Retrieving atmospheric CO_2 concentrations from satellite observations based on AI techniques

Context

Atmospheric concentrations of CO_2 retrieved from space-borne observations are increasingly used to infer the space-time distribution of natural and even anthropogenic CO_2 fluxes at the Earth's surface. The constellation of space-borne sensors is evolving and new satellite missions (e.g. MicroCarb (CNES) and CO2M (ESA)) will soon join the precursor instruments GOSAT (JAXA) and OCO-2 (NASA). For these missions, the baseline approach to estimate the CO_2 column-averaged concentration (XCO_2) consists in the inversion of a full-physics radiative transfer model from the radiances measured at high spectral resolution in the *ad hoc* bands. Processing the large data volume provided by these sensors, while ensuring the necessary retrieval accuracy, is a major technical challenge due to the considerable computing resources that are required. In addition, the various uncertainties in the measurements and in the retrieval process (including the modeling and spectroscopy errors) induce significant XCO_2 biases as demonstrated by the comparisons against reference ground-based data, which need to be corrected *a posteriori*. These factors have so far hampered near-real-time (NRT) applications, such as day-to-day monitoring of the distribution of CO_2 in the atmosphere to be exploited by operational data assimilation centres.

We have recently developed a neural network (NNET) -based retrieval algorithm to overcome these limitations. It is able to estimate XCO_2 with negligible computing time and with accuracy and precision similar or better than that of the official bias-corrected product. This proof of concept approach has been applied so far to the observations of the existing OCO-2 instrument over land surfaces, and it could be applied for the operational processing of the upcoming MicroCarb and CO2M missions. The results are very promising but some features of the approach remain unsatisfactory or not well understood. For instance, the application of the NNET to recent data, which is required for NRT application, shows large retrieval biases. This is mostly due to the fact that recent observations may be associated to larger atmospheric CO_2 concentrations than those considered in the training database (relying on a representative dataset of *past* observations by construction). Innovative analyzes and developments are therefore needed to improve the XCO_2 retrieval accuracy.

Objectives

The aim of this internship will be to participate in the improvement of the NNET structure and its overall performance. In this context, different and complementary research and technical activities can be investigated, depending on the candidate's interest and progress.

In particular, improved AI techniques could be tested to address current shortcomings. Also, the complementarity and possible synergy between the classical full-physics inversion and the NNET approach could be explored. The use of radiative transfer simulations (collaboration with CNES) to complement the training dataset (solely based on observations so far) could indeed allow representing conditions consistent with recent and future atmospheric CO_2 concentrations, and hence improve the retrieval accuracy for near-real time processing.

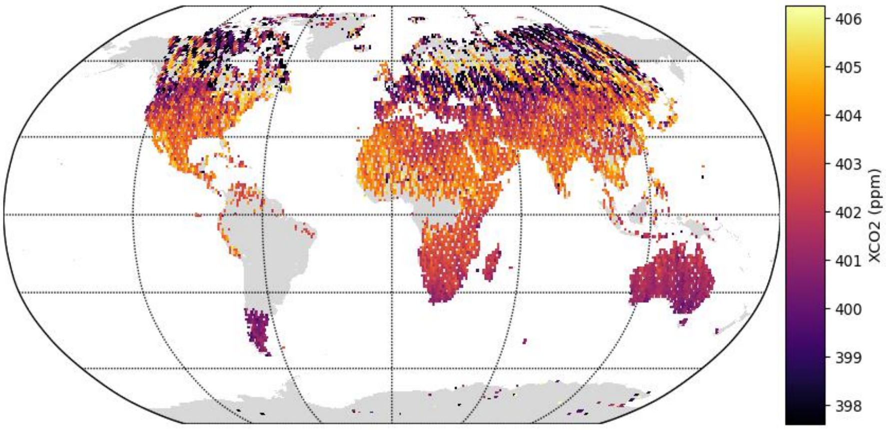

Figure: Yearly averaged XCO_2 values retrieved by a NNET approach (year 2016)

References:

David, L., Bréon, F. M., & Chevallier, F. (2021). XCO 2 estimates from the OCO-2 measurements using a neural network approach. Atmospheric Measurement Techniques, 14(1), 117-132.

Bréon, F. M., David, L., Chatelanaz, P., & Chevallier, F. (2022). On the potential of a neural-network-based approach for estimating XCO 2 from OCO-2 measurements. Atmospheric Measurement Techniques, 15(18), 5219-5234.

Keywords: Remote sensing, deep learning, atmospheric carbon dioxyde / CO_2

Candidate profile: Master or engineering degree in computer science or applied mathematics. The candidate should have prior knowledge of Python programming and of neural network modeling with deep learning frameworks such as Keras/Tensorflow (or PyTorch).

Supervisors: Cédric Bacour (cedric.bacour@lsce.ipsl.fr)
Co-supervisors: François-Marie Bréon (breon@lsce.ipsl.fr), Frédéric Chevallier (frederic.chevallier@lsce.ipsl.fr)

Where: LSCE - Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France (location)
Research team : SATINV
Duration and dates: 5-6 months starting in February or March 2023
Stipend: CNRS or CEA salary grid (~600 euros / month)
Hosting solution: 1 office to be shared
Possibility of extending the internship to a thesis? Yes