CO_2 Human Emissions

Impact of urban aerosols on satellite-retrieved CO2

Friedemann Reum

Sander Houweling

che-project.eu

Co-ordinated by ECMWF

D2.7 Impact of urban aerosols on satellite-retrieved CO2

Dissemination Level:	Public
Author(s):	Friedemann Reum (SRON), Sander Houweling (SRON, VUA)
Date:	21/12/2020
Version:	1.0
Contractual Delivery Date:	31/12/2020
Work Package/ Task:	WP2/ T2.5
Document Owner:	SRON
Contributors:	SRON
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

In this study, we investigate the impact of scattering-related XCO2 uncertainties of CO2M on retrieving CO2 emissions from cities. XCO2 uncertainties are provided by CHE deliverable D2.5 for a retrieval without multi-angle polarimeter (MAP) and are based on a neural network. We perturb synthetic CO2M observations according to these uncertainty estimates and assimilate them in a CO2 flux inversion system to update prior flux estimates. We analyse single overpasses of Berlin, Beijing and Shanghai in winter and summer in the year 2015. Scaling whole city emissions yields uncertainties of 11–25% in Berlin and -1–9% in the Chinese cities, suggesting promising performance for snapshots of emissions derived from single overpasses. We also try to constrain individual emission hotspots in Berlin based on the high-resolution flux dataset from D2.3. With flux errors of several tens of percent, this is less successful than constraining city totals. Reasons for the larger uncertainties of individual hotspots are conflation of several sources and sensitivity to spatial variations of XCO2 biases on the scale of individual CO2 plumes. The results for Berlin are similar to those from the H2020 project AeroCarb. However, in contrast to our study, AeroCarb found very large flux uncertainties in Beijing for the XCO2 retrieval without MAP. The main reason for this discrepancy is that CO2M uncertainties appear to have been overestimated in AeroCarb, as an update brings them closer in magnitude to the results of the neural network D2.5 uses. However, the two methods do not agree on the spatiotemporal distribution of systematic XCO2 errors, and XCO2 error variability was estimated higher outside of the small city domains we assimilated. Assimilating this higher error variability would likely yield larger flux errors than in our cases. Therefore, the differences between the XCO2 uncertainty estimation methods should be reconciled. While modelled aerosol optical depths, which are used to estimate CO2M uncertainties here, are not representative for heavily-polluted episodes in China, they are close to average conditions in the three cities and thus representative for a high number of days throughout the year 2015. Our modelled CO2 plumes from Beijing and Shanghai are on the high end of the few signals from these cities that are to date observed by OCO-2 and OCO-3. Therefore, these large signals may be rare, and inferring annual total emissions may be difficult. Overall, the ability of CO2M to constrain CO2 emissions from large cities in the presence of scattering uncertainties seems promising in favourable conditions. However, differences between XCO2 uncertainty estimation methods should be reconciled to boost confidence in flux uncertainty estimates. Constraining smaller sources, fluxes on regional scale and annual emissions likely requires minimizing the XCO2 uncertainties we investigated further, as well as addressing other error sources in CO2 flux estimation.

2 Introduction

2.1 Background

Anthropogenic greenhouse gas emissions are the main driver of modern climate change. Worldwide efforts to reduce those emissions require accurate knowledge of sources and sinks of these greenhouse gases. One of the pillars to achieve this goal is to infer the fluxes via inverse modelling based on observations of greenhouse gas concentrations in the atmosphere. One way the European Union supports these efforts is CO2M, a space-based CO2 imager. Globally, 70% of anthropogenic CO2 is emitted in cities (IEA, 2008). Therefore, the ability to accurately quantify urban CO2 emissions is key in monitoring greenhouse gas emission reductions. Here, we make a contribution towards quantifying how well CO2M can quantify urban CO2 emissions. CO2M is expected to be much better suited for this task than current CO2 monitoring satellites due to its vastly superior coverage. A special challenge for measuring CO2 from space in urban environments is the different character of aerosols in the atmosphere compared to biosphere-dominated environments. A first step towards analysing their impact on retrieving urban CO2 fluxes was done in the H2020 project AeroCarb. The project concluded that scattering-related errors in total column CO2 (XCO2) measurements were small enough for retrieving CO2 emissions from Berlin, but prohibitively large for

retrieving emissions from Beijing, which usually has much higher aerosol loads. AeroCarb also concluded that the large Beijing errors could be mitigated by carrying an auxiliary aerosol sensor (MAP, Multi-Angle Polarimeter) to augment the CO2 retrieval algorithm. Here, we expand on the analyses performed in AeroCarb by investigating cases in another period (2015 instead of 2013), and add the city of Shanghai. Furthermore, we analyse synthetic XCO2 uncertainties estimated with a new approach, the neural network described in CHE deliverable D2.5 by Johan Strandgren. They represent aerosol- and other scattering-related uncertainties anticipated for CO2M data (based on aerosol simulations with the LOTOS-EUROS model provided by TNO, D2.4), and we quantify their impact on flux retrievals.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

In this deliverable, we document the observation system simulation experiments we performed for quantifying the impact of anticipated CO2M XCO2 uncertainties on retrieving urban CO2 emissions. We introduce methods, quality control of inputs and simulations, results and conclusions.

2.2.2 Work performed in this deliverable

We simulated the capability of the CO2M observation system to constrain CO2 emissions from cities (“observation system simulation experiments”, OSSE). We focussed on single satellite overpasses over and emissions from Berlin, Beijing and Shanghai. For each city, simulations were performed for cases in summer and in winter of 2015. In addition to emissions from the whole city, we performed additional experiments where we attempted to retrieve CO2 emissions from individual hotspots in Berlin. The study relies on simulations of atmospheric transport of CO2 and aerosols. We critically evaluate the simulated aerosol loads on which XCO2 uncertainties provided to us are based, and CO2 signals, since they need to be realistic in order to yield realistic flux uncertainties. The aerosol evaluation is performed with data from AeroNet and MAIAC, and the CO2 signal evaluation with OCO-2 and OCO-3 data. We compare our results on flux uncertainties with those obtained in the H2020 project AeroCarb, of which this study is an expansion. Lastly, we draw conclusions with respect to the impact of CO2M uncertainties on retrieving CO2 emissions from cities.

2.2.3 Deviations and counter measures

The deliverable was delayed by two weeks to add an investigation of the difference and overall consistency with AeroCarb, in exchange with the radiative transfer experts that were involved in that project.

3 Forward model simulation: methods and setup

Atmospheric CO2 concentrations were modelled with WRF-CHEM with a resolution of 1.3 km over Berlin, Beijing and Shanghai. In the following sections, we give details about the model configuration and input data.

3.1 Model

Atmospheric CO2 transport was simulated with the GHG module (Beck et al., 2011) of WRF-Chem (Grell et al., 2005). We used ERA5 to initialize WRF and nudge it towards true weather conditions throughout the simulation.

3.2 Domains

The simulations focused on Berlin, Beijing and Shanghai. To downscale the driving meteorology, ERA5, the simulations covered large parent domains with 36 km resolution and nested domains with 12 km, 4 km and 1.333 km resolution over the cities (Figure 1).

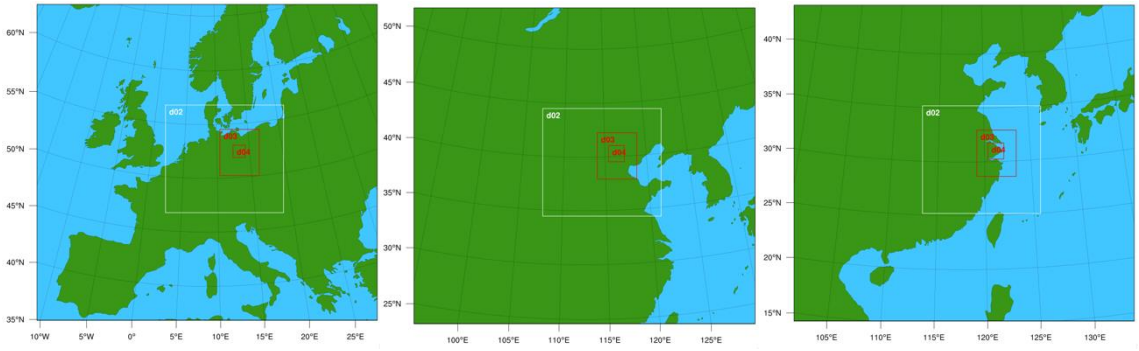

Figure 1: WRF domain configurations for Berlin (left), Beijing (centre) and Shanghai (right).

The central domain 4 spanned about 130 km in the Berlin simulations and about 160 km in both China simulations (Table 1).

Table 1: WRF domains

City	Domain	Resolution	Extent
Berlin	d01	36 km	3528 km x 3276 km
	d02	12 km	1224 km x 1116 km
	d03	4 km	408 km x 480 km
	d04	1.333 km	128 km x 128 km
Beijing	d01	36 km	3528 km x 3276 km
	d02	12 km	1224 km x 1116 km
	d03	4 km	408 km x 480 km
	d04	1.333 km	164 km x 168 km
Shanghai	d01	36 km	3528 km x 3276 km
	d02	12 km	1224 km x 1116 km
	d03	4 km	408 km x 480 km
	d04	1.333 km	164 km x 168 km

3.3 Simulation periods

In a preparatory step, two-month long WRF weather simulations were carried out to select dates with favorable conditions for CO2 measurements for each city and season. These criteria were low wind speeds of 1-3 ms^{-1} for several days prior to the overpass, and cloud-free scenes. Not many days fulfilled the requirements, and we included some days with partial cloud cover in the analyses. Forward simulations of CO2 were initialized 5 days prior to the overpass times to allow realistic 3D CO2 fields to build up. An overview is given in Table 2.

Table 2: Simulated periods.

City	Season	Initialization date	Sampled overpass time
Berlin	Winter	2015-02-09	2015-02-14 11:16 UTC
Berlin	Summer	2015-07-28	2015-08-02 11:16 UTC
Beijing	Winter	2015-01-28	2015-02-02 04:11 UTC
Beijing	Winter	2015-02-09	2015-02-14 04:11 UTC
Beijing	Summer	2015-07-04	2015-07-09 04:11 UTC
Beijing	Summer	2015-08-06	2015-08-11 04:11 UTC
Shanghai	Winter	2015-01-18	2015-01-23 03:46 UTC
Shanghai	Winter	2015-02-02	2015-02-07 03:46 UTC
Shanghai	Summer	2015-08-12	2015-08-17 03:46 UTC
Shanghai	Summer	2015-08-22	2015-08-27 03:46 UTC

3.4 Prior fluxes

For prior fluxes for Berlin, we made use of the work done by other work packages in CHE. Anthropogenic emissions in Berlin were taken from the TNO_GHGco_v1 emission inventory (Hugo Denier van der Gon, D2.3), pre-processed with code provided by Julia Marshall (MPI-BGC). Biospheric fluxes were taken from WRF-VPRM simulations (Julia Marshall, D2.3).

For the two China cases, anthropogenic emissions were taken from Edgar v4.2 FT2010. Subsequent analysis showed that this dataset produced XCO2 enhancements that were too large especially in Shanghai (Sect. 4), which is why Shanghai simulations were repeated with the MIX dataset (Li et al, 2017). Biosphere fluxes for the China cases were computed with WRF-VPRM with similar methods configuration as in D2.3 for Berlin/Europe.

Table 3: CO2 flux datasets.

City	Process	Dataset	Resolution	Reference
Berlin	Anthropogenic emissions	TNO_GHGco_v1	1 km (d04) 6 km (d01-d03)	D2.3
	Biosphere fluxes	WRF-VPRM	5 km	D2.3
Beijing and Shanghai	Anthropogenic emissions	Edgar v4.2 FT2010	0.1° x 0.1°	https://edgar.jrc.ec.europa.eu/overview.php?v= 42FT2010
	Biosphere fluxes	WRF-VPRM	d01-d04 native resolution	Own simulation
Shanghai	Anthropogenic emissions	MIX	0.25° x 0.25°	Li et al., 2017

Annual city fluxes are given in Table 4. Note that the Edgar estimate for Shanghai is much higher than the MIX estimate.

Table 4: Prior emissions of the simulated cities. For TNO and MIX we applied temporal scaling factors, the ranges cover the cases run in this study.

City	Emission dataset	Value [MtCO2/yr]
Berlin	TNO (winter case)	~23–30
Berlin	TNO (summer case)	~15–17
Beijing	Edgar	85.7
Shanghai	Edgar	310.9
Shanghai	MIX	~52–54

3.5 Boundary conditions

CO2 boundary conditions were taken from CAMS v18r1[1], applied at the outside boundaries of domain d01 during the simulation and as initial condition to all domains.

4 Forward simulation results and evaluation

4.1 Results

Forward runs of the WRF model were sampled at CO2M locations and convolved with the corresponding CO2M averaging kernel (taken from AeroCarb). Results for all simulations are shown in Figure 2. Except for the cloudy cases in Beijing (Feb 02 and Aug 11), the anthropogenic plumes are clearly visible, and in Berlin, individual power plants can be discerned. The maximum anthropogenic signals are given in Table 5. Beijing winter and all Shanghai Edgar cases stick out with very large XCO2 plume signals. Because we deemed these signals unrealistically high (based on an evaluation against real OCO-2 and OCO-3 measurements, Sect. 4.2), we repeated Shanghai simulations with anthropogenic emissions from the MIX dataset, which assigns much lower emissions to Shanghai (Table 4). Most Shanghai MIX simulations show large XCO2 signals over the mouth of the Yangtze river (Figure 2, fourth row). This is an artefact that may be caused by a combination of the low resolution of the MIX dataset of 0.25°, which places emissions into the river, and very low simulated planetary boundary height above water. This artefact is unlikely to affect our results, since we do not assimilate ocean observations.

[1] https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-greenhouse-gas-inversion

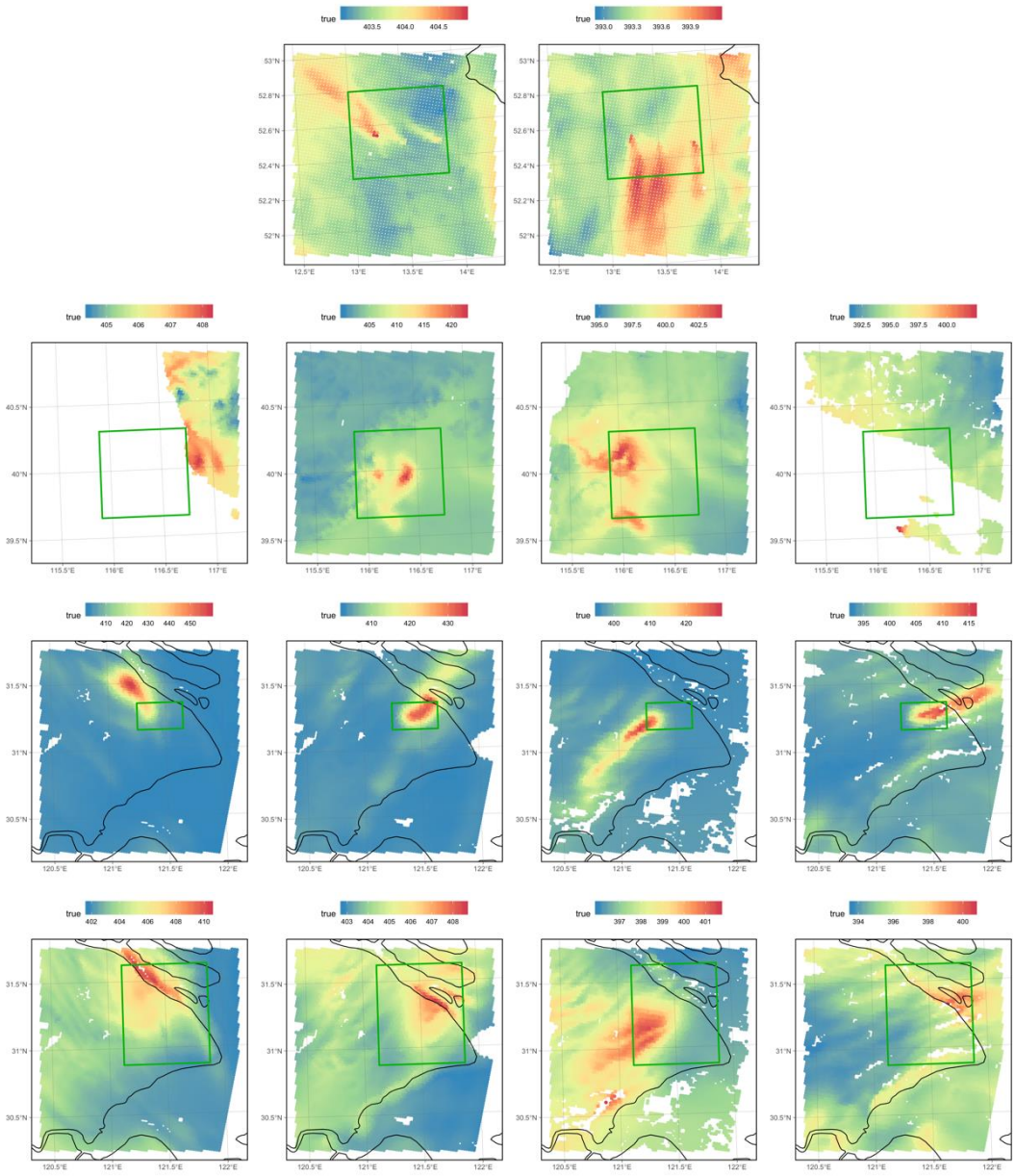

Figure 2: Forward simulations sampled at CO2M sounding locations and convolved with their averaging kernel. Here, ocean observations are included to show plume shapes. Green rectangles denote city outlines. The city definition of Shanghai depends on the resolution of the emission dataset. First row: Berlin. Left to right: Feb 14, Aug 02 2015. Second row: Beijing. Left to right: Feb 02, Feb 14, Jul 09, Aug 11 2015. Third row: Shanghai with Edgar emissions. Left to right: Jan 23, Feb 07, Aug 17, Aug 27 2015. Fourth row: Shanghai with MIX emissions. Left to right: Jan 23, Feb 07, Aug 17, Aug 27 2015.

Table 5: Signal size of simulated city plumes. The numbers here do not take into account smoothing error (averaging kernels), but are simple column averages. Ocean measurements are included. The 90th percentile refers to all pixels with a city signal above 0.5 ppm.

City	Emission data	Sampled overpass time	Max [ppm]	90th percentile [ppm]
Berlin	TNO	2015-02-14	2.5	1.6
Berlin	TNO	2015-08-02	1.1	0.7
Beijing	Edgar	2015-02-02	20.0	6.8
Beijing	Edgar	2015-02-14	35.0	12
Beijing	Edgar	2015-07-09	6.7	4.3
Beijing	Edgar	2015-08-11	7.2	3.3
Shanghai	Edgar	2015-01-23	100	36
Shanghai	Edgar	2015-02-07	42	26
Shanghai	Edgar	2015-08-17	38	16
Shanghai	Edgar	2015-08-27	25	19
Shanghai	Mix	2015-01-23	14	7.5
Shanghai	Mix	2015-02-07	8.5	4.8
Shanghai	Mix	2015-08-17	5.8	4.1
Shanghai	Mix	2015-08-27	7.8	4.4

4.2 Comparison to real OCO-2/OCO-3 data

While the city plumes for Berlin were of expected magnitude, we considered those obtained with Edgar emissions for Beijing (winter) and for Shanghai (winter and summer) very large (Table 5). Therefore, we searched the OCO-2 and OCO-3 data archives for overpasses (OCO-2) and Snapshot area maps ("SAM", OCO-3) of these cities. Very few suitable samples are available in both datasets. With OCO-2, some overpasses show clear enhancements downwind of the cities, some show none, and in some cases, data are missing (perhaps because of pollution). The largest XCO2 enhancements attributable to Beijing emissions that we found as of August 2020 were 6.9 ppm (OCO-2) / ~9 ppm (OCO-3) (Figure 3), and for Shanghai 4.6 ppm (OCO-2) / ~5 ppm (OCO-3) (Figure 4)[2]. The XCO2 enhancements observed by OCO-2 and OCO-3 indicate that our Beijing simulations in summer are in the observed range, but larger than typical observed enhancements in winter. Observations over Beijing in winter are too sparse to conclude with certainty that the simulations overestimate real XCO2 enhancements. Thus, we conclude that results from Beijing in winter may be optimistic. Shanghai plumes simulated with Edgar strongly overestimate typical real observations in all cases. To allow interpretation of Shanghai simulations, they were repeated with MIX as the anthropogenic emission dataset, which assigns much smaller emissions to Shanghai than Edgar (Table 4), yielding more realistic XCO2 enhancements (Table 5).

[2] Since only the magnitude of the signals mattered, the OCO-3 numbers were not estimated with a rigorous method, but by eye from plots from the OCO-3 data archive (Figure 3 and Figure 4)

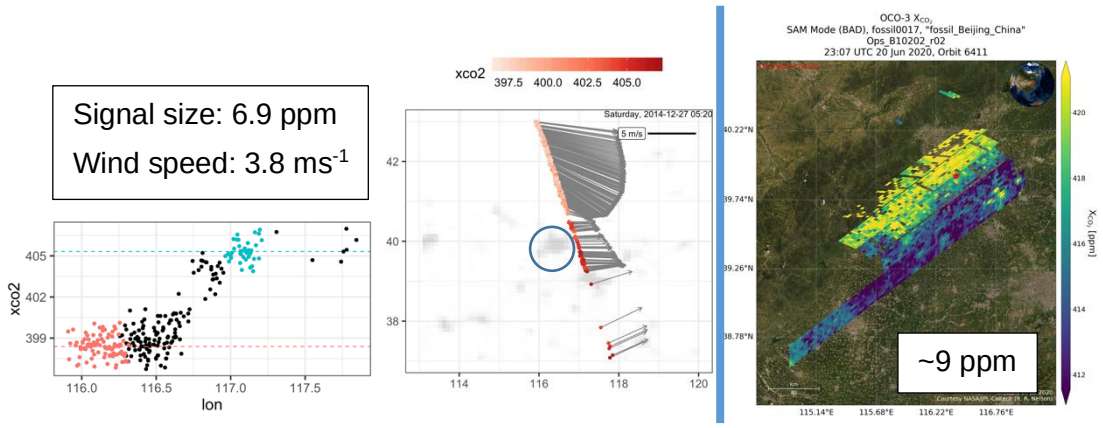

Figure 3: Largest XCO2 enhancement found observed by OCO-2 (left) and OCO-3 (right) downwind of Beijing as of August 2020.

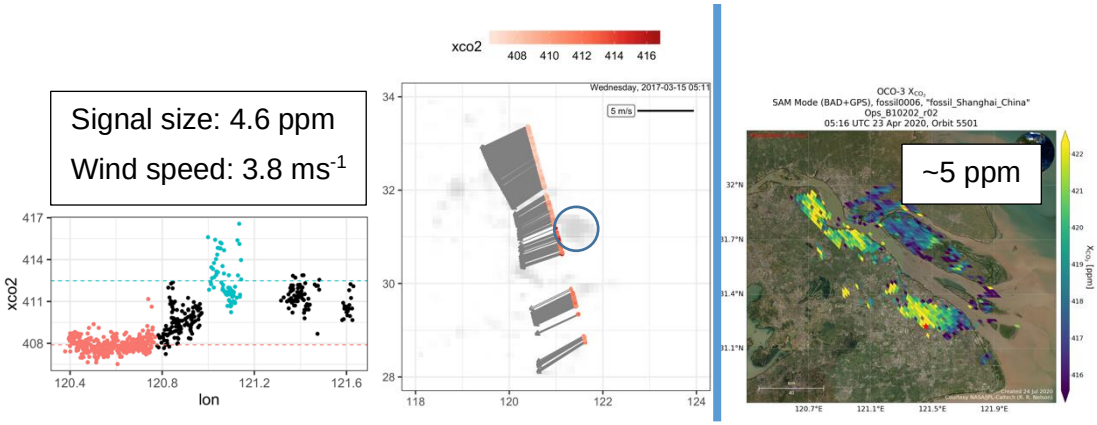

Figure 4: Same as Figure 3, but for Shanghai.

5 Aerosol-induced XCO2 uncertainties

5.1 Characterization

Systematic and random XCO2 uncertainties were computed by Johan Strandgren with the methodology described in CHE deliverable D2.5, applied to the cases analysed here. These XCO2 uncertainty estimates depend on albedo and aerosol optical thickness in three wavelengths (NIR, SWIR1, SWIR2), solar zenith angle and viewing zenith angle. Two neural networks that accept these quantities as input were trained to yield systematic and random XCO2 uncertainties, respectively (for further details on the method, see D2.5). A summary of the uncertainties is given in Table 6. We excluded ocean observations from the original dataset, since the nadir algorithm is known to be unsuitable for them. Random uncertainties are rather uniform (standard deviation ~0.1) and on average 0.3 ppm in most cases, with a few outliers. Systematic uncertainties vary on the order of 0.3 ppm as well, with some exceptions that are likely related to pockets of low simulated planetary boundary layer height (Beijing Aug 11, Shanghai Aug 27). Mean biases are -0.5–1.1 ppm. In comparison with average signal sizes (Table 5), the average XCO2 errors of the non–cloudy China cases are small, whereas those of the Berlin winter case may have a larger impact on the flux retrieval.

Table 6: Summary statistics of random and systematic XCO2 uncertainties for all scenes. Ocean observations are excluded. All values are in ppm.

		Random XCO2 error				Systematic XCO2 error			
City	**Case**	**Max**	**Mean**	**Med.**	**St.d.**	**Max**	**Mean**	**Med.**	**St.d.**
Berlin	2015-02-14	1.3	0.5	0.4	0.1	5.4	1.1	0.9	0.6
Berlin	2015-08-02	0.7	0.3	0.3	0.1	2.2	0.0	-0.1	0.3
Beijing	2015-02-02	0.7	0.3	0.3	0.1	0.9	0.4	0.5	0.3
Beijing	2015-02-14	0.6	0.3	0.3	0.0	1.9	0.1	0.1	0.3
Beijing	2015-07-09	0.8	0.3	0.3	0.0	1.8	-0.5	-0.5	0.3
Beijing	2015-08-11	0.9	0.3	0.3	0.0	29.2	-0.4	-0.3	1.5
Shanghai	2015-01-23	2.6	0.3	0.3	0.1	3.9	0.3	0.2	0.4
Shanghai	2015-02-07	5.5	0.3	0.3	0.1	3.4	0.0	-0.1	0.5
Shanghai	2015-08-17	1.0	0.3	0.3	0.0	7.0	0.0	0.0	0.4
Shanghai	2015-08-27	19.0	0.4	0.3	0.6	47.9	0.6	-1.2	5.5

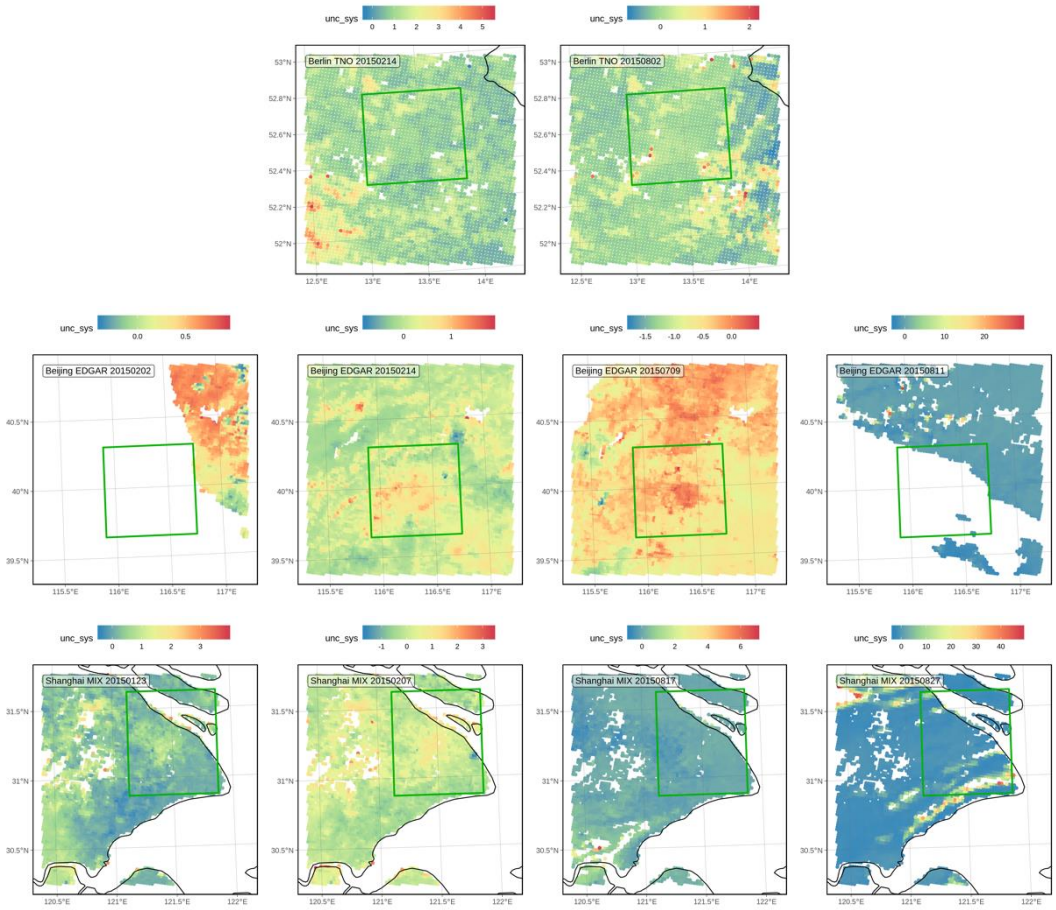

Figure 5: Estimates of systematic XCO2 uncertainties. Top row: Berlin simulations, left to right: Feb 14, Aug 02 2015. Middle row: Beijing simulations, left to right: Feb 02, Feb 14, Jul 09, Aug 11 2015. Bottom row: Shanghai simulations, left to right: Jan 23, Feb 07, Aug 17, Aug 27 2015.

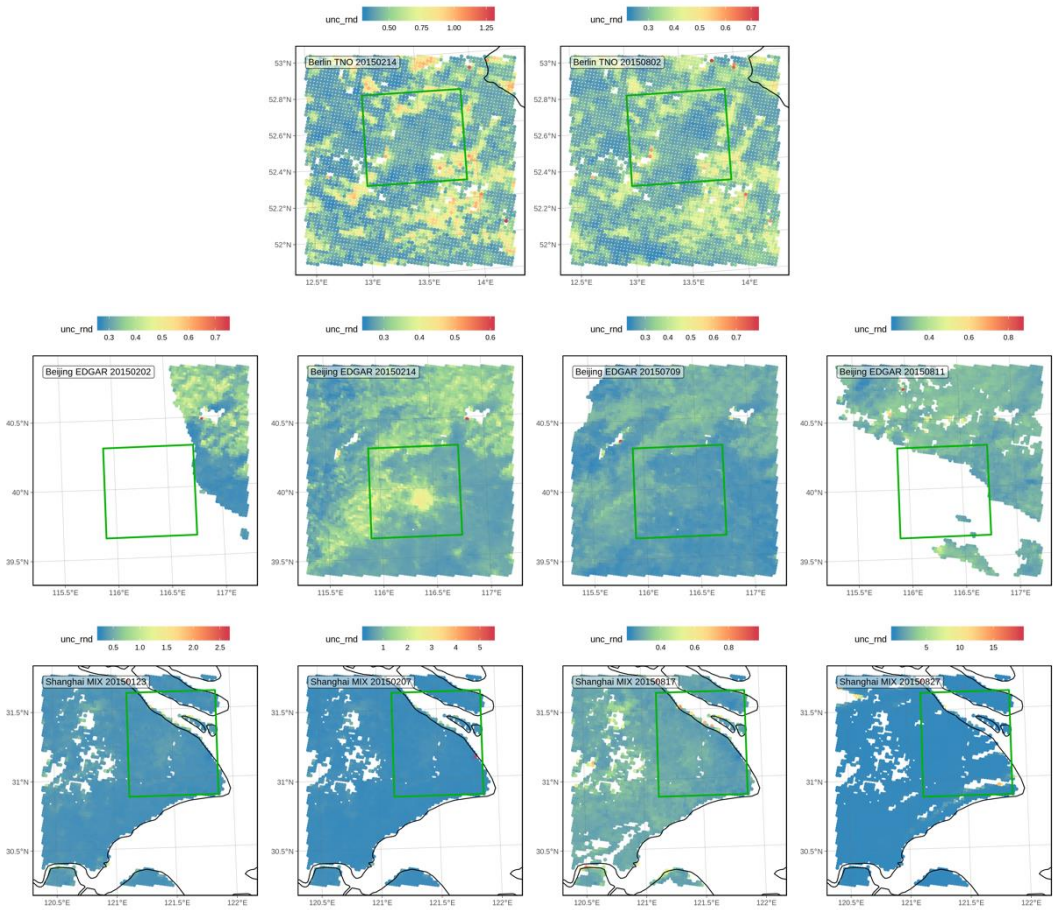

Figure 6: Estimates of random XCO2 uncertainties. Top row: Berlin simulations, left to right: Feb 14, Aug 02 2015. Middle row: Beijing simulations, left to right: Feb 02, Feb 14, Jul 09, Aug 11 2015. Bottom row: Shanghai simulations, left to right: Jan 23, Feb 07, Aug 17, Aug 27 2015.

5.2 Comparison with updated AeroCarb XCO2 uncertainty estimates

In the AeroCarb project, first steps were taken to investigate the impact of aerosols on urban CO2 emission quantification. For this purpose, CO2M XCO2 uncertainties were estimated for Berlin and Beijing, summer and winter, and for instrument configuration with and without a multi-angle polarimeter (MAP). In Berlin, the estimate of mean bias without MAP was on similar order of magnitude as here, but its variation was lower. In Beijing and Shanghai, the mean bias and its variation estimated here were about an order of magnitude smaller than without MAP in AeroCarb, and mostly slightly smaller than those with MAP in AeroCarb (compare Table 6 and Table 7). Therefore, we expect similar flux uncertainties as in the cases with MAP in AeroCarb.

The AeroCarb methodology has since been updated. The XCO2 error in AeroCarb was estimated by linearly mapping different error components onto XCO2. However, the aerosol model used to generate the synthetic measurement differs from what is assumed in the retrieval without MAP, which makes the aerosol-induced error difficult to estimate. Therefore, it was replaced by the difference between retrieved and true XCO2. Based on this update, we reprocessed synthetic XCO2 uncertainties for the AeroCarb cases. Especially in the Beijing cases without MAP, this update results in lower systematic uncertainties and variabilities (Trismono Krisna and Jochen Landgraf, personal communication, 2020).

We also computed XCO2 uncertainties for the AeroCarb cases with the trained neural network method from D2.5 (courtesy Johan Strandgren). The magnitude of systematic uncertainties estimated based on the neural network method is consistent with the updated AeroCarb method, with both yielding values similar to the ones with MAP obtained with the original

AeroCarb method (Table 7). However, the two methods do not agree on the distribution of uncertainties. For example, there are offsets and different dependencies on AOT (Figure 7). Explaining and reconciling these differences requires further investigation.

Table 7: Mean and variation of systematic CO2M XCO2 uncertainties estimated for the 2013 cases of the AeroCarb project. Shown here are original values from the AeroCarb project ("RemoteC", values from Table 4 in the final report of AeroCarb WP3), values from the updated AeroCarb methodology ("RemoteC updated", see text), and values from the neural network approach from D2.5 applied to the AeroCarb cases ("DLR NN"). Mean values for "RemoteC updated" and "DLR NN" are given as mean absolute values because the "RemoteC" method yielded only positive systematic uncertainties by design. All values are given in ppm.

Case	Mean RemoteC	Mean RemoteC updated	Mean DLR NN	St. dev. RemoteC	St. dev RemoteC updated	St. dev. DLR NN
Berlin summer with MAP	0.22			0.13		
Berlin summer without MAP	0.73	0.19	0.13	0.16	0.13	0.28
Berlin winter with MAP	0.66			0.47		
Berlin winter without MAP	3.63	0.34	0.51	1.31	0.21	0.40
Beijing summer with MAP	0.43			1.04		
Beijing summer without MAP	5.06	1.0	0.73	2.44	0.63	0.67
Beijing winter with MAP	0.82			1.47		
Beijing winter without MAP	18.44	0.90	1.87	13.83	0.84	1.31

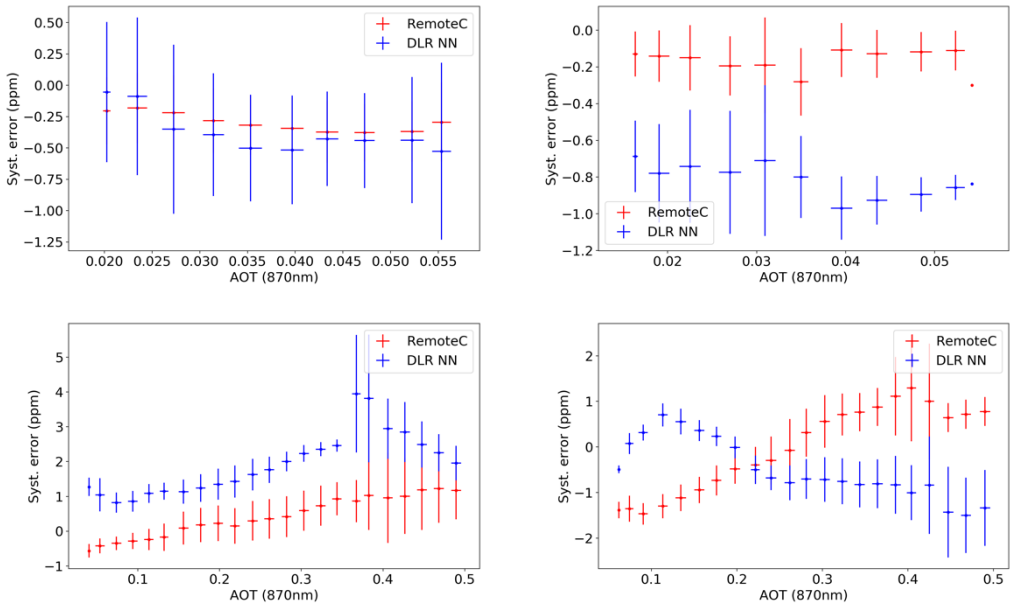

Figure 7: Difference between true and retrieved XCO2 (representing systematic XCO2 uncertainties) without MAP versus AOD for AeroCarb cases based on the AeroCarb method ("RemoteC") and the neural network method ("DLR NN"). Error bars represent the 1σ variation in systematic error across the measurements per AOD bin. Top: Berlin. Bottom: Beijing. Left: summer. Right: winter (all 2013).

5.3 Evaluation of simulated aerosol load with observations

The mismatch in simulated errors in comparison with AeroCarb, as well as (on first inspection) seemingly low aerosol loads simulated with LOTOS-EUROS (by Arjo Segers, TNO, as part of CHE task T2.3), which were used as input for CO2M XCO2 uncertainties in D2.5, prompted us to assess whether simulated aerosols represent realistic conditions. We used aerosol observations from AeroNet and MAIAC for this assessment. AeroNet is a ground-based aerosol observation network with several stations in the swaths of simulated CO2M data (Table 8), including two in the city centre of Beijing[3]. MAIAC is a product based on Terra/Aqua observations with a resolution of 1 km[4]. AeroNet stations report AOD at several wavelengths, including 870 nm (same as LOTOS-EUROS) and 440 nm (close to the 470 nm of MAIAC). Therefore, we compare LOTOS-EUROS results directly to AeroNet, and AeroNet observations to MAIAC. Due to the wavelength mismatch, we do not compare LOTOS-EUROS simulations directly to MAIAC.

Table 8: AeroNet stations used for comparison with CO2M[5].

Domain	Station name	Remark
Berlin	Leipzig	
	Lindenberg	
Beijing	Beijing	Beijing centre
	Beijing-CAMs	Beijing centre
	Xiang He	
Shanghai	Taihu	

[3] https://aeronet.gsfc.nasa.gov/cgi-bin/webtool_aod_v3
[4] doi:10.5067/MODIS/MCD19A2N.NRT.006
[5] https://aeronet.gsfc.nasa.gov/cgi-bin/webtool_aod_v3

For the comparisons, we sampled LOTOS-EUROS at AeroNet station locations (all cities), and downloaded MAIAC data for a small domain that covers the city centre (only Beijing). From AeroNet, we sampled observations during our simulated overpass times and the average for the year 2015. From MAIAC, we computed annual averages of daily mean and maximum AOD, an interval that should cover conditions in Beijing city centre. The results are shown in Table 9. While simulated AOD's were too low in Beijing at the time of simulation compared to AeroNet, they were representative for annual average AOD's in all cities. In addition, analysed MAIAC observations were similar to AeroNet averages.

Table 9: AOD averages as simulated for CO2M error models, and observed by AeroNet stations and MAIAC. The coloured outlines indicate which values can be directly compared (same wavelength).

	Berlin	Beijing	Shanghai
LOTOS-EUROS @ 870nm (at AeroNet station locations)	0.057	0.32	0.35
AeroNet during overpass @ 870nm	0.079	0.72	No observations
AeroNet mean @ 870nm	0.087	0.40	0.34
AeroNet mean @ 440 nm	0.22	0.85	0.77
MAIAC mean @ 470 nm	Not analysed	0.70	Not analysed
MAIAC max @ 470 nm	Not analysed	0.93	Not analysed

In Beijing, these observations cover 80% (AeroNet) and 70% (MAIAC) of the days of the year, and they include both clear and heavily polluted days (Figure 8).

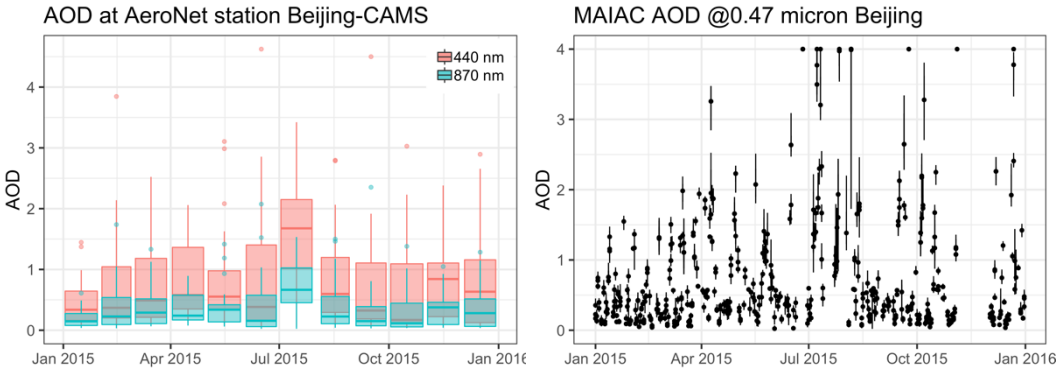

Figure 8: Observations of AOD in Beijing. Left: AeroNet station Beijing-CAMS; monthly values for both 440 nm (close to MAIAC wavelength) and 870 nm (LOTOS-EUROS wavelength). Right: MAIAC observations in a small domain around Beijing city centre.

We conclude that the aerosol loads simulated by LOTOS-EUROS, which were used to estimate XCO2 uncertainties, are representative for average real conditions in all domains.

6 Emission optimisation methods

We estimated emissions for whole cities with the analytical inversion method developed for the AeroCarb project. In a second experiment, we attempted to constrain large individual point sources in Berlin with a newly developed Ensemble Kalman Filter method. Both methods solve the Bayesian inference problem by minimizing the Bayesian cost function $J(\boldsymbol{f})$:

$$J(\boldsymbol{f}) = \big(\boldsymbol{c} - \boldsymbol{H}(\boldsymbol{f})\big)^T \boldsymbol{R}^{-1}\big(\boldsymbol{c} - \boldsymbol{H}(\boldsymbol{f})\big) + (\boldsymbol{f} - \boldsymbol{f_a})^T \boldsymbol{B}^{-1}(\boldsymbol{f} - \boldsymbol{f_a}) \quad \textbf{Eq. (1)}$$

Here, $\boldsymbol{f}$ are the unknown fluxes to estimate, $\boldsymbol{f_a}$ is a prior guess, $\boldsymbol{c}$ are observations and $\boldsymbol{H}$ is the forward model that links fluxes to observations (here, $\boldsymbol{H}$ comprises the atmospheric transport model and its sampling method for total column observations). $\boldsymbol{R}$ is the model-data mismatch covariance matrix, representing uncertainties in the observations and the transport model, and $\boldsymbol{B}$ is the prior covariance matrix, representing how uncertain the prior guess of the solution, $\boldsymbol{f_a}$, is.

If $\boldsymbol{H}(\boldsymbol{f})$ is linear, it can be expressed as a matrix ($\boldsymbol{z} = \boldsymbol{c} - \boldsymbol{c_{ini}} = \boldsymbol{H} \cdot \boldsymbol{f} + \boldsymbol{\varepsilon}$) and minimizing Eq. (1) yields optimal fluxes $\hat{\boldsymbol{f}}$ and their posterior covariance matrix $\boldsymbol{V}_{\hat{\boldsymbol{f}}}$ (e.g. Tarantola, 2005):

$$\hat{\boldsymbol{f}} = \boldsymbol{f_a} + \boldsymbol{B}\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T + \boldsymbol{R})^{-1}(\boldsymbol{z} - \boldsymbol{H}\boldsymbol{f_a}) \quad \textbf{Eq. (2)}$$

$$\boldsymbol{V}_{\hat{\boldsymbol{f}}} = \boldsymbol{B} - \boldsymbol{B}\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T + \boldsymbol{R})^{-1}\boldsymbol{H}\boldsymbol{B} \quad \textbf{Eq. (3)}$$

The optimization methods that were used differ in the approach to minimizing Eq. (1) and the state vector that is solved for. In the following two paragraphs, these aspects are explained for each method.

6.1 Analytical inversion for whole city emissions

In the first set of experiments, we estimated whole city emissions with an analytical inversion. It solved for three parameters:

- Anthropogenic CO2 emissions from a city (50% prior uncertainty)
- All other CO2 fluxes from the rest of the domain, including natural fluxes in the city (50% prior uncertainty)
- Background CO2 concentration (1 ppm prior uncertainty)

The state vector elements were assumed to be statistically independent (i.e. $\boldsymbol{B}$ was diagonal).

The model-data mismatch was computed by adding the variances of systematic (Figure 5) and random (Figure 6) XCO2 uncertainties, also assuming statistical independence (i.e. $\boldsymbol{R}$ was diagonal). Observations were computed by perturbing the known truth (Figure 2) with systematic uncertainties.

For this simple setup, no numerical approximations are needed to minimize the Bayesian cost function, so the computation was done analytically.

6.2 Ensemble Kalman Filter for distinguishing individual sources within Berlin

Additional experiments were carried out for Berlin with an Ensemble Kalman Filter. The assimilation system was developed with support from CHE and SCARBO, another H2020-project. It is based on the CarbonTracker Data Assimilation Shell (“CTDAS”, van der Laan-Luijkx et al., 2017), which solves the linearized Bayesian optimization problem via pseudorandom samples that represent fluxes and data (Peters et al., 2005). We developed a regional version of CTDAS by exchanging the global transport model TM5, which the original CTDAS relies on, for WRF-Chem (Grell et al., 2005), which can be run at much higher resolution. From WRF-Chem, we used the GHG module for computing CO2 transport (Beck et al., 2011).

In these experiments, we took advantage of the high-resolution anthropogenic emission inventory from TNO for Berlin to try and constrain large individual point sources within the city. The setup was similar as in the experiments with the analytical inversion method, except that the state vector, depicted in Figure 9, consisted of the following elements:

- Total CO2 fluxes from three pixels that contain power plants and a cement production plant (see

Table 10, 100% prior uncertainty)
- Total CO2 fluxes from the rest of Berlin (100% prior uncertainty)
- Total CO2 fluxes from outside of Berlin (100% prior uncertainty)
- Background CO2 concentration (0.5 ppm prior uncertainty)

Note that we did not distinguish anthropogenic emissions from natural fluxes within the city perimeter for this experiment, so the results are not directly comparable with those from the analytical method.

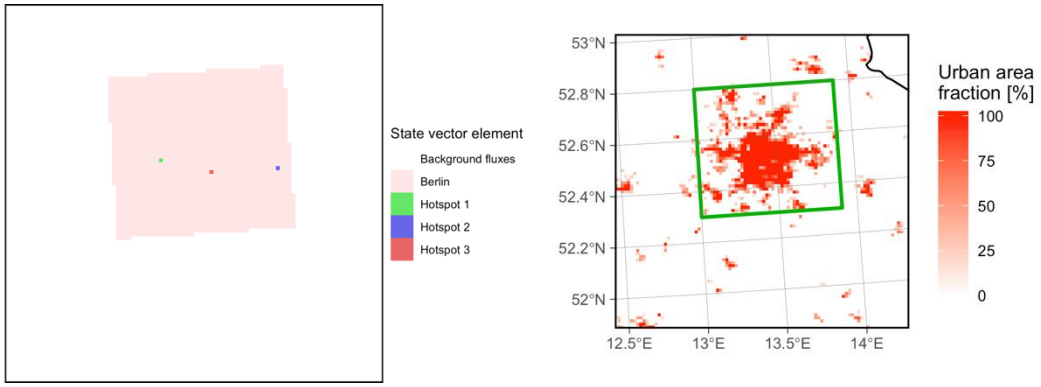

Figure 9: Left: state vector map for constraining individual CO2 hotspots in Berlin in the innermost WRF domain. Right: For reference, urban area fraction from WRF with the city outline in green, encircling the city area (area in light red in left plot).

Table 10: Description of CO2 emission hotspots in Berlin that we solved for in the hotspot experiment. The hotspots are single pixels in the WRF grid.

Hotspot	Notable CO2 emitters
HS1	Heizkraftwerk Reuter West Heizkraftwerk Reuter BSR Müllheizkraftwerk Ruhleben
HS2	Fels-Werke GmbH Kalkwerk Rüdersdorf CEMEX Zement GmbH IKW Rüdersdorf
HS3	Heizkraftwerk Klingenberg

7 Emission results and discussion

7.1 Whole city emissions

Inversion results for city emissions are shown in Figure 10 and in Table 11. Results for Berlin (~11–25% systematic flux uncertainty, ~3–8% random flux uncertainty) are similar to the results of AeroCarb (~5–17% systematic, ~5–9% random uncertainty).

In Beijing and Shanghai, emission uncertainties are slightly larger than those of Beijing with MAP in AeroCarb, but much smaller than those without MAP. In AeroCarb, systematic uncertainties of retrieved Beijing emissions without a MAP were extremely large at ~200–800%, whereas our clear-sky Beijing and Shanghai cases yield systematic flux errors of -1 to 9% (excluding Shanghai simulations with Edgar, which yielded unrealistically large XCO2 enhancements), which is closer to AeroCarb results with MAP (0.2–2%). The differences between the projects are similar for random flux uncertainties, which were 15–30% in the

Beijing cases in AeroCarb without MAP, whereas in our cases they are 0.5–2% in the Chinese cases, similar to the AeroCarb estimates of 0.4–1.2% with MAP. These results are consistent with our comparison of XCO2 uncertainties, which were similar to the case with MAP in AeroCarb (Sect. 5.2). More factors than differences in assimilated CO2 uncertainties may contribute to higher errors in AeroCarb. In that project, XCO2 retrievals in polluted conditions (i.e. over Beijing city centre) did not converge and were thus not assimilated, excluding the largest XCO2 signals. The neural network approach to quantifying XCO2 uncertainties used here does not allow to take this effect into account. Thus, in reality, some of the highest XCO2 signals assimilated here may not be observable. On the other hand, simulated AOD's in our study were on average well below 1 even in the city centre of Beijing, so this limitation may not be significant. The major driver is likely the differences in systematic XCO2 errors. This is supported by the flux results with updated XCO2 uncertainties, as they are close to our results (Sect. 7.3).

In the China cases, random flux errors are negligible, with almost all non-cloudy cases being below one percent (and below 10% in the cloudy cases).

In Beijing, the detrimental effect of partial cloud cover can be observed: the cloud cover masked the city plume (Figure 2, second row), and emissions obtained had systematic errors of ~36–57%, much larger than in the non-cloudy cases in Beijing (~3 and 8%).

In Shanghai, we can observe a direct effect of overestimating the city plume with the Edgar dataset (Sect. 4) on uncertainties of retrieved fluxes: the average of systematic flux errors with MIX was 7.5 times that obtained with Edgar (with large variations among overpasses), which is a similar factor as that of the average signal size of plumes, where the factor was 4.7 (Table 5). Still, all systematic uncertainties in Shanghai with MIX fluxes are low (-1–9%), similar to the Beijing results.

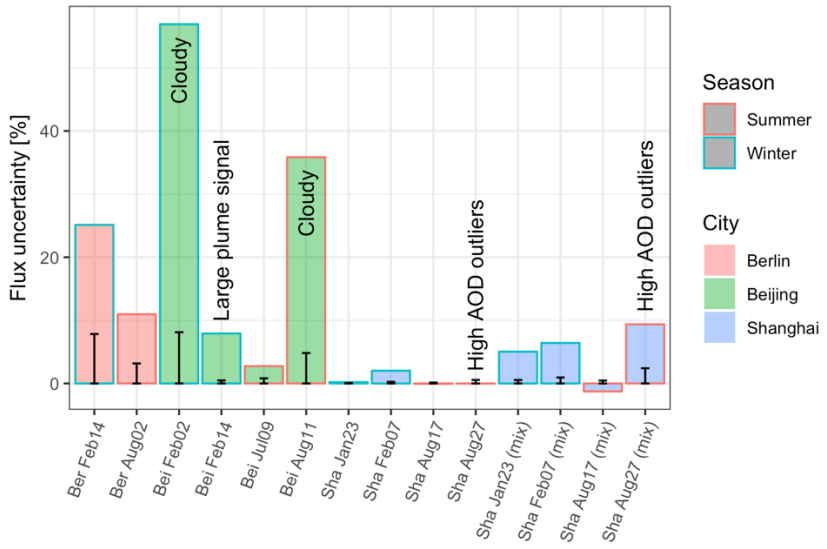

Figure 10: Uncertainties of retrieved city emissions with the analytical inversion method. Solid bars indicate deviation of retrieved emissions from truth (due to systematic XCO2 uncertainties). Error bars indicate posterior random uncertainty.

Table 11: Uncertainties of retrieved city emissions with the analytical inversion method in percent.

City	Case	Flux	Random	Systematic
Berlin	2015-02-14	TNO	8	25
Berlin	2015-08-02	TNO	3	11
Beijing	2015-02-02	EDGAR	8	57
Beijing	2015-02-14	EDGAR	0.5	7.9
Beijing	2015-07-09	EDGAR	0.8	2.7
Beijing	2015-08-11	EDGAR	5.	36
Shanghai	2015-01-23	EDGAR	0.1	0.2
Shanghai	2015-02-07	EDGAR	0.3	2.0
Shanghai	2015-08-17	EDGAR	0.1	-0.05
Shanghai	2015-08-27	EDGAR	0.6	-0.02
Shanghai	2015-01-23	MIX	0.6	5.0
Shanghai	2015-02-07	MIX	0.9	6.4
Shanghai	2015-08-17	MIX	0.5	-1.3
Shanghai	2015-08-27	MIX	2.0	9.4

7.2 Berlin hotspot experiment

Here, we take a closer look at CO2 fluxes in Berlin with an Ensemble Kalman Filter. The goal is to elucidate whether emissions from individual hot spots within the city can be constrained. Note that the estimated uncertainties are not directly comparable to those of the analytical inversion method because of differences in the state vector: while the analytical inversion focused on anthropogenic emissions only, we optimize total fluxes in the hotspot experiment.

7.2.1 Prior fluxes and uncertainties

Prior fluxes depend on the analysed period and are shown in Table 12. Particularly background and city (without hotspots) fluxes have large differences between the cases even of one city. This is because they refer to total fluxes, not anthropogenic emissions, and our rectangular city area definition encompasses significant biological activity that takes up significant amounts of CO2 in summer. In addition, anthropogenic emission estimates vary monthly.

Table 12: Berlin hotspot experiment: state vector and total city flux prior estimates and uncertainty.

State vector element	Prior (winter case)	Prior (summer case)	Prior uncertainty
Background flux	15.4 Mt CO2/yr	-31.4 Mt CO2/yr	100%
City flux	21.3 Mt CO2/yr	2.5 Mt CO2/yr	100%
Hotspot 1	4.2 Mt CO2/yr	2.9 Mt CO2/yr	100%
Hotspot 2	2.2 Mt CO2/yr	1.8 Mt CO2/yr	100%
Hotspot 3	1.8 Mt CO2/yr	1.2 Mt CO2/yr	100%
Mixing ratio offset	0 ppm	0 ppm	0.5 ppm
Combined city flux			
Total city flux	29.5 Mt CO2/yr	8.5 Mt CO2/yr	20.3 / 4.1 Mt CO2/yr

7.2.2 Covariance matrix of optimized fluxes and mixing ratio offset

One measure of the ability of the inversion to constrain emissions is the ratio of their prior and posterior uncertainties. These are shown in Figure 11 and Table 13. The uncertainty reduction is strong for total city fluxes (around 80% in both cases). But when disaggregating, the uncertainty reduction of the hotspots is moderate, around 20–60% (except hotspot 1 with 70% in the summer case). In other words, the information contained in the observations about the hotspots is significantly smaller than that about total city emissions.

Table 13: Berlin hotspot experiment: state vector and total city flux prior estimates and uncertainty.

State vector element	Posterior uncertainty (winter case)	Posterior uncertainty (summer case)
Background flux	8%	9%
City flux	17%	14%
Hotspot 1	55%	31%
Hotspot 2	50%	39%
Hotspot 3	78%	62%
Mixing ratio offset	0.11 ppm	0.08 ppm
Combined city flux		
Total city flux	12%	15%

Another measure for the ability of the inversion to constrain individual state vector elements is their correlation, which we show in Table 14. We consider correlations significant at a p-value below 0.05. The strongest correlations are between CO2 fluxes of background areas and the mixing ratio offset, which reveals that they can only be constrained together (the correlation is positive in summer where average background fluxes are negative). Significant correlations also exist between the city fluxes and the background fluxes and mixing ratio offset in winter, highlighting the need to estimate the background signal accurately. Lastly, the moderate information content in the observations on individual hotspots inferred above from uncertainty reductions is reflected here in negative correlations between hotspots and the remainder of the city, indicating that their sum is much better constrained than the individual components.

Table 14: Correlated optimized state vector elements in the Berlin hotspot experiment.

State vector elements	Correlation	p-value
Winter case		
Background – City	0.44	1.3e-08
City – Hotspot 1	-0.47	1.4e-09
City – Hotspot 3	-0.17	3.5e-02
Background – Offset	-0.96	9.0e-87
City – Offset	-0.49	1.2e-10
Summer case		
City – Hotspot 1	-0.40	4.3e-07
City – Hotspot 3	-0.33	3.0e-05
Background – Offset	0.93	1.8e-66

7.2.3 Deviations of optimized fluxes and mixing ratio offset from the truth

In addition to the random uncertainties of posterior fluxes, which depend solely on covariance ($\boldsymbol{R}$ and $\boldsymbol{B}$) and atmospheric transport ($\boldsymbol{H}$) matrices, we also analyse deviations of optimized fluxes from the truth. These are, in addition, influenced by measured values ($\boldsymbol{z}$), which represent systematic XCO2 error estimates in our setup. Here we report results for an additional experiment, which is a modification of the winter case where we did not optimize a mixing ratio offset. This is a common setup in larger scale CO2 flux estimation. The covariance matrix of the other optimized state vector elements was not strongly affected by this modification, which is why we do not show the results for this case in Sect. 7.2.2. Deviations from the truth are given in Figure 11 and Table 15.

Table 15: Deviations of posterior fluxes and mixing ratio offset from the prior.

State vector element	Deviation from prior		
	Winter case w/o offset	Winter case	Summer case
Background flux	+50%	-41%	-6%
City flux	+80%	-18%	-3%
Hotspot 1	+20%	+30%	-4%
Hotspot 2	-37%	-87%	+110%
Hotspot 3	+50%	+20%	+40%
Mixing ratio offset	-	+1.4 ppm	-0.1 ppm
Combined city flux			
Total city flux	+76% (+17 MtCO2/yr)	-14% (-4.2 MtCO2/yr)	+27% (+2.3 MtCO2/yr)

We find that optimizing the mixing ratio offset in the winter case is critical for retrieving the total flux of the city. The retrieved mixing ratio offset (winter 1.4 ppm, summer 0.1 ppm) is close to the assimilated average of systematic XCO2 error (winter 1.1 ppm, summer 0.0 ppm, Table

6). Without optimizing the offset in the winter case, the city total is strongly (+76%) overestimated, and with the optimization slightly (-14%) underestimated, a strong improvement and a difference between the estimates of a factor 2. We did not perform this experiment for the summer case, but since the average systematic XCO2 error and retrieved mixing ratio offset in this case are much smaller, the flux sensitivity to optimizing the offset is likely also much smaller than in the winter case.

Table 15 reveals that systematic XCO2 errors make it difficult to constrain fluxes of individual sources within Berlin. In the cases with offset optimization, the estimated fluxes of hotspot 2 deviate strongly from the truth (winter -87%, summer +110%), whereas the estimates of the others are -4% to +40% off, which indicates a slightly worse performance than that for total city fluxes (-14%, +27%). Hotspot 2 is similarly well constrained as the other hotspots (see posterior uncertainties), larger than hotspot 3, and unlike the other hotspots, it does not correlate with the city flux, as it is located outside of the city centre. Therefore, the deviations between retrieved and true fluxes for this hotspot are likely driven not by conflation with other signals or high sensitivity, but by the spatial distribution of systematic XCO2 errors. This is fairly obvious in the summer case, where the plume extends southward (Figure 2, top right) over an area with positive XCO2 errors (Figure 5, top right). We explore this case further in Sect. 7.2.4. In winter, the underestimation could be caused by a combination of positive XCO2 errors upwind of the hotspot (Figure 2 and Figure 5, top left) and the slight overestimation of the mixing ratio offset by 0.3 ppm. These results highlight that emission estimation of individual facilities can suffer from biases in XCO2 data that are coherent on the scale of CO2 plumes.

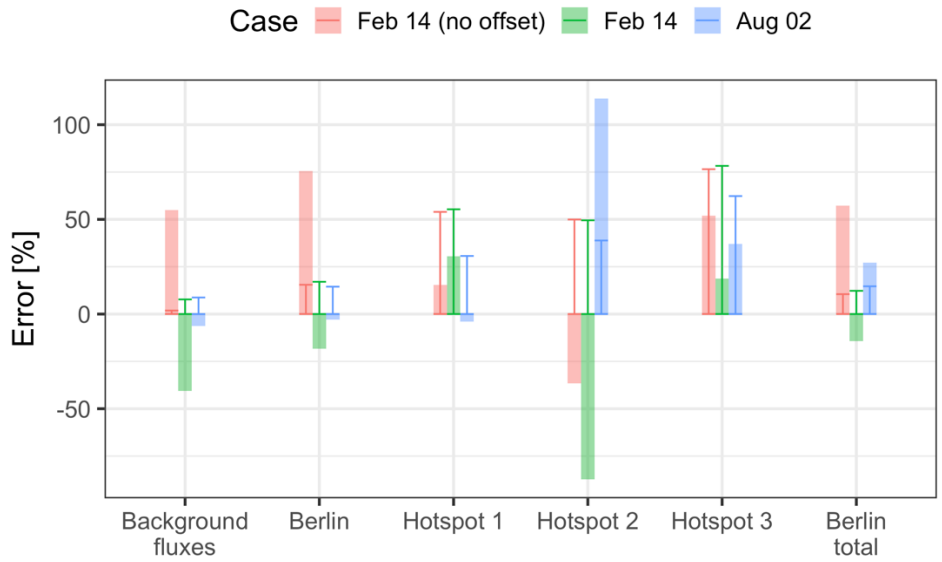

Figure 11: Uncertainties of retrieved emissions in the Berlin hotspot experiment. Bars are systematic uncertainties and error bars are random uncertainties (as in Figure 10).

7.2.4 Case study: Overestimation of Hotspot 2 in the August 02 case

Here we take a closer look at the reasons behind the largest single flux error in the Berlin hotspot experiment, i.e. the case of Hotspot 2 in summer, where the inversion yields twice the correct flux. In Sect. 7.2.3, we already mentioned that the systematic uncertainty has a pattern that correlates with the flux location. In Figure 12, we zoom in on the plume and plot the covariates that the XCO2 error model depends on. With a simple definition of the plume from Hotspot 2 (i.e. by defining a threshold for the XCO2 value and excluding a few pixels in the west that likely belong to other sources), we find that the plume signal is 0.2 ppm over background, and the systematic uncertainty in the plume area is 0.1 ppm higher than in the background. Keeping in mind that the plume definition is somewhat subjective, this pattern

likely explains the flux error of 110%. The spatial pattern of systematic uncertainties in this area and case correlates strongly with albedo, with largest errors in pixels with lowest albedo (Table 16). Note that we did not exclude observations over inland water bodies for this experiment, which may be responsible for a small number of pixels (three in our simple plume definition) with lowest albedos and thus largest errors. The error pattern correlates only weakly with aerosol optical depth. Nonetheless, it is the type of error that is expected to be mitigated somewhat by an additional aerosol sensor onboard, because it is caused by limitations of properly accounting for aerosols. Note that carefully filtering the data for water bodies, or scenes with low albedo in general, may partly mitigate this instance of measurement error, but this was not investigated further.

Table 16: Correlations of systematic XCO2 uncertainties and the covariates used to compute them in the area of Hotspot 2 in the Berlin August 02 simulation.

Covariate	Correlation with systematic XCO2 error (Hotspot 2, August 02)
albedo_nir	-0.80
albedo_swir1	-0.87
albedo_swir2	-0.80
aot_nir	0.26
aot_swir1	0.26
aot_swir2	0.26

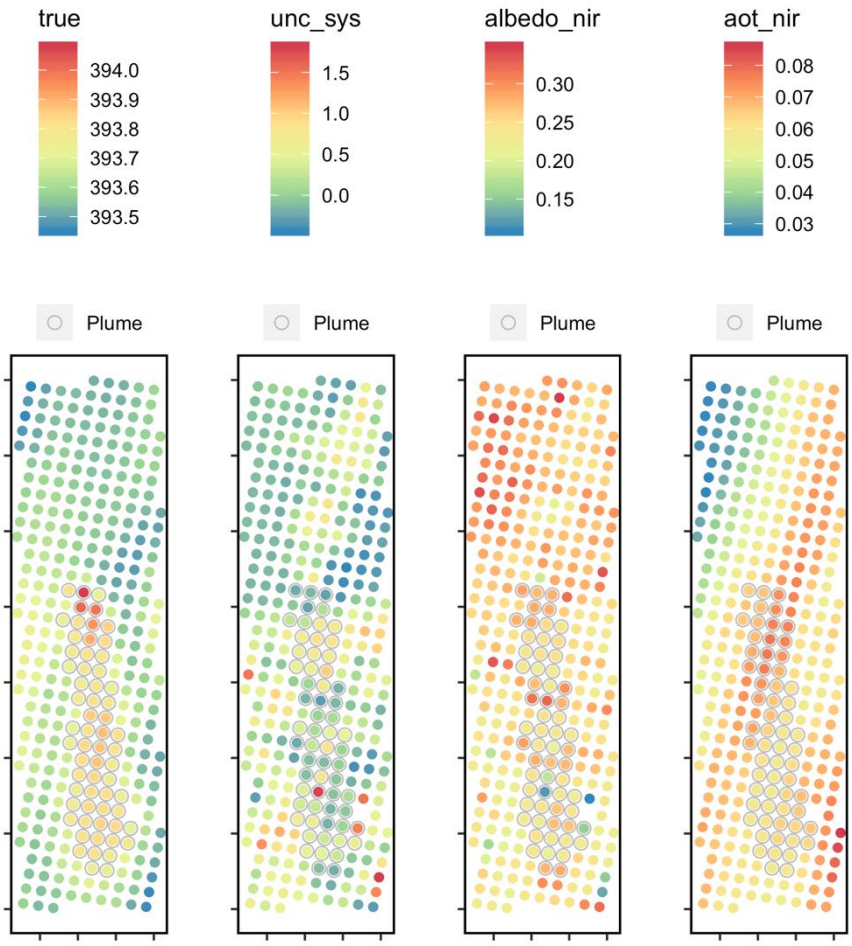

Figure 12: XCO2 data for Hotspot 2 in the August 02 case. The extent of each plot is 19 km by 70 km. From left to right: true XCO2, systematic XCO2 error, albedo (NIR) and AOT (NIR).

7.3 Updated AeroCarb flux results

We updated AeroCarb flux uncertainty results by repeating the flux optimization of that method with the updated synthetic XCO2 uncertainties without MAP presented in Sect. 5.2. City emission uncertainties are in the range of those estimated based on the neural network method for our 2015 cases and also dominated by systematic XCO2 uncertainties (Figure 13 and Table 17). The updated Berlin uncertainties are not very different, albeit slightly smaller than before the update and in our 2015 cases based on the neural network method for XCO2 uncertainties. In the Beijing cases, the smaller XCO2 uncertainties (Sect. 5.2) yield emission uncertainties that are much smaller than before the update (~5% vs ~200–800%).

Thus, updated AeroCarb emission uncertainties are similar to those we infer for the 2015 cases.

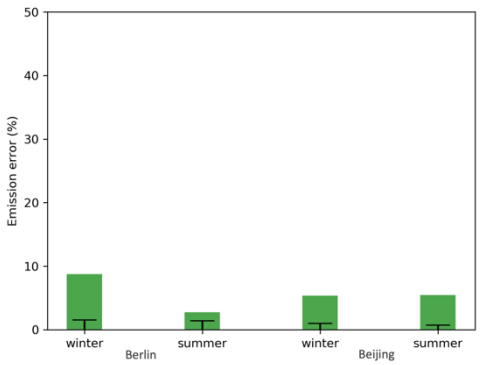

Figure 13: Emission uncertainties of the AeroCarb cases inferred from updated AeroCarb XCO2 uncertainties without MAP (presented in in Sect. 5.2). Solid bars show the deviation of retrieved to true flux (representing systematic XCO2 uncertainties), while error bars are posterior random flux uncertainties.

Table 17: Uncertainties of retrieved city emissions in AeroCarb cases without MAP. "RemoteC" denotes original values from the AeroCarb WP4 report. "RemoteC updated" was obtained with updated XCO2 uncertainties presented in in Sect. 5.2. All values are given in percent of the city emission.

AeroCarb case	RemoteC Random	RemoteC Systematic	RemoteC updated Random	RemoteC updated Systematic
Berlin summer without MAP	4.7	17.3	1.9	2.8
Berlin winter without MAP	8.9	5.2	2.4	8.8
Beijing summer without MAP	14.9	191	0.5	5.5
Beijing winter without MAP	30.2	837	1.0	5.4

8 Conclusions

In this study, we investigate how well CO2M is able to constrain anthropogenic CO2 emissions from cities given its observational coverage and anticipated scattering-related XCO2 uncertainties without a multi-angle polarimeter (MAP). Our analyses focus on Berlin, Beijing and Shanghai in summer and winter 2015. Using the WRF-Chem model, we model atmospheric transport of CO2 fluxes based on anthropogenic emission inventories and a biosphere CO2 flux model (D2.3 and others). The atmospheric CO2 fields are sampled to simulate CO2M observations (based on D2.5). We evaluate these synthetic observations with real measurements from OCO-2 and OCO-3, which shows that our simulations in Beijing in winter may slightly and in Shanghai with the EDGAR inventory are strongly overestimating the real city plume. On the other hand, Shanghai simulations with the MIX inventory yield realistic signal sizes on the high end of the few observations available from OCO-2 and OCO-3 to date, and we use these simulations to draw conclusions. We also evaluate modelled aerosol optical depths (part of T2.3), on which synthetic CO2M uncertainties from D2.5 are based (among other quantities), using observations from AeroNet and MAIAC. The simulations do not reflect conditions on a small percentage of heavily polluted days in China, but are nonetheless compatible with average conditions throughout the year in all cities. We then perturb the

simulated observations based on the synthetic uncertainties and use the perturbed observations to retrieve CO2 fluxes. First, with an analytical inversion method, we scale whole city emissions for all cases. In both Berlin and the Chinese cities, the impact of systematic XCO2 uncertainties is larger than that of random XCO2 uncertainties. They yield promising results with errors of 11–25% in Berlin and -1–9% in the China cases. The cloud-free Beijing case with ~8% systematic uncertainty may be optimistic due to a relatively large modelled XCO2 enhancement. Thus, results for the Chinese cities appear to be similar or slightly more accurate than for Berlin, which may be explained by smaller XCO2 enhancements in Berlin while XCO2 uncertainties are similar in most simulations. Partial cloud cover strongly increases systematic errors in retrieved fluxes in Beijing (36–57%), and overestimation of city XCO2 enhancements in Shanghai leads to underestimation of the systematic flux uncertainty (0–2%). Second, we try and constrain individual large emitters within the city of Berlin, where we make use of TNO's high-resolution flux dataset from D2.3. CO2M is less skilled at constraining individual emitters than scaling whole-city emissions, with both random and systematic errors on the order of several tens of percent. This is both because of overlap of multiple emitters (hotspots 1 and 3) and because the small and localized signals of individual emitters are sensitive to variations in CO2M biases on small scales of a few 10s of km (hotspot 2). We analyse the case with the largest deviation of emissions from a hotspot from the truth (110% offset), and it can be explained by correlations of the XCO2 error with albedo up- and downwind of the plume. This type of XCO2 error is expected to be reduced with a MAP.

Whole city emissions are obtained with the same flux inversion method as in the AeroCarb project, where it was employed for other periods and with synthetic uncertainties from another CO2M error model. The results for the whole city of Berlin are similar to those obtained in AeroCarb without a MAP. However, AeroCarb found very large flux errors in Beijing if no MAP is considered in the XCO2 retrieval, whereas our results are optimistic for both Beijing and Shanghai. The different results can be explained by differences in the magnitude of estimated XCO2 uncertainties, which were much larger in the AeroCarb cases without MAP. The estimates we received from D2.5 were closer to AeroCarb estimates with MAP, which is consistent with our flux uncertainty results. The XCO2 error model from AeroCarb has since been revised and now yields results close in magnitude to the uncertainties we received from D2.5 (Sect. 5.2). We repeated flux simulations of the AeroCarb study with the updated XCO2 uncertainties, and the flux results are similar to those of the neural network (Sect. 7.3).

While the flux uncertainties we find are promising, a number of restrictions apply to determining the ability of inverse modelling methods to constrain city emissions based on CO2M. We explain and discuss them in the following paragraphs.

While the magnitude of XCO2 errors is consistent among the updated AeroCarb and neural network methods, their distribution, e.g. with respect to AOT, is not (Sect. 5.2). While the flux uncertainties we obtain with either estimate are similar in our cases, this may not be the case in general, since the spatiotemporal distribution of XCO2 uncertainties has a direct impact on our flux results. Therefore, these disagreements should be reconciled. Recently, Rusli et al. (2020, in review) estimated that CO2M uncertainties without MAP have a spread of 2.07 ppm, based on a retrieval ensemble that spans a wide range of conditions. Applying the same metric to the neural network estimates that we assimilated yields a spread of only 0.58 ppm, much smaller than the estimate by Rusli et al. This number describes only the small regions of around 150km by 150km around Berlin, Beijing and Shanghai that we worked with. Additional XCO2 simulations were done for Europe in D2.5, and the spread of XCO2 errors of those is about 1–5 ppm, depending on the day. Therefore, the neural network does not seem to underestimate errors on continental scale. However, if this large XCO2 error spread occurs close to a city and is assimilated, the flux errors are likely larger than suggested by our results.

Next, like all observation system simulation experiments, we did not investigate all error sources known to affect CO2 emission estimation. In particular, we assumed perfect atmospheric transport, which can be a large error sources in flux estimation. The H2020

project CoCO2, the follow-on project of CHE that starts in 2021, dedicates resources to this challenge.

Our results are derived from single overpasses and thus only provide snapshots of temporally variable emissions. Constraining annual emissions will require sampling more often. For this project, we hand-picked large emitters and a few situations with favourable conditions, most notably cloud-free scenes and low wind speeds that allow build-up of the city plume. Those events were rare in the 4 months of meteorology simulations we performed to pick suitable scenes for each city. Thus, real XCO2 enhancements even from large emitters may often be smaller than those we investigated here, and large enhancements rare. The small number of scenes with large XCO2 enhancements over Beijing and Shanghai we found in the OCO-2 and OCO-3 records may support this hypothesis, but OCO-2 sampling is sparse and we expect the OCO-3 record to grow and improve, as it is still a fairly new instrument. Yet, most smaller cities will produce smaller XCO2 enhancements, likely resulting in larger flux errors. This may also apply to flux estimation on regional scale to infer national totals. For example, in the H2020 project SCARBO, an ongoing feasibility study about a fleet of CO2 detecting nanosatellites, we conclude that a MAP is indispensable for the SCARBO concept to infer CO2 fluxes on national scale.

Lastly, the achievable emission uncertainties have to be viewed in context to the emission reductions CO2M is built to detect. For example, the EU recently pledged to reduce emissions by 55% compared to 1990 levels by 2030, up from a 40% reduction goal submitted as Nationally Determined Contribution to the Paris agreement in 2015[6]. In 2018, emissions were already 23% below 1990 levels[7]. The remaining target of 32% over 12 years or about 13% over five years between planned CO2 stock takes is in the range of the uncertainties of fluxes we derived from single overpasses for Berlin. As described above, the uncertainty of annual emissions is different and needs to be investigated. Yet, these numbers suggest that likely all sources of error in flux estimation with CO2M need to be minimized to ensure its success. This includes the scattering errors investigated here which may be reduced with a MAP.

Overall, the ability of CO2M to retrieve CO2 emissions from cities in the presence of scattering errors seems promising in favourable conditions (large sources, clear sky, average AOD, large built-up city plumes). However, estimation methods for XCO2 uncertainties do not agree on their distribution, and updates to them may impact our flux results. Constraining smaller sources, individual ones within a city, or on regional scale likely require smaller XCO2 errors than those achieved without MAP investigated here, as well as minimizing other important sources of uncertainty in CO2 flux estimation.

[6] https://www4.unfccc.int/sites/submissions/INDC/Published%20Documents/Latvia/1/LV-03-06-EU%20INDC.pdf

[7] https://ec.europa.eu/clima/sites/clima/files/strategies/progress/docs/com_2019_559_en.pdf

9 References

Beck, V., T. Koch, R. Kretschmer, J. Marshall, R. Ahmadov, C. Gerbig, D. Pillai, and M. Heimann, (2011): The WRF Greenhouse Gas Model (WRF-GHG). Technical Report No. 25, Max Planck Institute for Biogeochemistry, Jena, Germany.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled “online” chemistry within the WRF model. *Atmospheric Environment*, *39*(37), 6957–6975. https://doi.org/10.1016/j.atmosenv.2005.04.027

International Energy Agency: World Energy Outlook 2008., 2008.

Li, M., Zhang, Q., Kurokawa, J. I., Woo, J. H., He, K., Lu, Z., et al. (2017). MIX: A mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. Atmospheric Chemistry and Physics, 17(2), 935–963. https://doi.org/10.5194/acp-17-935-2017

Peters, W., Miller, J. B., Whitaker, J. S., Denning, A. S., Hirsch, A., Krol, M. C., et al. (2005). An ensemble data assimilation system to estimate CO2 surface fluxes from atmospheric trace gas observations. *Journal of Geophysical Research Atmospheres*, *110*(24), 1–18. https://doi.org/10.1029/2005JD006157

Rusli, S., Hasekamp, O., aan de Brugh, J., Fu, G., Meijer, Y., & Landgraf, J. (2020). Anthropogenic CO2 monitoring satellite mission: the need for multi-angle polarimetric observations. Atmospheric Measurement Techniques Discussions, (June), 1–31. https://doi.org/10.5194/amt-2020-152

Tarantola, A. (2005). Inverse Problem Theory and Methods for Model Parameter Estimation. Methods for Data Fitting and Model Parameter Estimation. SIAM, Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898717921

van Der Laan-Luijkx, I. T., Van Der Velde, I. R., Van Der Veen, E., Tsuruta, A., Stanislawska, K., Babenhauserheide, A., Fang Zhang, H., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C. and Peters, W.: The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: Implementation and global carbon balance 2001-2015, Geosci. Model Dev., 10(7), 2785–2800, doi:10.5194/gmd-10-2785-2017, 2017.

Document History

Version	Author(s)	Date	Changes
1.0	Friedemann Reum (SRON) Sander Houweling (SRON, VU)	21/12/2020	Final submission

Internal Review History

Internal Reviewers	Date	Comments
Daniel Thiemert	21/12/0220	approved with minor edits

Estimated Effort Contribution per Partner

Partner	Effort
SRON	8
TNO	0.5
Total	**8.5**