# Remote sensing $CO_2$, $CH_4$ and CO emissions in a polluted urban environment

Denis M. O'Brien[1], Igor N. Polonsky[2], Steven R. Utembe[3], and Peter J. Rayner[3]

[1]Greenhouse Gas Monitor Australia, Melbourne, Australia
[2]RT & RS Solutions LLC, Tucson, USA
[3]School of Earth Sciences, University of Melbourne, Melbourne, Australia

*Correspondence to:* Denis O'Brien (Denis.OBrien@ggma.com.au)

**Abstract.**

This paper describes a numerical experiment to test the ability of the proposed geoCARB satellite to estimate emissions of trace gases ($CO_2$, $CH_4$ and CO) in the polluted urban environment of Shanghai. The meteorology over Shanghai is simulated with the Weather Research and Forecasting (WRF) model for a nine-day period in August 2010. The meteorology includes

5 water and ice clouds. The chemistry version of WRF (WRF-Chem V3.6.1) is used to predict the chemical composition, mass density and number density of aerosol species. Spectra in the bands measured by geoCARB are calculated, including the effects of polarisation and multiple scattering of radiation by clouds, aerosols and molecules. Instrument noise is added, and column-averaged trace gas mole fractions are estimated from the noisy spectra using an algorithm based on that for the Greenhouse Gases Observing Satellite (GOSAT) and the Orbiting Carbon Observatory (OCO2), but adapted to geoCARB. As expected,

10 the high aerosol loadings are challenging. However, when the retrieval algorithm is provided with regionally adjusted aerosol optical properties, as might be determined from observations of dark targets within the field of regard, the accuracies of retrieved concentrations are comparable to those reported earlier for geoCARB. Statistics of the errors in the retrieved column-averaged concentrations are used to predict the reduction in uncertainty of surface emissions possible with remotely sensed data.

## 1 Introduction

15 Emissions of $CO_2$, $CH_4$ and CO from cities and power plants potentially may be monitored from space, using column-averaged concentrations of the gases inferred from high resolution spectra of reflected sunlight in absorption bands of $CO_2$, $CH_4$, CO and $O_2$. Bovensmann et al. (2010b) examined CarbonSat (Bovensmann et al., 2010a), an instrument proposed to the European Space Agency, that would make spatially dense observations over power plants and cities from sun-synchronous orbit, and found that the observations could provide valuable emission estimates, especially under conditions of low wind speed.

20 Polonsky et al. (2014) examined geoCARB (Sawyer et al., 2013; Mobilia et al., 2013; Kumer et al., 2013b; Rayner et al., 2014), which would provide wall-to-wall coverage of continents with high spatial and temporal resolution, again with encouraging results. Gerilowski et al. (2011) flew an airborne $CH_4$ and $CO_2$ mapper over a German power plant, and successfully estimated emissions from the plant. Kort et al. (2012) estimated emissions from mega-cities using data from Japan's Greenhouse Gases Observing Satellite (GOSAT), even though GOSAT only provides sparse spatial and temporal coverage.

The ability to measure emissions relies on many factors:

1. the concentration measurements must be precise, because local sources generally will make only small perturbations to the column-averaged concentrations;

2. the observations must be spatially dense and highly resolved to capture plumes from local sources;

3. the measurements must be frequent in order to track varying winds;

4. the wind fields must be known well if plumes are to be traced reliably back to spatially diverse sources, although this requirement might be weakened if only the integrated emissions from cities are required;

5. the fluxes of trace gases across the boundary of the city must be known, because the boundary fluxes control the background level upon which the contributions from local sources are imposed.

The principal impediment to accurate concentration measurements using reflected sunlight comes from scattering of photons by aerosols and clouds, because scattered photons travel different paths through the atmosphere, and therefore suffer varying degrees of absorption. Thus, urban complexes, where emissions are large and observations would be most valuable, are locations where measurements are especially difficult because city atmospheres are polluted. This point was recognised clearly by Bovensmann et al. (2010b) and Polonsky et al. (2014), and they attempted to incorporate particulate emissions in their studies.

To some extent, obscuration by clouds and degraded accuracy caused by aerosols can be overcome if observations are spatially dense and frequent, as is possible with an instrument such as geoCARB observing from geostationary orbit. The spatial density allows the instrument to peek through gaps in clouds, while periods with cleaner air are captured by the frequent measurements. However, quantifying the extent is a complex task that must address all the issues listed above.

This paper is a first step, using the city of Shanghai as an example (Fig. 1). First the urban environment and the concentration measurements possible with geoCARB are simulated. Next the impact of those measurements is investigated in the context of perfect meteorology using the methods developed by Rayner et al. (2014). As a measure of impact, we use the reduction of the prior uncertainty of the surface emission fluxes gained from the (pseudo) satellite observations. In a later paper the errors caused by imperfect knowledge of the meteorology will be examined.

The structure of the paper is as follows. Section 2 describes the simulation of the atmosphere over Shanghai, including carbon fluxes, clouds and aerosols, for a nine-day period in August 2010. Section 3 outlines the optical properties of clouds, aerosols and the surface assumed in generating geoCARB spectra, while Sect. 4 discusses the noise model for geoCARB. Section 5 describes the inversion procedure by which column-averaged concentrations are derived from the noisy spectra. Section 6 estimates the reduction of uncertainty in surface fluxes gained from the relatively sparse set of observations that survive after rigorous screening for cloud and aerosol, and discusses the implications. Finally, Sect. 7 presents the results.

## 2 Simulation of the atmosphere over Shanghai

### 2.1 Time window

Observations of cloud and aerosol optical depth from MODIS were used to select a suitable period for the study in the northern hemisphere summer of 2010, that year being chosen because greenhouse gas emissions at high spatial resolution were readily

5 available. This preliminary step simply ensured that Shanghai was not obscured by cloud for the selected period. The period selected was from 2010-08-01 through 2010-08-09.

### 2.2 Meteorology

The chemistry version of the Weather Research and Forecasting model (WRF-Chem) was run for a total of ten days, allowing the extra day for model spin-up, to simulate the winds and trace gas concentrations over the target. A three-nested domain

10 (18 km, 6 km and 2 km resolution) was used with the innermost domain covering an area $200 \times 200$ km$^2$ centred close to Changzhou on the west of Shanghai. The innermosts domain is entirely over land, apart from the Yangtze River, Lake Taihu in the south and numerous small lakes. Output files were written at 30 minute intervals. The number of vertical layers was limited to 29 to reduce the numerical cost of the simulations.

The radiative transfer requires temperature and gas profiles to the top of the atmosphere, whereas the top of the grid used by

15 WRF-Chem was located at $p_t = 50$ hPa. Therefore, one more layer, extending from 50 hPa to the top of the atmosphere, was added to the output generated by WRF-Chem. The temperature and trace gas concentrations for this extra layer were set to those for the top WRF-Chem layer. The extra layer was assumed to be free of cloud and aerosol.

### 2.3 Surface and boundary carbon fluxes

Surface fluxes of $CO_2$ and CO were derived from the Fossil Fuel Data Assimilation System (FFDAS) (Rayner et al., 2010);

20 surface fluxes for $CH_4$ came from the Emission Database for Global Atmospheric Research (EDGAR, Olivier et al., 2005). Boundary fluxes of $CO_2$ were derived from the ECMWF Monitoring Atmospheric Composition and Climate (MACC) program (Panareda et al., 2014), while the global Model for OZone And Related chemical Tracers (MOZART-4, Emmons et al., 2010) provided fluxes for the other trace gases.

Because FFDAS uses night-lights to help distribute regional $CO_2$ fluxes at finer spatial resolution, it can shift emissions

25 from point sources like power stations to users who consume the electrical power. For example, the large power station (red dot) on the Yangtze River near the centre of Fig. 2 does not coincide with the peaks in the surface fluxes from FFDAS (green dots). For the purpose of the numerical experiment described in this paper, the mislocation of the emissions is not important, but for an analysis of real geoCARB data the locations of power plants would have to be specified correctly.

## 2.4 Aerosol

Primary aerosol emissions over Shanghai were derived from the Emission Database for Global Atmospheric Research (Olivier et al., 2005), which provides fluxes of organic carbon (OC), black carbon (BC), particulates less than $2.5\,\mu$m in diameter ($PM_{2.5}$) and particulates less than $10\,\mu$m in diameter ($PM_{10}$). No other primary emissions of particulates were included. In particular, sulphate and biomass-burning aerosols were advected from the boundary, and were not generated locally. This restriction will be removed in later work.

Because the spatial resolution of EDGAR ($0.1° \times 0.1°$) is roughly five times coarser than the spatial resolution of the inner domain of the WRF-Chem simulations, a method to distribute the EDGAR particulate fluxes over the finer WRF-Chem grid was required. Following Asefi-Najafabady et al. (2014), the down-scaling of particulate fluxes was based on the Defense Meteorology Satellite Program (DMSP) night-lights product, whose native resolution is 30 arc seconds. For each emitted species the high-resolution (30 arc second) flux $F_{\text{high}}$ was assumed to be

$$F_{\text{high}} = F_{\text{low}} L/\bar{L}, \tag{1}$$

where $F_{\text{low}}$ is the low resolution ($0.1°$) flux, $L$ the 30 arc second DMSP night-light field and $\bar{L}$ the same field averaged over the $0.1°$ grid cell of the low-resolution flux. This approach was applied to all fluxes except $CO_2$ emissions themselves. The low resolution flux $F_{\text{low}}$ was taken from the EDGAR V4.2 emissions (Olivier et al., 2005). For $CO_2$ emissions, the high-resolution version of the fossil-fuel data assimilation (FFDAS) dataset was used (Asefi-Najafabady et al., 2014). The 'other' sector of Asefi-Najafabady et al. (2014) was downscaled according to Eq. (1), while the power-plants were located in the appropriate 30 arc second cell. Power-plant locations were taken from the CARMA dataset, as updated by Asefi-Najafabady et al. (2014, Supporting information), who noted that many locations of power-plants in the underlying dataset are approximate.

EDGAR does not provide size-resolved particulate fluxes (apart from the broad $PM_{2.5}$ and $PM_{10}$ categories), whereas the MOdel for Simulating Aerosol Interactions and Chemistry (MOSAIC) embedded in WRF-Chem uses eight size bins

$$r_0 < r_1 < \ldots < r_i < \ldots r_s \text{ with } r_i = 2^{i-s} r_s \text{ and } s = 8.$$

In practice the maximum bin limit is $r_s = 5\,\mu$m, leading to $r_0 \approx 19$ nm.[1] The $i$-th bin represents the combined mass of particulates with radii in the range $[r_{i-1}, r_i]$. Because $r_0$, the lower limit of the smallest bin, exceeds the upper limit for Aitken particles (5 nm). MOSAIC accounts only for accumulation and coarse mode particulates.

Particulate fluxes from EDGAR were partitioned between Aitken, accumulation and coarse modes using the *ad-hoc* factors listed in Table 1; real data would assist if it were available. The fraction of the EDGAR flux assigned to the Aitken mode is lost from the subsequent calculations, because Aitken particles are smaller than the smallest MOSAIC bin. Next the accumulation and coarse modes were partitioned between the MOSAIC size bins with a fixed assignment for each mode, as specified in Table 2.

---

[1]Whereas this paper consistently uses radius, MOSAIC mostly uses diameter to characterise particles.

Lateral boundary conditions for aerosol species were obtained from MOZART-4 (Emmons et al., 2010). Boundary values and fluxes for sea-salt and dust were carried in sectional size bins, while those for organic carbon, black carbon, sulphate and ammonium nitrate were not size-resolved.

### 2.5 Cloud

WRF-Chem generated three-dimensional cloud fields. In each cell of the model domain, the number density and mass density were predicted for each cloud type (water or ice). The effective radius was derived using conservative models for the size distributions for water and ice clouds, as described in Appendix C.

## 3 Simulation of geoCARB spectra

### 3.1 Scan box and sample times

GeoCARB is assumed to be in geostationary orbit at longitude 110° E. When viewing to nadir, the east-west scan step is 3.0 km, and the north-south spacing of pixels is 2.7 km. These spacings increase when viewing away from nadir. For the Shanghai simulations, the optic axis, after reflection in the east-west and north-south scan mirrors, hits the geoid at 31.8° N latitude and 120.0° E longitude, just to the west of Shanghai and close to the city of Changzhou. The simulated data consists of 51 frames, each with 51 pixels, covering the scan box shown in Fig. 1. The time between frames was taken to be 4.625 s.[2]

GeoCARB observations of Shanghai were simulated three times per day at 00:15, 04:15 and 08:15 UTC, corresponding to 08:15, 12:15 and 16:15 local time in Shanghai.

### 3.2 Optical properties of aerosol and cloud

Size-resolved mass mixing ratios of nine aerosol species were predicted by MOSAIC during the simulations. The species and the procedures by which the optical properties of the aerosol mixtures were computed in the geoCARB spectral bands are detailed in Appendix A.

Figure 3 illustrates the aerosol optical depth (AOD) computed for the target region in the geoCARB spectral bands. The simulation is for 2010-08-09 at 08:15 UTC, corresponding to 16:15 local time in Shanghai. All four panels use the same dynamic range, which makes the differences between bands easier to see. There is a steady decrease in optical depth with wavelength, indicating small aerosol particles.

Histograms of AOD in the $O_2$ A-band predicted by WRF-Chem and observed by MODIS are compared in Fig. 4. The MODIS data were adjusted from the measured optical thickness at 550 nm to 770 nm using MODIS Ångström coefficients. The histograms present the probabilty of each optical thickness bin without regard for the spatial distributions. Only the near-noon WRF-Chem simulations are included, because these provide an approximation to the MODIS Terra observations at 10:30 and Aqua at 13:30 local time. The peaks of the two histograms coincide at approximately 0.3, but the MODIS observations

---

[2]The latest instrument model for geoCARB has 4.4625 s between frames.

contain a longer tail extending to 1.0, whereas the WRF-Chem tail extends only to 0.8. Generally the correspondence is good, so the scenarios encountered in the simulated geoCARB are not unrealistic.

Three-dimensional distributions of the mass and number of cloud liquid water and ice particles, all per unit mass of moist air, were predicted by WRF-Chem. In order to estimate the corresponding effective radii, needed to compute the optical properties

5 of the clouds, the particle sizes were assumed to follow modified gamma distributions (Deirmendjian, 1969). Optical properties in the geoCARB bands were calculated using Mie theory for spherical water drops, while for ice cloud the extinction coefficient and phase matrix were interpolated from data published by Baum (2007) and Baum et al. (2005a, b).

### 3.3 Surface properties

Surface properties, such as polarized bidirectional reflectance distribution function (from MODIS and POLDER), were as-

10 signed and interpolated to the frequencies of the geoCARB spectral bands. The method was similar to that described by O'Brien et al. (2009) and Polonsky et al. (2014). Surface heights from the WRF-Chem topography were interpolated to the centres of the geoCARB pixels.

### 3.4 Radiative transfer calculations

Using the three-dimensional fields of $CO_2$, $CH_4$, CO, clouds, aerosols and meteorology, spectra in the geoCARB bands were

15 simulated for all geoCARB pixels within the target area for early morning, noon and late afternoon observations on each of the non-spin-up days selected for the study. These simulations used the correct observation geometry. However, neither variations of trace gas concentrations nor variations of surface properties were modelled within the geoCARB pixels. Instead, the atmosphere above each pixel was assumed to consist of plane parallel layers, each uniform horizontally. Thus, the concentrations and optical properties of the atmosphere at any point along the line of sight from the centre of a pixel to the satellite were

20 assumed to be the same as those at the same height vertically above the centre of the pixel. Similarly, the surface properties were assumed to be horizontally uniform with values set at the centre of the pixel. These approximations are expedient, because they permit the use of plane-parallel radiation codes, but their justification is questionable for sensors (like geoCARB) with moderate spatial resolution, comparable to that of WRF-Chem. Fortunately, the approximation is immaterial for the numerical experiments over Shanghai, because the same approximation is employed in the forward simulation and the inverse retrieval,

25 but for geoCARB in flight this issue should be addressed.

An outline of the radiative transfer calculations is provided by O'Brien et al. (2015). Polarisation by the surface and by scattering within the atmosphere is included, but the polarisation effects of geoCARB's gratings are bypassed, because they may be corrected during analysis of the spectra using radiometric and polarimetric calibration data acquired before launch (O'Brien et al., 2015). Thus, the Stokes vector $(I, Q, U, V)$ was computed at the top of the atmosphere, and the intensity $I$ was

30 assumed to be the measured signal.

## 4  Simulation of instrument noise

The noise model for geoCARB is based on both laboratory characterisation and airborne trials with the Tropospheric Mapping Imaging Spectrometer (TIMS) developed by Lockheed Martin (Kumer et al., 2009, 2011). TIMS was used to demonstrate the feasibility of measuring $CH_4$ and CO in the 2.323 $\mu$m band. Details of the noise model and further information on the detector performance are provided by Kumer et al. (2013a).

The noise equivalent spectral radiance $N$ (in units nW/(cm$^2$ sr cm$^{-1}$)) for a representative spectral sample consists of shot noise due to the observed signal radiance $I$ and a signal independent noise floor radiance $N_0$. It is given by

$$N = \sqrt{N_0^2 + N_1 I},\tag{2}$$

where $N_0$ and $N_1$ are constants for each band.[3] Their numeric values, listed in Table 3, were derived from geoCARB instrumental and operational parameters by Kumer et al. (2013a).[4]

For each spectrum, the root-mean-square noise was computed using Eq. (2), and was provided to the inversion algorithm. In addition, for each spectrum a random sample of noise was drawn from a gaussian distribution with zero mean and the appropriate standard deviation at each frequency within each band. The noise sample was added to the spectrum prior to inversion.

## 5  Inversion of noisy spectra

The inversion algorithm developed by the Jet Propulsion Laboratory and Colorado State University for OCO, GOSAT and OCO2 was adapted to the spectral bands of geoCARB and applied to the simulated spectra. The results reported in Sect. 7 were obtained by fitting the intensity component of the Stokes vector, rather than the signal that includes the polarising effects of the gratings.

### 5.1  Instrument

The instrument line shape function (ILS) in each band was assumed to be gaussian in wavelength (and therefore slightly non-gaussian in frequency). The band limits ($\lambda_-$ and $\lambda_+$), dispersion per pixel ($\delta$) and resolution are listed in Table 4.

The number $n$ of pixels in each band was defined to be

$$n = [(\lambda_+ - \lambda_-)/\delta] + 1,$$

where $[x]$ denotes the integer part of $x$ The spectral sample points were

$$\lambda_p = \lambda_- + (p-1)\delta, \qquad p = 1, 2, \ldots, n,$$

---

[3]In practice $N_0$ and $N_1$ will vary from pixel to pixel, and their values will be determined by pre-flight calibration. However, in this paper average values were used for each band.

[4]The parameters in Table 3 differ slightly from those given by Kumer et al. (2013a). However, the differences are small, and they have negligible impact on the results.

with corresponding frequencies

$$\nu_p = 10^7/\lambda_p,$$

where $\lambda_p$ is expressed in nanometres and $\nu_p$ in $\text{cm}^{-1}$. Lastly, $\nu_p$ was represented as a cubic dispersion polynomial,

$$\nu_p = \sum_{k=0}^{3} a_k(p-1)^k.$$

The dispersion coefficients are listed in Table 5. With these definitions, the inversion code works entirely in frequency, rather than wavelength.

'Observed' and calculated spectra were compared in the reference frame of geoCARB, after allowing for the Doppler shifts arising from the motion of the target relative to the sun.

## 5.2 Priors

**Prior gas profiles**

The prior gas vertical profiles were based on the simulated profiles for the morning of the first day. Once computed, the same prior profiles were used for all subsequent times and days. The prior profiles are shown in Fig. 5.

The prior for the $CO_2$ field was taken to be the average of the background $CO_2$ field across the target region. In calculating the average, the background values from WRF-Chem on 29 layers firstly were interpolated to the 20 levels of the retrieval algorithm. The background represents the $CO_2$ advected into the domain, but does not include the contributions from local sources and sinks. To the extent that the broad scale features of the global $CO_2$ field can be predicted reliably by transport models, this is a reasonable prior for $CO_2$.

The prior for CO was averaged similarly over the Shanghai domain. However, the logic was convoluted, because the background field for CO initially was not available in the WRF-Chem output. Instead the 'true' CO field was taken to be the field generated with the FFDAS emission sources but evolved with chemistry disabled, while the prior was generated with FFDAS emissions and active chemistry. Thus, the prior and the truth had the same emission sources, and differed only in the chemistry. However, it emerged that the prior profile of CO determined at the first observation time on the first day was on average 40% higher than the profiles on other days.

For $CH_4$ the background field was not available initially, so the average of the true $CH_4$ field over the domain was used for the prior.

For $H_2O$, $CH_4$ and CO, the retrieval algorithm optimised scalar multiples of the prior profiles, while for $CO_2$ a full vertical profile was retrieved. The prior uncertainties for the scaling factors are shown in Table 6. The uncertainties for the scaling factors for $H_2O$ and $CH_4$ are very large, imposing almost no restraint. The uncertainty for CO is much tighter, probably unrealistically so, but it was necessary because with larger uncertainties the retrieval algorithm frequently failed to converge. This issue is discussed later.

**Prior meteorology**

Previous studies (O'Dell et al., 2012; Polonsky et al., 2014) suggest that the accuracy of the meteorology prior does not have a large impact upon the retrieved trace gas concentrations. Therefore, in this preliminary study, the prior for the meteorology was taken to be the truth, apart from the minor errors introduced by interpolating from the 29 layers of WRF-Chem to the 20 levels
5 of the retrieval algorithm. The prior uncertainty assumed for surface pressure was 4 hPa, a relatively large value, so the prior has only a secondary effect upon the retrieval. Similarly, the uncertainty assumed for the temperature offset was large (5 K).

**Prior aerosol**

The algorithm to retrieve concentration profiles is based on that used for GOSAT and OCO2. It assumes four particulate types, water drops with effective radius 8 $\mu$m, ice crystals with effective radius 70 $\mu$m and two aerosol types, Kahn 2b and Kahn 3b, as
10 described by Kahn et al. (2001). For each particulate, the extinction coefficient, single scattering albedo and phase matrix were precalculated and provided to the retrieval algorithm. The underlying assumption is that a mixture of the two aerosol types can represent the optical properties of real aerosol with sufficient accuracy. However, in practice this assumption frequently fails; the mixture cannot reproduce the optical properties, the retrieval algorithm fails to converge or produces estimates seriously in error, and the fits to the spectra (measured by the reduced $\chi^2$) are poor. The obvious remedy is to use more appropriate aerosol
15 models.

After inversion of the simulated spectra with the standard algorithm, the inversions were repeated with the wavelength dependence of the aerosol extinction coefficients adjusted to represent the regional aerosols more accurately. The adjustment was performed for each scene using the minimum and maximum of the simulated aerosol optical depth over the whole scene.

1. The minimum and maximum AOD over the scene, denoted $\tau_{\min}(\lambda)$ and $\tau_{\max}(\lambda)$, were computed for the geoCARB
20 bands.

2. The extinction coefficient, single scattering albedo and phase matrix of Kahn type 3b aerosol were replaced by those for type 2b.

3. For Kahn aerosol type 2b, the extinction coefficient $\gamma(\lambda)$ was replaced by

$$\gamma'(\lambda) = \gamma(\lambda_0)\tau_{\max}(\lambda)/\tau_{\max}(\lambda_0), \tag{3}$$

25 where $\lambda_0$ denotes the wavelength at the blue end of the O$_2$ A-band.

4. For Kahn aerosol type 3b (initially equal to type 2b), $\gamma(\lambda)$ was replaced by

$$\gamma'(\lambda) = \gamma(\lambda_0)\tau_{\min}(\lambda)/\tau_{\min}(\lambda_0). \tag{4}$$

Neither the single scattering albedo nor the phase matrix was changed.

As justification for this replacement, other sensors (or perhaps geoCARB itself) are likely to provide independent information
30 on aerosols. Furthermore, when there is a dark target in the region, such as Lake Taihu near Shanghai, robust measurements of

aerosol properties are possible, because photons scattered by aerosols over the dark target are not masked by photons reflected from the surface.

Thus, two options were considered, the first with standard aerosol and the second with regionally adjusted aerosol. In addition, the threshold for retrieved aerosol optical thickness in the post-processing filter (PPF) was increased from 0.1 to 0.15

5 for the experiment with regionally adjusted aerosol.

## 6  Estimation of surface fluxes

To test the capacity of the retrieved concentrations to constrain surface fluxes we extended the work of Rayner et al. (2014). That paper calculated the reduction in uncertainty for fluxes in an idealised urban box with a single isolated power-plant. Here we replaced the idealised environment with a realistic representation of the Shanghai-Nanjing region. The emissions are the

10 same as that used to generate the atmospheric state used in the trace-gas retrieval calculations. Six days of observations were considered. The diurnal cycle of emissions was assumed to be the same for each day, and was represented in four six-hour blocks. Fluxes were retrieved at $5 \times 5\mathrm{km}$ resolution. Every pixel within the $200 \times 200\mathrm{km}$ target area was considered as a source, so there were $4 \times 40 \times 40 = 6400$ unknown flux elements for $CO_2$. Following Rayner et al. (2014) we also solved for a constant emission factor for CO for each pixel. For methane we assumed no diurnal cycle of emissions, yielding 1600

15 unknown flux components. For each trace gas we also solved for a constant bias across the region.

If concentrations are linearly related to source fluxes and gaussian statistics are assumed, then the posterior flux error covariance $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^{\mathrm{t}}\mathbf{R}^{-1}\mathbf{H}, \tag{5}$$

where $\mathbf{B}$ is the prior flux error covariance, $\mathbf{H}$ is the jacobian matrix mapping fluxes to observed concentrations, and $\mathbf{R}$ is the

20 error covariance of the observations. This model predicts the reduction in flux errors brought about by the observations. In this form it does not address the impact of transport errors. The percentage uncertainty reduction for component $j$ of the flux is taken to be

$$E_j = 100(1 - \sqrt{A_{jj}/B_{jj}}), \tag{6}$$

where $A_{jj}$ and $B_{jj}$ are the corresponding diagonal elements of the posterior and prior flux error covariances respectively.

25 Following Rayner et al. (2014) we use the simplified transport model SatPlume to calculate $\mathbf{H}$. Justification of the SatPlume jacobian depends on whether or not it overestimates the structure of the retrieved field; the more structured the plumes, the stronger the error reduction because sources are less smeared together. Rayner et al. (2014) argued that the statistics of column dilution from SatPlume are plausible.

Apart from $\mathbf{H}$, the other two ingredients for the error reduction calculation are $\mathbf{B}$ and $\mathbf{R}$. Prior flux uncertainties for $CO_2$ are

30 set at 25% of the fluxes themselves. For the calculation using CO measurements we also require an emission factor for CO. Its uncertainty is set at 30% following Rayner et al. (2014). The posterior errors in $X_{CO_2}$ and $X_{CH_4}$ measurements are taken

directly from the L2 retrieval algorithm. For $X_{CO}$ the L2 retrieval algorithm underestimates the errors, as will be discussed later in Fig. 11, so experiments were conducted with values 15, 10 and 2.5 ppb.

## 7 Results

### 7.1 Trace gas concentrations

5 The optical depths of cloud water, cloud ice and aerosol in the $O_2$ A-band are plotted in Fig. 6. These histograms include all samples over Shanghai in the nine-day period of the simulation. Although the period selected generally had skies that a casual observer would call clear, relatively few pixels had optical depths sufficiently low to permit reliable retrieval of trace gases. Thus, even an instrument like geoCARB with its high spatial density of samples cannot provide wall-to-wall coverage of a polluted urban environment, and retrieval of accurate surface fluxes is challenging.

10 Table 7 lists the success rate of the retrieval algorithm, presenting the fraction of soundings for which the retrieval algorithm converged and the fraction that passed the PPF. The total number of soundings is the same for all the experiments (70,227), being the product of the number of days (9), the number of times per day (3), the number of geoCARB frames (51) and the number of pixels per frame (51). The two experiments with cloud enabled had approximately the same numbers of converged pixels, but experiment 3 with adjusted aerosol properties had many more pass the PPF. There are two reasons. First, the threshold for aerosol optical depth was increased in experiment 3 from the default of 0.10 to 0.15. Probably more important were the superior aerosol optical properties, which permitted better fits to the spectra.

Histograms of errors in the retrieved surface pressure $p_s$, $X_{CO_2}$, $X_{CH_4}$ and $X_{CO}$ are presented in Fig. 7 for experiment 1 with standard aerosol but without cloud. In each panel, the grey histogram (partly obscured) describes the errors with the PPF disabled; the foreground coloured histogram has the PPF active. Generally the results are poor.

1. The histograms without the PPF are very broad, far exceeding the range thought necessary to retrieve surface fluxes.
2. In the case of $X_{CO}$, the histogram has multiple peaks, the origin of which is unknown. They indicate that the retrieval algorithm is failing in several distinct modes.
3. Even after the PPF, the surface pressure is biased low by 1.7 hPa, while $X_{CO_2}$ is high by 2.2 ppm. This effect of anti-correlated errors is often observed in retrievals with simulated spectra, though the underlying cause is poorly understood.
25 It seems to occur in the presence of high aerosol loadings, particularly when the optical properties (single scattering albedo and phase matrix) assumed for the aerosol in the retrieval algorithm do not match those in the simulated data.

When cloud is enabled in experiment 2 with standard aerosol (Fig. 8), the histograms broaden, even those with the PPF. The low bias in surface pressure and high bias in $X_{CO_2}$ persist, and the distributions are not bell-shaped, suggesting that competing factors contribute to the error.

30 Figure 9 shows the corresponding results for experiment 3 in which the retrieval algorithm was provided with regionally adjusted optical properties for aerosol (but not cloud). Although the histograms after the PPF are slightly skewed, they are

more bell-shaped, and the mean values are close to zero. These results are comparable to those found in earlier studies for less polluted environments (Polonsky et al., 2014).

CO is particularly difficult to retrieve, because the absorption lines of CO are masked by lines of $H_2O$ and $CH_4$. Only with spectral resolution much finer than that provided by geoCARB is it possible to probe CO lines without interference from $H_2O$ and $CH_4$. Compounded with this spectrographic difficulty is an algorithmic limitation. In order to estimate the atmospheric state, the current version of the L2 retrieval algorithm performs unconstrained optimisation of a cost function that measures the difference between the observed and computed spectra, consistent with the instrument noise. However, in the course of the optimisation, the algorithm can produce unphysical states, a common example being negative values for optical depth. In an attempt to control this behaviour for aerosol, the optimised variable is the logarithm of the optical depth, which may take any value, but which compresses the dynamic range. Consequently the retrieval algorithm frequently aborts because the state vector has strayed into an unphysical region of state space. Further soundings are lost due to lack of convergence, a problem exacerbated by logarithms of optical depth. For example, several iterations may be wasted with tiny steps in aerosol optical depth, even though the steps in the logarithm are significant. In order to secure convergence of the algorithm, a tight prior error for the CO scaling factor was imposed. While that probably is acceptable if the prior CO profile is close to the truth, it clearly is not if the initial guess is poor. In that case, the profile is clamped to the prior profile by the tight prior error, and the final answer for $X_{CO}$ will be biased towards that of the prior. This might explain the positive bias in $X_{CO}$ in experiment 3, because the prior CO profile was higher than the truth for most pixels.

These problems might be alleviated by constraining state variables to ensure that all internally generated variables have physically reasonable values. For example, optical depth should be positive. When the state vector encounters a constraint boundary, rather than crossing the boundary and aborting, the optimisation algorithm should require the state to move along the constraint boundary as it continues its search. At each iteration, the optimisation algorithm should check whether a better reduction in the cost function could be achieved by releasing the state from the boundary or by retaining it there. Robust code to implement constrained optimisation is available.

Of particular interest are the histograms of errors for retrieved aerosol optical depth (AOD), shown in Fig. 10 for experiments 1, 2 and 3. In each case the mean error is negative, so the retrieval algorithm underestimates the true AOD. Furthermore, the standard deviation is close to the magnitude of the mean, suggesting that the prior value for the AOD is too low and that the prior uncertainty is too tight.

## 7.2 Errors versus SNR

Statistics of the errors as functions of the signal-to-noise ratio (SNR) are needed to conduct observational system simulation experiments (OSSEs) for geoCARB. They allow the measurement error to be predicted from the surface albedo, which is known by the OSSE, because to a good approximation the albedo and the solar zenith angle determine the SNR. The small dots in Fig. 11 are the posterior errors estimated by the L2 retrieval algorithm for the ensemble of soundings over Shanghai that passed the PPF in experiment 3. The range of SNRs is wide (although not as wide as that expected by geoCARB), because the ensemble includes morning and afternoon observations when the solar zenith angle is large and the incident solar flux density

is low. The solid curves have the form

$$\sigma = a/(1 + bx^c), \tag{7}$$

where $\sigma$ denotes the uncertainty (in ppm for $CO_2$ and ppb for $CH_4$ and CO) and $x$ denotes the SNR. Over the range of SNR for Shanghai, many functional forms could have been chosen to fit the data. However, it was decided to choose a functional

5 form that would be conservative should it be abused by applying it to SNR outside the range over Shanghai. The form in Eq. (7) decreases montonically from $a$, its value when $x$ is zero, so parameter $a$ should represent the prior uncertainty. For large $x$ the function decays to zero at a rate determined by the exponent $c$. Conservative values of 10 ppm, 100 ppb and 100 ppb for $X_{CO_2}$, $X_{CH_4}$ and $X_{CO}$ were assigned to $a$, and parameters $b$ and $c$ were determined by fitting the model to the posterior errors from the L2 algorithm. Table 8 lists the parameters so determined.

10 The points in Fig. 11 with error bars represent the actual error, defined to be the retrieved value minus the truth, with the latter weighted by the averaging kernel. A bootstrap procedure was used to estimate the error bars. These errors are larger than the estimates of posterior error from the retrieval algorithm, because they include systematic errors and not just the random errors from instrument noise. For $X_{CO_2}$ and $X_{CH_4}$ the actual errors are about 1.5 times larger than the estimated errors, but for $X_{CO}$ the factor is closer to 8. Prime examples of systematic errors are those arising from imperfect spectroscopy and from incorrect

15 optical properties of aerosol. When analysing spectra from GOSAT and OCO2, the ACOS team introduced the concept of 'empirical noise', which is treated like instrument noise but is intended to account for all sources of error in the retrievals. A similar approach could be adopted for geoCARB.

### 7.3 Reduction of uncertainty in surface fluxes

Of the 3,107 soundings that converged and passed the PPF in the full nine-day simulation for experiment 3, only 2,751 fell

20 within the selected six-day window and also had rays from the target to the sun and satellite passing through the top rather than the sides of the model domain. With only 2,751 observations for 6,401 flux components, any attempt to determine the posterior fluxes is under-constrained. However the observations will not be uniformly distributed through the day, so fluxes at some times may be well-determined.

Figure 12 shows the reduction of error in the 00:00 UTC time period for the case using both $X_{CO_2}$ and $X_{CO}$ measurements,

25 the latter with uncertainty 10 ppb. The spatial structure is quite similar to that of the prior fluxes. This is a common result following from Eq. (5) in which larger prior uncertainties are preferentially reduced. This relationship weakens towards the east of the domain. Here the Sun-Earth ray is likely to pass through the eastern wall of the domain, and hence be excluded (observation time is 08:15 local time). The maximum error reduction is 30% corresponding to one of the strongest prior sources.

Two other periods of the day are quite well constrained. This is not surprising for the local afternoon (14:00–20:00 local time)

30 since there is an observation made during that period. The block from 02:00–08:00 local time is also reasonably constrained by the observation made at 08:15. As in Rayner et al. (2014), the block corresponding to the unobserved night is completely unconstrained by observations.

We conducted four experiments, the first without $X_{CO}$ observations and the remainder including $X_{CO}$ with errors of 15, 10 and 2.5 ppb. The largest value (15 ppb) is approximately the mean of the true error, defined to be the difference between $X_{CO}$ retrieved and $X_{CO}$ simulated. The middle value (10 ppb) is a slightly optimistic estimate, while the lowest (2.5 ppb) is included to explore benefits that might accrue from improved spectral resolution in the CO band. For each experiment we performed four variations:

1. $\sigma(X_{CO_2})$ reduced by a factor of five;
2. $\sigma(X_{CO_2})$ inflated by a factor of five;
3. observations thinned by a factor of two;
4. observations thinned by a factor of four.

Spatial patterns of error reduction are similar, so we summarize the statistics of the results in Table 9. The row labelled 'Base' describes the four experiments; the following four rows cover the variations on the experiments.

The base case demonstrates the importance of the $X_{CO}$ observations. For the $X_{CO_2}$-only experiment, the maximum reduction of uncertainty is 17%, and occurs at 00:00 UTC (locally 08:00). The maximum rises with the addition of $X_{CO}$ observations, reaching 24%, 30% and 58% for the three levels assumed for $\sigma(X_{CO})$.

The variations explore trade-offs among precision and data density. For example, adding $X_{CO}$ observations with accuracy 2.5 ppb outweighs improving the $X_{CO_2}$ accuracy by a factor of five in the $X_{CO_2}$-only case. Similarly, $X_{CO}$ observations with 15 ppb error more than compensate for loss of half the observations in the $X_{CO_2}$-only case. Lastly, $X_{CO}$ observations with 15 ppb error roughly compensate for a five-fold loss of accuracy in the $X_{CO_2}$-only case.

## 8 Discussion

This work is intended as part of an evolution from the notional study of Rayner et al. (2014) towards a mission definition. It parallels the development from Rayner and O'Brien (2001) to the definitions of global missions. The main step we have made here is a more serious treatment of the concentration retrieval errors that can be expected from a real instrument, informed by previous experience with low-earth orbiting missions OCO2 and GOSAT.

The results are less dramatic than those of Rayner et al. (2014). This is particularly clear if one notes that the flux retrieval uncertainties here are at a spatial resolution of $5 \times 5$ km while those of Rayner et al. (2014) for $CO_2$ sources were at $3 \times 3$ km. The main reason for the weakened constraint of fluxes is the reduction in the number of available soundings. Detailed simulations including aerosol and cloud give a more pesimistic picture of sampling density than the simple statistical approach of Rayner et al. (2014).

Given the sensitivity of the results to contamination by clouds and aerosols we need to improve our confidence in these fields. Our simulations seem to perform quite well against available data, but there are many aspects we cannot test. Data on the diurnal cycle of cloudiness at this resolution is surprisingly lacking for example.

As with the study of Rayner et al. (2014), $X_{CO}$ measurements provide much information on fluxes. It is important, therefore, to understand their error characteristics. We noted already that the uncertainties generated by the L2 algorithm were a

poor guide to the errors (retrieval minus truth). The differences between experiments with 15 ppb and 2.5 ppb for $\sigma(X_{CO})$ in Table 9 underline the importance of the question; if we could generate measurements with the characteristics implied by the L2 algorithm, the flux constraint would be far stronger. There is interference between the CO lines used in this retrieval and some $H_2O$ lines. These interferences increase the posterior error from the L2 algorithm (since they introduce water as a nuisance

5   variable), and also lead to broader distributions of posterior $X_{CO}$ around the true value. These should be consistent. If they are not it suggests that the prior uncertainties in water vapour may not describe the statistics of the difference between the prior value and the truth (defined here as the forward WRF-Chem run).

Finally we must ask what the more physically-based calculations of this paper tell us about the value of a future mission with the characteristics described here. The weaker constraint on fluxes means that longer averaging periods are required before a

10   given signal (such as an emission reduction) can be detected. However, these times remain short compared to the required time-scale for policy. More important may be the reduction in spatial resolution from three to five kilometres. This takes the results further from the scale of most businesses.

## 9   Conclusions

WRF-Chem was used to model the dynamics and composition of the atmosphere over Shanghai for a nine-day period. Models

15   were developed for the optical properties of aerosols and clouds in the spectral bands of geoCARB, and the predicted distributions of aerosol optical depth were broadly consistent with those from MODIS. The simulated urban environment should provide a rigorous case on which to test retrievals of surface fluxes.

GeoCARB observations were simulated at three times per day for each of the nine days. The simulated spectra were degraded with noise expected for geoCARB, and column-averaged concentrations of $CO_2$, $CH_4$ and CO were retrieved from the spectra.

20   Experiments were conducted to assess the accuracy of the retrieved concentrations.

1. Even though a comparatively cloud-free period was selected, many soundings were too contaminated by aerosol or cloud to permit accurate determination of the column-averaged concentrations. The experiments suggest that the number of successful retrievals might be as low as 5%. This places a strong constraint on the complexity of the surface flux model.

2. Retrievals using preassigned aerosol optical properties, similar to those used for GOSAT and OCO2, gave poor results.

25   Surface pressure was biased high, while $X_{CO_2}$ was low. The distributions were non-normal. Retrieval of $X_{CO}$ had several unidentified modes of failure.

3. When the aerosol optical properties provided to the retrieval algorithm were regionally adjusted, the results improved dramatically. Surface pressure and $X_{CO_2}$ no longer had large, anti-correlated biases; distributions were closer to normal; the throughput of the retrieval algorithm improved. As justification for the regional adjustment, dark targets in the

30   vicinity of the urban complex (such as Lake Taihu near Shanghai) could be used by other sensors (or even geoCARB) to characterise the wavelength dependence of aerosol optical thickness. This complex experiment suggests the advantages would be significant. Further examination of this idea is a priority.

4. Despite the polluted environment over Shanghai, and the consequent difficulty in retrieving column-averaged concentrations accurately, the reduction in uncertainty from six days of observations is comparable to that found by Rayner et al. (2014) in an idealised scenario. Assuming $X_{CO}$ measurements with accuracy 10 ppb, they found a reduction of uncertainty for $CO_2$ fluxes of approximately 35%; in the this work the corresponding reduction was 30%. This is an encouraging result.

## Appendix A: Optical properties of aerosols

Within the WRF-Chem model domain, MOSAIC tracked the aerosol chemistry and predicted the spatial distributions of size-resolved mass concentrations of the aerosol species listed in Table 10. Because refractive indices are not known (or are unreliable) for some of the species, proxies were nominated, as indicated in the third column of Table 10. For example, the refractive index for sulphate ions was assumed to be the same as that for ammonium sulphate ($NH_4SO_4$), for which laboratory data is available. The real and imaginary parts of the refractive indices of the proxies are listed in Table 11 and Table 12.

Given an ensemble of aerosols, including any water coating the aerosols, the refractive index $n$ of the composite was assumed to be the volume-weighted average of the constituents,

$$n = \sum_a v_a n_a / \sum_a v_a,$$

where $n_a$ and $v_a$ denote the refractive index and volume occupied by aerosols of the species indexed by $a$. The volumes $v_a = c_a/\rho_a$ were obtained from MOSAIC's mass mixing ratios $c_a$ ($\mu$g of aerosol per kg of dry air) using densities $\rho_a$ listed in Table 13. The density of black carbon is the mid-point of the values quoted by Bond and Bergstrom (2006).

Because the advection scheme in WRF-Chem may cause the mass and number prognostic equations to give inconsistent values, usually at sharp gradients of these quantities, MOSAIC checks internally that the predicted number density of aerosol lies between bounds determined by the limits of the size bin. For particles in the $i$-th size bin, whose lower and upper radii limits are

$$r_- = r_{i-1} \qquad \text{and} \qquad r_+ = r_i,$$

let $N_-$ denote the number of particles per unit mass of dry air consistent with the computed volume $v_d$ if all the particles have the maximum size $r_+$,

$$v_d = (4/3)N_-\pi r_+^3 = (4/3)N_+\pi r_-^3$$

Similarly, let $N_+$ denote the number if the particles have the smallest size $r_-$. In this calculation, $v_d$ is the total volume occupied by all aerosol species with the exception of condensed water; thus, $v_d$ is the total dry volume. If $N$, the number of particles per

unit mass of dry air, falls outside the range $[N_-, N_+]$, then it is clamped;

$$N' = \begin{cases} N_- & \text{if } N \leq N_-, \\ N & \text{if } N_- < N < N_+, \\ N_+ & \text{if } N_+ \leq N. \end{cases}$$

In subsequent calculations of the optical properties, $N$ is replaced by $N'$.

Lastly, the wet radius $r_w$ is estimated from

5    $v_w = (4/3)N'\pi r_w^3,$

where $v_w$ is the wet volume, given by

$$v_w = \sum_a v_a.$$

The calculations above were performed for each size bin, providing a volume-averaged refractive index $n$ and wet radius $r_w$, henceforth denoted simply by $r$. The refractive index could vary from band to band, but within each band it was assumed

10    to be constant. Thus, the aerosol in size bin $j$ was characterised by one real and four complex numbers, $(r, n_1, n_2, n_3, n_4)$.

**Appendix B: Quantisation of aerosols**

In order to avoid the numerical cost of Mie calculations for every combination of $(r, n_1, n_2, n_3, n_4)$, the radius and refractive index were quantised as follows.

Firstly, $r$ was assigned to one of the MOSAIC bins by finding the bin index $j$ (zero-based) such that

15    $j = \begin{cases} 0 & \text{if } r \leq r_0, \\ k & \text{if } r_k \leq r < r_{k+1}, \\ s-1 & \text{if } r_s \leq r. \end{cases}$

Secondly, for the purpose of the Mie calculations, $r$ was replaced by $r_j'$, the effective radius of a uniform distribution of radii between the bin limits $r_j$ and $r_{j+1}$,

$$r_j' = \frac{3}{4}\left(\frac{r_{j+1}^4 - r_j^4}{r_{j+1}^3 - r_j^3}\right).$$

Quantisation of the real part $\mu$ of the refractive index $n$ used a grid

20    $1 = \mu_0 < \mu_1 < \ldots < \mu_P.$

In each band, from $\mu$ the corresponding integer $p$ was found such that

$\mu_p \leq \mu < \mu_{p+1},$

and $\mu$ then was replaced for the purpose of Mie calculations by the average of the bin limits,

$$\mu' = (\mu_p + \mu_{p+1})/2.$$

In practice the grid spanned the range $[1,2]$ in 20 equally spaced steps.

The imaginary part $\nu$ of the refractive index $n$ was quantised similarly; a grid was defined,

$$0 = \nu_1 < \nu_2 < \ldots < \nu_Q,$$

$\nu$ was located in the grid,

$$\nu_q \leq \nu < \nu_{q+1},$$

and $\nu$ was replaced by

$$\nu' = (\nu_q + \nu_{q+1})/2.$$

The grid for the imaginary part of the refractive index was quasi-logarithmic, as specified in Table 14.

Thus, aerosols were characterised by a list of quantum numbers,

$$(j, p_1, q_1, p_2, q_2, p_3, q_3, p_4, q_4),$$

and Mie calculations were performed for radii and refractive indices

$$(r'_j, \mu'_1, \nu'_1, \mu'_2, \nu'_2, \mu'_3, \nu'_3, \mu'_4, \nu'_4).$$

Fortunately the number of combinations of quantum numbers was manageable, because certain combinations kept recurring. Whenever a new combination arose, Mie calculations were performed and the extinction efficiency, scattering efficiency and extinction coefficient for all bands were written to a 'mie-file'. Similarly, a 'mom-file' was written that contained the corresponding phase matrices. These files were used subsequently in the radiative transfer calculations.

The Mie calculations used the SPHER code from Mishchenko and Travis. For each size bin, rather than a single Mie calculation with the nominal radius for the bin, a power law distribution of sizes was assumed with effective variance fixed at 0.1, a value broadly representative of many aerosol types. The reason for assuming a distribution was to reduce the risk that spikes in the optical properties, associated with resonances, might distort the results.

Each quantised aerosol was considered a separate species. If $\alpha_k$, $\beta_k$ and $\gamma_k$ denote the volume coefficients for absorption, scattering and extinction of species $k$, and $\Phi_k$ denotes the phase matrix, then the corresponding quantities for the composite aerosol were computed in the usual way:

$$\alpha = \sum_k \alpha_k, \qquad \beta = \sum_k \beta_k, \qquad \gamma = \sum_k \gamma_k$$

and

$$\beta \Phi = \sum_k \beta_k \Phi_k.$$

## Appendix C: Optical properties of clouds

WRF-Chem predicts three-dimensional distributions of the mass $m$ of cloud liquid water, cloud ice and the number $N$ of cloud particles, all per unit mass of moist air. In order to compute the optical properties of clouds, the effective radii of the cloud particles are needed. If all the particles share the same radius $r$ (a delta distribution), then the volume $v$ occupied by the

5    particles and the corresponding mass $m$ will be

$$v = 4N\pi r^3/3, \qquad \text{and} \qquad m = \rho v,$$

where $\rho$ denotes the density of water or ice. Because WRF-Chem provides both $m$ and $N$, $r$ can be calculated.

However, because clouds contain a wide distribution of particle sizes, the delta distribution was replaced by a modified gamma distribution, with

10    $$N(r) = a r^\alpha e^{-br^\gamma},$$

following the notation of Deirmendjian (1969). For this distribution, the total number of particles (per unit mass of moist air) is

$$N = \frac{a}{\gamma} b^{-(1+\alpha)/\gamma} F_1,$$

the effective radius is

15    $$r_e = b^{-1/\gamma} F_4/F_3$$

and

$$v/N = (4/3)\pi b^{-3/\gamma} F_4/F_1, \tag{C1}$$

where

$$F_n = \Gamma[(n+\alpha)/\gamma].$$

20    If parameters $\alpha$ and $\gamma$ are specified, then $b$ (and hence the effective radius) may be computed from Eq. (C1) using $v = m/\rho$ and $N$. The constant $a$ serves only to normalise the distribution, and its value is not needed.

Lastly, the effective radii derived from the mass and the modified gamma distributions were quantised and truncated, with $r_e'$ below replacing $r_e$:

$$r_e' = \begin{cases} r_{e-} & \text{if } r_e \leq r_{e-} \\ a\{r_e/a\} & \text{if } r_{e-} \leq r_e < r_{e+} \\ r_{e+} & \text{if } r_{e+} \leq r_e \end{cases} \tag{C2}$$

where $r_{e-}$ and $r_{e+}$ are the lower and upper limits set for cloud particles, and $\{x\}$ denotes the nearest integer to $x$. Values of the parameters are specified in Table 15.

Optical properties in the geoCARB bands were calculated using Mie theory for spherical water drops, while for ice cloud the extinction coefficient and phase matrix were interpolated from data published by Baum (2007) and Baum et al. (2005a, b).

5   *Acknowledgements.* The authors express their gratitude to Jack Kumer of the Lockheed-Martin Advanced Technology Center (ATC) for many helpful discussions of the geoCARB instrument and the interpretation of its spectra. Utembe acknowledges support from the GORGE project, jointly funded by the Australian Research Council (ARC) and Lockheed Martin under ARC Linkage Grant number LP130100404. O'Brien and Polonsky also acknowledge direct support from Lockheed-Martin ATC. The retrieval algorithm was run on the OCO cluster at Colorado State University, thanks to support from Chris O'Dell. The code that generated geoCARB spectra from WRF fields was run in

10   Melbourne on the 16-node cluster at Greenhouse Gas Monitor Australia Pty. Ltd.

# References

Asefi-Najafabady, S., Rayner, P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., Huang, J., Elvidge, C., and Baugh, K.: A multiyear, global gridded fossil fuel CO2 emission data product: Evaluation and analysis of results, J. Geophys. Res.: Atmos., doi:10.1002/2013JD021296, http://dx.doi.org/10.1002/2013JD021296, 2014.

Baum, B. A.: http://www.ssec.wisc.edu/~baum/Cirrus/IceCloudModels.html, 2007.

Baum, B. A., Heymsfield, A. J., Yang, P., and Bedka, S. T.: Bulk scattering properties for the remote sensing of ice clouds. Part I: microphysical data and models, J. Appl. Meteorol., 44, 1885–1895, doi:10.1175/JAM2308.1, 2005a.

Baum, B. A., Yang, P., Heymsfield, A. J., Platnick, S., King, M. D., Hu, Y.-X., and Bedka, S. T.: Bulk scattering properties for the remote sensing of ice clouds. Part II: narrowband models, J. Appl. Meteorol., 44, 1896–1911, doi:10.1175/JAM2309.1, 2005b.

Bond, T. C. and Bergstrom, R. W.: Light absorption by carboneceous particles: an investigative review, Aerosol Sci. and Tech., 40, 27–67, doi:10.1080/02786820500421521, http://dx.doi.org/10.1080/02786820500421521, 2006.

Bovensmann, H., Buchwitz, M., and Burrows, J. P.: Carbon monitoring satellite (CarbonSat) as an Earth Explorer Opportunity Mission: mission overview, http://www.iup.uni-bremen.de/carbonsat, 2010a.

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, www.atmos-meas-tech.net/3/781/2010/, 2010b.

CARMA: Carbon Monitoring for Action, http://carma.org, 2012.

Deirmendjian, D.: Electromagnetic scattering on spherical polydispersions, American Elsevier, New York, 1969.

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J. F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of the Model for OZone And Related chemical Tracers, version 4 (MOZART-4), Geosci. Model Dev., 3, 43–67, doi:10.5194/gmd-3-43-2010, 2010.

Gerilowski, K., Tretner, A., Krings, T., Buchwitz, M., Bertagnolio, P. P., Belemezov, F., Erzinger, J., Burrows, J. P., and Bovensmann, H.: MAMAP — a new spectrometer system for column-averaged methane and carbon dioxide observations from aircraft: instrument description and performance analysis, Atmos. Meas. Tech., 4, 215–243, doi:10.5194/amtd-4-215-2011, www.atmos-meas-tech.net/4/215/2011/, 2011.

Kahn, R., Banerjee, P., and McDonald, D.: Sensitivity of multiangle imaging to natural mixtures of aerosols over ocean, J. Geophys. Res., 106, 18 219–18 238, 2001.

Kort, E. A., Frankenberg, C., Miller, C. E., and Oda, T.: Space-based observations of megacity carbon dioxide, Geophys. Res. Lett., 39, doi:10.1029/2012GL052738, http://dx.doi.org/10.1029/2012GL052738, l17806, 2012.

Kumer, J., Rairden, R., and Roche, A.: Modeling the 2.33 $\mu$m TIMS spectrometer radiometric noise: Implications for space applications, in: Aerospace Conference, 2013 IEEE, pp. 1–20, doi:10.1109/AERO.2013.6497369, 2013a.

Kumer, J. B., Roche, A. E., Rairden, R. L., Desouza-Machado, S., Blatherwick, R., Hawat, T., and Chatfield, R.: NASA ESTO Instrument Incubator Program (IIP) Tropospheric Infrared Mapping Spectrometers (TIMS): demonstration of multi-layer CO retrieval from atmospheric data acquired simultaneously in the solar reflective region near 2330 nm and the thermal emissive region near 4680 nm, in: Hyperspectral Imaging and Sensing of the Environment, OSA Technical Digest, p. HMC5, Optical Society of America, http://www.opticsinfobase.org/abstract.cfm?URI=HISensE-2009-HMC5, 2009.

Kumer, J. B., Rairden, R. L., Roche, A. E., and Chatfield, R.: NASA ESTO Instrument Incubator Program (IIP) Tropospheric Infrared Mapping Spectrometers (TIMS) demonstration first deployment on an airship: preliminary results, in: Hyperspectral Imaging and Sensing of the Environment, OSA Technical Digest, p. HTuD2, Optical Society of America, http://www.opticsinfobase.org/abstract.cfm?URI=HISE-2011-HTuD2, 2011.

Kumer, J. B., Rairden, R. L., Roche, A. E., Chevallier, F., Rayner, P. J., and Moore, B.: Progress in development of Tropospheric Infrared Mapping Spectrometers (TIMS): geoCARB green house gas (GHG) application, Proc. SPIE, 8867, 88 670K, doi:10.1117/12.2022668, 2013b.

Mobilia, J., Kumer, J. B., Palmer, A., Sawyer, K., Mao, Y., Katz, N., Mix, J., Nast, T., Clark, C. S., Vanbezooijen, R., Magoncelli, A., Baraze, R. A., and Chenette, D. L.: Determination of technical readiness for an atmospheric carbon imaging spectrometer, Proc. SPIE, 8867, 88 670L, doi:10.1117/12.2029634, 2013.

O'Brien, D. M., Polonsky, I., O'Dell, C., and Carheden, A.: Orbiting Carbon Observatory (OCO), algorithm theoretical basis document: The OCO simulator, Technical report ISSN 0737-5352-85, Cooperative Institute for Research in the Atmosphere, Colorado State University, 2009.

O'Brien, D. M., Polonsky, I. N., and Kumer, J. B.: Sensitivity of remotely-sensed trace gas concentrations to polarisation, Atmos. Meas. Tech., 7, accepted, 2015.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS $CO_2$ retrieval algorithm, Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99–121, doi:10.5194/amt-5-99-2012, www.atmos-meas-tech.net/5/99/2012/, 2012.

Olivier, J. G. J., Aardenne, J. A. V., Dentener, F. J., Pagliari, V., Ganzeveld, L. N., and Peters, J. A. H. W.: Recent trends in global greenhouse gas emissions:regional trends 1970-2000 and spatial distributionof key sources in 2000, Environmental Sciences, 2, 81–99, doi:10.1080/15693430500400345, http://dx.doi.org/10.1080/15693430500400345, 2005.

Panareda, A., Massart, S., Chevallier, F., Boussetta, S., Balsamo, G., Beljaars, A., Ciais, P., Deutscher, N. M., Engelen, R., Jones, L., Kivi, R., Paris, J.-D., Peuch, V.-H., Sherlock, V., Vermeulen, A. T., Wennberg, P. O., and Wunch, D.: Forecasting global atmospheric $CO_2$, Atmos. Chem. Phys., 14, 11 959–11 983, doi:10.5194/acp-14-11959-2014, 2014.

Polonsky, I. N., O'Brien, D. M., Kumer, J. B., O'Dell, C. W., and the geoCARB Team: Performance of a geostationary mission, geoCARB, to measure $CO_2$, $CH_4$ and CO column-averaged concentrations, Atmos. Meas. Tech., 7, 959–981, doi:10.5194/amt-7-959-2014, http://www.atmos-meas-tech.net/7/959/2014/, 2014.

Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed $CO_2$ concentration data in surface source inversions, Geophys. Res. Lett., 28, 175–178, 2001.

Rayner, P. J., Raupach, M. R., Paget, M., Peylin, P., and Koffi, E.: A new global gridded data set of $CO_2$ emissions from fossil fuel combustion: Methodology and evaluation, J. Geophys. Res.: Atmos., 115, doi:10.1029/2009JD013439, http://dx.doi.org/10.1029/2009JD013439, d19306, 2010.

Rayner, P. J., Utembe, S. R., and Crowell, S.: Constraining regional greenhouse gas emissions using geostationary concentration measurements: a theoretical study, Atmos. Meas. Tech., 7, 3285 3293, doi:10.5194/amt-7-3285-2014, www.atmos-meas-tech.net/7/3285/2014/, 2014.

Sawyer, K., Clark, C., Katz, N., Kumer, J., Nast, T., and Palmer, A.: GeoCARB design maturity and geostationary heritage, Proc. SPIE, 8867, 88 670M, doi:10.1117/12.2024457, 2013.
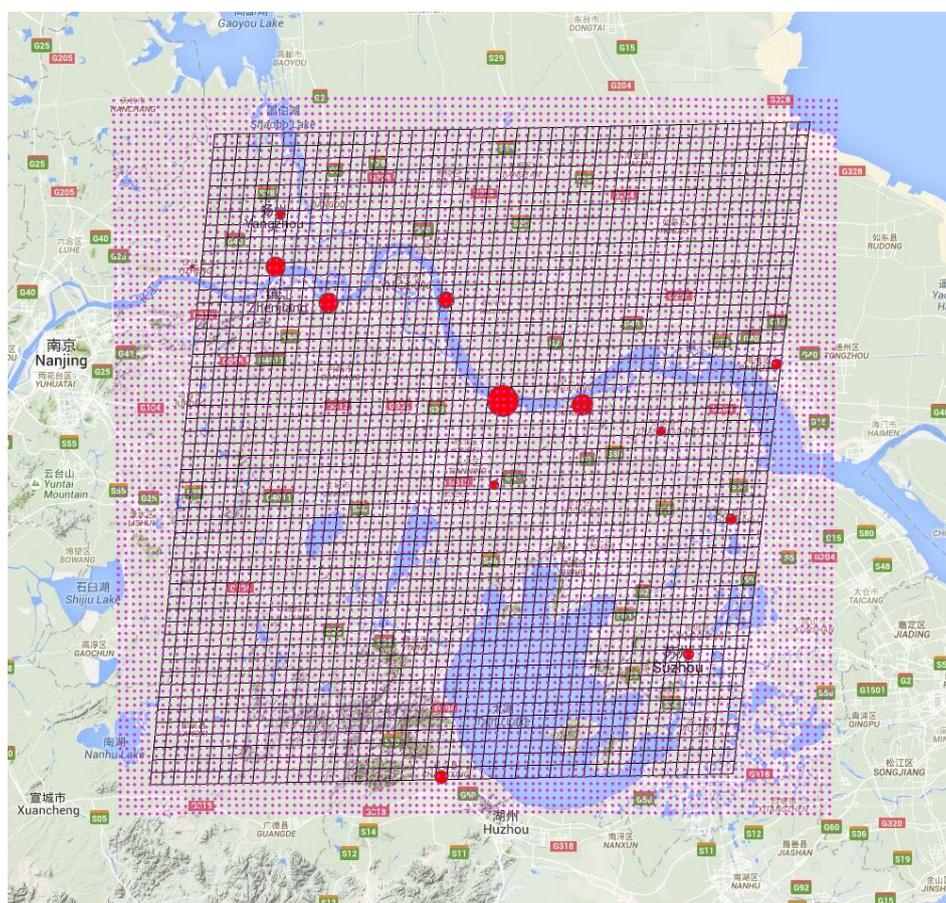
**Figure 1.** Target area showing the geoCARB footprints (skewed grid) and the centres of the Weather Research and Forecasting model cells (small red dots). The large red dots indicate power stations as listed by CARMA (2012), with the sizes of the dots proportional to the annual $CO_2$ emissions from the plants. The distance from the south-west to the south-east corner of the geoCARB grid is approximately 158 km, while from the south-east to the north-east corner is approximately 176 km.
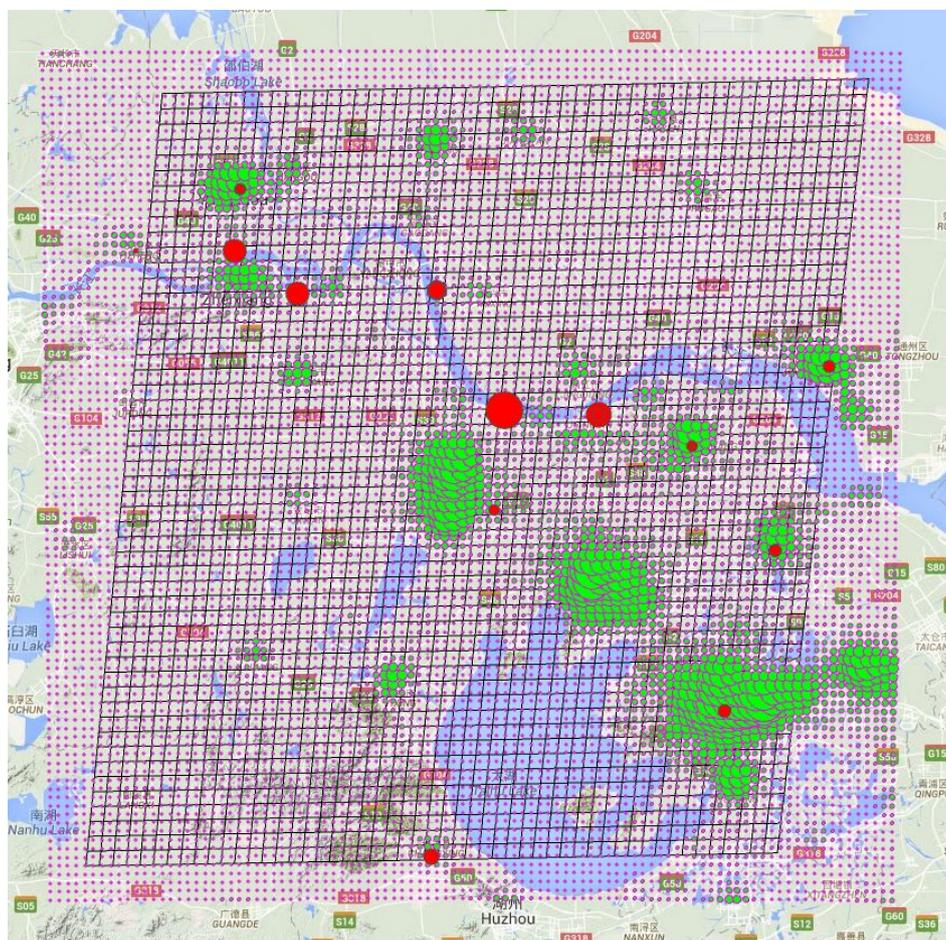
**Figure 2.** The red dots show the locations of power plants as listed by CARMA (2012). The sizes of the dots are proportional to the annual $CO_2$ emissions of the plants. The green dots are the emissions used in the WRF simulations, redistributed using FFDAS.
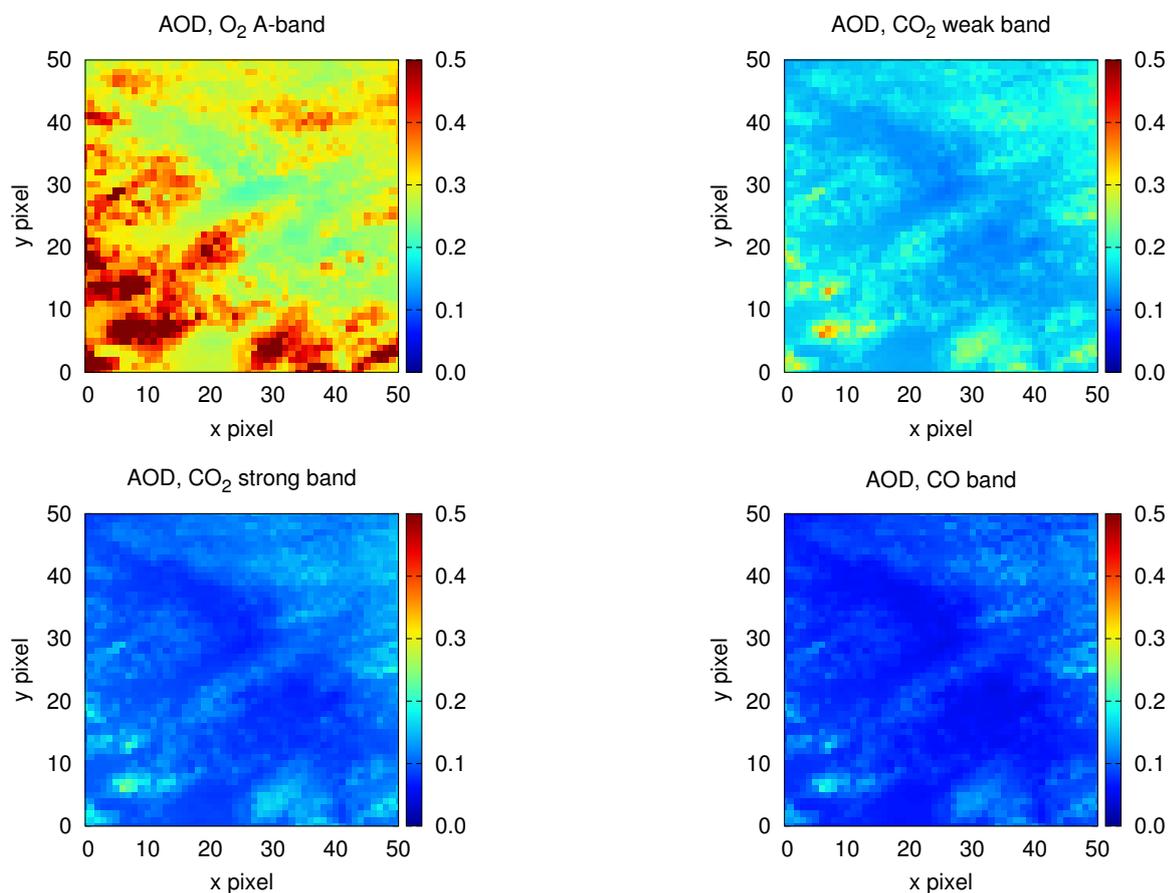
**Figure 3.** Aerosol optical depth over Shanghai on 2010-08-09 at 08:15 UTC in the geoCARB spectral bands. The horizontal and vertical axes are labelled by geoCARB frames and pixels.
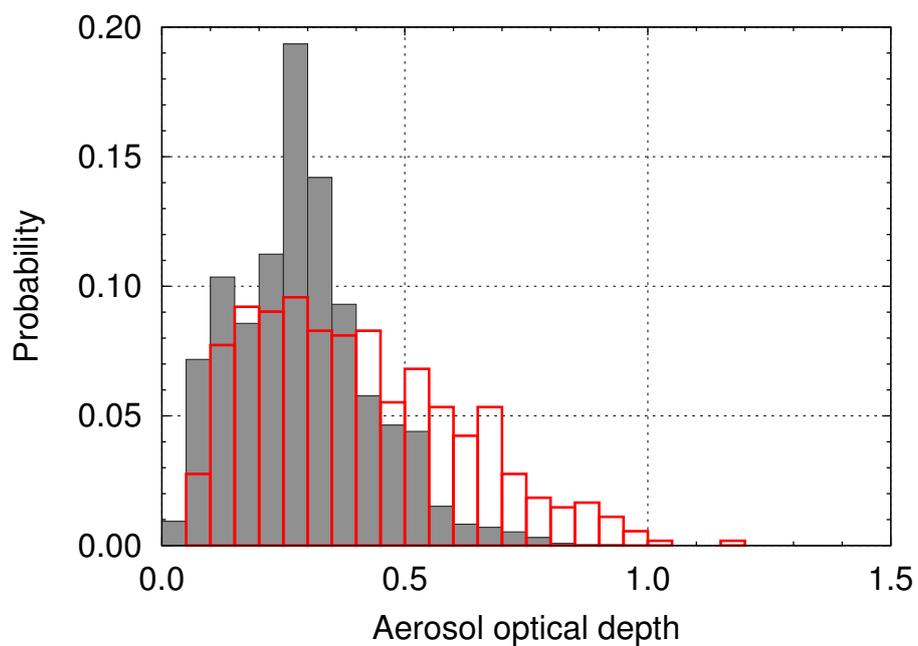
**Figure 4.** Histograms of aerosol optical depth over Shanghai from the WRF-Chem simulations (grey) and from MODIS (red). The wavelength is 770 nm.

.
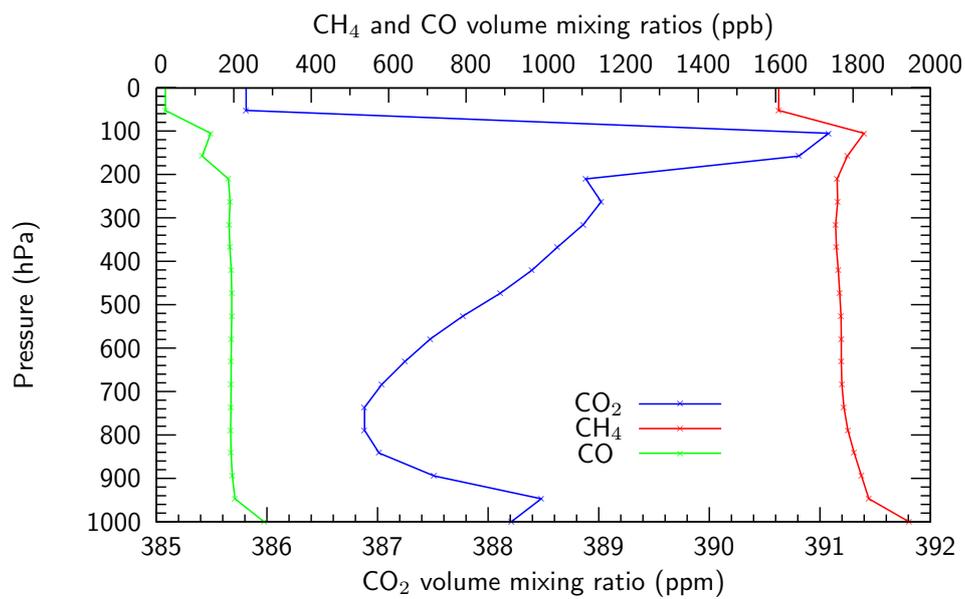
**Figure 5.** Prior vertical profiles for volume mixing ratios of $CO_2$, $CH_4$ and CO.
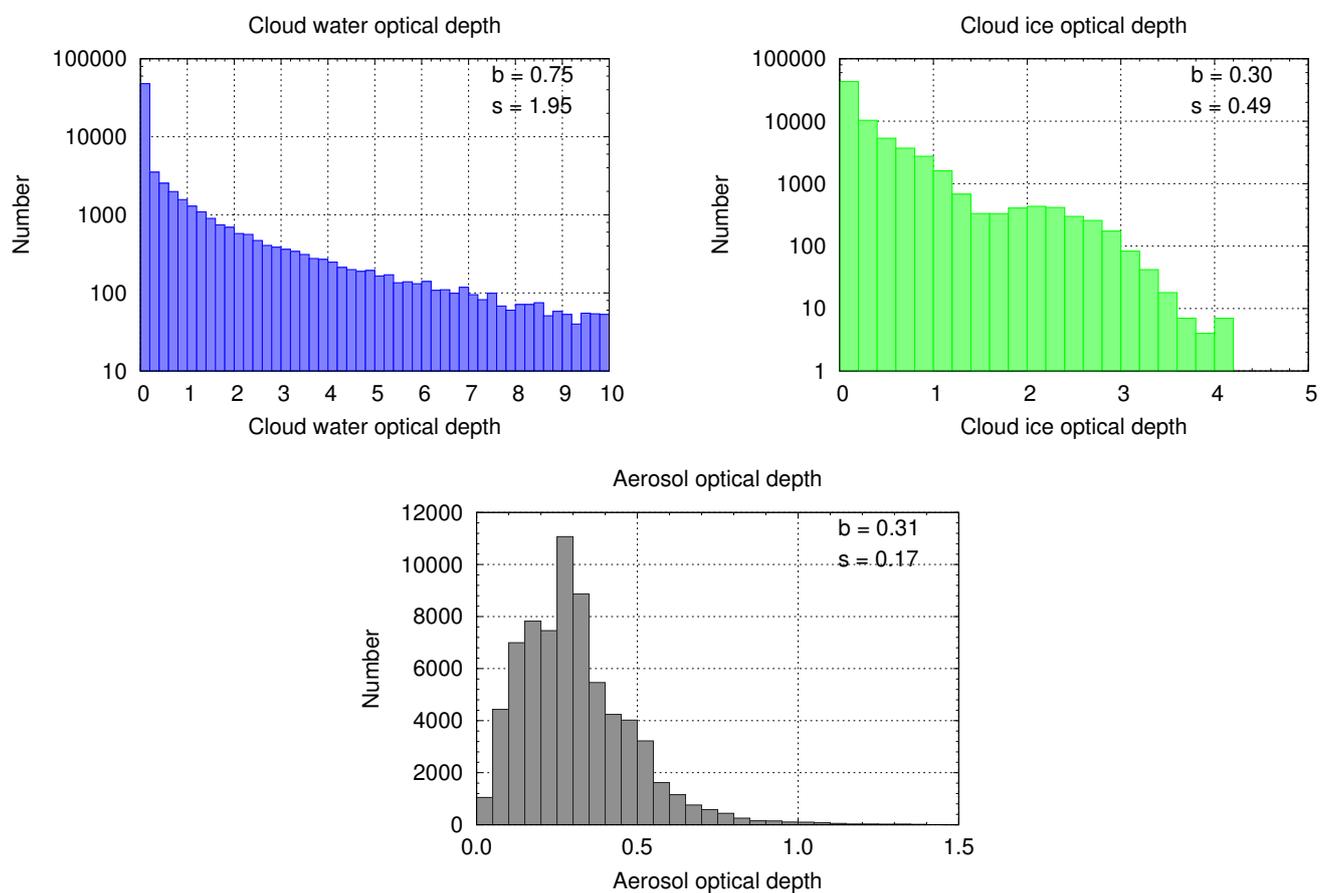
**Figure 6.** Histograms of the optical depths of cloud water, cloud ice and aerosol in the $O_2$ A-band for the cloudy simulations over Shanghai. The histogram of the optical thickness of cloud water extends to 40, but only the range $[0, 10]$ is plotted for clarity. The biases ($b$) and standard deviations ($s$) quoted are for the ensemble of soundings.
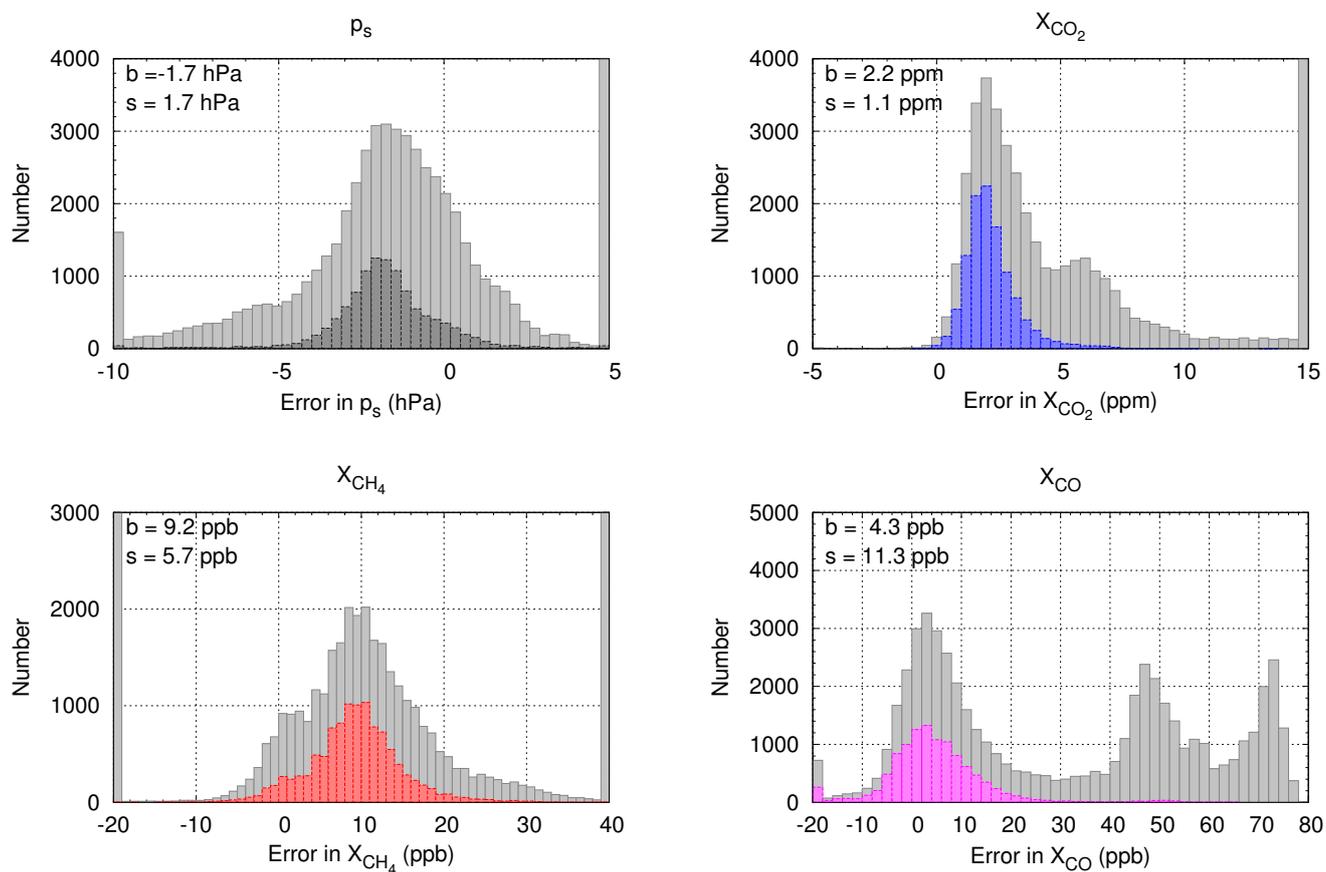
**Figure 7.** Histograms of errors in retrieved surface pressure $p_s$, $X_{CO_2}$, $X_{CH_4}$ and $X_{CO}$ for experiment 1 with cloud disabled and standard aerosol. The PPF has been applied to the coloured histograms, but not to those in light grey. The biases ($b$) and standard deviations ($s$) quoted are for the histograms after the PPF. Of the total of 55,387 cases that converged, 10,995 passed the PPF (19.9%).
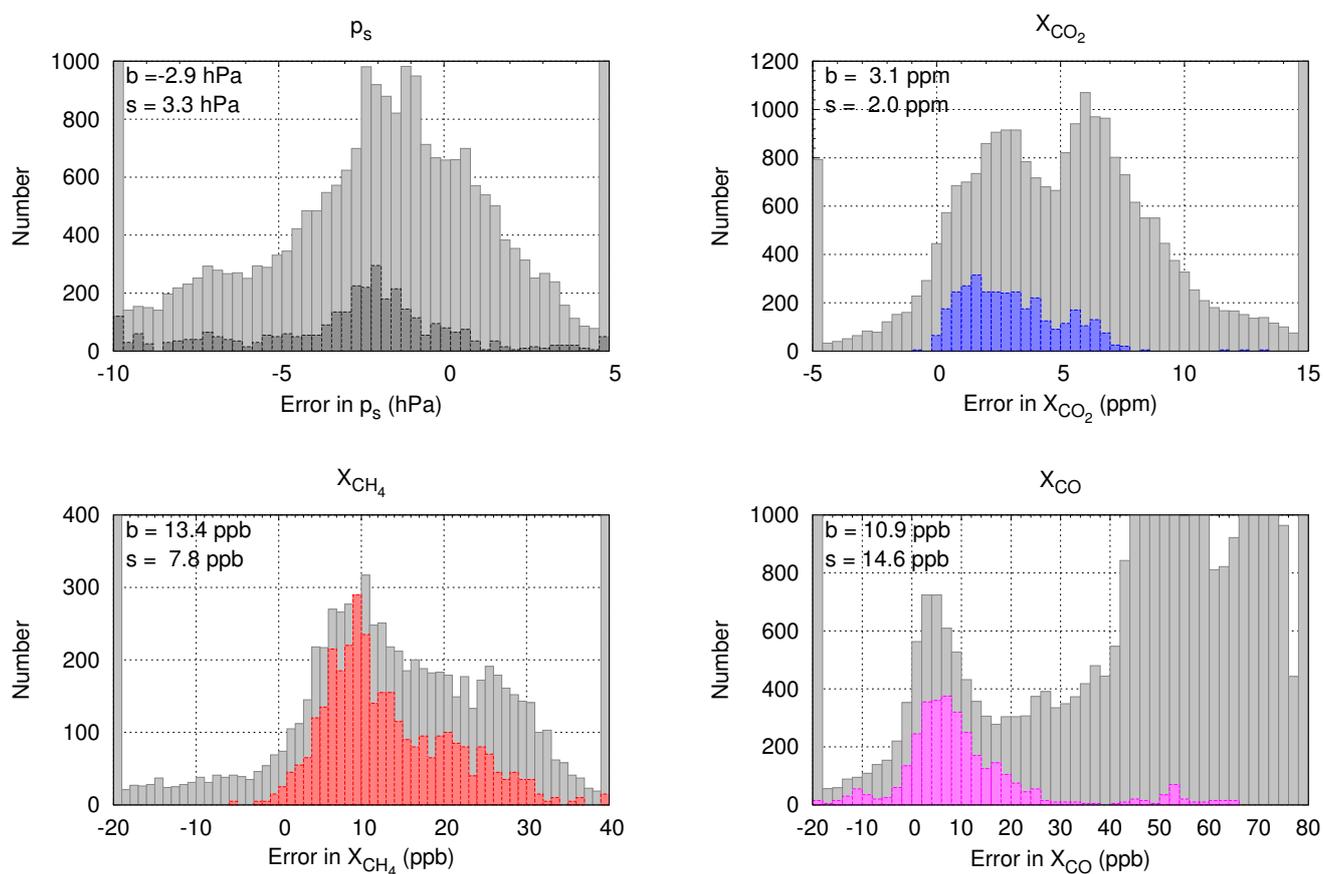
**Figure 8.** Histograms of errors in retrieved surface pressure $p_s$, $X_{CO_2}$, $X_{CH_4}$ and $X_{CO}$ for experiment 2 with cloud enabled and standard aerosol. The PPF has been applied to the coloured histograms, but not to those in light grey. The biases ($b$) and standard deviations ($s$) quoted are for the histograms after the PPF. Of the total of 36,000 cases that converged, 666 passed the PPF (1.9%). For clarity, the vertical scale in the coloured histograms has been multiplied by 5.
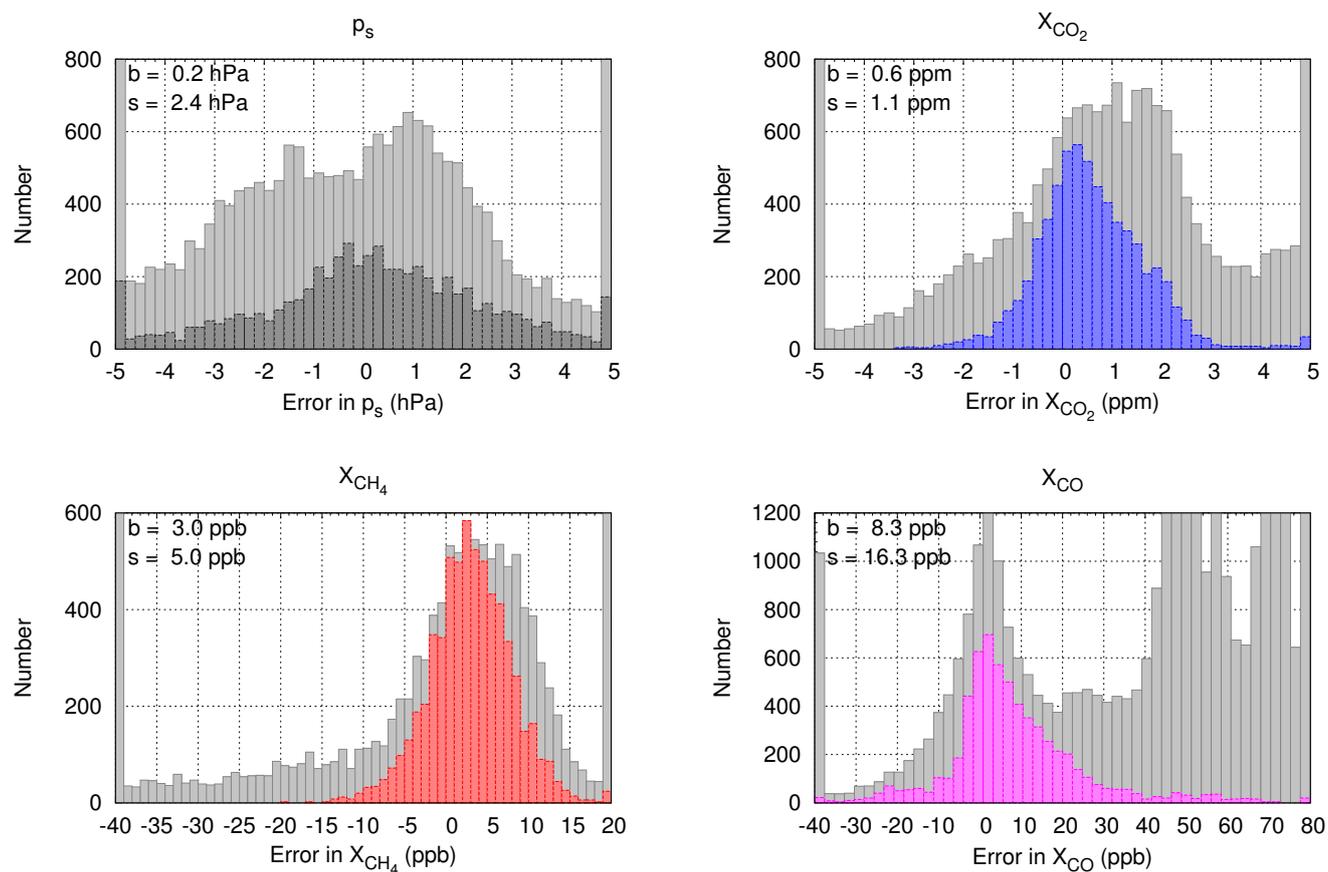
**Figure 9.** Histograms of errors in retrieved surface pressure $p_s$, $X_{CO_2}$, $X_{CH_4}$ and $X_{CO}$ for experiment 3 with cloud enabled and regionally adjusted aerosol. The PPF has been applied to the coloured histograms, but not to those in light grey. The biases ($b$) and standard deviations ($s$) quoted are for the histograms after the PPF. Of the total of 34,664 cases that converged, 3,107 passed the PPF (9.0%). For clarity, the vertical scale in the coloured histograms has been multiplied by 2.
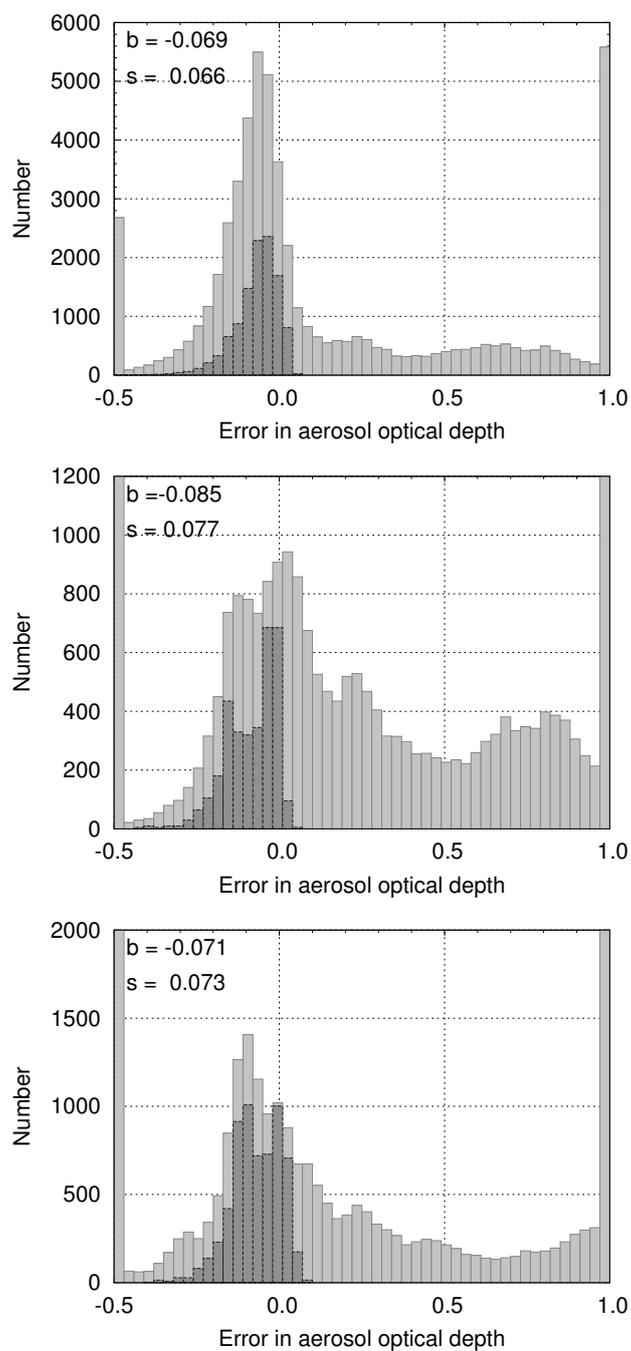
**Figure 10.** Histograms of the errors in retrieved aerosol optical depth in the $O_2$ A-band for experiments 1, 2 and 3 (top, middle and bottom). The PPF has been applied to the dark grey histograms, but not to those in light grey. The bias ($b$) and standard deviation ($s$) quoted are for the histogram after the PPF.
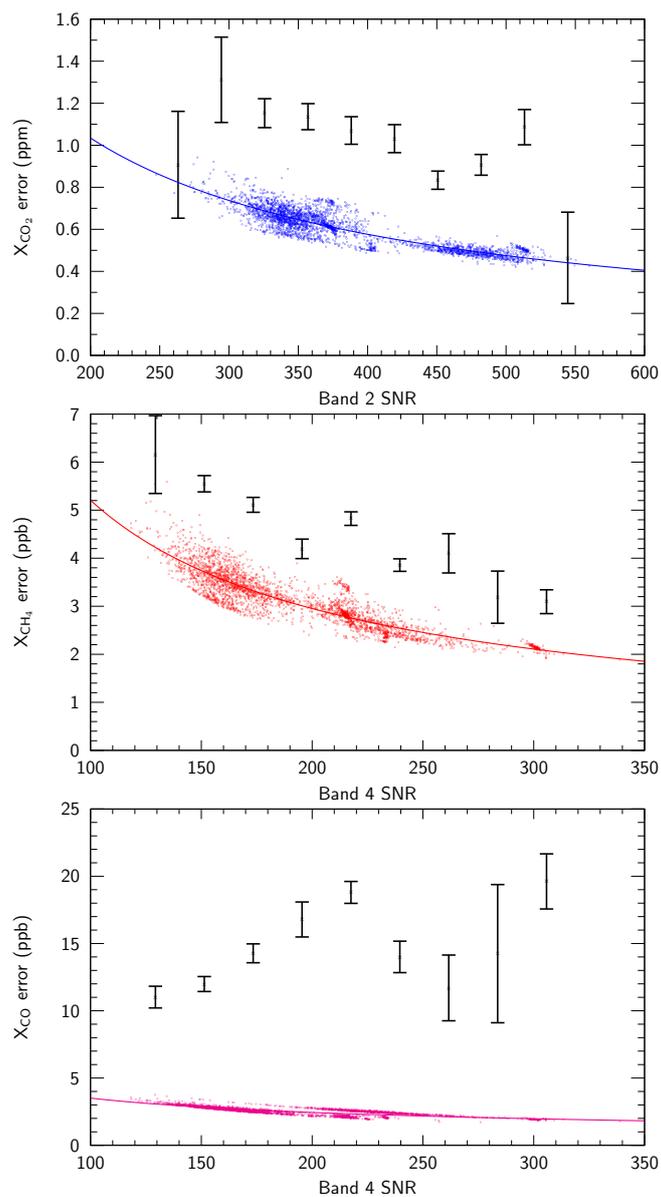
**Figure 11.** Posterior errors estimated by the retrieval algorithm (small dots) and actual errors (retrieved minus the truth weighted by the averaging kernel) plotted as functions of signal-to-noise ratio for the ensemble of soundings in the Shanghai data set that passed the PPF in experiment 3.
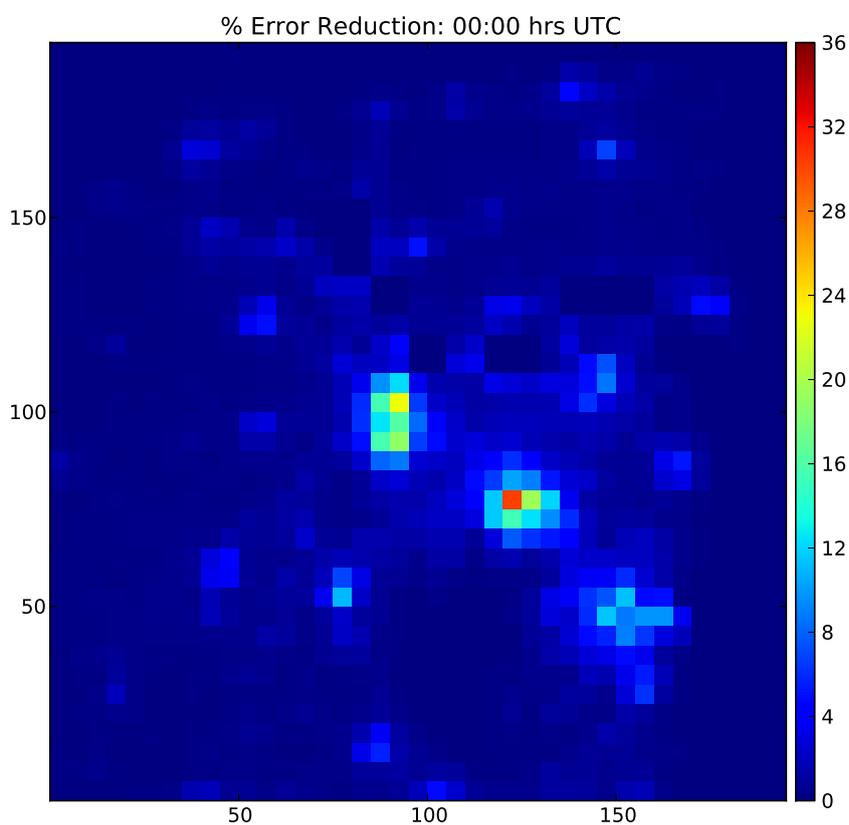
**Figure 12.** Spatial distribution of $E$, the reduction in uncertainty in the $CO_2$ flux, at 00:00 UTC as determined using both $X_{CO_2}$ and $X_{CO}$ data.

**Table 1.** Assignment of EDGAR particulates to Aitken, accumulation and coarse modes.

| EDGAR flux | Aitken | Accumulation | Coarse |
|---|---|---|---|
| Black carbon (BC) | 0.25 | 0.75 | 0.0 |
| Organic carbon (OC) | 0.20 | 0.80 | 0.0 |
| $PM_{2.5}$ | 0.20 | 0.80 | 0.0 |
| $PM_{10}$ | 0.00 | 0.00 | 1.0 |

**Table 2.** Assignment of EDGAR particulate fluxes to MOSAIC bins.

| EDGAR mode | Accumulation | | | | | | Coarse | |
|---|---|---|---|---|---|---|---|---|
| MOSAIC bin | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Fraction $f_j$ | 0.060 | 0.045 | 0.245 | 0.400 | 0.100 | 0.150 | 0.300 | 0.700 |
| | | | $\sum f_j = 1$ | | | | $\sum f_j = 1$ | |

**Table 3.** Noise parameters for the geoCARB bands. The units of $N_0$ and $N_1$ are $nW/(cm^2\,sr\,cm^{-1})$.

| Band ($\mu$m) | $N_0$ | $N_1$ |
|---|---|---|
| 0.765 | 0.1819 | 0.003295 |
| 1.606 | 0.1172 | 0.002107 |
| 2.065 | 0.0814 | 0.001452 |
| 2.323 | 0.0811 | 0.001303 |

**Table 4.** Spectral ranges and resolutions of the geoCARB bands for the baseline configuration.

| Band | Wavelength range | | Band centre | Band width | Dispersion per pixel | Resolution | Resolving power | Pixels per band |
|---|---|---|---|---|---|---|---|---|
| $(\mu m)$ | (nm) | | (nm) | (nm) | (nm) | (nm) | | |
| | $\lambda_-$ | $\lambda_+$ | $\lambda_c$ | | $\delta$ | $\Delta\lambda$ | $\lambda_c/\Delta\lambda$ | $n$ |
| 0.765 | 757.9 | 772.0 | 764.95 | 14.1 | 0.0149 | 0.0475 | 16,104 | 947 |
| 1.606 | 1591.6 | 1621.2 | 1606.4 | 29.6 | 0.0314 | 0.1018 | 15,780 | 943 |
| 2.065 | 2045.0 | 2085.0 | 2065.0 | 40.0 | 0.0424 | 0.1361 | 15,173 | 944 |
| 2.323 | 2300.6 | 2345.6 | 2323.1 | 45.0 | 0.0476 | 0.1531 | 15,174 | 946 |

**Table 5.** Dispersion coefficients $a_k$ for the geoCARB bands.

| Band | $a_0$ | $a_1$ | $a_2$ | $a_3$ |
|------|-------|-------|-------|-------|
| 0.765 | $0.12953445 \times 10^5$ | $0.25001014 \times 10^{+0}$ | $0.48237819 \times 10^{-5}$ | $0.95913210 \times 10^{-10}$ |
| 1.606 | $0.61683511 \times 10^4$ | $0.11947268 \times 10^{+0}$ | $0.23132703 \times 10^{-5}$ | $0.46159484 \times 10^{-10}$ |
| 2.065 | $0.47962017 \times 10^4$ | $0.97535260 \times 10^{-1}$ | $0.19827083 \times 10^{-5}$ | $0.41662801 \times 10^{-10}$ |
| 2.323 | $0.42633342 \times 10^4$ | $0.86518030 \times 10^{-1}$ | $0.17550788 \times 10^{-5}$ | $0.36796186 \times 10^{-10}$ |

**Table 6.** Prior uncertainties $\sigma$ (standard deviations) for the scaling factors applied to the prior profiles for $H_2O$, $CH_4$ and CO.

| Gas | $\sigma$ |
|-----|----------|
| $H_2O$ | $\sqrt{0.25} = 0.5$ |
| $CH_4$ | $\sqrt{0.1} = 0.3162$ |
| CO | $\sqrt{0.001} = 0.03162$ |

**Table 7.** Success rates for converged soundings and soundings that passed the PPF.

| Experiment | | | Number | | | Fraction | |
|---|---|---|---|---|---|---|---|
| Label | Aerosol | Cloud | Total | Converged | Passed PPF | Converged | Passed PPF |
| 1 | standard | disabled | 70,227 | 55,387 | 10,995 | 78.8% | 15.6% |
| 2 | standard | enabled | 70,227 | 36,000 | 666 | 51.2% | 0.9% |
| 3 | adjusted | enabled | 70,227 | 34,664 | 3,107 | 49.3% | 4.4% |

**Table 8.** Parameters for the model of posterior error in terms of signal-to-noise ratio.

| Gas | Parameters | | |
|---|---|---|---|
| | $a$ | $b$ | $c$ |
| $CO_2$ | 10 ppm | 0.068 | 0.915 |
| $CH_4$ | 100 ppb | 0.360 | 0.852 |
| CO | 100 ppb | 2.321 | 0.537 |

**Table 9.** Maximum error reduction (in percent). The labels 00, 06, 12 and 18 refer to the UTC time slot.

| Configuration | CO$_2$ only | | | | CO$_2$ and CO | | | | | | | | | | | |
| | | | | | $\sigma(X_{CO}) = 15\,ppb$ | | | | $\sigma(X_{CO}) = 10\,ppb$ | | | | $\sigma(X_{CO}) \approx 2.5\,ppb$ | | | |
| | 00 | 06 | 12 | 18 | 00 | 06 | 12 | 18 | 00 | 06 | 12 | 18 | 00 | 06 | 12 | 18 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Base | 16.8 | 15.7 | 0.0 | 10.9 | 24.4 | 22.8 | 0.0 | 16.0 | 30.3 | 28.3 | 0.0 | 20.0 | 58.0 | 53.2 | 0.0 | 43.9 |
| $0.2\,\sigma(X_{CO_2})$ | 51.7 | 47.3 | 0.0 | 37.0 | 52.2 | 47.8 | 0.0 | 37.5 | 52.8 | 48.3 | 0.0 | 38.1 | 61.4 | 56.2 | 0.0 | 47.9 |
| $5.0\,\sigma(X_{CO_2})$ | 3.5 | 3.3 | 0.0 | 2.3 | 19.3 | 18.2 | 0.0 | 12.6 | 27.2 | 25.5 | 0.0 | 17.9 | 57.8 | 53.0 | 0.0 | 43.7 |
| Obs $N/2$ | 11.1 | 10.7 | 0.0 | 7.4 | 18.1 | 18.4 | 0.0 | 12.8 | 23.4 | 21.8 | 0.0 | 15.4 | 51.7 | 48.0 | 0.0 | 37.5 |
| Obs $N/4$ | 8.2 | 8.5 | 0.0 | 5.7 | 11.7 | 12.9 | 0.0 | 9.1 | 17.7 | 15.0 | 0.0 | 9.9 | 41.1 | 43.9 | 0.0 | 31.7 |

**Table 10.** MOSAIC aerosol species.

| Species name | Symbol | Refractive index proxy |
|---|---|---|
| Black carbon | BC | black carbon |
| Chloride ions | Cl | sea salt NaCl |
| Sodium ions | Na | sea salt NaCl |
| Ammonium ions | $NH_4$ | $NH_4NO_3$ |
| Nitrate ions | $NO_3$ | $NH_4NO_3$ |
| Organic carbon | OC | organic carbon |
| Other inorganic mass | OIN | dust |
| Sulphate ions | $SO_4$ | $NH_4SO_4$ |
| Water | $H_2O$ | water |

**Table 11.** Real parts of the refractive indices of the proxies listed in Table 10 at the centres of the geoCARB spectral bands.

| Proxy | Band | | | |
|-------|------|------|------|------|
|       | 1    | 2    | 3    | 4    |
| black carbon | 1.850 | 1.850 | 1.850 | 1.850 |
| sea salt NaCl | 1.485 | 1.460 | 1.450 | 1.435 |
| $NH_4NO_3$ | 1.500 | 1.500 | 1.500 | 1.500 |
| organic carbon | 1.750 | 1.770 | 1.800 | 1.815 |
| dust | 1.530 | 1.400 | 1.260 | 1.220 |
| $NH_4SO_4$ | 1.520 | 1.490 | 1.470 | 1.450 |
| water | 1.330 | 1.318 | 1.306 | 1.280 |

**Table 12.** Imaginary parts of the refractive indices of the proxies listed in Table 10 at the centres of the geoCARB spectral bands.

| Proxy | Band | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| black carbon | 0.710 | 0.710 | 0.710 | 0.710 |
| sea salt NaCl | $2.0 \times 10^{-6}$ | $7.0 \times 10^{-4}$ | $1.0 \times 10^{-3}$ | $5.5 \times 10^{-3}$ |
| $NH_4NO_3$ | 0.0 | 0.0 | 0.0 | 0.0 |
| organic carbon | $4.3 \times 10^{-1}$ | $4.6 \times 10^{-1}$ | $4.9 \times 10^{-1}$ | $5.0 \times 10^{-1}$ |
| dust | $8.0 \times 10^{-3}$ | $8.0 \times 10^{-3}$ | $8.0 \times 10^{-3}$ | $9.0 \times 10^{-3}$ |
| $NH_4SO_4$ | $1.0 \times 10^{-7}$ | $7.7 \times 10^{-5}$ | $1.0 \times 10^{-3}$ | $1.0 \times 10^{-3}$ |
| water | $1.0 \times 10^{-7}$ | $9.0 \times 10^{-5}$ | $1.1 \times 10^{-3}$ | $1.0 \times 10^{-3}$ |

**Table 13.** Densities assumed for MOSAIC aerosol species.

| Species name | Symbol | Density (kg/m$^3$) |
|---|---|---|
| Black carbon | BC | $1.8 \times 10^3$ |
| Chloride ions | Cl | $2.2 \times 10^3$ |
| Sodium ions | Na | $2.2 \times 10^3$ |
| Ammonium ions | NH$_4$ | $1.8 \times 10^3$ |
| Nitrate ions | NO$_3$ | $1.8 \times 10^3$ |
| Organic carbon | OC | $1.0 \times 10^3$ |
| Other inorganic mass | OIN | $2.6 \times 10^3$ |
| Sulphate ions | SO$_4$ | $1.8 \times 10^3$ |
| Water | H$_2$O | $1.0 \times 10^3$ |

**Table 14.** Bin limits for the imaginary part of the refractive index.

| $q$ | $\nu_q$ | $q$ | $\nu_q$ | $q$ | $\nu_q$ |
|---|---|---|---|---|---|
| 0 | 0.0000 | 5 | 0.0100 | 10 | 0.2500 |
| 1 | 0.0010 | 6 | 0.0250 | 11 | 0.5000 |
| 2 | 0.0025 | 7 | 0.0500 | 12 | 0.7500 |
| 3 | 0.0050 | 8 | 0.0750 | 13 | 1.0000 |
| 4 | 0.0075 | 9 | 0.1000 | | |

**Table 15.** Parameters used to quantise the cloud particle radius according to Eq. (C2).

| Parameter | Water cloud | Ice cloud | Unit |
|:---:|:---:|:---:|:---:|
| $\alpha$ | 6 | 6 | — |
| $\gamma$ | 1 | 1 | — |
| $r_{e-}$ | 1 | 10 | $\mu$m |
| $r_{e+}$ | 60 | 90 | $\mu$m |
| $a$ | 1 | 5 | $\mu$m |