Synthetic satellite datasets

Johan Strandgren

che-project.eu

Co-ordinated by ECMWF

D2.5 Synthetic satellite datasets

Dissemination Level: Public

Author(s): Johan Strandgren (DLR)
Date: 31/04/2020
Version: 0.3
Contractual Delivery Date: 31/12/2019
Work Package/ Task: WP2/ T2.4
Document Owner: DLR
Contributors: DLR
Status: Final

CHE: CO2 Human Emissions Project

Coordination and Support Action (CSA)
H2020-EO-3-2017 Preparation for a European capacity to monitor CO2 anthropogenic emissions

Project Coordinator: Dr Gianpaolo Balsamo (ECMWF)
Project Start Date: 01/10/2017
Project Duration: 39 months

Published by the CHE Consortium

Contact:
ECMWF, Shinfield Park, Reading, RG2 9AX,
gianpaolo.balsamo@ecmwf.int

The CHE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776186.

Table of Contents

Figures

Tables

1 Executive Summary

To aid the analysis of the effect of atmospheric aerosol for the detection and quantification of CO_2 plumes from space-borne observations, datasets of synthetic satellite observations are generated. To this end, a new parameterization is developed that estimates expected random and systematic errors for the retrieval of the column-averaged dry-air mole fraction of CO_2 (XCO_2) for CO_2 instrument flying in constellation aboard six satellites in a sun-synchronous orbit. The parameterization takes the sun-satellite geometry as well as spectrally resolved surface albedo and aerosol optical thickness as input and estimates the corresponding XCO_2 retrieval errors using two artificial neural networks (ANN). The parameterization has been trained using a global dataset of simulated satellite measurements and corresponding XCO_2 retrievals representing a large variety of geophysical scenarios. Random (noise) XCO_2 errors are parameterized with almost perfect precision (R=0.99). For the systematic errors, representing the deviation between retrieved and true XCO_2, the precision is slightly lower (R=0.87), given the many aerosol properties that contribute to the systematic XCO_2 errors like aerosol type, shape, amount, size distribution and vertical distribution.

The new XCO_2 error parameterization is used to generate two datasets of synthetic satellite observations for two domains. The first domain, focusing on Europe, covers a geographical area ranging from approx. 33°N to 66°N and -26°E to 53°E (see e.g. Figure 5). The second domain, focusing on the city of Berlin and its surroundings, covers an area ranging from approx. 49°N to 55°N and 7.6°E to 19°E (see e.g. Figure 9). For the European domain, synthetic satellite observations are generated for the entire year of 2015, whereas for the Berlin domain, observations are limited to February and July 2015. Using a satellite orbit simulator developed by SRON, satellite orbits that intersect with the two domains are simulated for the six satellites. With a satellite Level-2 product generator developed by EMPA, high-resolution aerosol (and XCO_2) data simulated with the LOTOS-EUROS model, as well as MODIS surface albedo data are projected onto the simulated satellite grids such that the corresponding random and systematic XCO_2 errors can be estimated for each orbit and satellite pixel using the XCO_2 error parameterization. For each satellite orbit intersecting the respective domains a netcdf-file is generated containing the synthetic satellite observations for the given orbit segment including also the input data (albedo and aerosol optical thickness) used by the XCO_2 errors parameterization. The collection of these orbit files, ca. 7800 and 270 orbits for the European and Berlin domains, respectively, constitute the deliverable *D2.5 Synthetic satellite datasets* of the CO_2 Human Emissions (CHE) project.

2 Introduction

Within the CO_2 Human Emissions (CHE) project, the effect of atmospheric aerosol when detecting and quantifying CO_2 plumes from future satellite observations is studied. To this end, synthetic satellite observations with realistic estimates of random and systematic errors are needed. By analysing CO_2 plumes in such synthetic satellite observations, one can learn about the expected CO_2 flux estimate errors for a given instrument and mission design and atmospheric scenario and hence, aid the decision making in optimizing the satellite mission.

2.1 Background

The Copernicus CO_2 Monitoring (CO2M) mission will enable space-borne monitoring of CO_2 plumes through observations of enhancements in the column-averaged dry-air mole fraction of CO_2 (XCO_2). The ability to inversely estimate corresponding CO_2 fluxes from such observations is limited by errors in the retrieved XCO_2. Such errors arise due to instrument noise and inadequate information about the light path through the atmosphere as a result of scattering particles like aerosol. To allow for the investigation about how such random and systematic errors will affect the ability to detect and quantify CO_2 plumes from XCO_2 observations with the CO_2 instrument aboard CO2M, realistic synthetic observations and XCO_2 retrievals with corresponding XCO_2 errors for scenes with known atmospheric composition are necessary. To this end, three-dimensional fields of aerosol properties have been simulated at a high spatial resolution of 0.05° × 0.05° (European domain) and 0.01° × 0.01° (Berlin domain) using the LOTOS-EUROS chemistry transport model (Manders et al. 2017). These simulations were conducted by TNO as part of task T2.3 in CHE. Corresponding fields of XCO_2 have also been simulated.

The Multi-Angle Polarimeter (MAP) instrument and its expected error reduction is not considered in this study, since the code development work, required to account for the additional MAP data in the radiative transfer simulations, is to comprehensive to be conducted in this work package, given the allocated resources. Hence, the systematic errors reported in this deliverable reflect those expected for the CO_2 instrument alone, and not for the CO2M mission as a whole.

Performing radiative transfer simulations for all modelled scenes within CHE in order to quantify the random and systematic XCO_2 errors is computationally too expensive. Hence, a parameterization that can reasonably estimate the XCO_2 retrieval errors is developed and used to generate the datasets of synthetic satellite observations.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

The objective of this deliverable is to provide two datasets of synthetic observations of the CO_2 instrument planned to fly aboard the CO2M satellites. The first dataset will contain synthetic observations over Europe during the course of 2015, whereas the second dataset will focus on a smaller domain around the city of Berlin for the months of February and July in 2015.

The deliverable provides simulated retrievals of XCO_2 (taken from LOTOS-EUROS model output) and, more importantly, the corresponding random and systematic XCO_2 errors projected onto the satellites' observation grids.

2.2.2 Work performed in this deliverable

The main part of this deliverable has been to develop a new parameterization for the estimation of random and systematic XCO_2 errors of synthetic satellite observations. This also included extensive radiative transfer simulations using a global ensemble of various geophysical scenarios that can be expected in terms of sun-satellite geometry, surface albedo and aerosol properties.

Furthermore, maps of surface albedo data have been generated for the two domains and a satellite Level-2 product generator software provided by EMPA (including an orbit simulator from SRON) has been adapted in order to read the albedo maps as well as the LOTOS-EUROS model output and project the data onto the simulated satellite observation grids.

2.2.3 Deviations and counter measures

The initial idea was to use a subset of the model data from LOTOS-EUROS to simulate synthetic satellite observations that could be used to develop the error parameterization that could then be applied to the full time series of LOTOS-EUROS data for the two domains. But in order to be able to spend more time on the development of the new XCO_2 error parameterization, it was early on decided to instead use a global ensemble of geophysical data already implemented in the radiative transfer code to generate a dataset of synthetic satellite observations that could be used to develop the XCO_2 error parameterization. This approach, using a global ensemble for the development, also has the advantage that the parameterization is less limited to the two domains simulated and analysed in this study.

3 XCO_2 error parameterization

For this deliverable a new parametrization has been developed. The parameterization estimates the random and systematic errors that are expected to accompany XCO_2 retrievals from the CO_2 spectrometer aboard CO2M. The parameterization is based on an artificial neural network (ANN) approach and consists of two ANNs. The first ANN estimates the random error σ_{rdm}, representing the instrument's signal-to-noise-ratio (SNR) propagated into a statistical error estimate according to the rules of Gaussian error propagation. The second ANN estimates the systematic error σ_{sys}, representing the difference between retrieved XCO_2 (retr) and true XCO_2 (true).

$$\sigma_{sys} = \frac{XCO_2(retr) - XCO_2(true)}{XCO_2(true)} \cdot 100\%. \qquad (1)$$

Both ANNs take eight input variables: solar zenith angle (SZA) and viewing zenith angle (VZA) as well as surface albedo and aerosol optical thickness (AOT) at three different wavelengths representing three spectral windows near 760 nm, 1600 nm and 2000 nm (hereafter referred to as NIR, SWIR-1 and SWIR-2, respectively). A schematic overview of the parameterization can be seen in Figure 1.

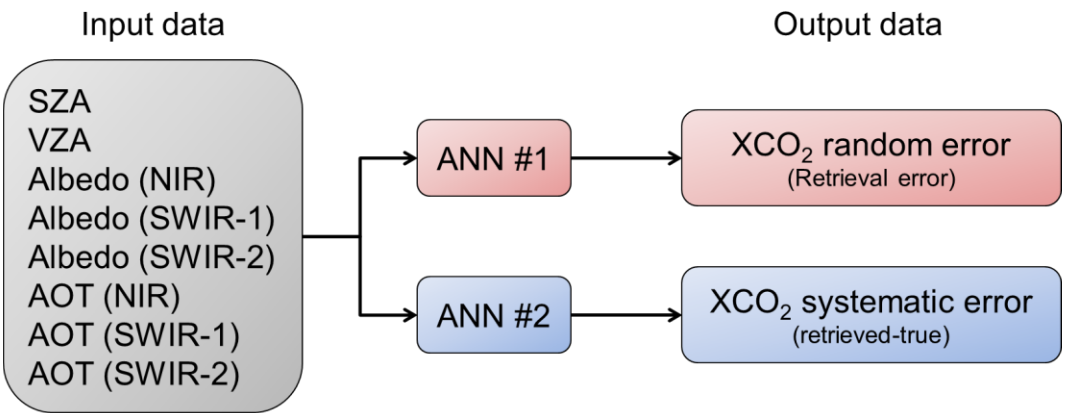

Figure 1: Schematic overview of XCO_2 error parameterization.

In order for the parameterization to learn and understand the relationship between input data and corresponding XCO_2 errors, the ANNs have to be trained. This is done using a supervised learning approach where the ANNs are provided with pairs of input data and known XCO_2 errors. The weights (free parameters) of the ANN are then continuously updated and optimized such that the error between the reference XCO_2 error (true) and the error estimated by the ANN is minimized. To this end a training dataset with the relevant input data as well as the corresponding random and systematic XCO_2 errors is required.

3.1 Training dataset

To generate the training dataset, a global ensemble with a large collection of geophysical scenarios is used. Vertical fields of meteorology and trace gas concentrations stem from the ECHAM5-HAM model (Stier et al. 2005), CarbonTracker model (Peters et al. 2007) and Tracer Model 4 (Meirink et al. 2006). Surface albedo representative for the NIR, SWIR-1 and SWIR-2 windows stem from the MODIS product MCD43A4 (Schaaf et al. 2002). Finally aerosol optical properties are calculated for an aerosol size distribution superimposed from five aerosol types and seven log-normal size distributions at 19 vertical layers, read from the ECHAM5-HAM model (Stier et al. 2005). This global ensemble has been used in several previous studies (where further details on the dataset can be found) in order to estimate the random and systematic greenhouse gas retrieval errors of both operational and proposed satellite instruments (Butz et al. 2009, 2010, 2012, 2015, Strandgren et al. 2020).

The global ensemble is used to simulate synthetic satellite measurements (spectral radiances) using the radiative transfer software RemoTeC (Butz et al. 2011, Schepers et al. 2014). For each measurement the expected measurement noise is computed assuming the satellite orbit configuration described in Kuhlmann et al. (2019) and coefficients for instrument SNR computations provided by SRON (Hein van Heck, personal communication, 2019). These are the orbit and SNR data currently assumed for the CO_2 spectrometer aboard the CO2M.

The simulated spectra of spectral radiances and the corresponding measurement noise are then used to retrieve fields of XCO_2 using the RemoTeC retrieval algorithm (e.g. Butz et al. 2011). For the retrieval, a comparatively simple forward model is used where e.g. only three aerosol parameters are fitted (amount, the size parameter of a single mode power-law size distribution and the centre height of a Gaussian aerosol height distribution). Such differences in the aerosol representation compared to the forward model used to simulate accurate synthetic measurements lead to forward model errors that, alongside the instrument noise induced errors propagate into the retrieved quantity XCO_2. Previous studies have shown that this approach gives a good estimate of the random and systematic errors of greenhouse gas concentration retrievals under scattering conditions (e.g. Butz et al. 2009, 2010).

In total, synthetic satellite measurements and corresponding XCO_2 retrievals are simulated for approx. 250 000 scenes representing different days throughout the year as well as varying satellite viewing zenith angles, alongside the varying geophysical scenarios within the global ensemble itself.

3.2 Training the ANNs

To construct and train the ANNs, a similar approach as in Strandgren et al. (2017) is used, but adapted and optimized for the task of this study. The two ANNs that constitute the XCO_2 error parameterization share the same topology and consist of four layers. One input layer with eight neurons (one for each input variable), two hidden layers with 16 hidden neurons each and one output layer with one neuron (random and systematic XCO_2 errors for the two ANNs, respectively). The ANNs are trained using 80% of the training dataset described above; the remaining 20% are used for internal validation (10%) during the training as well as final validation (10%) of the ANNs (see Section 3.3).

During the training, the ANN is provided with a set of input-output pairs. Using the input data and the current weights (randomly initialized) one output value (either random or systematic XCO_2 error depending on which ANN is being trained) is computed for each set of input data. Each weight in the ANN is then updated such that the total error between the estimated and reference XCO_2 error is minimized. This procedure is repeated until the total error does no longer improve. To make sure that the ANNs do not overfit and loose their ability to generalize, the error against the internal validation dataset is also monitored during the training.

3.3 Validation of the ANNs

Next, the performance of the ANNs is evaluated in order to analyse to which extent the relationship between the input data and the corresponding XCO_2 errors have been understood and learnt by the ANNs. To this end, the 10% of the training dataset that was never used while training the ANNs, is used. Figure 2 shows two density scatter plots with the random and systematic XCO_2 errors in the left and right panels, respectively. The x-axes show the reference XCO_2 errors computed through radiative transfer simulations, while the y-axes show the corresponding parameterized XCO_2 errors estimated by the ANNs.

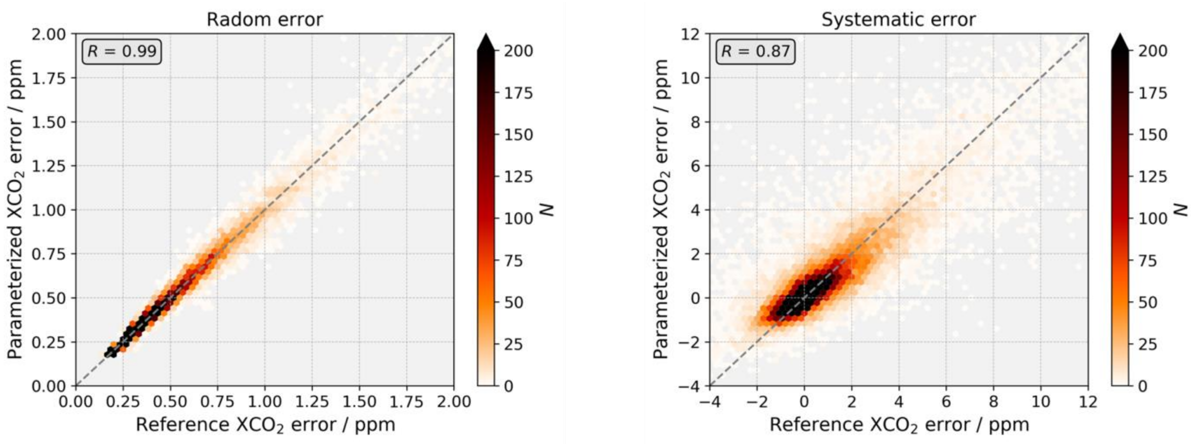

Figure 2: Density histograms showing the relationship between the true reference XCO_2 errors (x-axes) and the corresponding XCO_2 errors estimated by the ANNs (y-axes). Left: Random XCO_2 error. Right: Systematic XCO_2 error. The dashed line shows the 1-1 relationship.

The parameterization can model the random errors almost perfectly (*R*=0.99). This is expected since the random errors to a large degree are directly related to the signal strength and hence the surface albedo and SZA, information which is provided as input data. For the systematic errors the precision is slightly lower (*R*=0.87), since the errors are not only related to the signal strength, but also the amount, type, shape, size distribution and vertical distribution of the atmospheric aerosol, information that can not be fully represented in the AOTs provided as input data.

For a reference scene with SZA of 50 degrees and surface albedo equal to 0.2, 0.1 and 0.05 for the NIR, SWIR-1 and SWIR-2 spectral bands respectively (referred to as *VEG_50* in Bovensmann et al. 2010) the error parameterization estimates a random XCO_2 error of 1.0 ‰ (i.e. around 0.4 ppm), assuming clear sky and nadir view. This is well below the XCO_2 precision requirement of 0.7 ppm specified for CO2M (ESA 2019). The random XCO_2 errors derived in this deliverable can potentially be scaled to represent also other noise scenarios (e.g. 0.5, 0.7 and 1.0 ppm @ VEG_50), if such scenarios are to be investigated (see e.g. Kuhlmann et al. 2019).

4 MODIS surface albedo

In order to apply the parameterization for all simulated satellite orbits over Europe in 2015, corresponding surface albedo data are required. To this end, the Nadir Bidirectional Reflectance Distribution Function (BRDF)-Adjusted Reflectance (NBAR) data (MCD43A4; Schaaf and Wang (2015)) for the MODIS bands 2 (≈860 nm), 6 (≈1600 nm) and 7 (≈2100 nm) are used. This is the same albedo data used to develop the parameterization. The albedo data for 2015 have been downloaded and projected onto the European (and Berlin) domain, with a spatial resolution of 0.01° × 0.01°. Figure 3 shows the surface albedo near 860 nm (NIR, MODIS band 2) and 2100 nm (SWIR-2, MODIS band 7) for the European domain on 2015-02-15 (top) and 2015-07-15 (bottom). In winter time, there are areas of high albedo in the NIR spectral region, especially over Ukraine, Russia and Turkey, but also to some extent over the Nordic countries, as a result of snow cover. These areas are characterized by low surface albedo in the SWIR-2 spectral region. In summer time, the albedo is comparatively homogeneous across the European domain (especially for NIR), with highest albedo for barren-type surfaces e.g. in Spain, Turkey and Northern Africa. Due to persistent cloud cover and low-quality data there are gaps in the surface albedo maps (e.g. oceans as well as Ireland and Scotland on 2015-07-15), which consequently can lead to data gaps also in the parametrized XCO_2 errors later on.

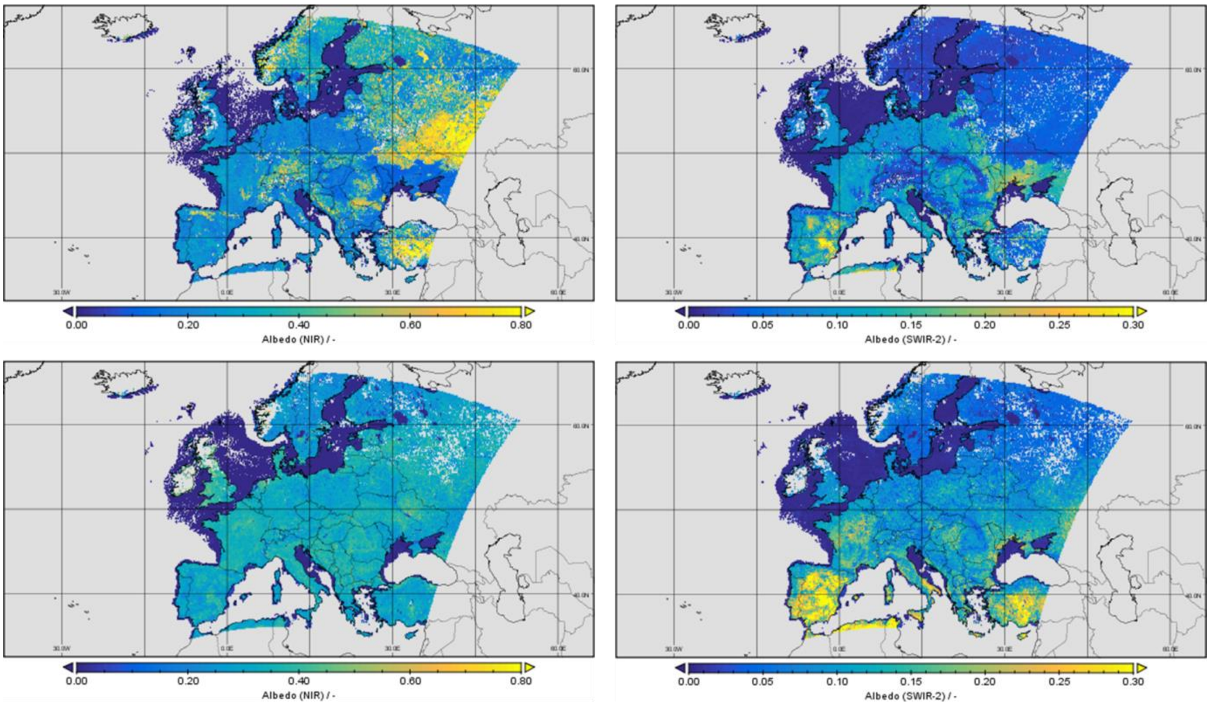

Figure 3: Left: Surface albedo near 870 nm (NIR, MODIS band 2) projected onto the European domain. Right: Surface albedo near 2100 nm (SWIR-2, MODIS band 7) projected onto the European domain. Top: 2015-02-15, boreal winter. Bottom: 2015-07-15, boreal summer.

5 Synthetic satellite observation datasets

The final product of this deliverable consists of a set of data files representing simulated orbits of synthetic satellite observations over the European and Berlin domains for the two respective time periods. Hence, satellite orbits are computed for the given satellite orbits (again using the CO2M orbit configuration described in Kuhlmann et al. 2019). The modelled AOT and XCO_2 read from LOTOS-EUROS output as well as the MODIS surface albedo maps are then re-projected onto the satellite observation grids. To this end a python package provided by EMPA was used (ESA Project SMARTCARB, 2018). The software package generates Level-2 satellite products by simulating given satellite orbits using an orbit simulator developed by SRON (SRON, 2017) and re-projecting external model/observational data onto the satellite grids. Adaptions to the software package have been done in order to read and process the LOTOS-EUROS model and MODIS albedo data used in this study. Furthermore, the new XCO_2 error parameterization was integrated in the software package in order to compute the XCO_2 errors and include them to the output files.

Synthetic observations of six satellites are simulated; SAT-1, SAT-2, SAT-3, SAT-4, SAT-5 and SAT-6. The six satellites are assumed to fly evenly distributed (60 degree displacement) in the same sun-synchronous orbit, leading to a temporal offset of approx. 16 minutes between the satellites. A set of six satellites is simulated such that different possible constellations can be investigated with e.g. two satellites flying 180 degrees apart, 3 satellites flying 120 degrees apart etc. In total ca. 7800 orbits that intersect with the European domain during the course of 2015 have been simulated; of which ca. 270 of the orbits intersect with the Berlin domain during February and July 2015.

Although, the simulated orbits do not differ between the European and Berlin domains, the synthetic satellite observations are different since LOTOS-EUROS model data at a higher resolution of 0.01° × 0.01° have been used for the Berlin domain, compared to the 0.05° × 0.05° available for the European domain.

The orbit files follow the following naming convention:

`SAT-X_dom_YYYYMMDDTHHMMSS_oNNNN.nc`

where '`X`' represents the six satellites ('1', '2', '3', '4', '5' and '6'), '`dom`' the domain ('eur' or 'ber'), '`YYYYMMDD`' the date, '`HHMMSS`' the acquisition time in UTC at the orbit centre and '`NNNN`' the orbit number (starting at 0000 for the first orbit on 2015-01-01). Each netcdf-file contains 17 variables as listed and explained in Table 1. The surface albedo and AOT included in the orbit files are the input data used to compute the XCO_2 errors. These data fields should not be considered part of the *observational product*, but are kept in the orbit files as reference for a better understanding of the XCO_2 errors.

Table 1: Variables included in the simulated satellite orbit files.

Variable	Description
albedo_nir	Surface albedo from MODIS band 2 (MCD43A4)
albedo_swir1	Surface albedo from MODIS band 6 (MCD43A4)
albedo_swir2	Surface albedo from MODIS band 7 (MCD43A4)
aot_nir	Aerosol optical thickness at 870 nm from LOTOS-EUROS model
aot_swir1	Aerosol optical thickness at 1650 nm from LOTOS-EUROS model
aot_swir2	Aerosol optical thickness at 2060 nm from LOTOS-EUROS model
orbit	Satellite orbit settings
latitude	Pixel center latitude
latitude_corners	Pixel corners latitude
longitude	Pixel center longitude
longitude_corners	Pixel corners longitude
sza	Solar zenith angle
time	Time of measurement since 2015-01-01 00:00:00 UTC
vza	Viewing zenith angle
xco2	Column mixing ratio of CO_2 in dry air from LOTOS-EUROS model
xco2_err_rdm	XCO_2 random error
xco2_err_sys	XCO_2 systematic error

5.1 Synthetic satellite observations over the European domain

Figure 4 shows the MODIS surface albedo for the three spectral windows; NIR, SWIR-1 and SWIR-2, for a satellite orbit (SAT-1) crossing the European domain on 2015-07-07. Albedo data for the entire domain is shown in a shaded nuance. Similarly, the modelled AOT for the three spectral windows for the same orbit is shown in Figure 5. Finally, Figure 6 shows the modelled XCO_2 together with the corresponding XCO2 random and systematic errors computed with the XCO_2 error parameterization.

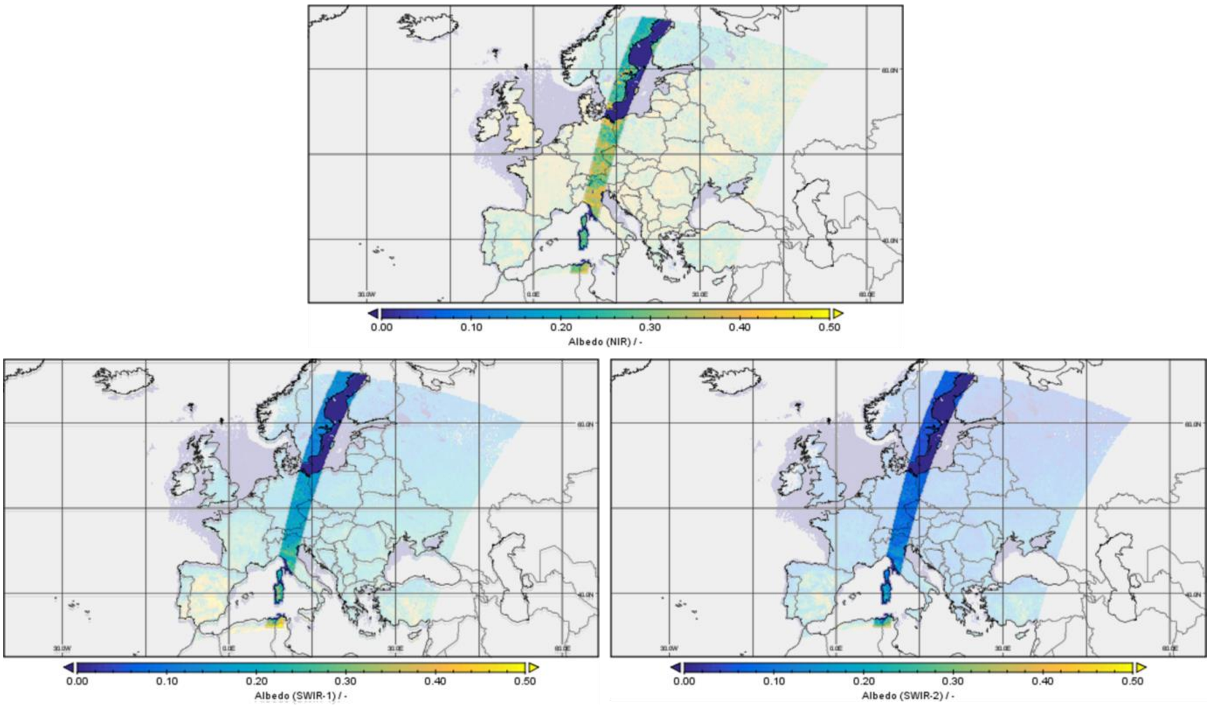

Figure 4: MODIS surface albedo in three spectral regions (NIR, SWIR-1 and SWIR-2) projected on a satellite orbit crossing the European domain on 2015-07-07. Shaded areas show the surface albedo for the entire domain as reference.

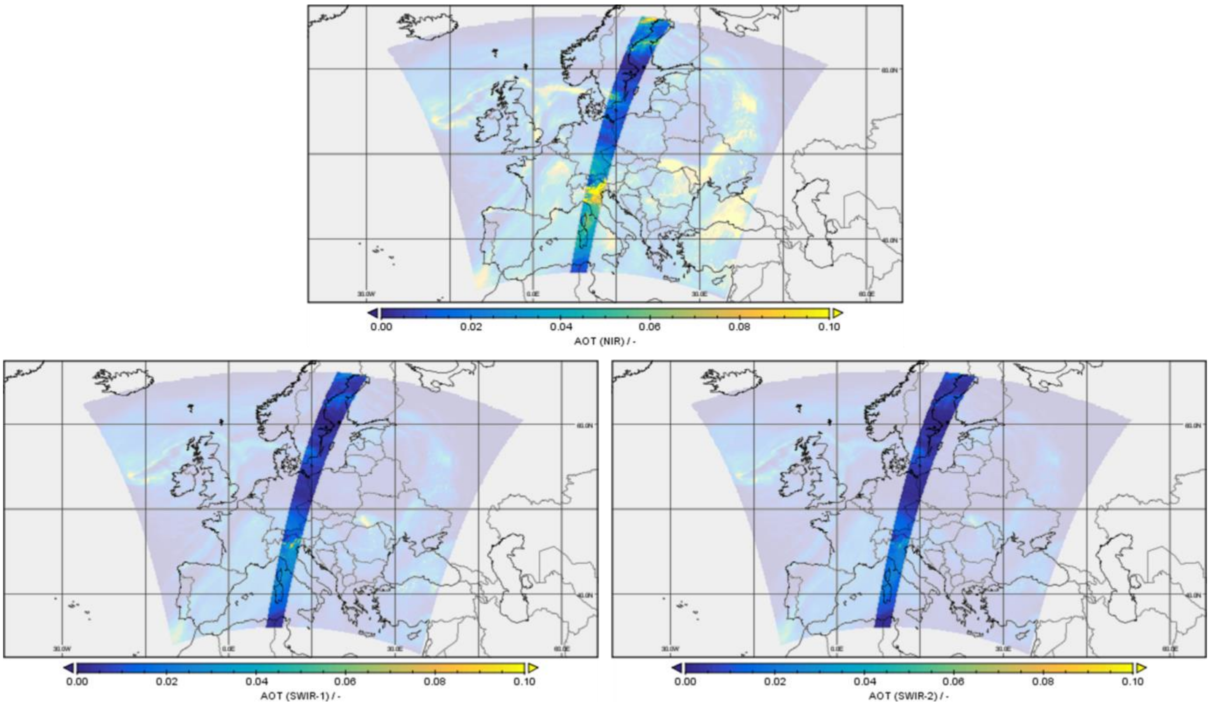

Figure 5: Modelled AOT in three spectral regions (NIR, SWIR-1 and SWIR-2) projected on a satellite orbit crossing the European domain on 2015-07-07. Shaded areas show the modelled AOT for the entire domain as reference.

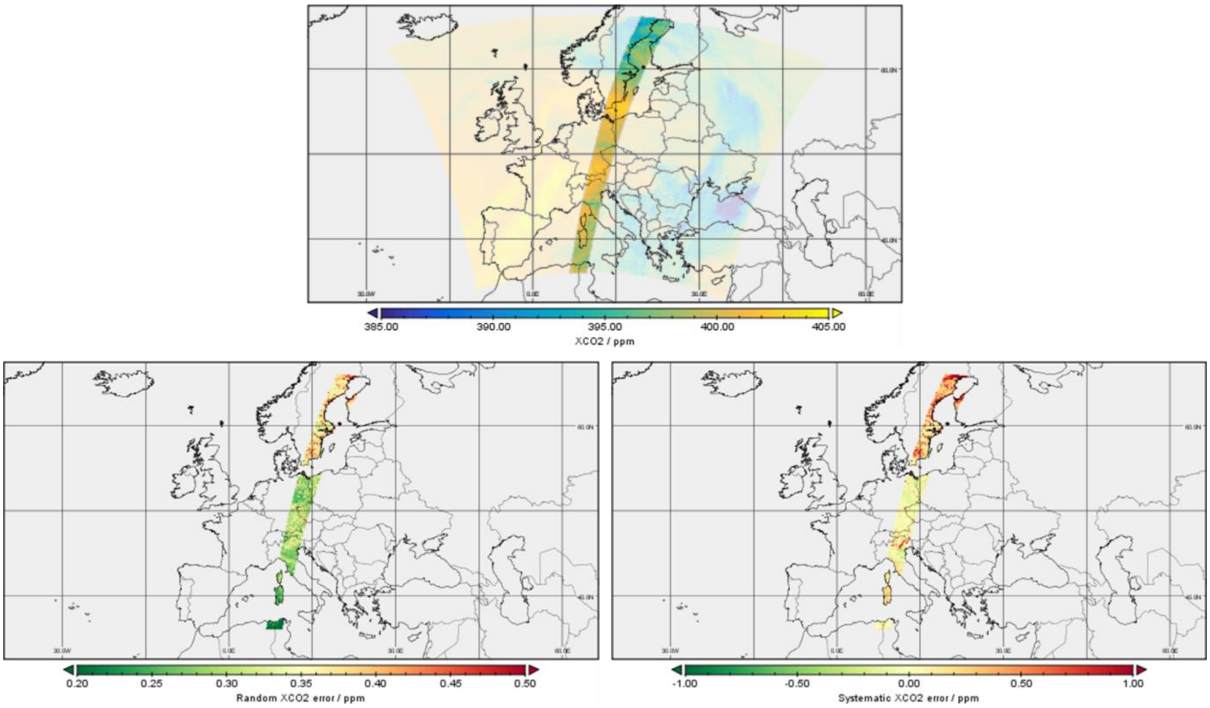

Figure 6: Modelled XCO_2 (top) projected on a satellite orbit crossing the European domain on 2015-07-07 together with the corresponding random (bottom left) and systematic (bottom right) XCO_2 retrieval errors computed with the XCO_2 error parameterization. Shaded areas show the modelled XCO_2 for the entire domain as reference.

From Figure 6, it is clear that the random XCO_2 errors generally increase with latitude. This is a result of the corresponding increase in SZA and thus the gradual decrement in signal strength. Additional variability in the random XCO_2 errors can be seen as a result of the varying surface albedo in the short-wave infrared spectral regions, where lower albedo infers larger random errors. Atmospheric aerosol and AOT has little effect on the random errors. The systematic XCO_2 errors varies between approx. -0.2 and 1.0 ppm, with the largest deviations from the true XCO_2 occurring either for comparatively thick aerosol layers (AOT (NIR) > 0.2) like in the Po Valley or intermediate aerosol layers in combination with large SZA and/or low albedo in the SWIR spectral region (and thus weak signal), like in northern Sweden.

For a better understanding of how the six simulated satellites fly in constellation, Figure 7 shows the parameterized XCO_2 errors for all six satellites as they fly over the European domain on 2015-05-03. For each satellite, two consecutive orbits are shown, meaning that the easternmost orbit is the first orbit of SAT-1 (crossing domain centre at 09:48 UTC) whereas the westernmost orbit is the second orbit of SAT-6 (crossing domain centre at 12:45 UTC). Again a clear latitudinal gradient, with some additional albedo related variability is seen in the random XCO_2 errors, ranging from approx. 0.2 to 1.0 ppm. Although the systematic errors are generally small (around 0.0 ppm i.e. no bias), there are also scenes that exhibit considerably larger errors ($\gtrsim$ 2.0 ppm, with maximum values exceeding 12 ppm) and variability. Again, the large systematic errors are triggered by thicker aerosol layers and weak signal, as a result of large SZA and low surface albedo in the SWIR spectral region.

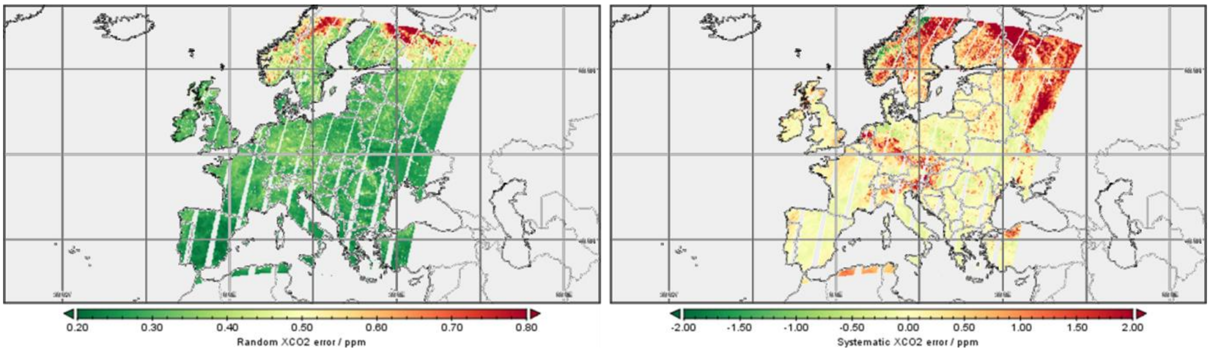

Figure 7: Parameterized random (left) and systematic (right) XCO_2 errors over the European domain on 2015-05-03. Errors from two orbits each of the six simulated satellites are shown.

5.2 Synthetic satellite observations over Berlin domain

In additional to the European domain, additional aerosol and CO_2 simulations have been performed at a higher spatial resolution of 0.01° × 0.01° for a second domain, focusing on the city of Berlin and its surroundings. Corresponding synthetic satellite observations have been simulated using the high-resolution model data and XCO_2 error parameterization, assuming the same six satellites flying in the same sun-synchronous orbit. The high-resolution model data for the Berlin domain, and hence corresponding synthetic satellite observations, are available for February and July 2015. Figure 8 to Figure 10 show the surface albedo, modelled AOT and XCO_2 as well as the corresponding XCO_2 random and systematic errors computed with the XCO_2 error parameterization for a satellite orbit (SAT-1) crossing the Berlin domain on 2015-02-14.

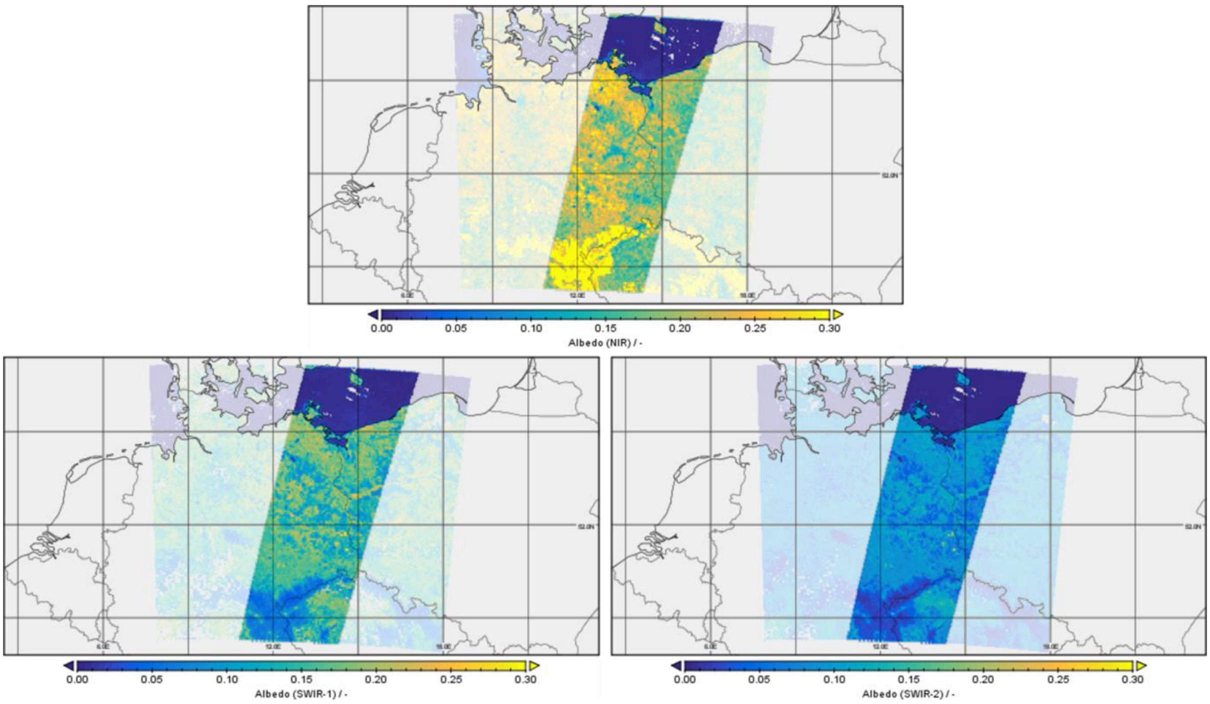

Figure 8: MODIS surface albedo in three spectral regions (NIR, SWIR-1 and SWIR-2) projected on a satellite orbit crossing the Berlin domain on 2015-02-14. Shaded areas show the surface albedo for the entire domain as reference.

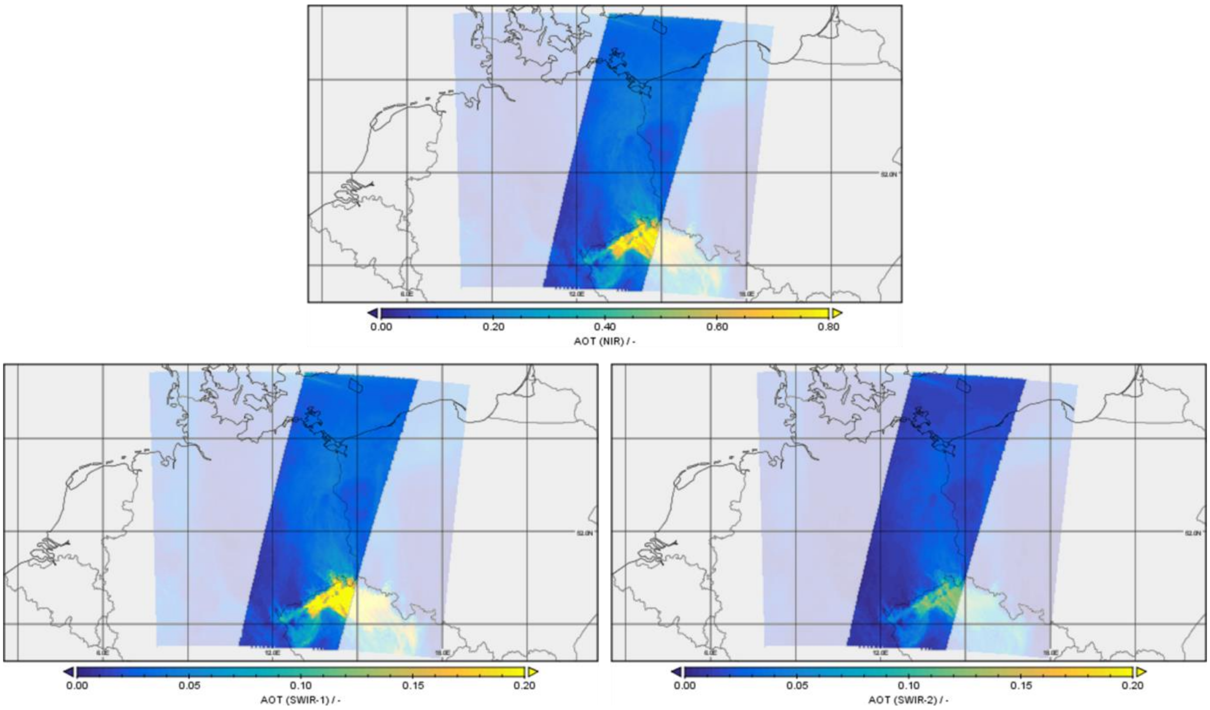

Figure 9: Modelled AOT in three spectral regions (NIR, SWIR-1 and SWIR-2) projected on a satellite orbit crossing the Berlin domain on 2015-02-14. Shaded areas show the modelled AOT for the entire domain as reference.

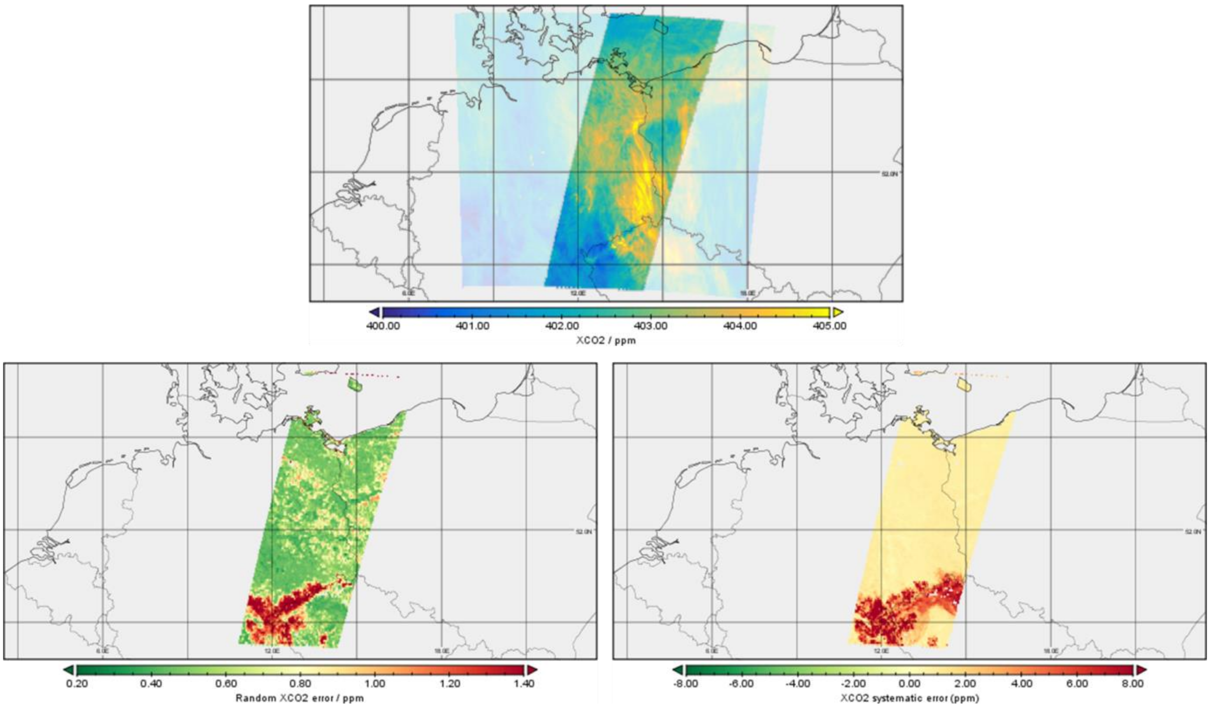

Figure 10: Modelled XCO_2 (top) projected on a satellite orbit crossing the Berlin domain on 2015-02-14 together with the corresponding random (bottom left) and systematic (bottom right) XCO_2 retrieval errors computed with the XCO_2 error parameterization. Shaded areas show the modelled XCO_2 for the entire domain as reference.

The random XCO_2 errors are generally considerably higher in Figure 10, compared to Figure 6 and Figure 7. This is directly related to the lower SZA and hence weaker signal in February, compared to July and May. Largest random errors (1-2 ppm) are observed over snow covered areas, around the border between Germany and Czech Republic, characterized by low surface albedo and hence weak signal in the SWIR spectral region, where the absorption by CO_2 takes place. Also over small isolated areas in Germany and Poland, with low surface albedo (Albedo (SWIR-2) $\lesssim 0.05$), the random XCO_2 errors are comparatively large. Also the largest systematic errors, ranging from approx. 2.0 to 12.0 ppm, are observed over snow covered areas, as a result of the low albedo and weak signal in the SWIR spectral region. Large systematic errors around 4.0 to 8.0 ppm are also observed over the Czech Republic where a thick aerosol layer with AOT (NIR) > 0.5 is modelled. For scenes with neither particularly low surface albedo (in the SWIR spectral region) nor thick aerosol layers, the systematic errors are around 0.7 ppm.

6 Data access

The synthetic satellite datasets can be downloaded from the ECMWF ftp server

```
ftp che-project@ftp.ecmwf.int
```
[1]

```
cd data-exchange/WP2/Synthetic_satellite_datasets
```

The single orbit files have been archived and compressed into daily .tar.gz files that contain the simulated satellite orbits of a given satellite (e.g. SAT-1) that intersects with a given domain at a given date. The files are stored according to the following folder hierarchy and naming convention:

```
dom
  SAT-X
    YYYY
        SAT-X_dom_YYYYMMDD.tar.gz
```

where '`dom`' represents the domain ('eur' or 'ber'), `X` the six simulates satellites ('1', '2', '3', '4', '5' and '6'), '`YYYY`' the year, '`MM`' the month and '`DD`' the day. Each .tar.gz file typically contains one to four netcdf-files with simulated synthetic satellite observations.

7 Conclusion

This document describes the synthetic satellite observations prepared as part of the deliverable *D2.5 Synthetic satellite datasets*. The task of this deliverable is to generate datasets of synthetic satellite observations with realistic estimates of random and systematic XCO_2 errors for various geophysical scenarios that can be used to analyse the effect of atmospheric aerosol for the detection and quantification of CO_2 plumes from space-borne instruments. To this end, a new error parameterization has been developed that estimates random and systematic XCO_2 retrieval errors, expected for the CO_2 instrument aboard the Copernicus CO_2 Monitoring (CO2M) mission. The parameterization has been developed using a global ensemble dataset covering a large variety of geophysical scenarios for which synthetic satellite measurements and corresponding XCO_2 retrievals have been simulated using radiative transfer calculations. The parameterization takes the sun-satellite geometry as

[1] Please contact the CHE-Coord che-coord@lists.ecmwf.int to obtain the FTP password access

well as spectrally resolved surface albedo and aerosol optical thickness (AOT) as input in order to compute the random (noise) and systematic (retrieved – true XCO_2) XCO_2 errors using two artificial neural networks. The random errors can be estimated with almost perfect precision (R=0.99), whereas the precision is slightly lower for the systematic errors (R=0.87) as a result of the many aerosol parameters (type, shape, amount, size distribution, vertical distribution) that contribute to the systematic errors.

The XCO_2 error parameterization is used to estimate the expected XCO_2 retrieval errors of six identical CO_2 instruments flying in constellation over two domains focusing on Europe (for the year 2015) and Berlin and its surroundings (for February and July 2015), respectively. AOT data computed from three-dimensional fields of aerosol properties at a high resolution of 0.05° × 0.05° (European domain) and 0.01° × 0.01° (Berlin domain) simulated with the LOTOS-EUROS model are used as input for the parameterization together with MODIS surface albedo and sun-satellite geometry computed with a satellite orbit simulator developed by SRON. In total ca. 7800 orbits that intersect with the European domain are simulated (of which 270 intersect with the Berlin domain during February and July 2015).

To simulate the synthetic satellite observations, the orbit and instrument SNR of the CO_2 spectrometer planned to fly aboard the CO2M have been used. While the XCO_2 precision requirement of CO2M (< 0.7 ppm for a vegetation scenario and SZA of 50 degrees) is met in this study, the systematic XCO_2 errors do not meet the requirement of CO2M: < 0.5 ppm up to an AOT of 0.5. This is because the Multi-Angle Polarimeter (MAP) instrument, planned to fly alongside the CO_2 instruments aboard CO2M, could ***not*** be considered when computing the systematic errors in this study. Hence, the systematic errors provided in this deliverable reflect the systematic errors expected when using data from the CO_2 instrument alone. As have been shown in previous studies, the additional information from the MAP can reduce the systematic XCO_2 errors significantly, such that the above mentioned requirement can be met (ESA, 2019). Consequently, one should include the additional data from the MAP in future studies in order to not overestimate the systematic XCO_2 errors and fully understand how the systematic XCO_2 errors are expected to propagate into CO_2 flux estimate uncertainties. The same methodology used in this study could be used to develop an improved XCO_2 error parameterization that also accounts for the additional aerosol information retrieved by the MAP. This would, however, require new radiative transfer simulations and an extended retrieval algorithm that also utilizes the MAP data when retrieving the XCO_2. This code development required to extend the retrieval algorithm was not feasible within this work package.

Since the MAP is not considered, the synthetic satellite datasets, in particular the combination of the high-resolution aerosol simulations and the XCO_2 error parameterization, can, however, also be used to study the aerosol induced systematic XCO_2 errors for other satellite instruments like Microcarb and OCO-2/3 as well, that do not have a MAP instrument aboard. Given that the aerosol error parameterization has been developed using a global ensemble of geophysical scenarios, synthetic satellite observations could also be simulated for other domains around the world, given that spectrally resolved AOT data are available.

8 List of abbreviations

ANN	Artificial Neural Network
AOT	Aerosol Optical Thickness
BRDF	Bidirectional Reflectance Distribution Function
CHE	CO_2 Human Emissions
CO_2	Carbon dioxide
CO2M	Copernicus CO2 Monitoring mission
DLR	German Aerospace Center
ECMWF	European Centre for Medium-Range Weather Forecasts
EMPA	Swiss Federal Laboratories for Materials Science and Technology
ESA	European Space Agency
MAP	Multi-Angle Polarimeter
MODIS	Moderate Resolution Imaging Spectroradiometer
NBAR	Nadir BRDF-adjusted Reflectance
NIR	Near Infrared
SNR	Signal-to-Noise Ratio
SRON	Netherlands Institute for Space Research
SWIR	Shortwave Infrared
SZA	Solar Zenith Angle
VZA	Viewing Zenith Angle
XCO_2	Column-averaged dry-air mole fraction of CO_2

9 References

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications, Atmos. Meas. Tech., 3, 781–811, doi:10.5194/amt-3-781-2010, 2010.

Butz, A., Hasekamp, O. P., Frankenberg, C., and Aben, I.: Retrievals of atmospheric CO_2 from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects, Appl. Opt., 48, 3322–3336, doi:10.1364/AO.48.003322, 2009.

Butz, A., Hasekamp, O., Frankenberg, C., Vidot, J., and Aben, I.: CH_4 retrievals from space-based solar backscatter measurements: Performance evaluation against simulated aerosol and cirrus loaded scenes, J. Geophys. Res. Atmos., 115, doi:10.1029/2010JD014514, 2010.

Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A., Aben, I., Frankenberg, C., Hartmann, J.-M., Tran, H., Kuze, A., Keppel-Aleks, G., Toon, G., Wunch, D., Wennberg, P., Deutscher, N., Griffith, D., Macatangay, R., Messerschmidt, J., Notholt, J., and Warneke, T.: Toward accurate CO_2 and CH_4 observations from GOSAT, Geophys. Res. Lett., 38, doi:10.1029/2011GL047888, 2011.

Butz, A., Galli, A., Hasekamp, O., Landgraf, J., Tol, P., and Aben, I.: TROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH_4 retrievals for aerosol and cirrus loaded atmospheres, Rem. Sens. Envir., 120, 267–276, doi:10.1016/j.rse.2011.05.030, 2012.

Butz, A., Orphal, J., Checa-Garcia, R., Friedl-Vallon, F., von Clarmann, T., Bovensmann, H., Hasekamp, O., Landgraf, J., Knigge, T., Weise, D., Sqalli-Houssini, O., and Kemper, D.: Geostationary Emission Explorer for Europe (G3E): mission concept and initial performance assessment, Atmos. Meas. Tech., 8, 4719–4734, doi:10.5194/amt-8-4719-2015, 2015.

ESA Project SMARTCARB: SMARTCARB – Use of satellite measurements of auxiliary reactive trace gases for fossil fuel carbon dioxide emission estimation, Documentation of deliverable data1 of ESA study contract n°4000119599/16/NL/FF/mg, v0.1, 2018

ESA: Copernicus CO_2 Monitoring Mission Requirements Document, EOP-SM/3088/YM-ym, v2.0, 2019

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., and Brunner, D.: Detectability of CO_2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO_2 Monitoring (CO2M) mission, Atmos. Meas. Tech., 12, 6695–6719, https://doi.org/10.5194/amt-12-6695-2019, 2019.

Manders, A. M. M., Builtjes, P. J. H., Curier, L., Denier van der Gon, H. A. C., Hendriks, C., Jonkers, S., Kranenburg, R., Kuenen, J. J. P., Segers, A. J., Timmermans, R. M. A., Visschedijk, A. J. H., Wichink Kruit, R. J., van Pul, W. A. J., Sauter, F. J., van der Swaluw, E., Swart, D. P. J., Douros, J., Eskes, H., van Meijgaard, E., van Ulft, B., van Velthoven, P., Banzhaf, S., Mues, A. C., Stern, R., Fu, G., Lu, S., Heemink, A., van Velzen, N., and Schaap, M.: Curriculum vitae of the LOTOS–EUROS (v2.0) chemistry transport model, Geosci. Model Dev., 10, 4145–4173, doi:10.5194/gmd-10-4145-2017, 2017.

Meirink, J., Eskes, H., and Goede, A.: Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modelling, Atmos. Chem. Phys., 6, 1275–1292, doi:10.5194/acp-6-1275-2006, 2006.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proc. Natl. Acad. Sci. (USA), 104, 18 925–18 930, doi:10.1073/pnas.0708986104, 2007.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R. P., Hu, B., Liang, S., Privette, J. L., and Roy, D.: First operational BRDF, albedo nadir reflectance products from MODIS, Rem. Sens. Envir., 83, 135–148, doi:10.1016/S0034-4257(02)00091-3, 2002.

Schaaf, C., Z. Wang: MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006, distributed by NASA EOSDIS Land Processes DAAC, doi:10.5067/MODIS/MCD43A4.006, 2015.

Schepers, D., aan de Brugh, J., Hahne, P., Butz, A., Hasekamp, O., and Landgraf, J.: LINTRAN v2.0: A linearised vector radiative transfer model for efficient simulation of satellite-born nadir-viewing reflection measurements of cloudy atmospheres, J. Quant. Spectrosc. Radiat. Transfer, 149, 347–359, doi:10.1016/j.jqsrt.2014.08.019, 2014.

SRON: Orbit simulator, Technical note, SRON-CSS-TN-2016-01, v0.2, 2017

Stier, P., Feichter, J., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Ganzeveld, L., Tegen, I., Werner, M., Balkanski, Y., Schulz, M., Boucher, O., Minikin, A., and Petzold, A.: The aerosol-climate model ECHAM5-HAM, Atmos. Chem. Phys., 5, 1125–1156, doi:10.5194/acp-5-1125-2005, 2005.

Strandgren, J., Bugliaro, L., Sehnke, F., and Schröder, L.: Cirrus cloud retrieval with MSG/SEVIRI using artificial neural networks, Atmos. Meas. Tech., 10, 3547–3573, doi:10.5194/amt-10-3547-2017, 2017.

Strandgren, J., Krutz, D., Wilzewski, J., Paproth, C., Sebastian, I., Gurney, K. R., Liang, J., Roiger, A., and Butz, A.: Towards space-borne monitoring of localized CO_2 emissions: an instrument concept and first performance assessment, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2019-414, in press, 2020.

Document History

Version	Author(s)	Date	Changes
V0.1	Johan Strandgren (DLR)	03/03/2020	First draft
V0.2	Johan Strandgren (DLR)	10/03/2020	New figures as a result of new model (AOT) data being provided. The parameterization was tuned. Corresponding text and figure (Figure 2) was updated accordingly. General improvements of the text, also taking comments from project partners into account.
V0.3	Johan Strandgren (DLR)	30/03/2020	The number of simulated orbits was added. Document was revised based on reviewer comments from JRC. Instances with 'CO2' were replaced with 'CO_2'. Added new section about data access.
V1.0	Johan Strandgren (DLR)	21/04/2020	Final version after review

Internal Review History

Internal Reviewers	Date	Comments
Greet Janssens-Maenhout (JRC)	26/03/2020	Approved with comments

Estimated Effort Contribution per Partner

Partner	Effort
DLR	4.5
Total	**4.5**