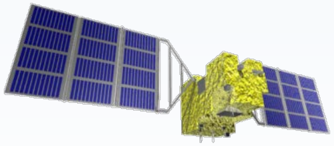
The 8th GOSAT RA PI Meeting /IWGGMS-12, Kyoto, June 6-9

EOF-based regression algorithm for the fast retrievals of XCO_2 from the GOSAT observations

Andrey Bril[1] , Shamil Maksyutov[2] , Dmitry Belikov[2], Sergey Oshchepkov[1]



[1]Institute of Physics of National Academy of Sciences of Belarus
[2]National Institute for Environmental Studies, Tsukuba, Japan

project

Analysis of PPDF-based XCO_2 and XCH_4 retrievals from GOSAT TANSO-FTS and further development of PPDF-S retrieval algorithm

within

The 9th Research Announcement (RA)

Contents of the talk

- Brief resume of PPDF-algorithm and the ways of its improvement
- Constrains on PPDF from ground-based observations and their extrapolation to global scale using EOF/PCA regressions
- Implementation of XCO_2 regression-based retrieval algorithm

  - constructing of EOF reference basis

  - training

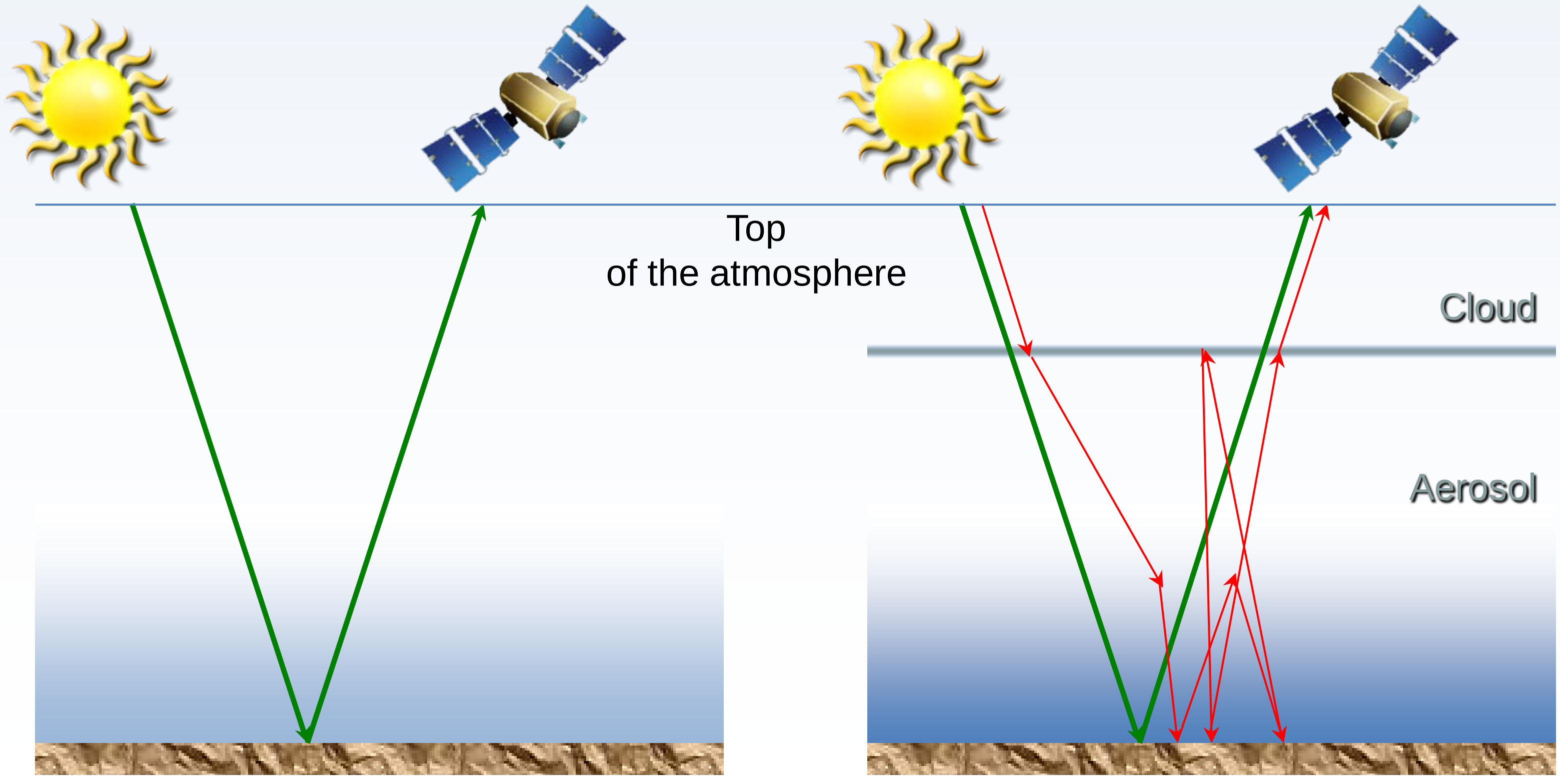
  - XCO_2 retrieving/validation

Basic steps of PPDF approach

Photon Path-length probability Density Function (PPDF)-based approach combines
•Differential Optical Absorption Spectroscopy (DOAS);
•Equivalence theorem; and
•Statistical description of the optical path modification

$$R_\nu \sim \exp\{-(k_\nu \mathbf{L})\}$$

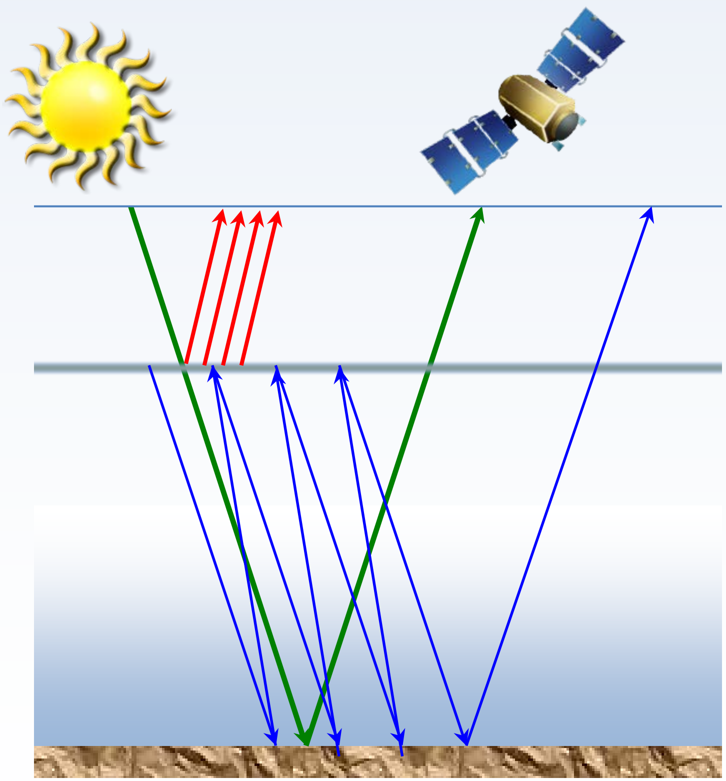
$$R_\nu \sim \int \exp\left(-k_\nu \vec{\mathbf{L}}\right) P\left(\vec{\mathbf{L}}\right) d\vec{\mathbf{L}}$$



We were first who have introduced light path terminology in the GOSAT data processing

The parameterization of the PPDF for “cirrus-like” localization of the scattering particles

We have shown that PPDF under different combinations of aerosol and cloud optical characteristics could be parameterized by four parameters for each atmospheric each layer:

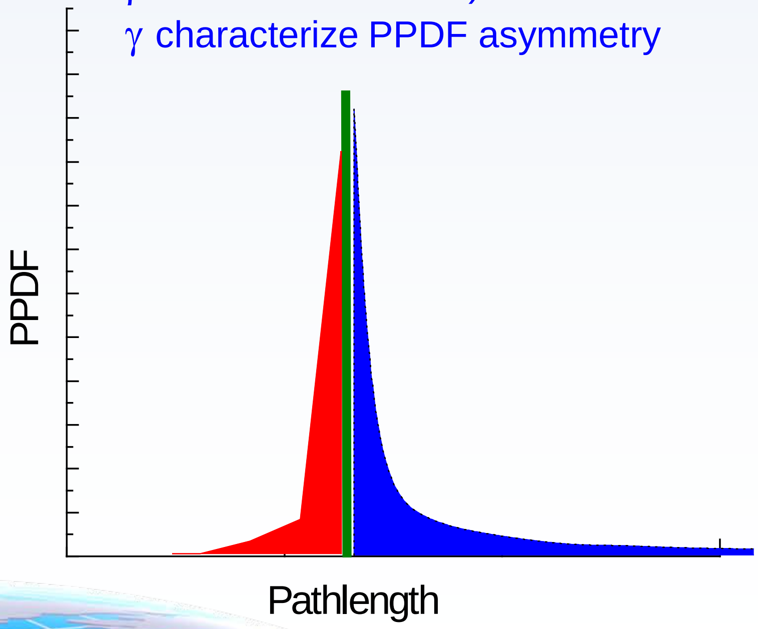


h is the effective layer altitude

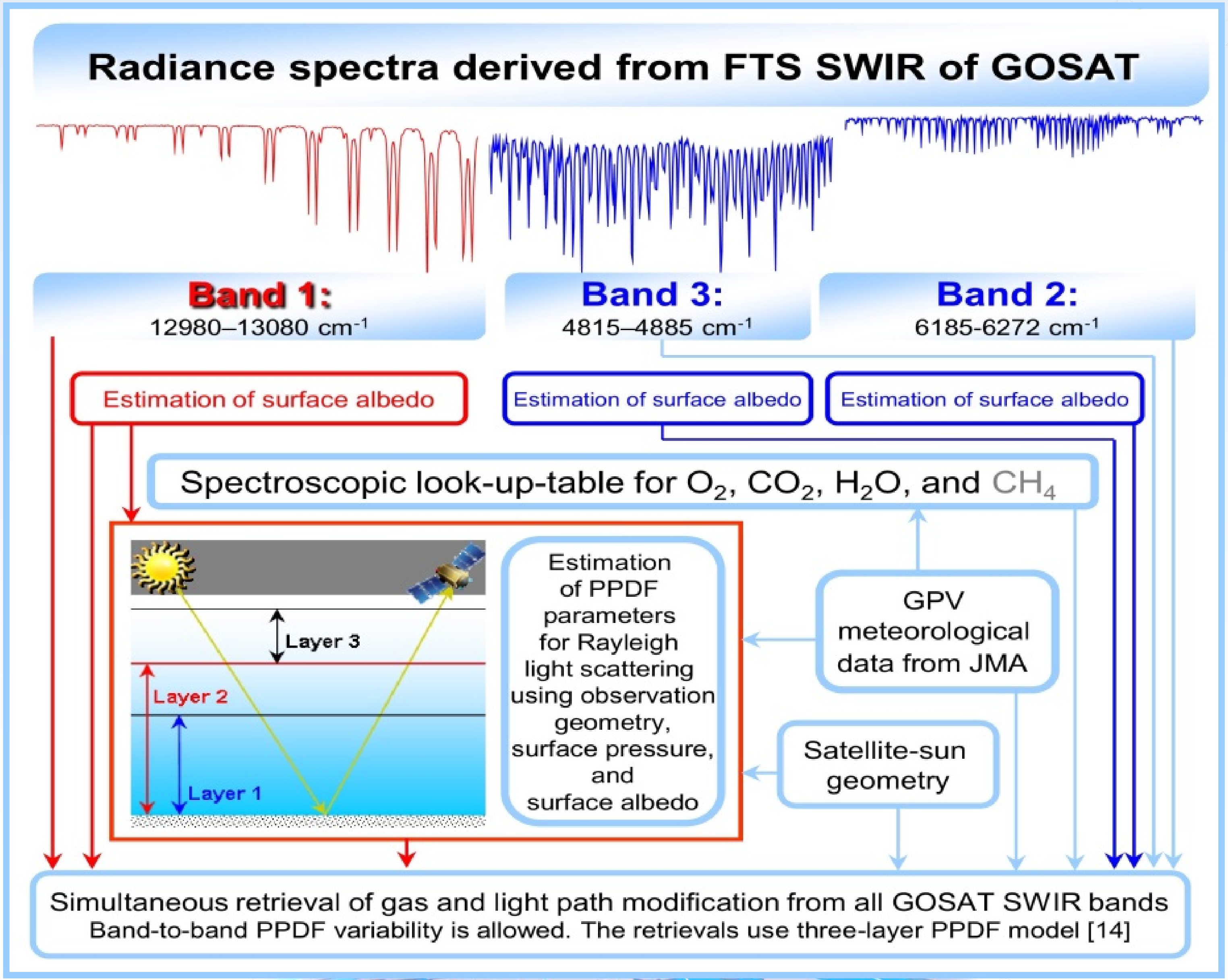
α is the relative cloud reflection

ρ is PPDF half-width)

γ characterize PPDF asymmetry



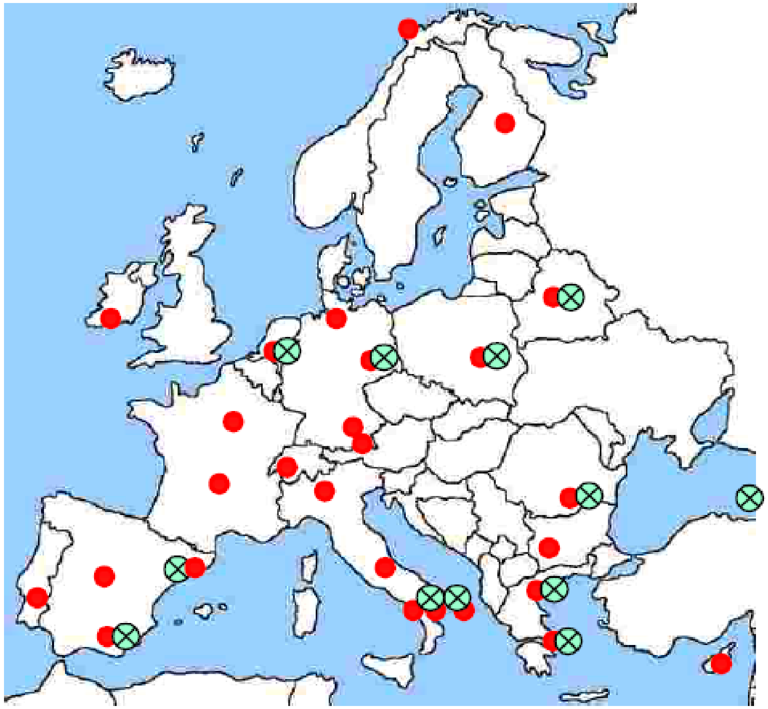
Implementation of the of PPDF-approach: PPDF-S



2 ways to improve the algorithm:

- **Generalized PPDF parameterization for aerosols**
- **Using additional *a priori* information on PPDF/XCO_2 and imposing stronger constrains when retrieving these variables**

Ground-based observations to constrain PPDF



EARLINET stations (red dots). Green dots indicate the stations where LIRIC program package has been implemented.
from
A. Chaikovsky, O. Dubovik, B. Holben, A. Bril, et al., "*Lidar-Radiometer Inversion Code (LIRIC) for the retrieval of vertical aerosol properties from combined lidar/ radiometer data*", **AMT 9, 2016**

- GOSAT signals are synthesized using measured aerosol profiles
- PPDF parameters are retrieved from the synthesized data

TCCON stations: PPDF parameters within footprints are estimated under fixed XCO_2 at TCCON values

EOF-based XCO_2 retrieval algorithm

Main steps

•EOF-based decomposition of the measured spectral radiance

•Combination of limited number of the decomposition coefficients (principle components) with *a priori* information (e.g. airmass, surface pressure)

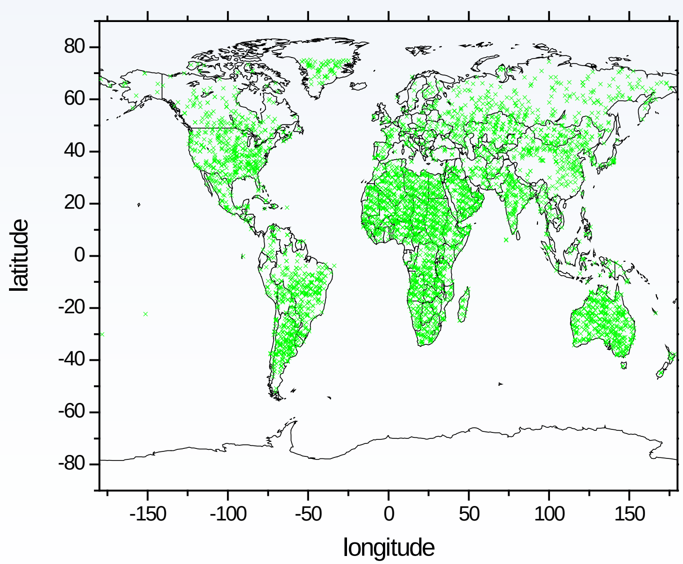
•Derivation of regression formulae to relate the combined information with target gas amounts by using training sets of collocated GOSAT and ground-based observations.

EOF-based XCO_2 retrieval algorithm: EOF basis

$$R = E \bullet \Psi \qquad R_{l,\nu} = \sum \varepsilon_{l,m} \Psi_{m,\nu}$$

To perform EOF decomposition we derived reference EOF basis using standard subroutine for Singular Value Decomposition (SVD) from the IMSL library

$$U^T R V = \Sigma$$



Over-land scalar radiance from NIES operational algorithm for January, April, July, and October of 2010 and 2012 (~ 5000 scans) was used to create EOF basis for three spectral regions

6180 cm^{-1} – 6270 cm^{-1} ,	Band 2
4815 cm^{-1} – 4885 cm^{-1} ,	Band 3
13000 cm^{-1} – 13090 cm^{-1} ,	Band 1

EOF-based XCO_2 retrieval algorithm

Now any spectral signal can be expressed in terms of reference EOF with weighting coefficients defined by

$$\mathrm{E}_{(k)} = R_{(k)} \cdot \Psi_{(k)}^{T} = R_{(k)} \cdot V_{(k)}$$

Generalized vector of weighting coefficients includes limited number of PC and *a priori* info – airmass, surface pressure *a priori XCO2*

$$\tilde{\mathrm{E}} = \left\{ E_{(1)}^{1}, \ldots, E_{(1)}^{M(1)}, E_{(2)}^{1}, \ldots, E_{(2)}^{M(2)}, E_{(3)}^{1}, \ldots, E_{(3)}^{M(3)}; \Pi_{1}, \ldots \Pi_{P} \right\}$$

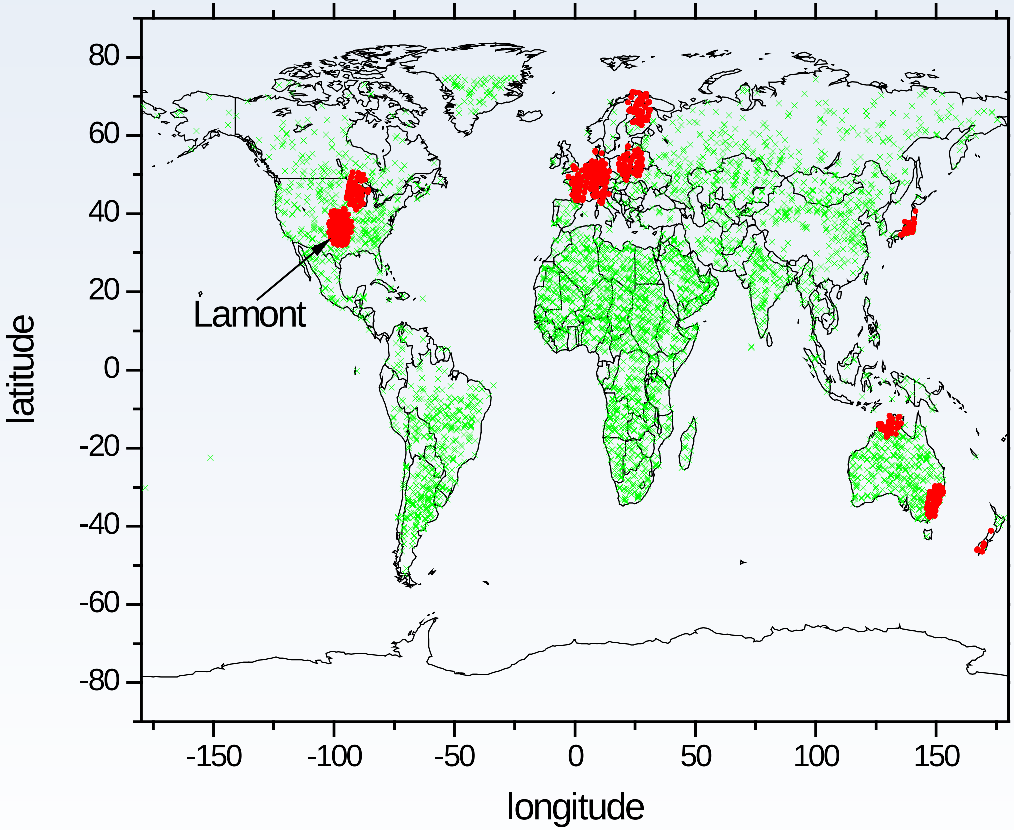
This generalized vector is expected to include necessary information on XCO_2, which is extracted using “transformation vector”

$$X_{CO2} = G \cdot \tilde{\mathrm{E}}$$

G is determined from the condition of the best fit of XCO_2 over the “training subset” of the observations for which values are somehow known

$$G = X_{*,CO2} \cdot \mathrm{E}_{*}^{T} \cdot \left(\mathrm{E}_{*} \cdot \mathrm{E}_{*}^{T} \right)^{-1}$$

Training set selection



12 TCCON stations:
Bialystok, Bremen, Darwin, Garmisch, Karlsruhe, Lamont, Lauder, Orleans, Park Falls, Sodankyla, Tsukuba, and Wollongong

Collocation criteria:
± 1h of the GOSAT overpass time
GOSAT observation is located within 5° latitude-longitude circle around the site.

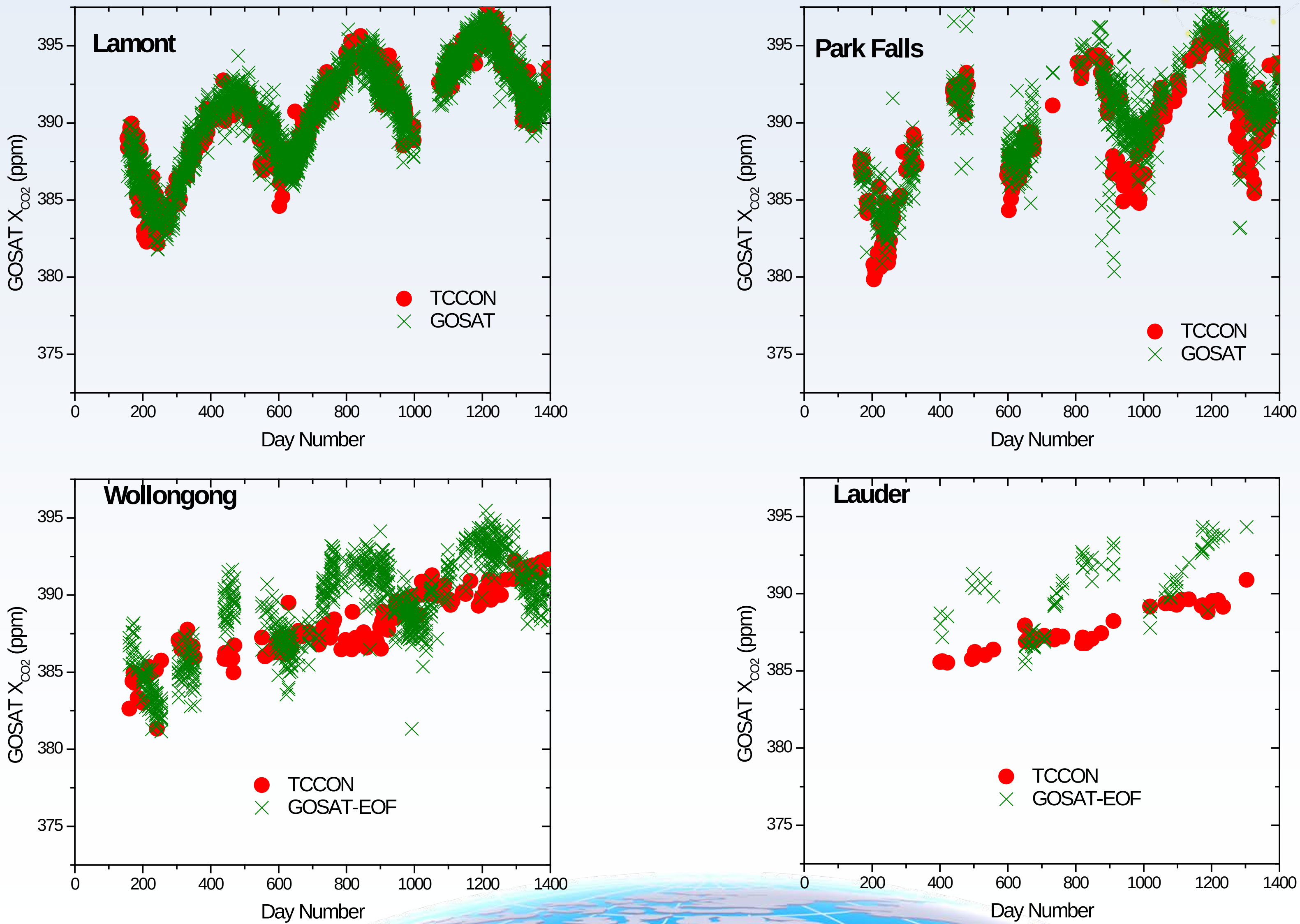
About ~ 12 000 collocated observations

Two sets were tested

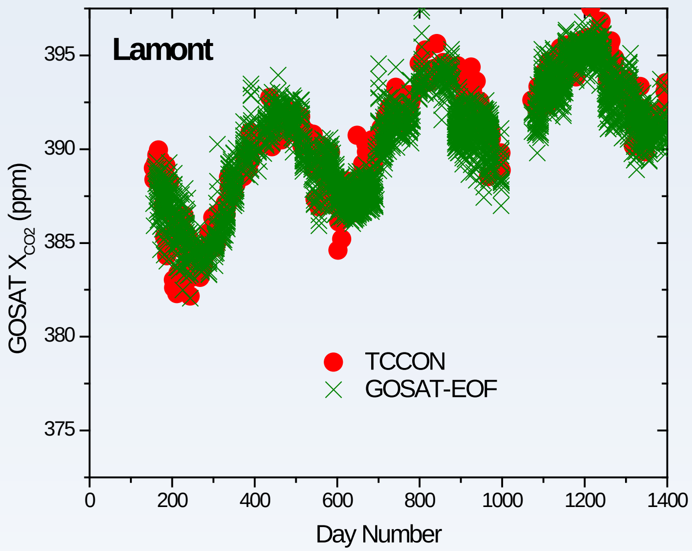
1. Lamont data only
2. Balanced data from all 12 sites (20% of Lamont data; 30% of Park Falls, …)

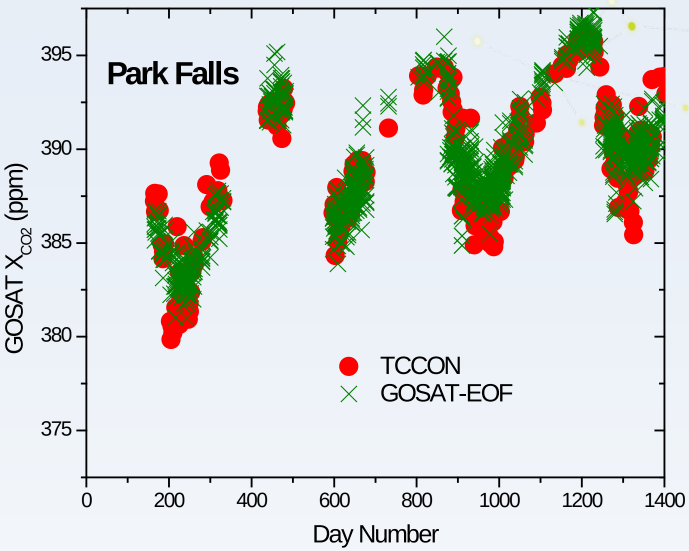
~ 3000 observations/set (to have the rest for the validation)

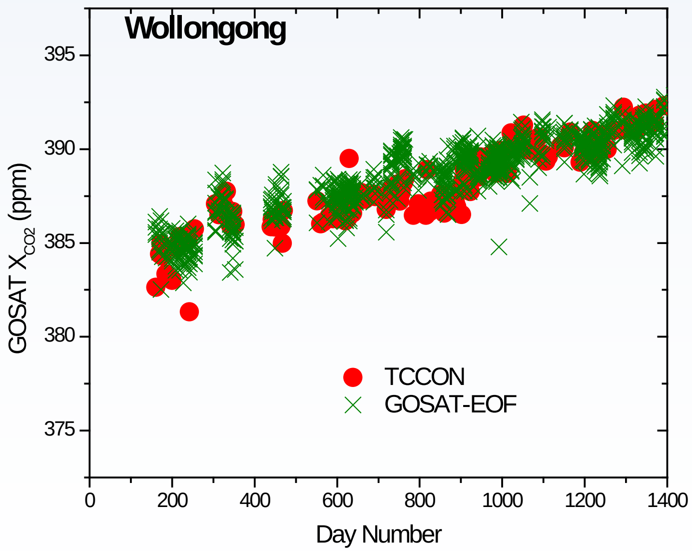
Training set 1 (Lamont only)

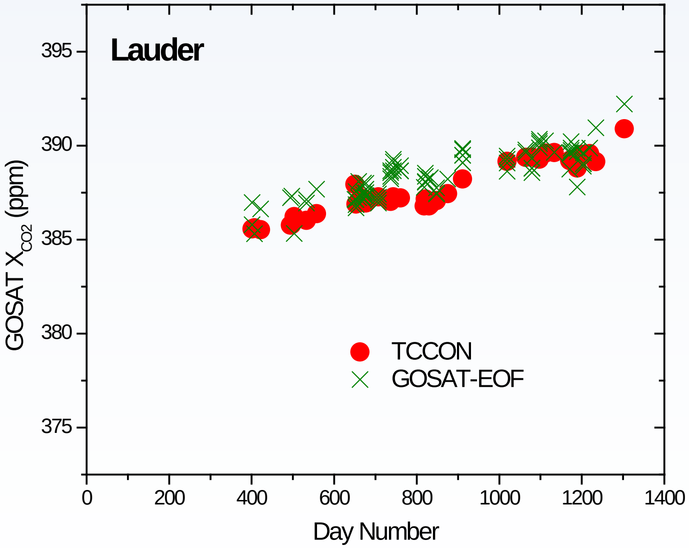


Training set 2

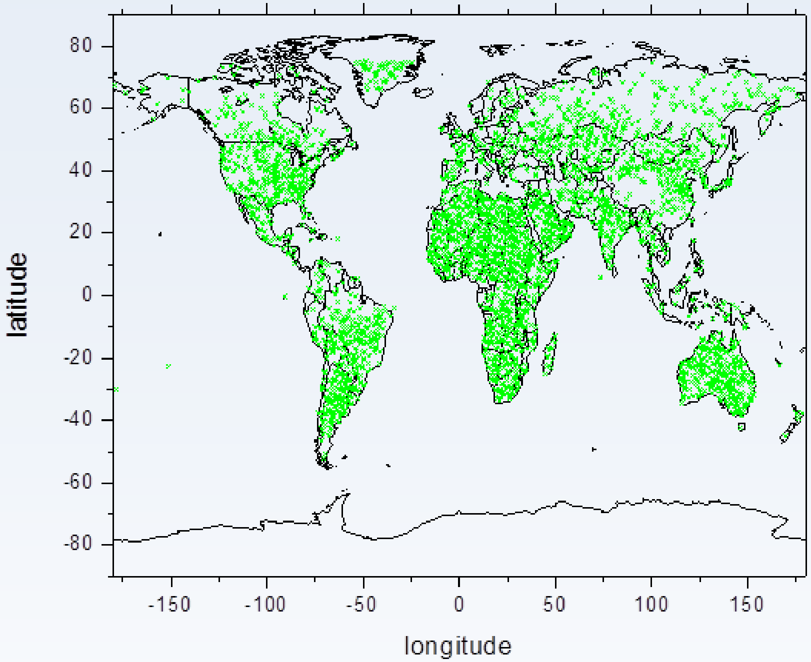


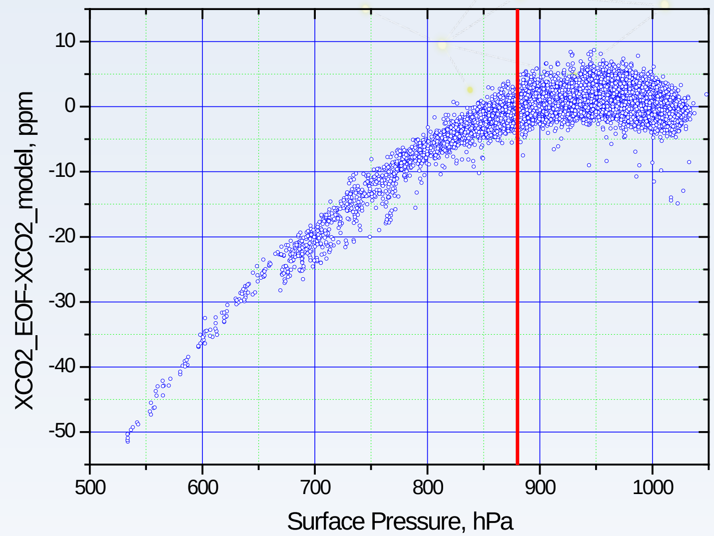




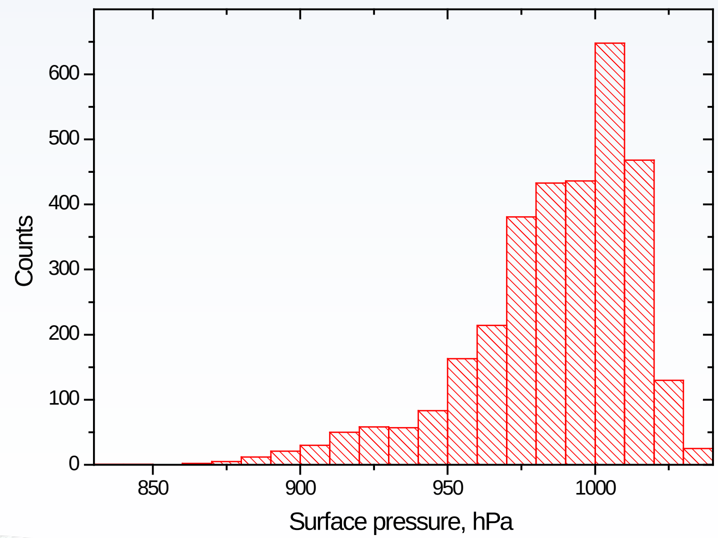


Additional tests using modeled data





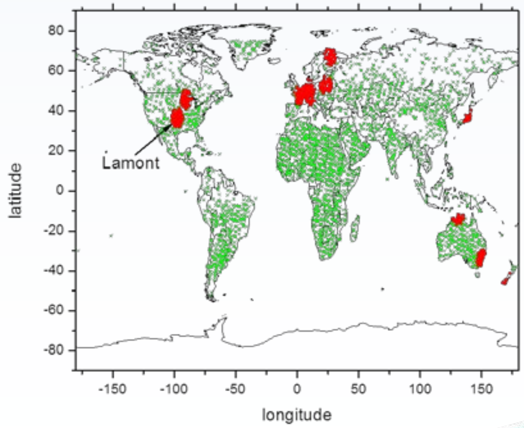
Over-land NIES TM data for January, April, July, and October of 2010 and 2012
(~ 25000 scans)



Additional tests using modeled data

	N	Bias	σ	Slope	r
All observations	22602	0.93	1.48	1.00	0.86
North, latitude >23.5°	8940	0.59	1.45	1.05	0.90
South, latitude < - 23.5°	3436	0.74	0.96	0.87	0.91
Tropics, - 23.5° <latitude < 23.5°	10226	1.29	1.56	0.94	0.81

Statistical characteristics of the EOF-model intercomparison



Conclusions

- EOF/PCA-based regressions proved to be effective tool to extrapolate local observations to global scale
- Special case : rapid and accurate XCO_2 retrieval algorithm (precision/accuracy appear to be similar to the ones of full-physics algorithms)
- Further improvements of the EOF/PCA- algorithm requires elaborations of the training sets (additional TCCON stations, advanced GOSAT-TCCON collocation criteria)