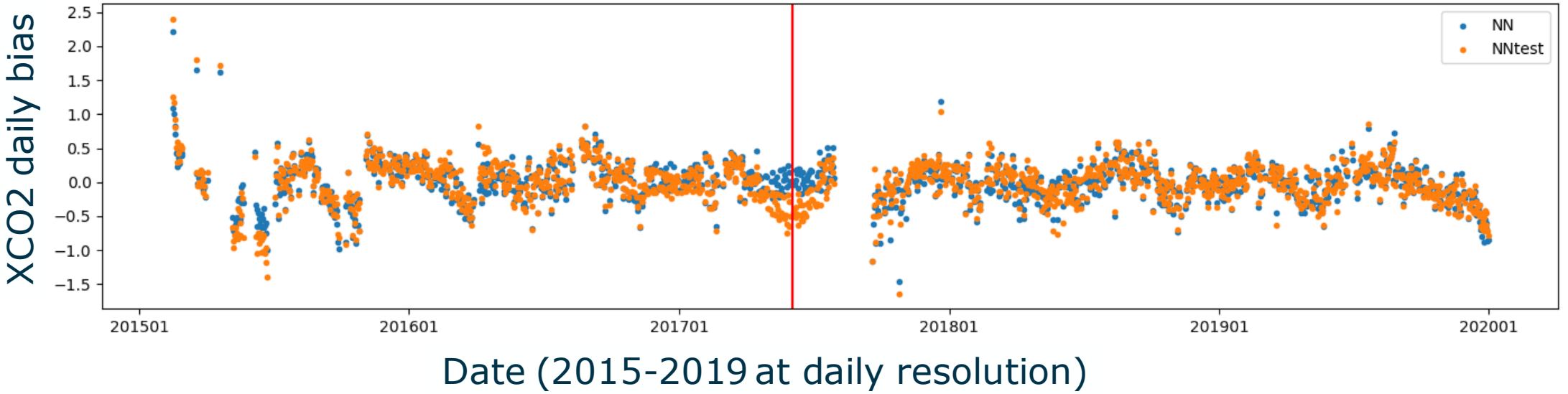
- The NN product does not detect small scale increases in XCO2
- It does detect large scale variations
- All observations are processed independently

How is that possible ???

Back to basics

If one puts a wrong feature in the training, does one gets it in the results?

1 ppm bias in a single month of the training



Mean differences NN-CAMS (daily) for a non-biased training and a training with a 1 ppm bias over on month

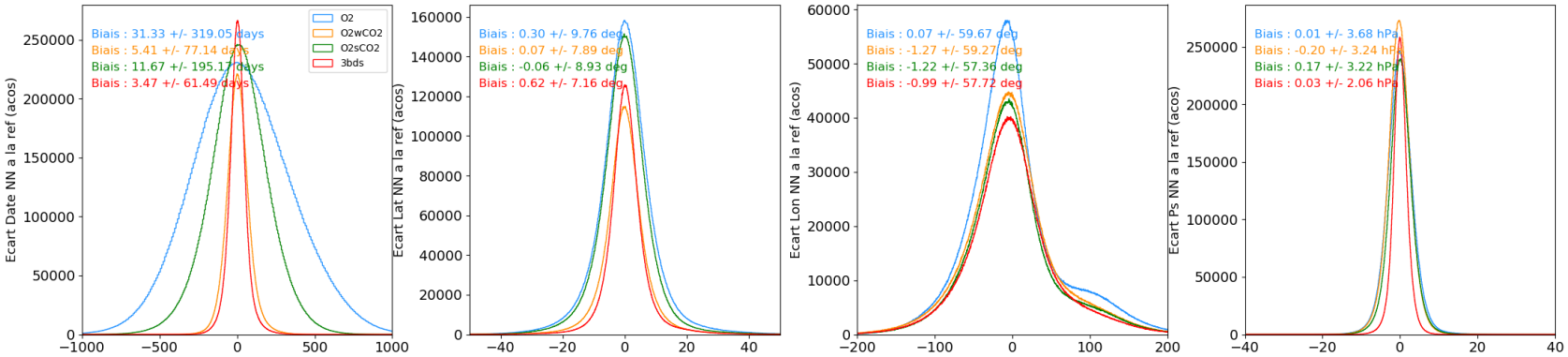
The bias in the training finds its way to the results, for a few weeks before and after !

This strongly indicates that the NN has some indirect information on the date of observation

How is that possible ???

Where is the information on the observation date ?

NN training and test with various combinations of the OCO-2 bands:
Band O2 single, O2+weak CO2, O2+strong CO2, All
The angles (SZA, VZA, Azimuth) are also included in the training



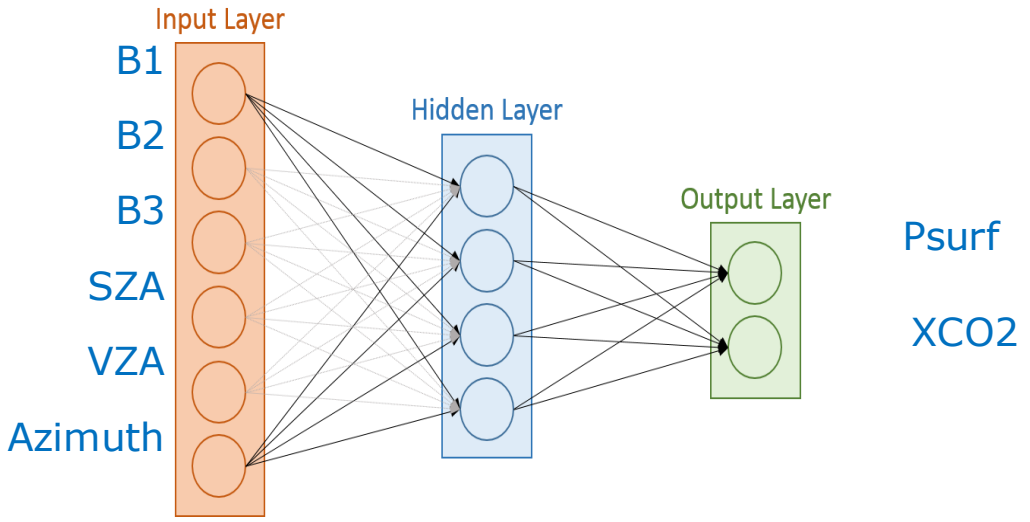
Date Latitude Longitude Surface pressure

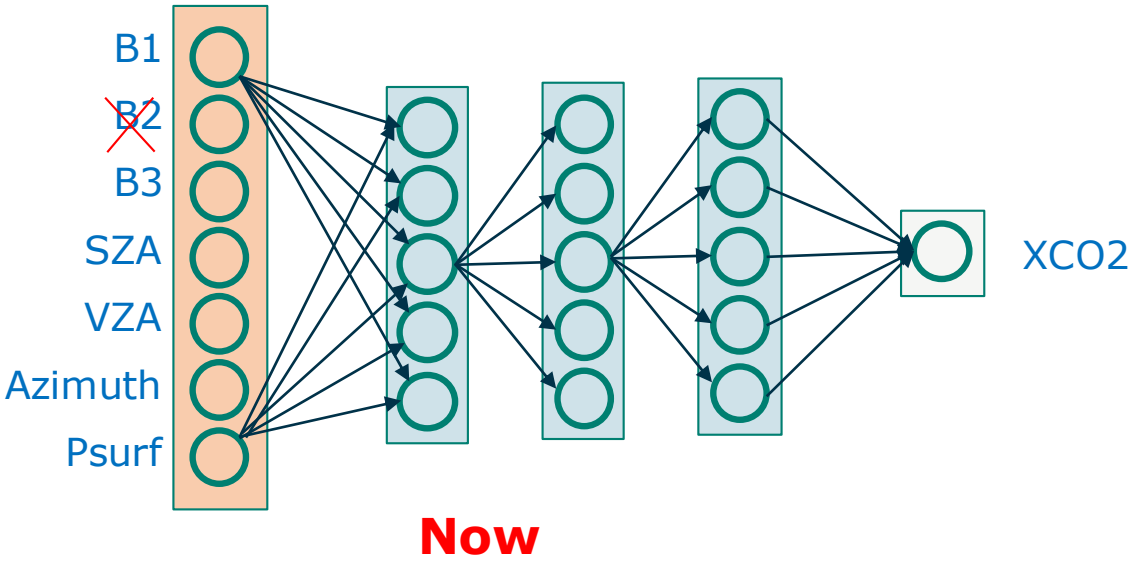
Note that the uncertainty on the longitude is much larger than that on the latitude (information from the obs. geometry

Information on the observation date appears to lie in the **"Weak CO2" band**

Hypothesis (Non demonstrated, but with no other today) : The weakCO2 band is sensitive to the stratospheric CO2 that increases regularly. This stratospheric CO2 provides an indicated information on the observation date.

New version(s) of the Neural Network





We do not use B2 to avoid that the NN use the “date” information to infer XCO2

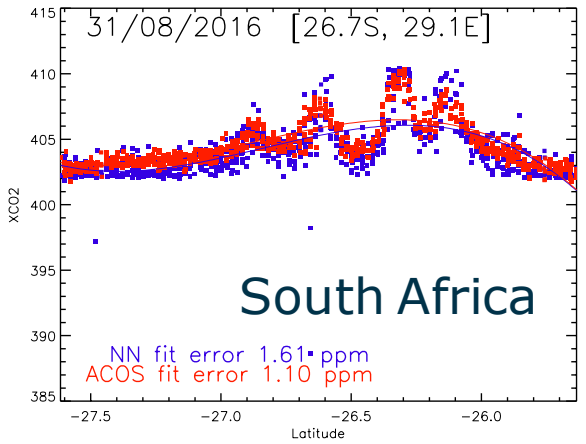
Our experience indicates that, with a deeper network, there is less “over-fitting”. I.e. the “training” dataset has a similar performance as the “test” dataset.

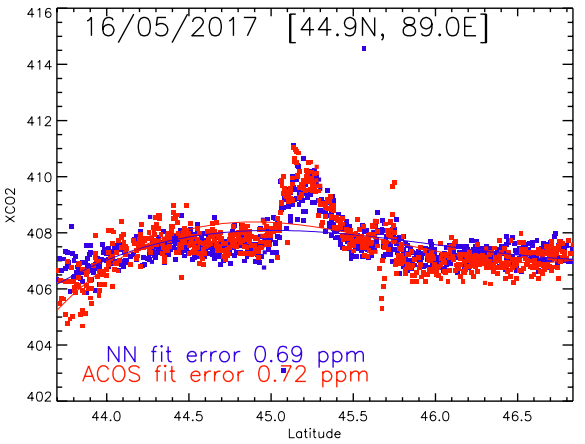
Psurf (from Met) adds useful information that is available. No reason not to use it

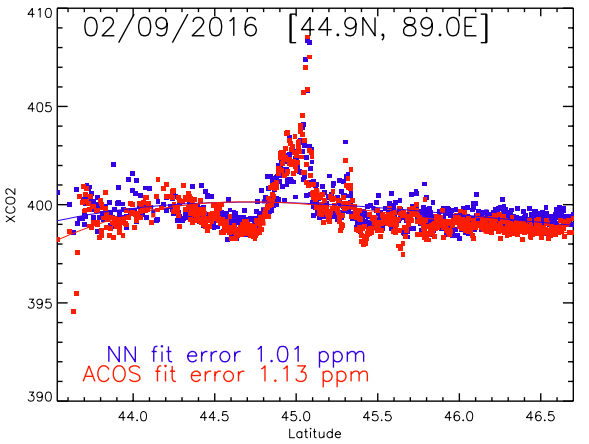
Same technique applied to Glint data. Shows similar performance.

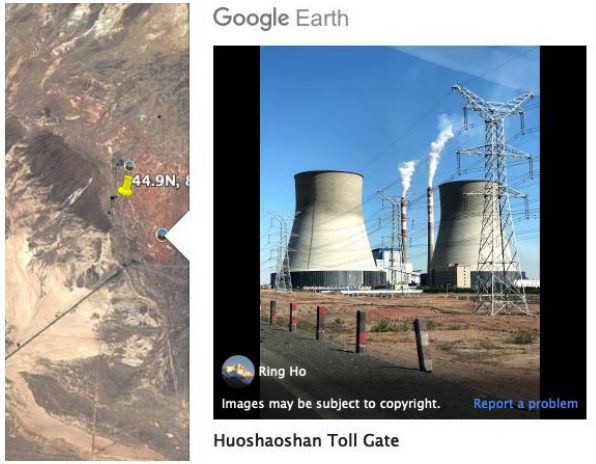
Observation of “plumes”

NN
ACOS





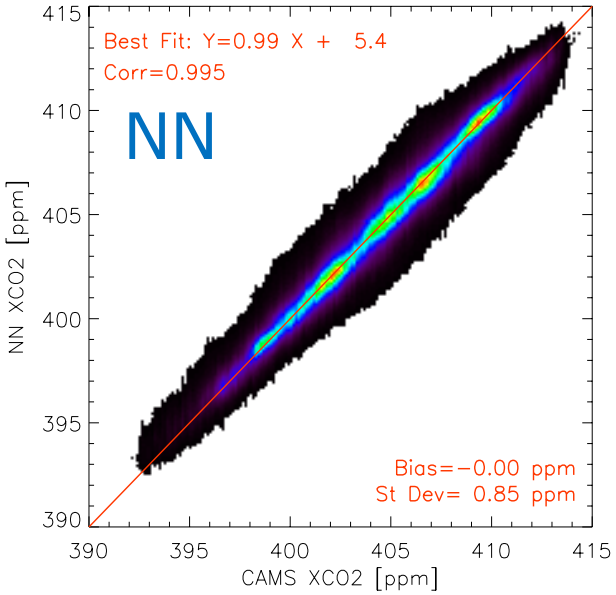


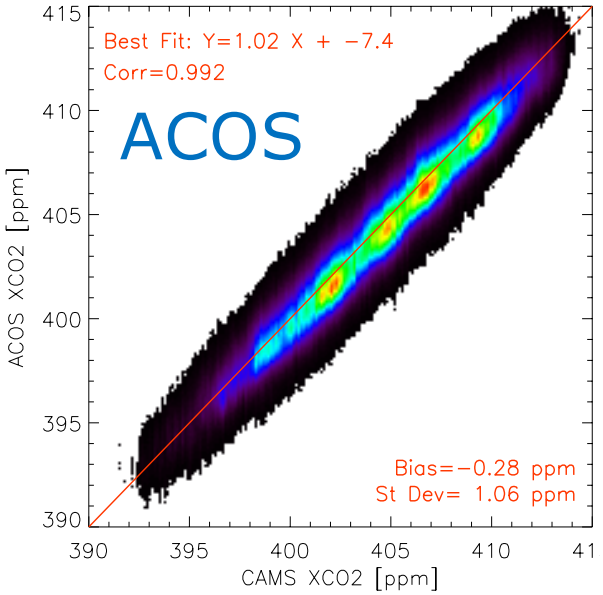


The new version of the NN **sees XCO2 plume**, similarly as the ACOS algorithm product

This demonstrates that the NN is able to retrieves XCO2 features that are not in the training dataset

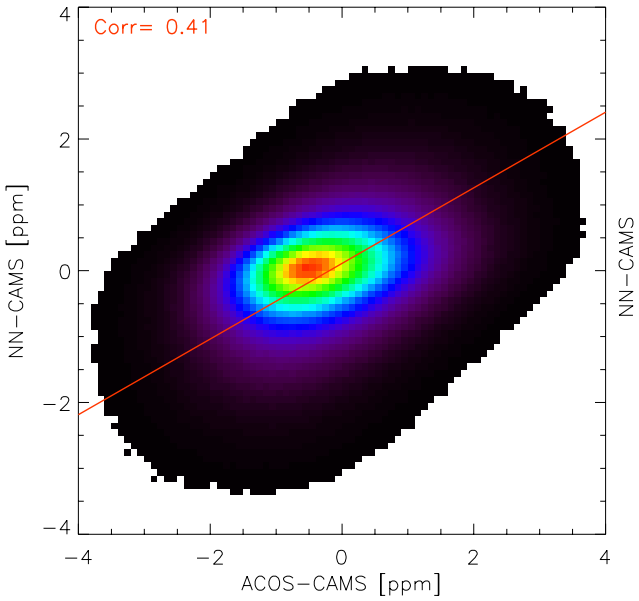
Evaluation agains CAMS

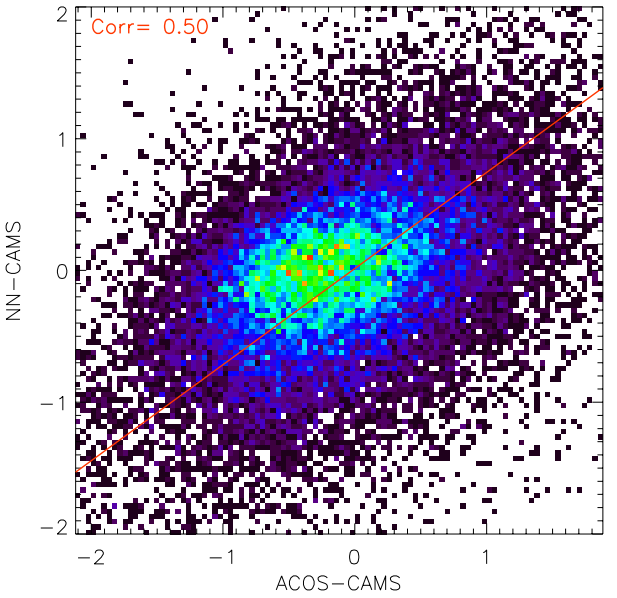




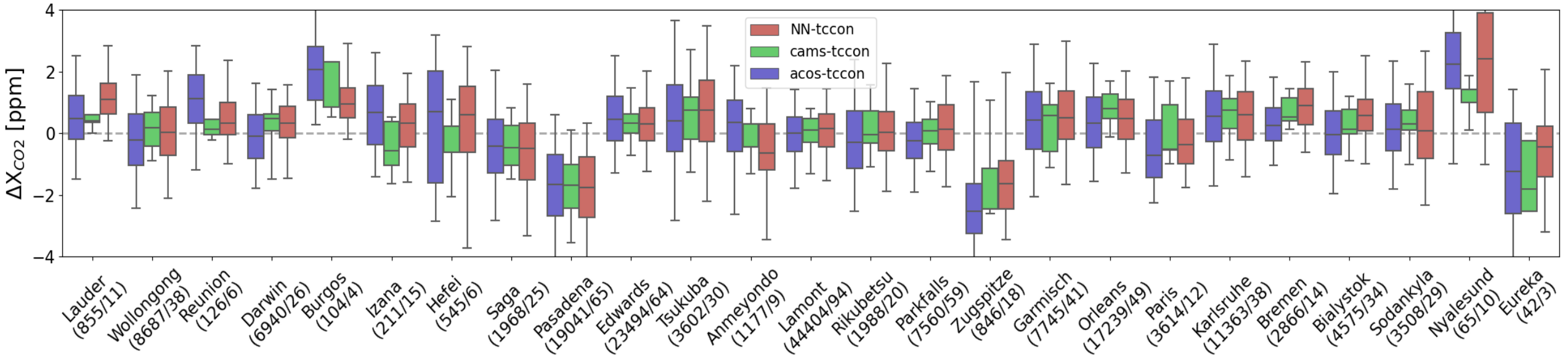
The agreement with CAMS is slightly better with the NN than it is for ACOS

Significant correlation on the innovation brought by the satellite products with respect to CAMS, both at the individual FOV and 5x5-monthly scales





Validation against TCCON



The TCCON sunphotometer network provides an **independent validation dataset**

A few stations (Pasadena, Zugspitze) provide data that are not representative of the surrounding

For other stations, **similar accuracy and precision** for the ACOS operational algorithm and the alternative NN approach

Note also that the model (CAMS) remains better than both satellite estimates

New paper submitted

Refer to this discussion paper for details on this work

https://doi.org/10.5194/amt-2021-313

27 Oct 2021

Review status: this preprint is currently under review for the journal AMT.

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, F-91191 Gif-sur-Yvette, France

Received: 05 Oct 2021 – Accepted for review: 25 Oct 2021 – Discussion started: 27 Oct 2021

Summary and Conclusions

Expectation on the potential of the NN approach has changed with time

Previous version stick to the training. For reasons that remain unclear, it is able to estimate the date and location from the input data with fair accuracy and delivers the corresponding CAMS product. No innovation.

Current version uses only the O2 and sCO2 bands. Little information on the observation date in the input. The NN behaviour is then markedly different and estimates XCO2 from the physics. **The innovation is demonstrated as the NN retrieves XCO2 features that are not in the training dataset**.

There is a **potential for a fully different approach for the retrieval of XCO2** from the satellite data. Advantages are (i) much easier to develop and (ii) huge reduction in computation requirements. Impacts in terms of accuracy remain to be demonstrated

Disadvantages : (i) no information on vertical weighting function (ii) No easy way to identify good and bad retrievals (no mod-obs residus) (ii) need for a “good” training dataset