

UNIVERSITY OF WOLLONGONG

DOCTORAL THESIS

Thesis Title

Author:

Supervisors:
Dr. Jenny FISHER

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in the

Centre of Atmospheric Chemistry
Chemistry Department

June 1, 2018

Declaration of Authorship

I, Jesse GREENSLADE, declare that this thesis titled, "Thesis Title" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

“Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism.”

Dave Barry

Contents

Declaration of Authorship	iii
Abstract	vii
Acknowledgements	ix
1 Introduction and Literature Review	1
1.1 The atmosphere	1
1.1.1 Structure	2
1.1.2 Composition and chemistry	3
1.1.3 Radiative Forcing	3
1.2 Ozone	4
1.2.1 Stratospheric ozone	4
1.2.2 Tropospheric ozone	7
1.2.3 Stratosphere to troposphere transport	9
1.2.4 Chemical production	10
1.3 Volatile Organic Compounds	11
1.3.1 Emissions	13
1.3.2 Isoprene	14
1.3.3 Isoprene chemistry	15
Oxidation	15
High NO _x pathway	17
Low NO _x pathway	17
Night time processes	18
1.4 Formaldehyde	19
1.4.1 Sources and sinks	19
1.4.2 Measurement techniques	20
Satellite measurements	21
1.5 Atmospheric Chemistry Modelling	23
1.5.1 Box models	23
1.5.2 Chemical transport models	23
1.5.3 Emissions	25
1.5.4 Uncertainties	25
Emissions Inventories	26
Resolution	26
Chemistry mechanisms	26
Clouds	27
Soil Moisture	27
1.6 Australia and the southern hemisphere	27

1.6.1	Ozone	29
1.6.2	VOCs	30
1.6.3	Measurements	30
1.7	Aims	31
2	Data and Modelling	33
2.1	Introduction	33
2.2	Datasets	33
2.2.1	Daintree	34
2.2.2	Marine and Urban MBA ? (MUMBA)	34
2.2.3	Sydney Particle Studies (SPS1, SPS2)	34
2.2.4	Satellite	36
OMNO ₂	38	
OMHCHO	38	
OMAERUVd	41	
MOD14A1	41	
2.2.5	Drought Index	41
2.2.6	NPI	42
2.3	GEOS-Chem	42
2.3.1	Outline	42
2.3.2	Running GEOS-Chem (before isop?)	42
Installation and requirements	42	
2.3.3	GEOS-Chem isoprene modelling	43
Outline	43	
2.3.4	Chemical Mechanisms	44
2.3.5	Emissions from MEGAN	45
2.3.6	Rescaling NO _x	47
2.3.7	GEOS-Chem simulations	47
Run comparisons	59	
2.3.8	GEOS-Chem outputs	59
2.4	Measurement Techniques	63
2.4.1	DOAS	63
2.4.2	Satellites	64
LIDORT	65	
OMI	65	
Air mass factor (AMF)	66	
Uncertainties	68	
2.5	Recalculation of OMI HCHO	70
2.5.1	Outline	71
2.5.2	Reading satellite data	74
2.5.3	Scattering weight and the shape factor	75
Normalised vertical shape factor	77	
Sigma coordinate conversion	78	
2.5.4	Recalculating the AMF using PP code	78
2.5.5	Reference sector correction	79
2.5.6	Binning the results daily	83
2.5.7	Difference between new and old OMI HCHO columns	84

2.6 Filtering Data	84
2.6.1 Fire and smoke	86
Checking that fire masks are influencing pyrogenic HCHO	87
2.6.2 NO _x	92
2.6.3 Summary of filters effects on HCHO	92
2.7 Data Access	92
3 Biogenic Isoprene emissions in Australia	87
3.1 Introduction	87
3.1.1 Top-down emissions estimates	88
Linear	88
Bayesian	88
3.1.2 Aims	89
3.2 Methods	89
3.2.1 Outline	89
3.2.2	90
3.2.3 Satellite inversion	90
3.2.4 Calculation of Emissions	93
3.2.5 Emissions drivers	94
3.2.6 HCHO Products and yield	94
3.2.7 Accounting for smearing	97
3.3 Results	100
3.3.1 Emissions comparisons	100
3.3.2 Emissions affect on GEOS-Chem	101
3.4 Uncertainty	101
3.4.1 Model Uncertainty	101
3.4.2 Satellite Uncertainty	101
3 Stratospheric ozone intrusions	63
3.1 Introduction	63
3.2 Data and Methods	65
3.2.1 Ozonesonde record in the Southern Ocean	65
3.2.2 Model description	68
3.2.3 Characterisation of STT events and associated fluxes	69
3.2.4 Biomass burning influence	71
3.2.5 Classifying synoptic conditions during STT events	72
3.3 STT event climatologies	72
3.4 Simulated ozone columns	75
3.5 Stratosphere-to-troposphere ozone flux from STT events	80
3.5.1 Method	80
3.5.2 Results	82
3.5.3 Comparison to literature	82
3.6 Sensitivities and limitations	87
3.6.1 Event detection	87
3.6.2 Flux calculations	88
3.7 Conclusions	89
3.8 Contributions and Acknowledgements	90

A Supplementary Notes	135
A.1 Chemistry	135
SOA	135
A.1.1 Relationship to Glyoxyl TODO: remove if never used	136
A.2 CAABA/MECCA	139
B Appendix A	141
B.1 Grid Resolution	141
C Frequently Asked Questions	143
C.1 How do I change the colors of links?	143
Bibliography	145

Chapter 2

Data and Modelling

2.1 Introduction

In this thesis the word model is most often used to represent a chemical transport model (CTM), which simulates chemistry and chemical transport through the atmosphere. Models of the atmosphere can be used to improve measurements, estimate concentrations in regions not sampled, and allow predictions of atmospheric composition in the future. Models of ozone in the atmosphere are used broadly for international assessments of ozone related emissions and processes which involve ozone (such as radiation) (Young et al. 2017). Models can provide an estimate of trace gas concentrations without requiring measurements, however verification is required.

In-situ measurements from campaigns or measurement stations can be used to examine what is happening at a particular (or several) location(s). These data are useful to determine how accurate models or estimates are - however the utility is limited to where the measurements have been made. In this thesis several campaigns are compared against model outputs.

Satellite datasets give us access to lots of data over large areas. However, satellite measurements can have high amounts of uncertainty due to instrument limitations. Many datapoints can be averaged in order to reduce uncertainty. In this thesis several satellite datasets are combined to provide biogenic HCHO amounts over Australia. In order to compare satellite data with other datasets, some work must be undertaken to avoid introducing bias (eg. Palmer et al. 2001; Eskes and Boersma 2003; Marais et al. 2012; Lamsal et al. 2014).

In this chapter, modelled data are compared against both campaign measurements and satellite data in order to analyse Australia specific HCHO and isoprene sensitivities and to determine where the GEOS-Chem model needs the most improvement. Section 2.2 lists and describes the campaign datasets, section 2.3 describes the GEOS-Chem model, the creation and effects of filters are described in section 2.6 TODO: rest of section outlined.

2.2 Datasets

Here I will describe the various datasets I've used to analyse GEOS-Chem output. I will also give details on filtering and groupings which are undertaken when reading

data, as each dataset has its own resolution. These datasets are used my thesis to determine isoprene emissions estimations in chapter 3 and ozone transport extrapolations in chapter 3.

Figure 2.1 shows the locations of each of the campaigns I mention in this text. These took place over disparate times, and are in-situ measurements which require filtering and interpolation in order to compare against GEOS-Chem output which is averaged over a large horizontal space.

TODO: these summaries.

2.2.1 Daintree

Daintree summary (P. Nelson)

2.2.2 Marine and Urban MBA ? (MUMBA)

2.2.3 Sydney Particle Studies (SPS1, SPS2)

Two VOC and other trace gas measurement campaigns took place at the Westmead Air Quality Station scientists from CSIRO, OEH, and ANSTO. Stage 1 (SPS1) was from 5 February to 7 March in 2011, while stage 2 (SPS2) ranged from 16 April to 14 May 2012.

Two instruments measured VOC concentrations: one was a Proton transfer reaction mass spectrometer (PTR-MS), the other a gas chromatographer (GC) with an equipped flame ionisation detector (FID). The PTR-MS uses chemical ionisation mass spectrometry and can quantify VOCs at high temporal resolution (< 1 s). It was calibrated several times per day against hcho, isoprene, α -pinene, and several more VOCs. Further details can be found in Dunne2012; Dunne et al. (2017) (TODO: Check papers). The output lists hourly averaged ppbv concentrations of trace gases based on the mass to charge ratio (m/z), which for isoprene is 69. It's possible that other chemicals (such as Furan, with the same m/z) interfered with this value, especially at low ambient isoprene concentrations and towards the end of autumn (SPS2) when wood fires usage starts to become frequent (TODO cite something). The GC-FID analysed samples collected in multi-absorbent tubes, with lower temporal resolution but no interferences. GC-FID data is averaged from 0500-1000 LT, and 1100-1900 LT. Further details for this method can be found in TODO: cite Min et al 2016.

Figure 2.2 shows isoprene and formaldehyde over the course of these two campaigns, as well as the detection limits (dashed lines), as measured by PTR-MS. In order to compare with GEOS-Chem output a daily average and an overpass time (1200-1300 LT) average are both created from these data. In averaging, any measurements below the machine detection limit are set to half of the detection limit, as done in (TODO: doi:10.5194/acp-15-223-2015, 2015) which should minimise any introduced bias.

Figure 2.3 shows GEOS-Chem output in the gridsquare containing Sydney overlaid on SPS measurement data. Superficially the comparison is not too bad between these two datasets, however GEOS-Chem output is daily averaged over $2 \times 2.5^\circ$ (latitude by longitude) The SPS data is point-source and taken during the daytime when isoprene is higher, so it is very likely that GEOS-Chem HCHO and isoprene output is in fact too high.



FIGURE 2.1: Locations of Australian campaigns which are analysed within this thesis

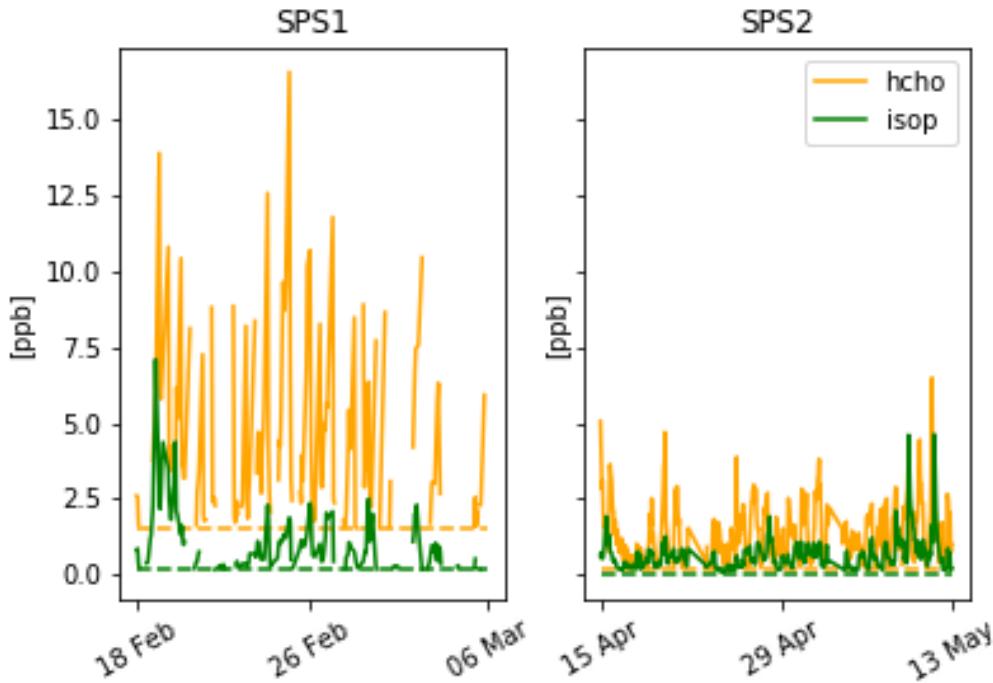


FIGURE 2.2: SPS 1 and 2 HCHO (yellow) and isoprene (green) time series, along with detection limits (dashed).

2.2.4 Satellite

Satellite data products are generally classed into several categories, level 0 through to level 3. Level 0 products are sensor counts and orbital swath data, level 1B data calibrates and geo-locates the level 0 data. Level 2 products additionally have temporal, spatial, solar, and viewing geometry information, as well as quality flags. To create level 2 data slant column density is determined and then translated into vertical column density using an AMF calculated through radiative transfer models. Level 3 data is a temporally aggregated version of the this, for instance monthly or yearly averages.

One satellite is NASA's Earth Observing System's "Aura", which provides several useful datasets, often called products. Aura orbits the earth in a polar sun-synchronous pattern, circling the earth on a plane coincident with the sun and the poles. OMI measurements are used to map several atmospheric trace gases, including NO₂, SO₂, BrO, HCHO, O₃, and aerosols. OMI measurements occur from right to left on a band covering 115°, resulting in swaths of around 2600 km, with pixel sizes from 13x24 km² at nadir to 26x135 km² at the swath edges (Gonzalez Abad et al. 2015). The swaths cover Earth daily, although half of these are at night time and contain no useful near-UV/Visible information.

Aura houses the Ozone Monitoring Instrument (OMI), a near-UV/Visible Charged Coupled Device (CCD) spectrometer. From here on the word pixel is used to describe one data point retrieved by OMI, each pixel includes a latitude and longitude within OMI's data product. OMI spectra are used in several products used in this thesis, including OMNO2d, OMHCHO, and OMAERUVd.

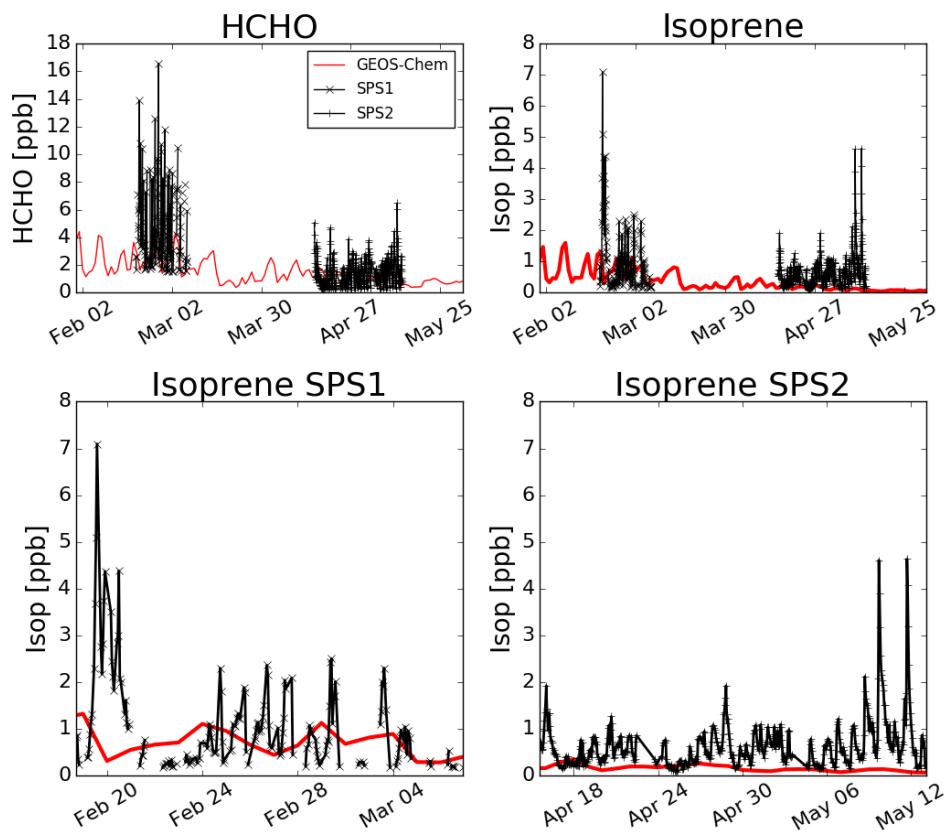


FIGURE 2.3: Comparison between GEOS-Chem HCHO concentrations in the gridsquare containing Sydney for the duration of the SPS 1 and 2 campaigns

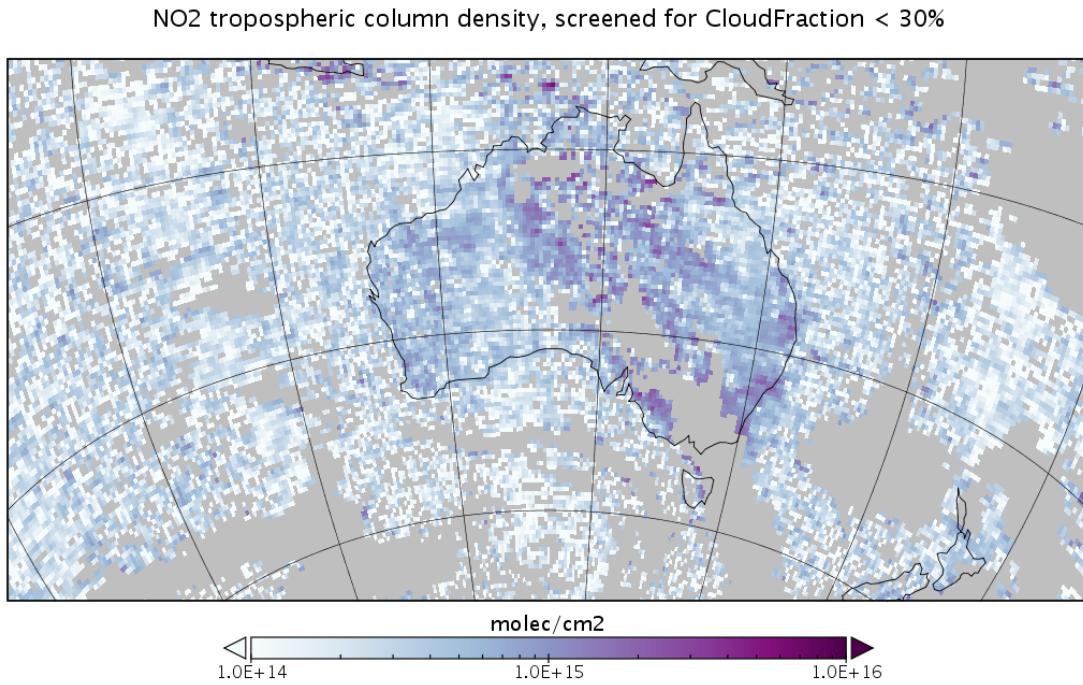


FIGURE 2.4: Example of NO₂ tropospheric columns taken from the OMNO2d product.

OMNO2

NO₂ measured by OMI is used to check whether NO₂ is well represented by GEOS-Chem. OMNO2d is a gridded daily level three product with good satellite pixels averaged into 0.25x0.25° horizontally resolved bins. An example figure from Jan 29, 2005 is shown in figure 2.4, while an average for 2005 (global) is shown in figure 2.5.

See section ?? for the comparison between this product and GEOS-Chem calculations.

OMHCHO

Atmospheric HCHO can be measured using Differential Optical Absorption Spectroscopy (DOAS), as long as trace gases with similar features near the same wavelength are accounted for. A DOAS fit determines the total column amount of a trace gas along the path that the instrument views. This uses the Beer-Lambert law where radiance is reduced as light travels through a medium. I use the NASA OMHCHOv003 data product (Gonzalez Abad et al. 2015), with HCHO determined using the spectral window 328.5 nm–356.5 nm. The algorithm used is based on direct fitting of radiances, and accounts for competing absorbers, under-sampling, and Ring effects. An OMI radiance measurement over the remote Pacific ocean is used instead of an irradiance measurement. This means that the slant columns (Ω_S) are actually the difference with respect to the radiance reference column (Ω_{S_0}). The full method details for slant column retrieval by OMI are outlined in section ??.

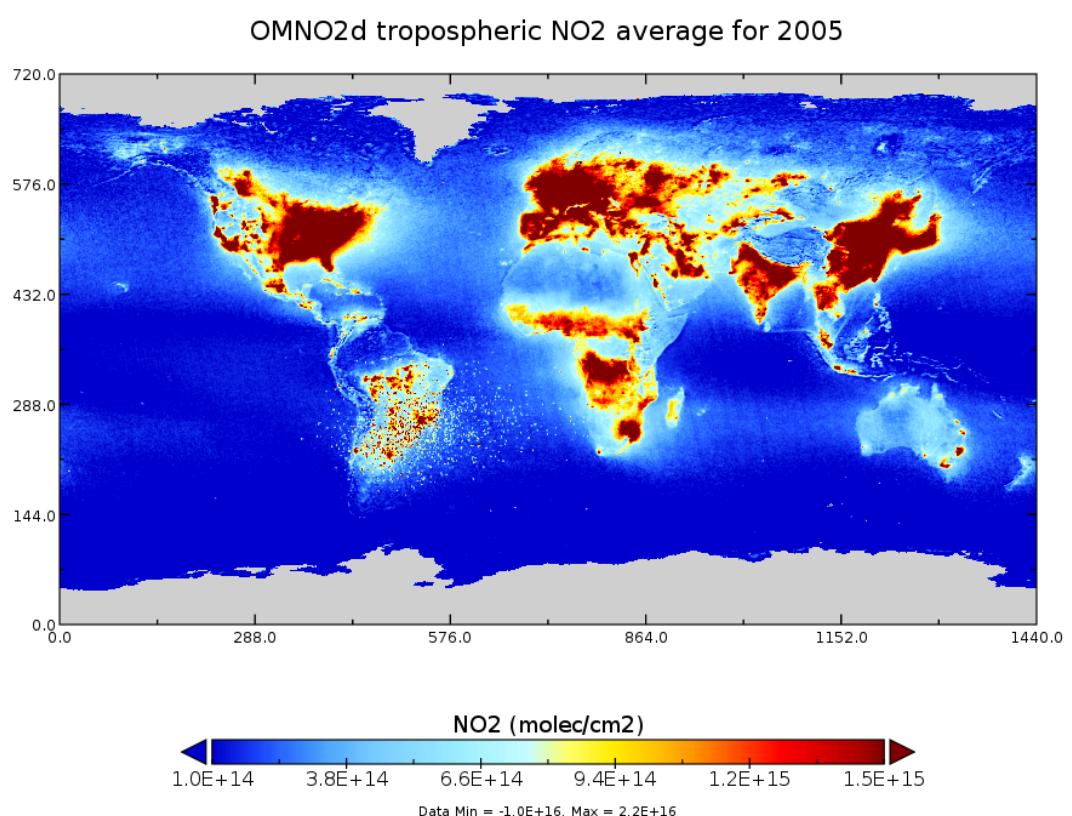


FIGURE 2.5: Average 2005 tropospheric NO₂ from OMNO2d with pixels screened for < 30% cloud cover.

TABLE 2.1: OMI quality flag values table from Kurosu and Chance (2014)

Value	Classification	Rational
0	Good	Column value present and passes all quality checks; data may be used with confidence.
1	Suspect	Caution advised because one or more of the following conditions are present: <ul style="list-style-type: none"> • Fit convergence flag is < 300 but > 0: Convergence at noise level • Column $+2\sigma$ uncertainty $< 0 <$ Column $+3\sigma$ uncertainty • Absolute column value $>$ Maximum column amount ($1e19$ molec cm^{-2})
2	Bad	Avoid using as one of the following conditions are present: <ul style="list-style-type: none"> • Fit convergence flag is < 0 : No convergence, abnormal termination • Column $+3\sigma$ uncertainty < 0
< 0	Missing	No column values have been computed; entries are missing

$\sim 4 \times 10^{15}$ to $\sim 6 \times 10^{16}$ molec cm^{-2} , with uncertainties from 30% (larger columns) to over 100% (smaller columns) (Gonzalez Abad et al. 2015).

OMHCHO level two data includes 14-15 daily swaths of measurements provided by NASA. Each swath contains roughly 9×10^4 pixels, each of which contains various data including latitude, longitude vertical column HCHO, etc. The OMHCHO dataset has a quality flag which can be used to remove unlikely or poor satellite measurements. The states represented by this quality flag are shown in table 2.1 which is taken from Kurosu and Chance (2014). Filtering bad or missing measurement pixels is preformed prior to any other filtering, this includes the datapoints affected by the row anomaly. This anomaly (<http://projects.knmi.nl/omi/research/product/rowanomaly-background.php>) affects radiance data at particular viewing angles, corresponding to a row on the CCD detectors, and is dynamic over time. The slant columns affected are flagged and easy to remove before further processing.

The cloud fraction with each pixel is provided with the OMHCHO dataset, however its source is the OMI cloud product, OMCLDO2. To give an idea of how much data is filtered out, around 30% of the pixels which remain after filtering out the bad or missing data are subsequently removed due to cloudiness.

Each ~ 90 minutes the AURA satellite sweeps over the sunny side of the planet, with OMI recording roughly 90 k pixels, of which around 50 k – 80 k are classified as good. Each pixel contains several important pieces of data which are needed for recalculation of the HCHO vertical column: the total column of HCHO (Ω molec cm^{-2}),

cloud fraction, associated shape factor, AMF, geometric AMF, scattering weights and their vertical altitudes (hPa), viewing zenith angle, solar zenith angle, latitude, longitude, OMI sensor track, main data quality flag, cross track flag, and total column uncertainty. All of these data are needed in order to reconstruct the total vertical column using a modelled a priori shape factor rather than NASA's included a priori shape factor.

Recalculated OMI formaldehyde columns are used as a basis for estimating isoprene emissions in Chapter 3.

OMAERUVd

Aerosols in the atmosphere can be seen through their affects on light. Smoke and dust can be seen as an increase in AAOD (see section 2.4.1). This is due these particles scattering and absorbing UV radiation (Ahn2008).

OMAERUVd (DOI: 10.5067/Aura/OMI/DATA3003) provides a useful dataset allowing us to filter gridsquares which may be smoke affected. OMI aerosol extinction and absorption optical depths (AOD, AAOD respectively) at three wavelengths (354, 388, and 500 nm), along with UV aerosol index (UVAI), are available publicly from Earthdata: https://disc.gsfc.nasa.gov/datasets/OMAERUVd_V003/summary. The OMAERUVd product is level three, gridded daily data, based on quality filtered level two swath pixels which are then gridded by averaging. The product is most sensitive to error in the form of subpixel scale cloud interference, so I select AAOD as the basis for my smoke filter as it is least affected by clouds (Ahn2008).

Gridded smoke AAOD is read from OMAERUVd at 1x1°resolution daily, and mapped to finer resolution using the nearest value for each gridsquare. I use the AAOD at 500 nm wavelength, which is blocked by fire smoke plumes. This daily AAOD is used to mask fire smoke plume influence, by masking gridsquares with higher AAOD > 0.03.

MOD14A1

Daily gridded fire counts compiled from Terra and Aqua satellite into 1x1 km² resolution.

2.2.5 Drought Index

The S Precipitation Evapotranspiration Index (SPEI) is a measure of drought using various parameters such as TODO. (Wang et al. (2017)). SPEI will be compared against the difference between top-down estimated emissions and MEGAN bottom up estimated emissions. This is used to determine whether there are biases in the MEGAN calculations due to the GEOS-Chem implementation ignoring soil moisture. It is downloaded from TODO and holds monthly averaged values at 0.5° horizontal resolution. When comparing against the emissions estimates this is interpolated linearly onto the same grid as that of GEOS-Chem output at 2x2.5°.

2.2.6 NPI

One possible solution to anthropogenic filtering is the national pollution index (TODO:cite:<http://www.npi.gov>) which contains industrial HCHO and NO_X emissions from 2003 to 2014.

2.3 GEOS-Chem

2.3.1 Outline

GEOS-Chem is an atmospheric chemical model (ACM), using a 3-D grid of boxes with transport driven by the GEOS meteorological model and chemistry calculated in each box independently. Within each box chemistry, transport, and meteorology are simulated at 15 minute time steps. GEOS-Chem has a meteorological model coupled to a chemical model, which simulates the world in a three dimensional grid of connected boxes.

GEOS-Chem is a well supported global, Eulerian CTM with a state of the science chemical mechanism, with transport driven by meteorological input from the Goddard Earth Observing System (GEOS) of the NASA Global Modeling and Assimilation Office (GMAO). GEOS-Chem simulates more than 100 chemical species from the earth's surface up to the edge of space (0.01 hPa) and can be used in combination with remote and in-situ sensing data to give a verifiable estimate of atmospheric gases and aerosols. It was developed, and is maintained, by Harvard University staff as well as users and researchers worldwide. In this thesis I use version 10.01 of GEOS-Chem, which simulates and records up to 66 chemical species (tracers) in the standard run, at 2 by 2.5° horizontal resolution, with 47 levels up to the top of the atmosphere (TOA at 0.01 hPa).

Global CTMs are often run using one or several emission models (or the output from them) to determine boundary conditions for many gridboxes. Some of the inventories are described here. Meteorological fields are taken from NASA's GEOS-5 dataset (0.5° x 0.666°) (Chen et al. 2009), which exists up to April, 2013. GEOS-5 meteorological fields are used as the boundary conditions driving transport. Fire emissions come from the GFED4 product (Giglio, Randerson, and Van Der Werf 2013). Anthropogenic VOC emissions come from the EDGAR inventory, while biogenic VOC emissions are simulated using the MEGAN model. MEGAN is used to determine biogenic emissions for our default GEOS-Chem simulation, described in more detail in section TODO. The estimated biogenic VOC emissions are important for accurately simulating chemistry within models, as discussed in Section 1.1.2.

2.3.2 Running GEOS-Chem (before isop?)

Installation and requirements

GEOS-Chem instructions for download, compilation, and running can be found in the user guide provided by Harvard: <http://acmg.seas.harvard.edu/geos/doc/man/>. In order to build and run GEOS-Chem a high-speed computing system is optimal, as globally gridded chemical calculations can take a long time to perform. I installed GEOS-Chem onto a suitably configured workspace on the National Computational

Infrastructure (NCI, <http://nci.org.au/>). This workspace included access to compilers and libraries which are needed to build the Fortran based GEOS-Chem source code, and IDL, Python, and various editors and scripting languages to read, run, edit, and analyse both GEOS-Chem and its output.

After downloading GEOS-Chem, the code can be compiled with different options for resolution and chemical mechanisms.

2.3.3 GEOS-Chem isoprene modelling

Outline

The isoprene reactions simulated by GEOS-Chem were originally based on Horowitz et al. (1998). This involved simulating NO_x, O₃, and NMHC chemistry in the troposphere at continental scale in three dimensions, with detailed NMHC chemistry with isoprene reactions and products. The mechanism was subsequently updated by Mao et al. (2013), who change the isoprene nitrates yields and add products based on current understanding as laid out in Paulot et al. (2009a) and Paulot et al. (2009b). Further mechanistic properties, like isomerisation rates, are based on results from four publications: citeCrounse2011,Crounse2012,Peeters2010,Peeters2011. (TODO: check abstracts Peeters papers).

Crounse et al. (2011) examines the isomerisations associated with the oxidation of isoprene to six different isomers of ISOPOO formed in the presence of oxygen through $ISOP + OH \rightarrow ISOPOO$. They determine rates and uncertainties involved in these reactions, and study the rate of formation of C₅-hydroperoxyaldehydes (HPALDs) by isomerisation. Methacrolein (MACR) is one of the products of isoprene oxidation. Prior to 2012 MACR oxidation chamber studies were performed in high NO or HO₂ concentrations, giving peroxy lifetimes of less than 0.1 s (Crounse et al. 2012). In most environments this is not the case, GEOS-Chem uses MACR production rates for different NO concentrations and peroxy radical lifetimes determined by Crounse et al. (2012). Additionally, photolysis of photolabile peroxy-acid-aldehydes creates OH and improved model agreement with continental observations. photolysis of hydroperoxy-methyl-buteneals (HPALDs, produced by isoprene isomerisation) regenerate OH levels in areas with high isoprene emissions (Peeters and Muller 2010). OH and HPALD interactions are central to maintaining the OH levels in pristine and moderately polluted environments, which makes isoprene both a source and a sink of OH (Peeters and Muller 2010; Taraborrelli et al. 2012).

Formation of isoprene nitrates have an effect on ozone levels through NO_x sequestration, and the yields and destinies of these nitrates is analysed in Paulot et al. (2009a). In a chamber with clean air and high NO concentrations, isoprene photooxidation is initially driven by OH addition, followed by NO_x chemistry (150 min - 600 min), and finally HO_x dominated chemistry. GEOS-Chem uses these the yields of various positional isomers of isoprene nitrates, and pathways of their oxidation products, and reactions within its suite of chemical mechanisms determined by Paulot et al. (2009a) and Mao et al. (2013).

In low NO_x conditions, isoprene oxidises to yield 70% hydroxyhydroperoxides (ISOPOOH), which then oxidises to create dihydroxyperoxides (IEPOX) with OH recycling maintaining the OH levels in the atmosphere (Paulot et al. 2009b). In older

models isoprene produced ISOPOOH which then titrated OH, however, the loss of OH has not been seen in measurements (Paulot et al. 2009b; Mao et al. 2013). The isoprene mechanism in GEOS-Chem has been updated to include OH regeneration from oxidation of epoxydiols and slow isomerisation of ISOPO₂ (Mao et al. 2013).

ISOPN can be oxidised (by OH) to form nitrated organic products (Paulot et al. 2009a). In low NO_x ISOPOO reacts with HO₂ (producing hydroxy hydroperoxides, ISOPOOH), RO₂ (producing mainly MACR, MVK, and HCHO), or isomerises (1,5-H shift producing MACR, MVK, HCHO, or 1,6-H shift producing hydroperoxyenals HPALDs). ISOPOOH can be oxidised (by OH) to produce epoxydiols (IEPOX), precursors to SOA (Paulot et al. 2009b). HPALDs can photolyse to regenerate OH and small VOCs (Crounse et al. 2011; Wolfe et al. 2012; Jozef et al. 2014) TODO: Check out crounse2011. See section 1.3.3 for more information.

Under high NO_x conditions, isoprene undergoes OH addition at the 1 and 4 positions, becoming β (71%) or δ (29%) hydroxyl peroxy radicals (ISOPO₂). The β -hydroxyl reacts with NO_x and produces HCHO (66%), methylvinylketone (40%) (MVK), methacrolein (26%), and β -hydroxyl nitrates (6.7%) (ISOPNB). The δ -hydroxyl reacts with NO to form δ -hydroxyl nitrates (24%) (ISOPND), and ISOPNB (6.7%). ISOPNB and ISOPND yield first generation isoprene at 4.7% and 7% respectively.

Under low NO_x conditions, ISOPO₂ may react with HO₂ to form ISOPOOH. In this case there is also production of HCHO (4.7%), MVK(7.3%), and MACR (12%). As stated in earlier; most ISOPOOH will form IEPOX (epoxydiols) after reacting with OH and lead to OH regeneration. The other mechanism in low NO_x environments is unimolecular isomerisation of ISOPO₂. This leads to production of hydroperoxyaldehydes (HPALDS), which generally photolyse and have an OH yield of 100%. Mao et al. (2013) show that a lower (factor of 50) rate constant for ISOPO₂ isomerisation leads to better organic nitrate agreements with ICARTT.

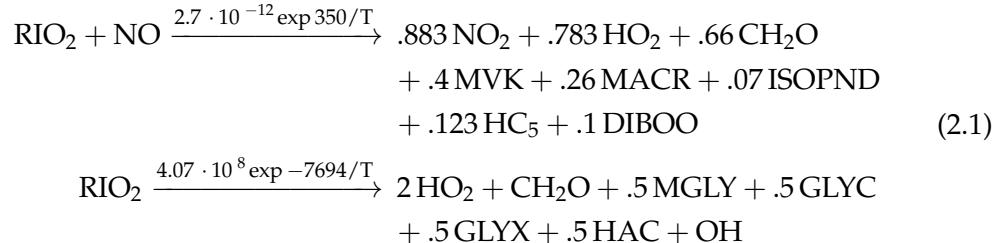
This update leads to more accurate modelling of OH concentrations, especially in low NO_x conditions common in remote forests. Prior to Mao et al. (2012), measurements of OH in high VOC regions may have been up to double the real atmospheric OH levels, due to formation of OH inside the instrument. Mao et al. (2012) examine an upgraded method of measurement, and compare these against a regional atmospheric chemistry model (RACM2), with the OH recycling updates from Paulot et al. (2009b) as discussed in prior paragraphs.

The updates to isoprene chemistry by Mao et al. (2013), and those shown in Crounse et al. (2011) and Crounse et al. (2012) are the last before version 11. The full current mechanism is described online at http://wiki.seas.harvard.edu/geos-chem/index.php/New_isoprene_scheme.

2.3.4 Chemical Mechanisms

Chemical reactions are turned into systems of differential equations (DEs) to be solved by the CPU for each gridbox in GEOS-Chem. A chemical mechanisms is the name for a closed system of chemical reactions and the rates of each reaction. Simplifications are required due to the massive amount of reactions which occur in the atmosphere, and the coupled and stiff nature of these reactions which serve to slow down computation of the solutions TODO: ref Brasseur Jacob book.

Some of the important ones involving isoprene are copied here, including reaction rates in the form $k = A \exp -ER/T$. The full list of chemical reactions can be found online at TODO: find list



2.3.5 Emissions from MEGAN

MEGAN is a global model with resolution of around 1 km, and is used to generate the BVOC emissions used in various global chemistry models such as GEOS-Chem. MEGAN uses leaf area index, global meteorological data, and plant functional types (PFTs) to simulate terrestrial isoprene emissions. The model includes global measurements of leaf area index, plant functional type, and photosynthetic photon flux density, from remote sensing databases (Kefauver, Filella, and Peñuelas 2014). The various PFTs are used to generate emission factors which represent quantities of a compound released to the atmosphere through an associated activity. For example, an emission factor for isoprene within a forest would include the requirement of sunshine and suitable temperature. The schematic for MEGAN, taken from Guenther (2016), is shown in figure 2.6

GEOS-Chem V10.01 uses MEGAN V2.1 with biogenic emissions from Guenther et al. (2012). It computes some emissions using predefined EF maps from MEGAN source code, and others using PFT maps and associated EFs. MEGAN “is a modelling framework for estimating fluxes of biogenic compounds between terrestrial ecosystems and the atmosphere to account for the major known processes controlling biogenic emissions.” (Guenther et al. 2012). It allows parameterisation of various BVOC emissions, with descriptions given in Guenther et al. (2012). Instructions to run version 2.1 are available at http://lar.wsu.edu/megan/docs/MEGAN2.1_User_GuideWSU.pdf, and a version using the Community Land Model (CLM) is available at <http://www.cesm.ucar.edu>. It uses meteorological fields from the Weather Research and Forecasting (WRF) modelling system. Version 2.1 (updated from 2.0 (Guenther et al. 2006)) includes 147 species, in 19 BVOC classes, which can be lumped together to provide appropriate output for mechanisms in various chemical models.

MEGAN was developed as a replacement for two earlier canopy-environment emission models (BIES and GEIA), and initially included a simple canopy radiative transfer model, which parameterised sun-lit and shaded conditions through a canopy. Early models didn’t account for abiotic stresses, such as drought, prior rainfall and development processes. These stresses influenced species-specific emissions by more than an order of magnitude (Niinemets et al. 1999). Isoprene emissions were based on temperature, leaf area, and light, but have since been updated to include leaf age activity (Guenther et al. 2000), and a leaf energy balance model (Guenther et al. 2006)

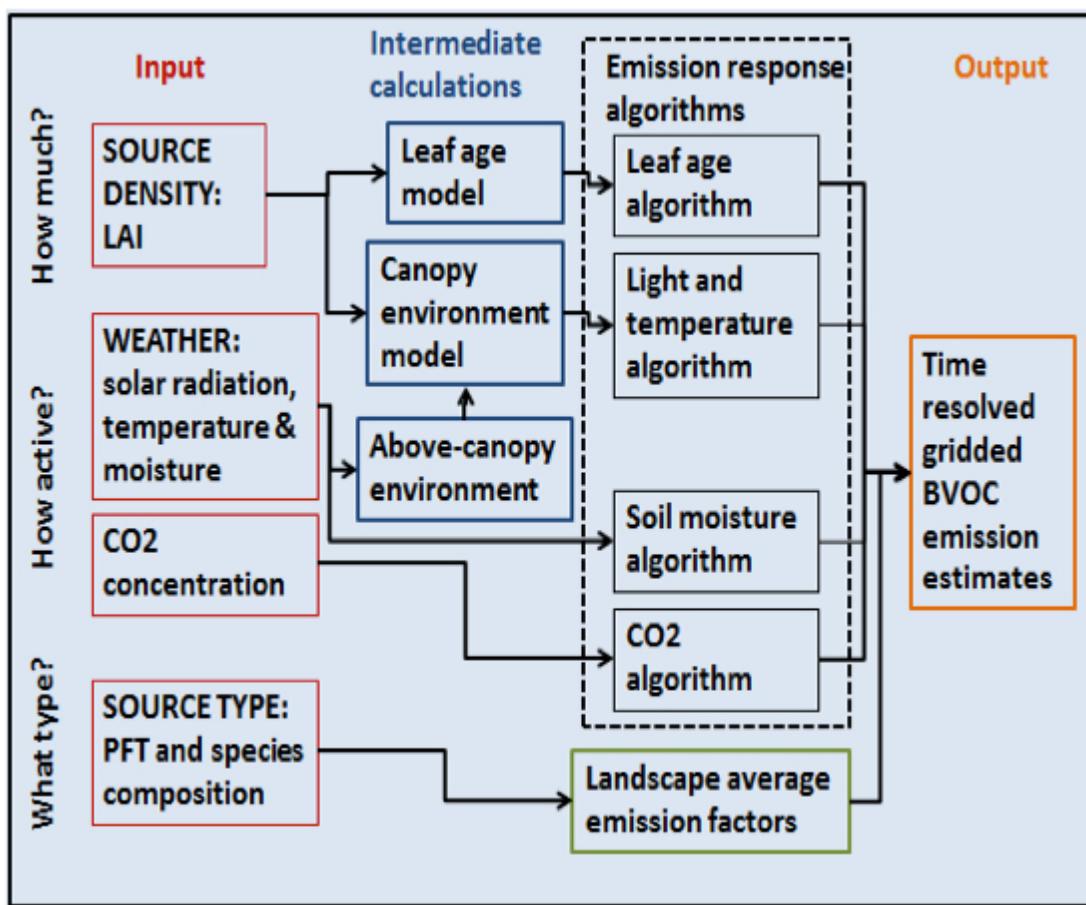


FIGURE 2.6: MEGAN schematic, copied from Guenther (2016)

in MEGANv2.0. This update included a parameter for soil moisture, to account for drought conditions, however this parameter is currently (as of version 2.1) not applied to isoprene (Sindelarova et al. 2014). Soil moisture effects on isoprene emission are very important, and can drastically affect estimates.

2.3.6 Rescaling NO_x

NO_x concentrations affect HCHO yield, isoprene lifetimes, and other things due to affects on the atmospheres oxidative capacity. This means that if the model is poorly simulating NO_x, isoprene to HCHO yield and transport (see 3.2.7) may be poorly estimated. In order to determine if rescaling the NO emissions over Australia is necessary in GEOS-Chem, I looked at modelled NO₂ amounts compared to satellite data for most of 2005.

Simulated GEOS-Chem tropospheric NO₂ columns averaged from 1300-1400 LT are compared against OMNO2d data (Sec. 2.2.4). Figure 2.7 shows the direct comparison between these datasets averaged over the month of January, 2005. It's clear that the OMNO2d product can pick out Sydney and Melbourne as NO₂ hotspots, which are underestimated by GEOS-Chem due to averaging over the 2x2.5°horizontal resolution. Over much of the country GEOS-Chem overestimates NO₂ by 10-60%, except in NA and northern Queensland where up to 50% underestimation occurs. The correlation between the bias (GEOS-Chem - OMNO2d) with anthropogenic and soil emissions is shown in the bottom rows. The comparison for January and February of 2005 in Figure 2.7, and winter (JJA) of 2005 in Figure 2.7. The poor correlations for anthropogenic NO suggest that blanket alterations over Australia would not lead to improved NO₂ fit.

This comparison is expanded, including a comparison against modelled emissions, and repeated for autumn (MAM), winter (JJA), and spring (SON) in figures 2.9 to 2.16. These show an analysis of GEOS-Chem NO emissions and their correlations with the bias between GEOS-Chem NO₂ mid-day columns and the OMNO2d product, averaged over each season in 2005. The scatter plots have one datapoint for each land square over Australia.

The correlation between model and satellite NO₂ columns is OK throughout the year over Australia, with some overestimation in the north during non-summer months. There is also slight underestimation over Sydney and Melbourne throughout the year. Figures 2.9 to 2.16 show that the visible biases are not over Australia are not driven by modelled emissions of NO. While the correlation between column NO₂ and emitted NO is clear, emissions do not appear to bias the model in either direction away from the satellite data.

The conclusion drawn is that modelled anthropogenic and soil NO emissions do not show sufficient evidence of biasing GEOS-Chem NO₂ columns away from satellite measurements over Australia. For this reason modelled NO emissions are not scaled in model runs in this thesis.

2.3.7 GEOS-Chem simulations

GEOS-Chem is run four independent times in this thesis, with different outputs from each simulation used to determine specific information.

GC NO vs OMNO2d 20050101-20050228

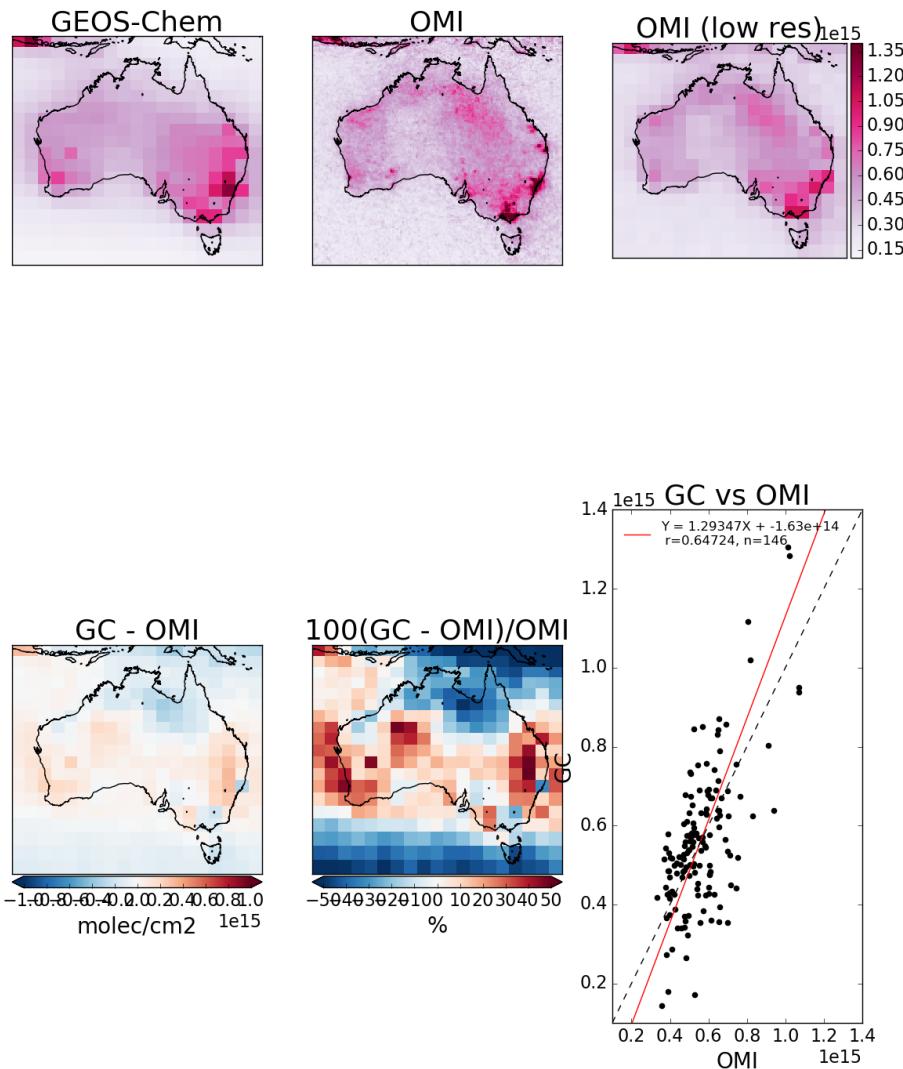


FIGURE 2.7: Row 1 shows the tropospheric columns in molec cm⁻², GEOS-Chem, OMNO2d, and OMNO2d averaged onto the lower resolution of GEOS-Chem from left to right. Row 2 shows the correlations of GEOS-Chem (X axes) between daily anthropogenic emissions, and mid-day OMNO2d columns. Row 3 shows the differences with OMNO2d columns averaged into the lower resolution of GEOS-Chem.

GC NO vs OMNO2d 20050101-20050228

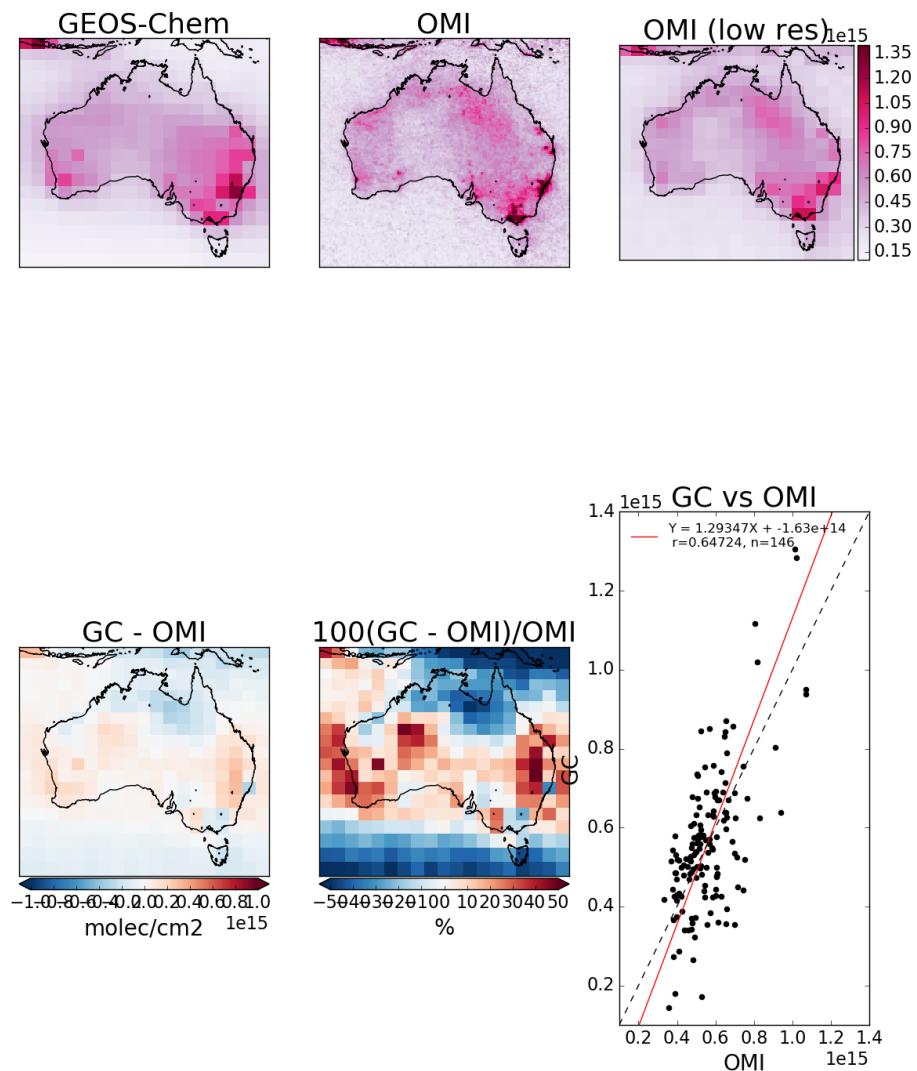


FIGURE 2.8: Row 1 shows the tropospheric columns in molec cm⁻², GEOS-Chem, OMNO2d, and OMNO2d averaged onto the lower resolution of GEOS-Chem from left to right. Row 2 shows the correlations of GEOS-Chem (X axes) between daily anthropogenic emissions, and mid-day OMNO2d columns. Row 3 shows the differences with OMNO2d columns averaged into the lower resolution of GEOS-Chem.

GEOS-Chem vs OMNO2d Jan-Feb 2005

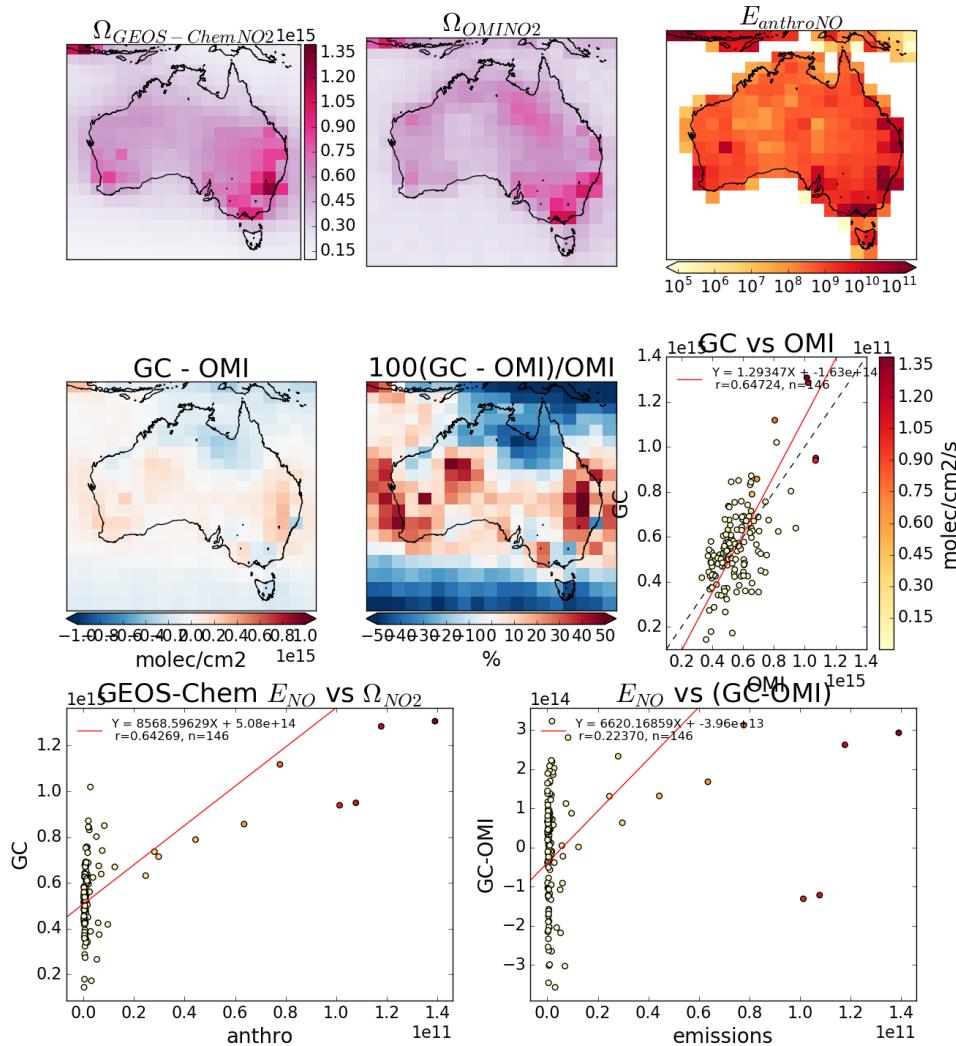


FIGURE 2.9: Top row (left to right): GEOS-Chem NO₂ mid-day tropospheric columns, OMNO2d NO₂ columns, modelled anthropogenic NO emissions. Second row: absolute and relative difference between GEOS-Chem and OMI NO₂ data, and the correlation. Third row: correlation between GEOS-Chem tropospheric column NO₂ and emitted NO, then between the model-satellite bias and the emissions. All correlation plots are coloured by emission rates.

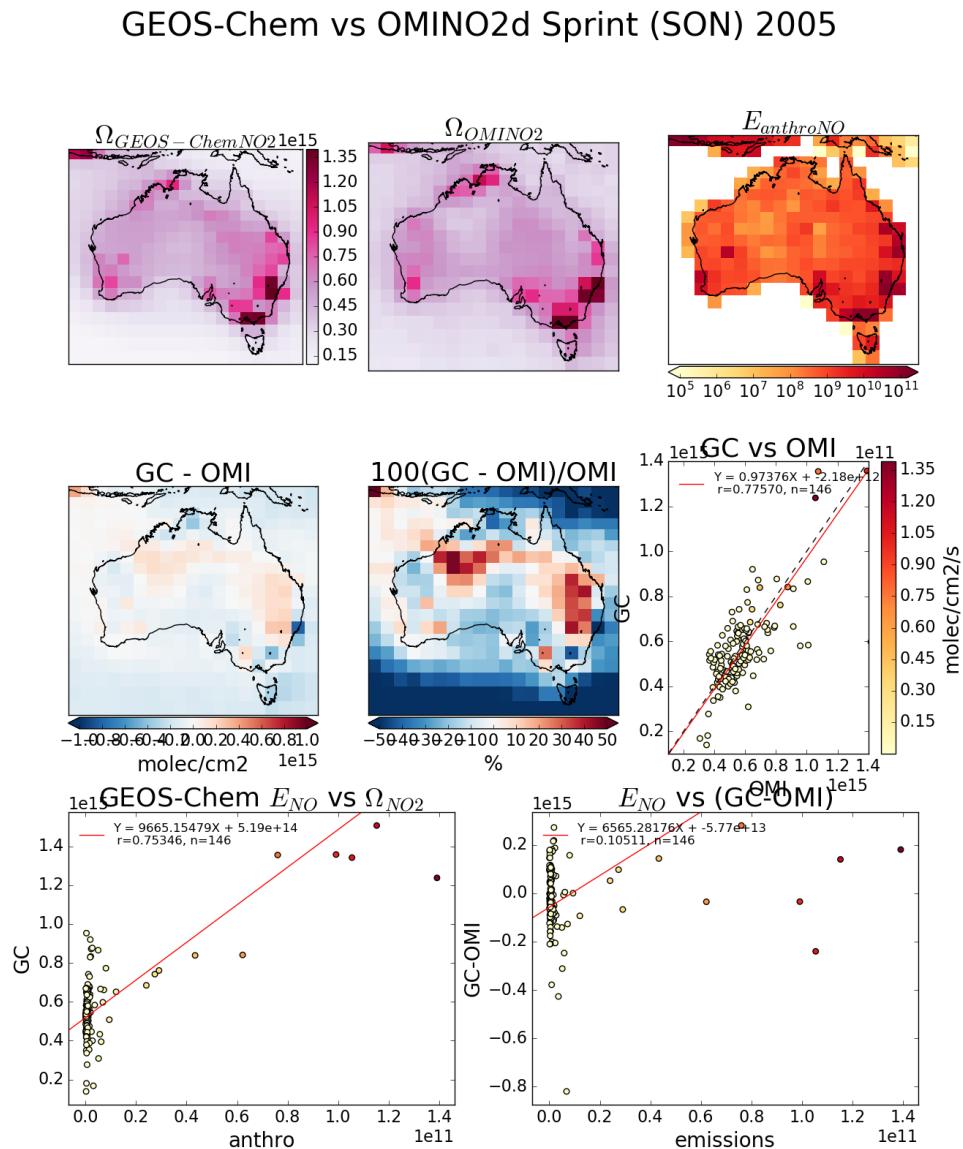


FIGURE 2.10: As figure 2.9, for Autumn 2005.

GEOS-Chem vs OMINO2d Winter (JJA) 2005

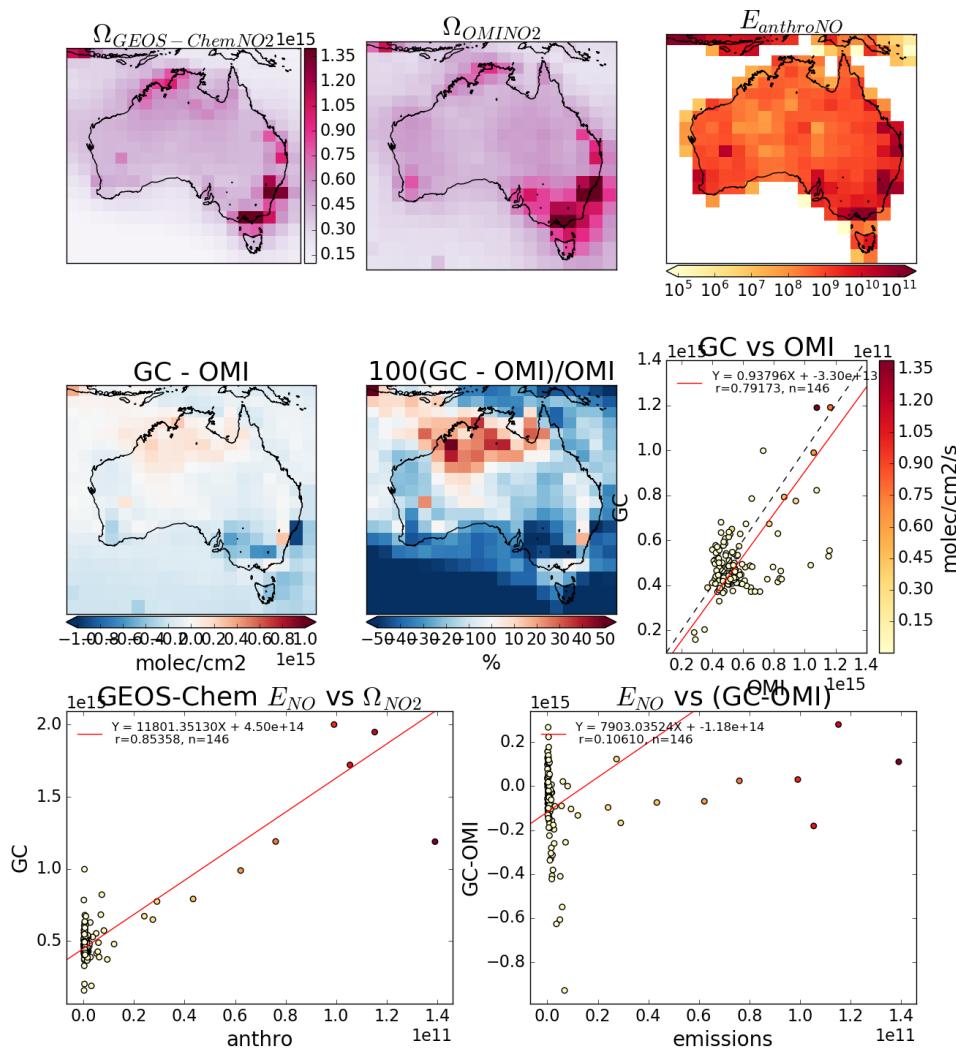


FIGURE 2.11: As figure 2.9, for Winter 2005.

GEOS-Chem vs OMINO2d Autumn (MAM) 2005

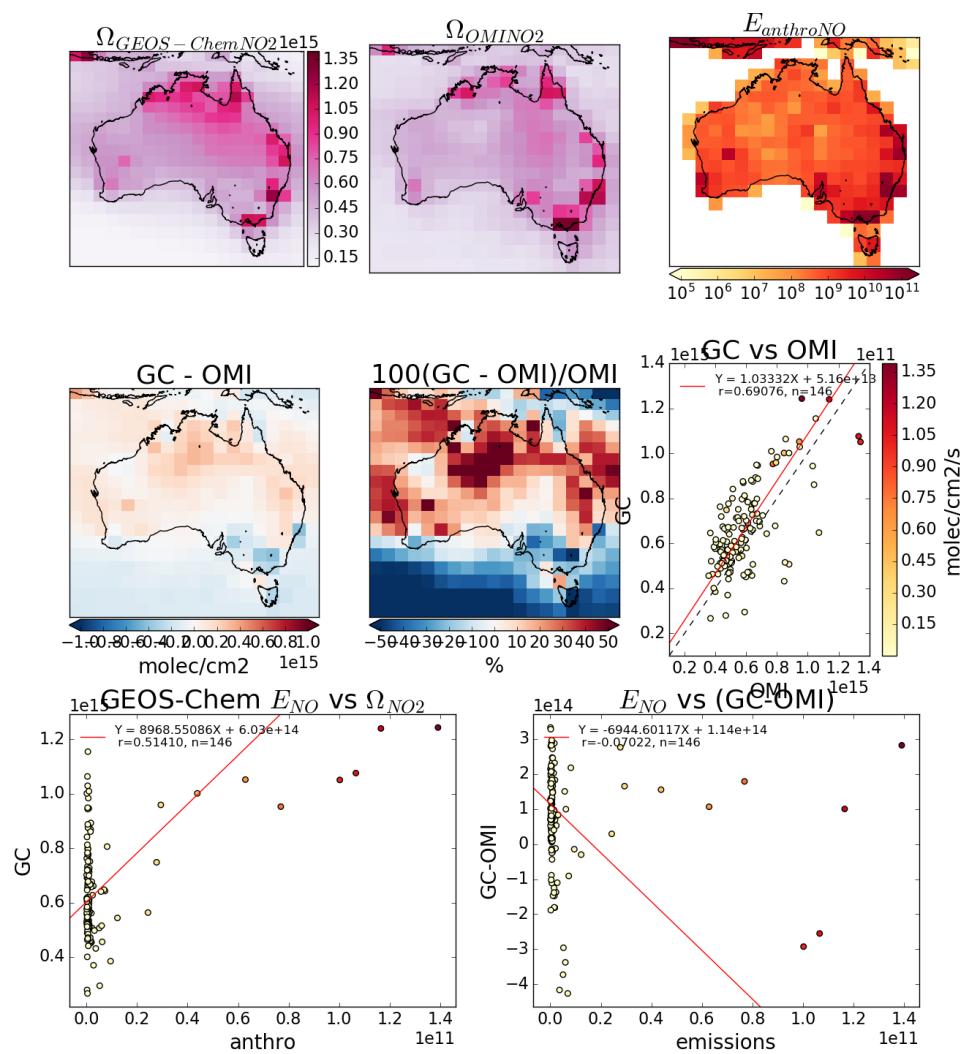


FIGURE 2.12: As figure 2.9, for Spring 2005.

GEOS-Chem vs OMINO2d 20050101-20050228

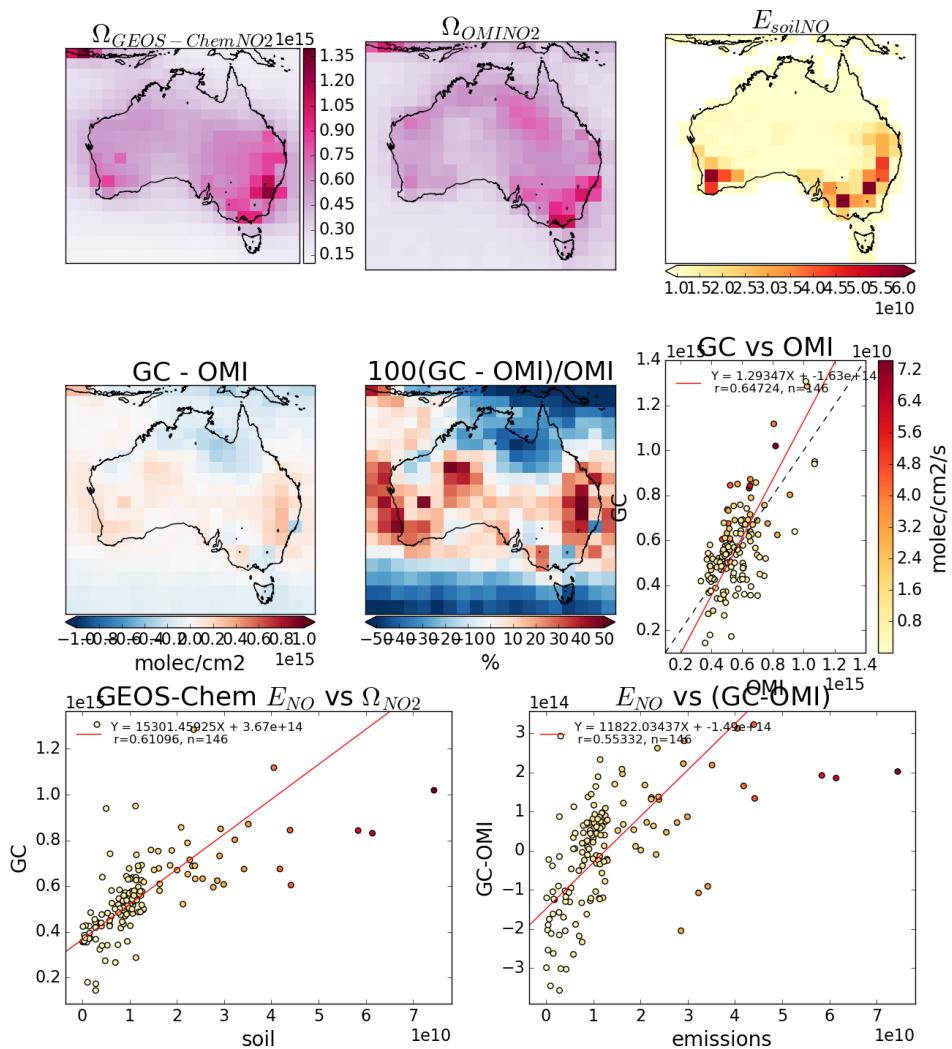


FIGURE 2.13: As figure 2.9, except anthropogenic NO emissions are replaced by soil NO emissions.

GEOS-Chem vs OMINO2d 20050301-20050531

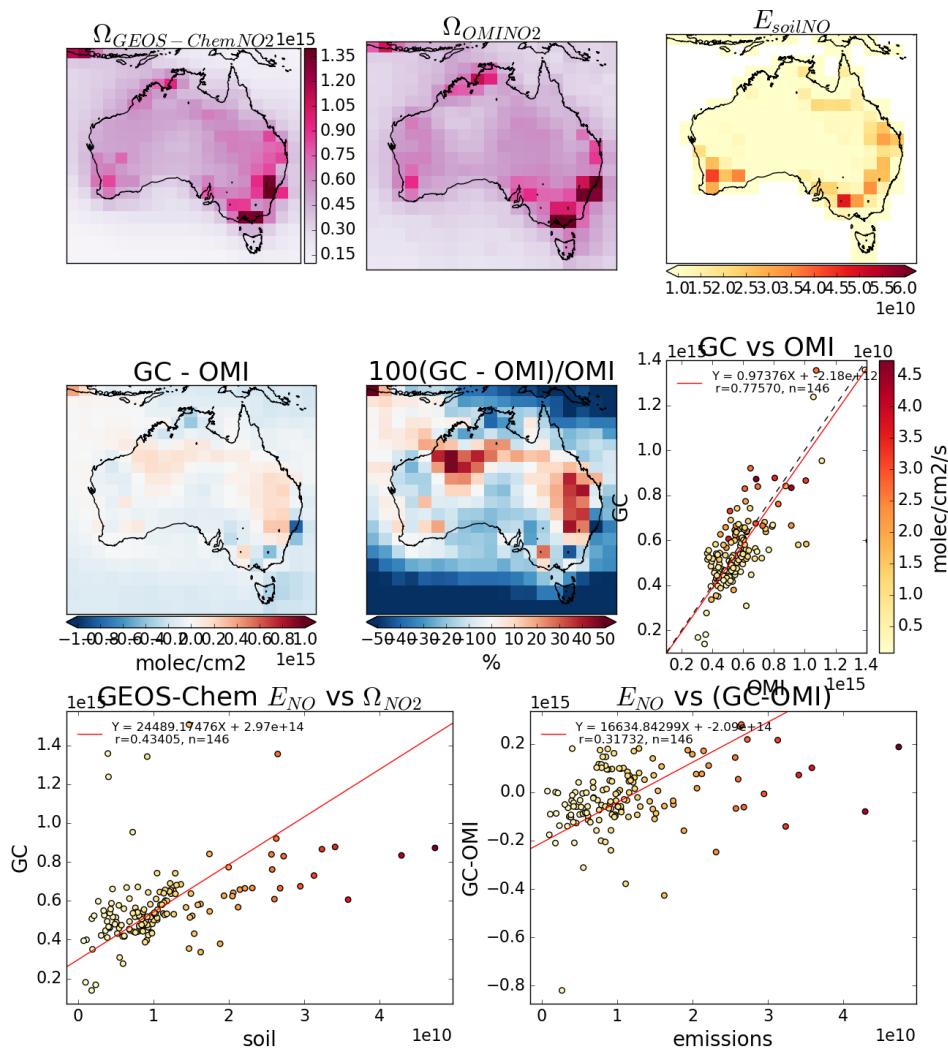


FIGURE 2.14: As figure 2.9, for Autumn 2005, with soil NO emissions replacing anthropogenic NO emissions.

GEOS-Chem vs OMINO2d Winter (JJA) 2005

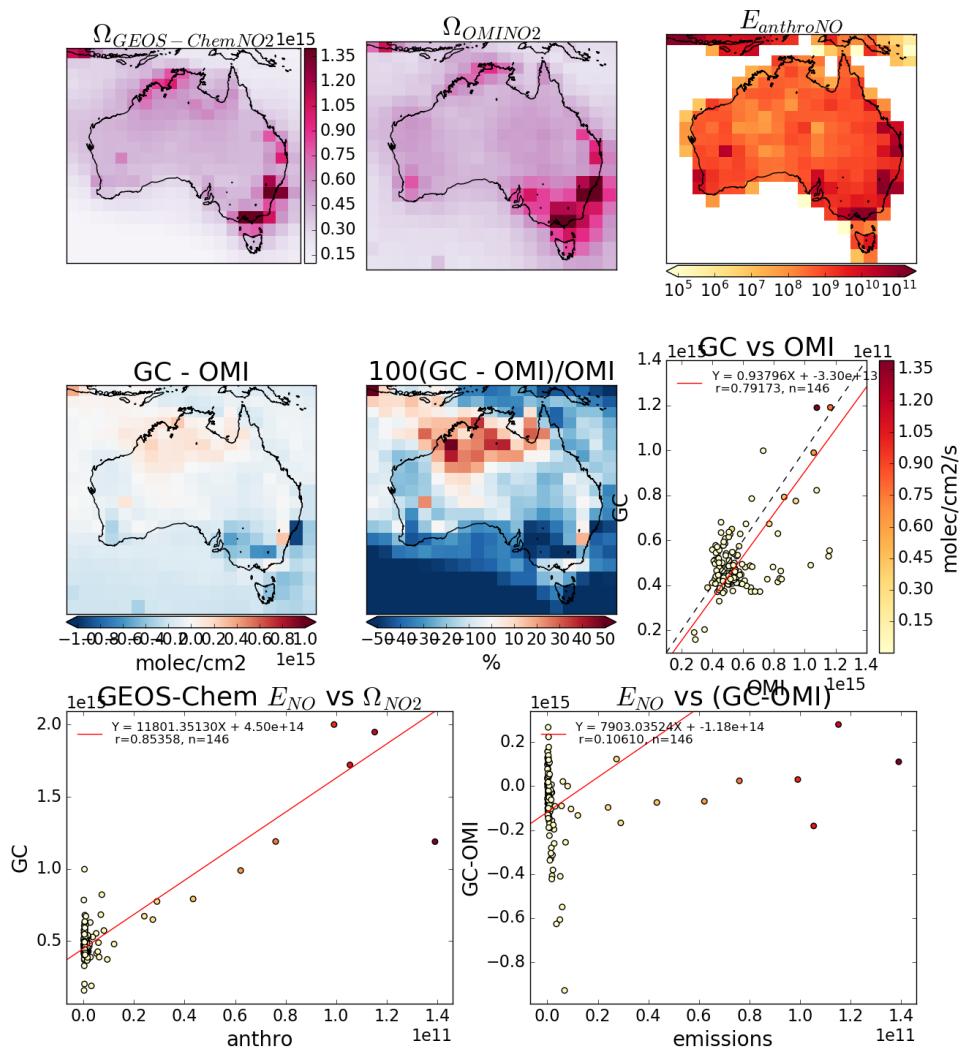


FIGURE 2.15: As figure 2.9, for Winter 2005, with soil NO emissions replacing anthropogenic NO emissions.

GEOS-Chem vs OMINO2d Autumn (MAM) 2005

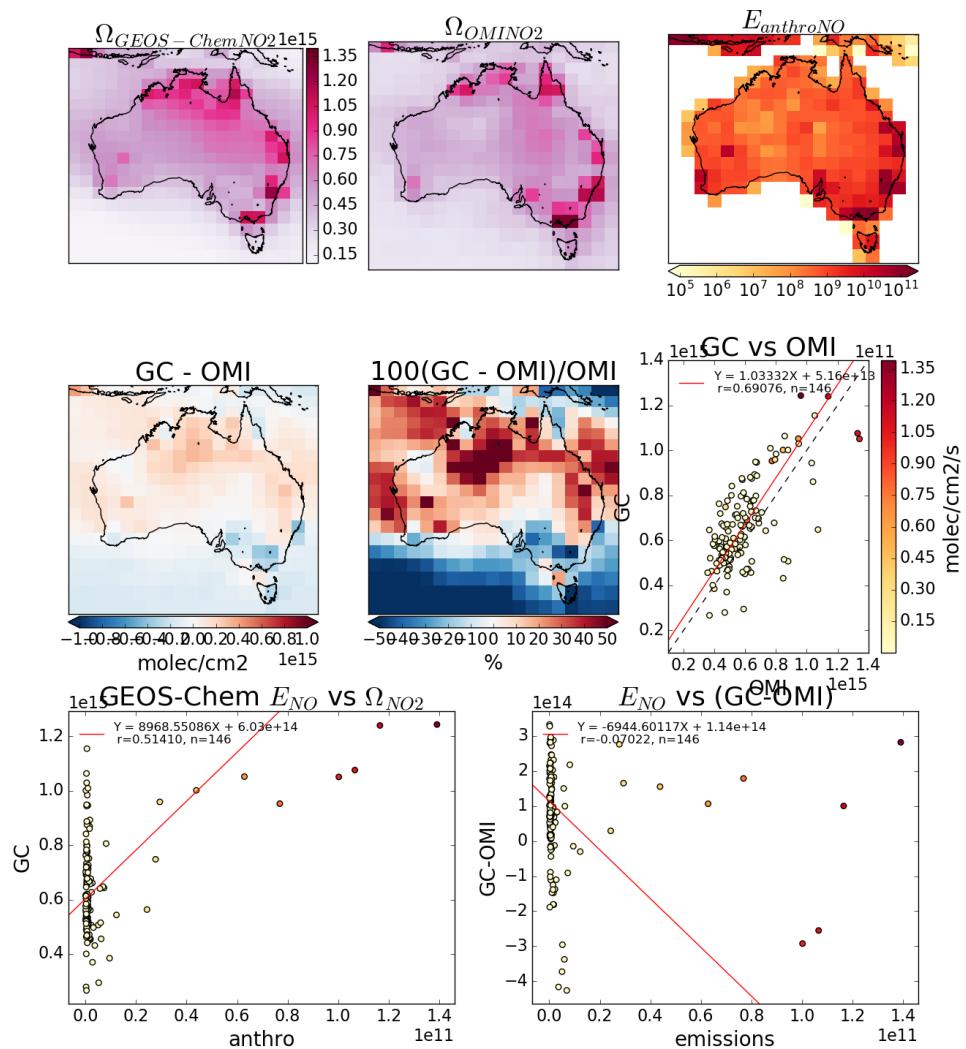


FIGURE 2.16: As figure 2.9, for Spring 2005, with soil NO emissions replacing anthropogenic NO emissions.

Output for an area averaged over 1300 - 1400 local time is saved for comparison and recalculation with satellite overpass records. These averages are used to calculate both the GEOS-Chem based AMF, and the modelled background HCHO over the remote pacific which is used in the reference sector correction for OMI column retrievals (see section 2.5.5). They are also used to determine isoprene to HCHO yield, after removing days with high biomass burning emissions.

TODO: Go through work process and clarify these items Run descriptions and their outputs are listed here, with outputs described in more detail in section :

1. UCX

- (a) Satellite output (1300-1400LT)
- (b) Create shape factors for AMF recalculation in OMI
- (c) This run was initially used to determine if the stratospheric chemistry had much influence over tropospheric isoprene or HCHO concentrations.

2. Tropchem (standard)

- (a) satellite output, daily tracer averages
- (b) Recreate the AMFs for OMI when running code from Dr. Paul Palmer, modified by Dr. Luke Surl.
- (c) Combined with an identical run where isoprene emissions are halved in order to determine smearing.
- (d) TODO: Compare total yearly isoprene emissions before and after new estimate.

3. Tropchem(isoprene emissions halved)

- (a) In GEOS-Chem the isoprene emissions can be globally multiplied by a constant factor.
- (b) An estimate of modelled HCHO sensitivity to isoprene emissions and transport (smearing) can be determined.
- (c) Smearing is determined by running the model with the biogenic isoprene emissions halved while other parameters remain unchanged.

4. Tropchem(biogenic emissions only, all other inventories turned off)

- (a) Satellite output, hourly biogenic emissions from MEGAN
- (b) The output from this simulation is used to determine the biogenic yield from isoprene to HCHO over Australia, described in section 3.2.
- (c) TODO: compared to run with updated emissions

NB: for non-UCX runs, satellite output was modified to include tropopause height

Run comparisons

There are many options available when running GEOS-Chem depending on the desired chemistry, resolution, meteorology, and boundary conditions. Here we compare the model output with and without enabling the Universal tropospheric-stratospheric Chemistry eXtension (UCX). Both runs use 2° latitude by 2.5° longitude, however the UCX mechanism is run with 72 vertical levels from the surface to the top of the atmosphere (TOA ~ 0.1 hPa), while the standard (tropchem) run uses 47 levels. The extra vertical levels are added in the stratosphere, providing finer vertical resolution from around 70 hPa to the top of the atmosphere. For both runs the input parameters such as MEGAN emissions and GEOS-5 meteorological fields are identical.

GEOS-Chem output of HCHO does not differ much between runs with or without the Unified Chemistry eXchange (UCX). Figure 2.17 shows an example of surface HCHO amounts with and without UCX turned on. The differences do not exceed 3% over Australia for the averaged month of January, 2005.

Figure 2.18 shows the differences in surface isoprene amounts over Australia. Here we start to see a higher relative difference in concentrations, although this is generally over the areas with less absolute concentrations. Very little isoprene is seen away from the continent (4-5 orders of magnitude less), due to the short lifetime of isoprene, and lack of emissions over the oceans. Generally isoprene is 0-30% higher over Australia when the UCX mechanism is turned on. This enhancement can be seen throughout the entire tropospheric column as shown by Figure TODO fix ref ??.

Figure TODO: shows the columns for isoprene and HCHO simulated by our two mechanisms over Australia in January of 2005. The differences are minimal compared to other uncertainties in both AMF calculation and emissions estimation.

TODO: The difference in isoprene between UCX and tropchem is likely caused by differences in the modelled radiation reaching the troposphere due to differences in simulated ozone in the stratosphere. With higher stratospheric ozone levels, less radiation would reach the troposphere, changing the photochemistry. Figure 2.19 shows the total column ozone between UCX and non-UCX runs, we can see that UCX has TODO: less or more ozone over Australia/USA in January.

2.3.8 GEOS-Chem outputs

There are various outputs available when running GEOS-Chem, which require understanding in order to compare with observations. GEOS-Chem in this thesis is run with a 15 minute time step for both chemistry and transport, at $2 \times 2.5^{\circ}$ horizontal resolution over 47 vertical levels. Generally output is the average of these time steps either over an entire month, or else per day.

In my work when estimating model yields of isoprene to HCHO, I use daily averaged HCHO columns and compare them to colocated isoprene emissions from MEGAN. Optionally one can save high temporally resolved data for a single (or list of) column(s). I've used this diagnostic to compare modelled ozone with ozonesonde profiles at three sonde release sites discussed in Chapter 3.

Satellite overpass is output from averaging over a window of local time for each grid-box. This output allows comparison with satellite measurements, which overpass at the same local time every day. This diagnostic allows easier analysis of

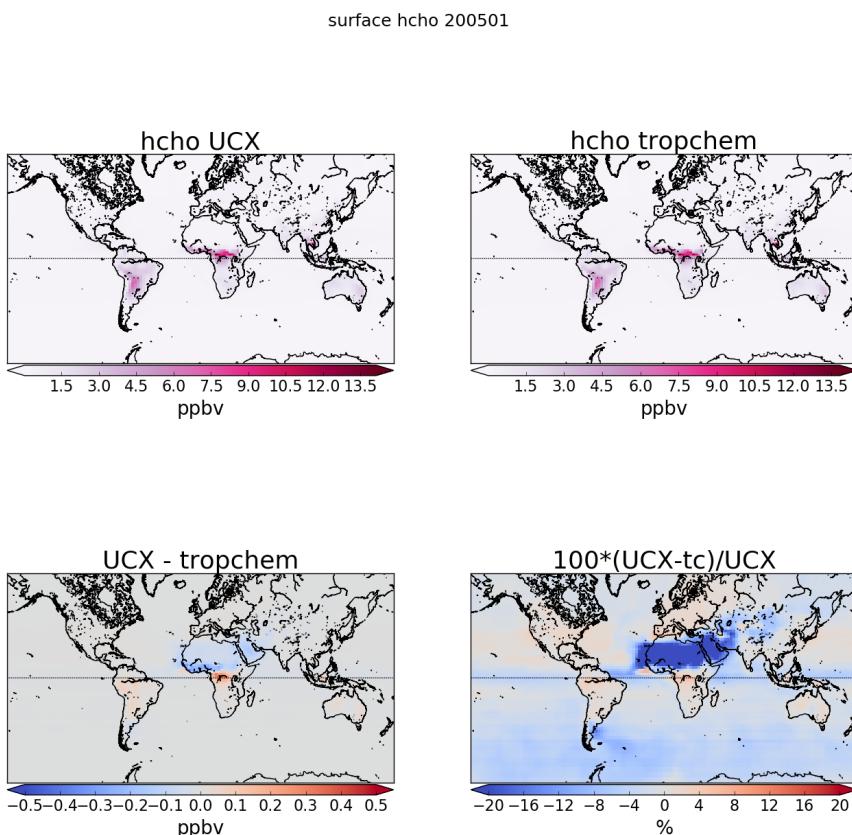


FIGURE 2.17: Surface HCHO simulated by GEOS-Chem with UCX (top left), and without UCX (top right), along with their absolute and relative differences (bottom left, right respectively). Amounts simulated by GEOS-Chem for the 1st of January, 2005.

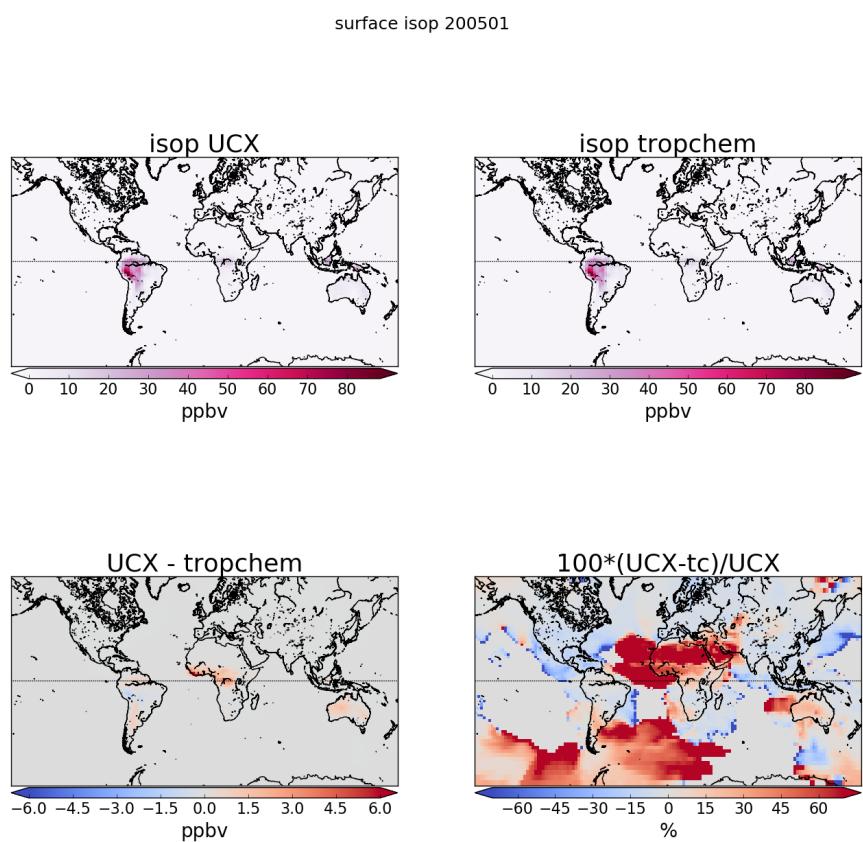


FIGURE 2.18: As figure 2.17, except looking at isoprene.

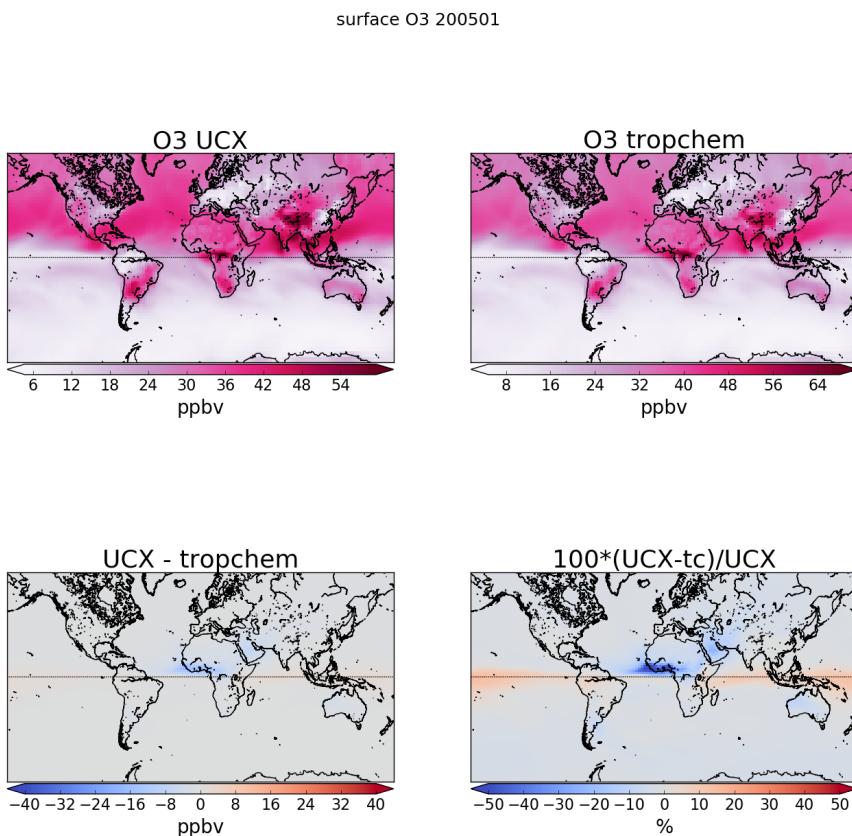


FIGURE 2.19: As figure 2.17, except looking at ozone.

model data against a satellite as one can match the output with the satellite's overpass time. Output averaged between 1300-1400 LT is saved to allow comparisons with Aura satellite measurements, as Aura overpasses at \sim 1330 LT each day. This has been performed by others (eg. Jin et al. 2017).

HEMCO diagnostics In order to get hourly MEGAN modelled isoprene emissions, HEMCO (the module of GEOS-Chem dealing with emissions inventories) diagnostic output was created. When working with globally gridded data, handling local time offsets becomes more important. The hourly output emissions of isoprene is saved using GMT, which needs to be offset based on longitude in order to retrieve local time. I do this by setting up a latitude by longitude array which matches the horizontal resolution of the data, filling each gridbox with its local time offset. This offset is determined as one hour per 15 degrees (since 360 degrees is 24 hours), and then used to retrieve global data at any specific local time. The retrieval of a daily local time global array is done by index matching the GMT+LT (modulo 24) with the desired hour on this grid over the 24 GMT hours. are the emissions TODO: averaged or instantaneous? in each gridbox, which I've stored for each 3 hours.

Tracer averages are daily or monthly averaged gridbox concentrations.

2.4 Measurement Techniques

While I have not made any measurements myself, it is important to understand the techniques used in datasets I have utilised in order to understand possible anomalous datapoints or trends. In-situ measurements contain errors, and depending on the device used and chemical being measured this error can be significant. The major sources of uncertainty in measurement techniques included interference from non-target compounds and under-reporting (eg. Dunne et al. 2017). Overall isoprene uncertainty in measurements analysed by Dunne et al. (2017) was a factor of 1.5 to 2. This can feed into uncertainties in modelling and satellite retrievals, as verification and correlations are affected.

2.4.1 DOAS

The DOAS technique uses solar radiation absorption spectra to measure trace gases through paths of light. Beer's law states that $T = I/I_0 = e^{-\tau}$ with T being transmittance, τ being optical depth, and I , I_0 being radiant flux received at instrument and emitted at source respectively. From $\tau_i = \int \rho_i \beta_i ds$ we get:

$$I = I_0 \exp \left(\sum_i \int \rho_i \beta_i ds \right)$$

Where i represents a chemical species index, ρ is a species density(molecules per cm^3), β is the scattering and absorption cross section area (cm^2), and the integral over ds represents integration over the path from light source to instrument. τ can be described using the attenuation cross section (the attenuation coefficient divided by its number

density), with the following relation:

$$\tau = \int_0^l \alpha(z) \eta(z) dz$$

where $\alpha(z)$ and $\eta(z)$ represent absorption cross section in $m^2 \text{ molecule}^{-1}$, and number density in molecules m^{-3} respectively, and l represents the length of the medium (in this case the path through which the light travels).

Another way of describing optical depth, also called optical thickness, is the natural logarithm of the ratio of incident radiant power to transmitted radiant power through a material. In the atmosphere we are interested in the optical depth of various chemical species, and we use incoming solar radiation to determine this. The difference between solar radiation at the top of the atmosphere and the Earth's surface defines the atmospheric optical depth along the path of observation.

$$\tau = \ln \frac{\phi_e^i}{\phi_e^t}$$

where ϕ_e^i is radiant flux seen at the earth surface, ϕ_e^t is the solar radiant flux which arrives at the top of the atmosphere. In the atmosphere, optical depth can be due to several factors including scattering, chemical absorbance, and aerosols.

2.4.2 Satellites

In order to detect trace gases such as HCHO, satellites use a DOAS based technique to detect concentrations along the path of light which reaches the satellite instrument. This requires chemical transport and radiative transfer models used to transform the non-vertical light path into vertical column amounts. Measurements done using DOAS often apply a forward radiative transfer model (RTM) such as LIDORT (see section ??) in order to determine a trace gas's radiative properties at various altitudes. The forward RTM used for satellite data products also involves functions representing extinction from Mie and Rayleigh scattering, and the efficiency of these on intensities from the trace gas under inspection, as well as accounting for (often estimated) atmospheric parameters such as albedo.

Rayleigh and Mie scattering describe two kinds of particle effects on radiation passing through a medium. Rayleigh scattering is heavily wavelength dependent, and is the dominant form of scattering from particles up to roughly one tenth of the wavelength of the light. Mie scattering is more dominant from larger particles, and has less wavelength dependence. The effects of scattering are what gives us the information about substances in the atmosphere. The different particles and gases in the air have various properties which affect remote sensing devices such as a satellite, making them more or less sensitive at certain altitudes for detecting various species (Martin et al. 2002b, e.g.).

Satellites record near nadir (vertical) reflected spectra between around 250-700 nm split into spectral components at around 0.3 nm in order to calculate trace gases including O₃, NO₂, and HCHO (eg: Leue et al. (2001)). Several public data servers

are available which include products from satellites, including NASA's Earthdata portal (<https://earthdata.nasa.gov/>) and the Belgian Institute for Space Aeronomy (IASB-BIRA) Aeronomie site (<http://h2co.aeronomie.be/>).

Satellite measurements are generally performed using spectral fitting followed by conversion to vertical column densities. Instruments including MODIS on board the AQUA and TERRA satellites are also able to determine aerosol optical depth (AOD), a measure of atmospheric scatter and absorbance. An AOD of under 0.05 indicates a clear sky, while values of 1 or greater indicate increasingly hazy conditions. This is important in order to determine where measurements from other instruments may be compromised by high interference. Satellite measured AOD requires validation by more accurate ground based instruments like those of AERONET which uses more than 200 sun photometers scattered globally.

Soon even more HCHO data will be available in the form of geostationary satellite measurements (Kwon et al. (2017)). Kwon et al. (2017) examine simulated geostationary measurements against GEOS-Chem column simulations to determine the most important instrument sensitivities. Geostationary satellites can provide temporally rich measurements over an area, as they are not sweeping around the earth but fixed relative to one latitude and longitude.

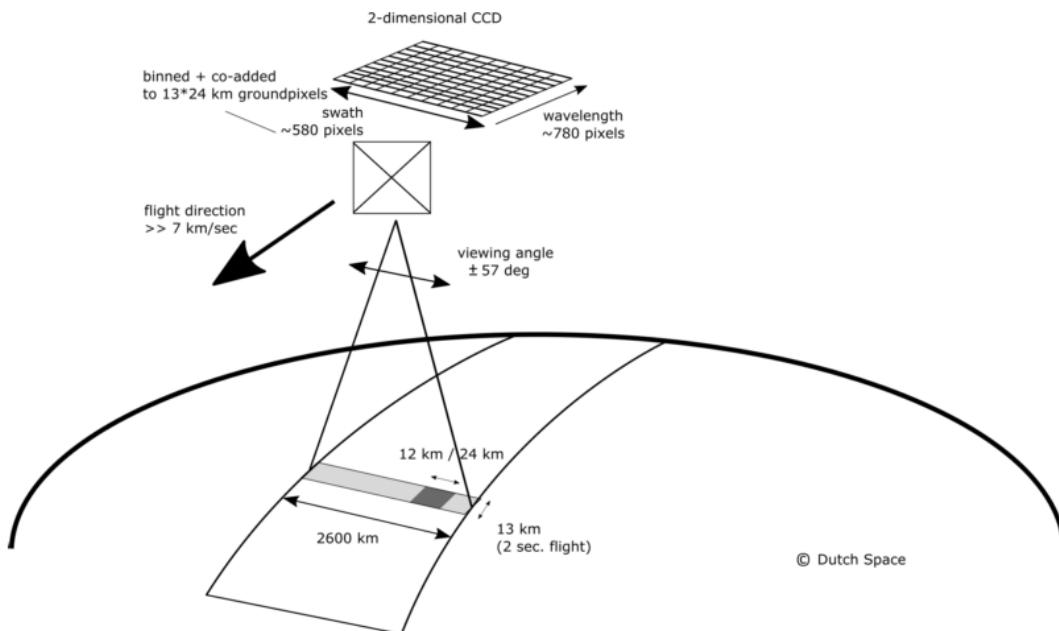
LIDORT

??

LIDORT is a model of Linearized Discrete Ordinate Radiative Transfer, used to determine backscatter intensities and weighting functions at arbitrary elevation angles (Spurr2001). The model solves radiative transfer equations and can be used to determine various atmospheric column measurement attributes such as optical depth, ring effects, and scattering. These radiative properties (or at least estimates thereof) are required when measuring trace gases in the atmosphere through a long path such as seen by satellites (eg. Palmer et al. 2001; Martin et al. 2002a; De Smedt et al. 2015; Gonzalez Abad et al. 2015).

OMI

The OMI instrument on board AURA has been active since July 2005, it records spectra from 264-504 nm using an array of 60 detectors with mid-resolution (0.4-0.6 nm). This band of wavelengths allows measurements of trace gases including O₃, NO₂, SO₂, HCHO, and various other quantities like surface UV radiation. Recently Schenkeveld et al. (2017) analysed the performance over time of the instrument and found irradiance degradation of 3-8%, changed radiances of 1-2%, and a stable wavelength calibration within 0.005-0.020 nm. They also provide a very nice summary of the OMI instrument copied here in Fig. 2.20, as it shows the instrument's spectral, temporal, and spatial resolutions. These changes are measured excluding the row anomaly (RA) effect, which is relatively stable since 2011, although it is still growing and remains the most serious concern. An analysis of the row anomaly by Huang et al. (2017) state that OMI ozone columns remain suitable for scientific use, with recommendation for further evaluation. And analysis of OMI output by Schenkeveld et al. (2017) concludes that data is still of high quality and will deliver useful information for 5-10 more years,



Channel	Wavelength range	Spectral resolution	Spectral sampling	Ground pixel size
UV1	264–311 nm	0.63 nm = 1.9 px	0.33 nm px ⁻¹	13 × 48 km
UV2	307–383 nm	0.42 nm = 3.0 px	0.14 nm px ⁻¹	13 × 24 km
VIS	349–504 nm	0.63 nm = 3.0 px	0.21 nm px ⁻¹	13 × 24 km

FIGURE 2.20: Figure 1 and Table 1 from Schenkeveld et al. (2017), with the following caption “An impression of OMI flying over the Earth. The spectrum of a ground pixel is projected on the wavelength dimension of the charge-coupled device (CCD; the columns). The cross-track ground pixels are projected on the swath dimension of the CCD (the rows). The forward speed of 7 kms^{-1} and an exposure time of 2 s lead to a ground pixel size of 13 km in the flight direction. The viewing angle of 114° leads to a swath width on the ground of 2600 km.” The table shows the optical properties for OMIs three channels.

with radiances only changing by 1 – 2% outside of RA impacted areas. The RA began in June 2007, with some cross-track rows seemingly blocked. The most likely cause is some instrument insulation partially obscuring the radiance port (Schenkeveld et al. (2017)).

Air mass factor (AMF)

An AMF characterises measurement sensitivity to a trace gas at various altitudes Palmer et al. 2001, e.g. Lorente et al. (2017) show that AMF calculations can be the largest source of uncertainty in satellite measurements. Another way of describing AMFs are as measures of how radiance at the top of the atmosphere (TOA) changes with trace gas optical depths at specific altitudes (Lorente et al. (2017)). Calculation

of the AMF is important as it is multiplied against the estimated slant columns in order to give vertical column amounts. To convert the trace gas profile from a reflected solar radiance column (slanted along the light path) into a purely vertical column requires calculations of an air mass factor (AMF). In satellite data, the AMF is typically a scalar value for each horizontal grid point which will equal the ratio of the total vertical column density to the total slant column density. This value requires calculations to account for instrument sensitivities to various wavelengths over resolved altitudes, and is unique for each trace gas under consideration.

DOAS retrieval columns are an integration of a trace gas over the instruments viewing path, in order to convert this total to a vertically distributed column a few assumptions and estimates are required. The vertical profile of a trace gas is assumed or estimated via a CTM, while its' scattering and radiative properties are calculated at all altitudes using an RTM. These properties are combined into a single array called the AMF. Two examples of this are GOME-2 products on the MetOp-A satellite (http://atmos.caf.dlr.de/gome/product_hcho.html) and OMI products which use IMAGESv2 combined with LIDORT and GEOS-Chem with LIDORT for product processing respectively (Instrument 2002; Gonzalez Abad et al. 2015). AMFs are unique to each trace gas and due to their complexity and the influence of cloud cover they remain one of the larger error sources in remote sensing of BVOCs (Palmer et al. 2001; Millet et al. 2006)).

Related to the AMF is the averaging kernal (AK), which is used to handle instrument measurements which are sensitive to concentrations at different altitudes in the atmosphere. DOAS methods can be heavily influenced by the initial estimates of a trace gas profile (the a priori) which is often produced by modelling, so when comparing models of these trace gases to satellite measurements extra care needs to be taken to avoid introducing bias from differing a priori assumptions. One way to remove these a priori influences is through the satellites AK (or by using AMFs), which takes into account the vertical profile of the modelled trace gas and instrument sensitivity to the trace gas (Eskes and Boersma (2003) and Palmer et al. (2001)). Lamsal et al. (2014) recommends that when comparing satellite data to models, the AMF should first be recalculated using the model as an a priori. This is in order to remove any a priori bias between model and satellite columns. Another way of removing this bias is through deconvolution ($\Omega = AK \times VC_{satellite} + \times(I - AK)VC_{apriori}$) of the averaging kernal (AK) of the satellite instrument. The AK represents sensitivities to each species at multiple altitudes through the atmosphere and in the case of OMI, can be approximated from the scattering weights ($\omega(z)$) function as follows:

$$AK(z) = \frac{\omega(z)}{AMF} \quad (2.2)$$

Note that this is an approximation for the OMI product, which does not include the AK but does include the ω and AMF, as explained in Gonzalez Abad et al. (2015).

The latest OMI algorithm uses a shape factor determined from GEOS-Chem using 47 vertical levels at monthly temporal resolution and 2° latitude by 2.5° longitude horizontal resolution (Gonzalez Abad et al. 2015). The GEOS-Chem model has been substantially updated since then, and using the more recent version V10.01 to recalculate the AMF is performed within this thesis, details are shown in section 2.5.

Uncertainties

While satellite data is effective at covering huge areas (the entire earth) it only exists at a particular time of day, is subject to cloud cover, and generally does not have fine horizontal or vertical resolution. Concentrations retrieved by satellites have large uncertainties, which arise in the process of transforming spectra to total column measurements, as well as instrument degradation (satellite instruments are hard to tinker with once they are launched). Uncertainty in transforming satellite spectra comes from a range of things, including measurement difficulties introduced by clouds, and instrument sensitivity to particular aerosols (Millet et al. 2006). Many products require analysis of cloud and aerosol properties in order to estimate concentration or total column amounts (Palmer et al. 2001; Palmer 2003; Marais et al. 2012; Vasilkov et al. 2017). The main source of error in satellite retrievals of HCHO are due to instrument detection sensitivities, and the vertical multiplication factor (Millet et al. 2006). Calculations of the AMF performed by different groups tend to agree fairly well, as long as all the a priori and ancillary data is similar. Large differences can occur depending on the a priori vertical profile, trace gas concentrations, and cloud properties (Lorent2017). Choice of RTM and interpolation operations have a relatively small affect compared to the assumed state of the atmosphere, with high structural uncertainty introduced at this stage of AMF calculation - as shown in Lorent2017

There are two types of measurement error, arguably the worst of these is systematic error (or bias) which normally indicates a problem in calculation or instrumentation. If the systematic error is known, it can be corrected for by either offsetting data in the opposite direction, or else fixing the cause. A proper fix can only be performed if the sources of error are known and there is a way of correcting or bypassing it. Random error is the other type (often reported as some function of a datasets variance, or uncertainty), and this can be reduced through averaging either spatially or temporally. By taking the average of several measurements, any random error can be reduced by a factor of one over the square root of the number of measurements. This is done frequently for satellite measurements of trace gases (which are often near to the detection limit over much of the globe). For example: Vigouroux et al. (2009) reduce the measurement uncertainty (in SCIAMACHY HCHO columns) by at least a factor of 4 through averaging daily over roughly 500km around Saint-Denis, and only using days with at least 20 good measurements.

Satellite measurements of HCHO are relatively uncertain, however this can be improved by averaging over larger grid boxes or longer time scales. An example of this can be seen in Dufour et al. (2008), where monthly averaging is used to decrease the measurements uncertainty. The finer nadir resolution of OMI (13 by 24 km²) compared to other satellites reduces cloud influence (Millet et al. 2006; Millet et al. 2008). Although the uncertainty in each pixel is $\sim 2 \times 10^{16}$, which is 5 \times higher than GOME, there are $\sim 100 - 200 \times$ as many measurements due to the smaller footprint and better temporal resolution of OMI, which allows a greater reduction of uncertainty with averaging (Instrument 2002; Millet et al. 2008). Uncertainty in a single pixel for OMI is roughly the same magnitude as HCHO background levels. The top row in figure 2.21 shows OMI HCHO columns binned to at 0.25°longitude by 0.3125°latitude averaged over one day, one month, and one month after filtering. Row two shows uncertainty

OMI HCHO and uncertainty for 200501

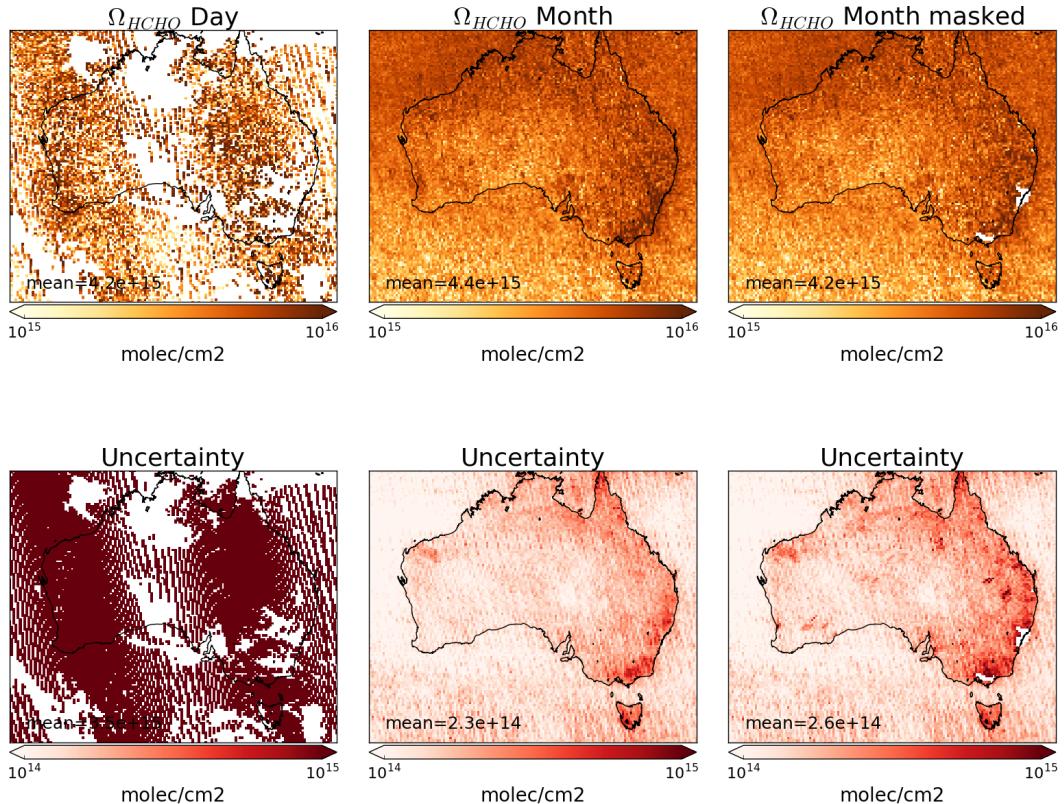


FIGURE 2.21

of the satellite data after averaging. It's clear that one day of satellite data is too uncertain when binned at $0.25 \times 0.3125^\circ$ horizontal resolution, however after a month with or without filtering the uncertainties become manageable. If we assume the uncertainty is random error, and not bias introduced through calculation techniques, then we are able to reduce the uncertainty through averaging. Random error can be reduced by temporal and/or spatial averaging, decreasing uncertainty by a factor of $1/\sqrt{N}$ where N is the number of observations being averaged. High resolution low detection limit estimates can be built up using "oversampling", which averages satellite measurements over time (eg. Zhu et al. 2014).

In cloudy, hazy or polluted areas measurements are more difficult to analyse (e.g. Palmer 2003; Marais et al. 2014). Recent work by Vasilkov et al. (2017) showed that updating how the surface reflectivity is incorporated into satellite measurements can change the retrievals by 50 % in polluted areas. With the satellite HCHO columns, we filter cloud fractions over 40%, which introduces a clear-sky bias. This bias has been measured as a 13% positive monthly mean bias by Palmer et al. (2001) and Surl, Palmer, and González Abad (2018).

In satellite HCHO products, concentrations over the remote pacific ocean are sometimes used to analyse faulty instrument readings. This is due to the expected invariance of HCHO over this region. For instance GOME (an instrument which measures trace gases on board the ERS-2) corrects for an instrument artifact using modelled HCHO over the remote pacific (Shim et al. 2005). OMI HCHO products use a similar technique to account for sensor plate drift and changing bromine sensitivity (Gonzalez Abad et al. 2015). Uncertainty in the OMI satellite instrument is calculated by the Smithsonian Astrophysical Observatory (SAO) group using the uncertainty in backscattered radiation retrievals (Gonzalez Abad et al. 2015; Abad et al. 2016). Another method of calculating the uncertainty is used by the Belgian Institute for Space Aeronomy (BIRA) group, who determine uncertainty from the standard deviation of HCHO over the remote pacific ocean (De Smedt et al. 2012; De Smedt et al. 2015).

For many places the tropospheric column HCHO measured by satellite is biased low, Zhu et al. (2016) examine six available datasets and show a bias of 20 - 51% over south east USA when compared against a campaign of aircraft observations (SEAC⁴RS). De Smedt et al. (2015) also found OMI and GOME2 observations were 20 - 40% lower than ground based vertical profiles, and Barkley et al. (2013) determine OMI to be 37% low compared with aircraft measurements over Guyana. These bias can be corrected by improving the assumed a priori HCHO profiles which are used to calculate the AMFs of the satellite columns. Millet et al. (2006) examine OMI HCHO columns over North America and determine overall uncertainty to be 40%, with most of this coming from cloud interference. Millet et al. (2008) shows that there also exists some latitude based bias, as well as a systematic offset between the OMI and GOME instruments. This does not appear to be due to the different overpass times of the two instruments.

A full analysis of the AMF uncertainty in OMI measurements, as well as the structural uncertainty (between different systems of calculations applied to the same data) is performed by Lorente et al. (2017). They determine the structural uncertainty on seven AMF calculation approaches used by different retrieval groups. They show that in scenarios where the gas is enhanced in the lower troposphere, AMF calculation is the largest uncertainty in satellite measurements. In polluted environments the structural uncertainty is estimated at 42 %, or 31 % over unpolluted environments. The importance of a priori and ancillary data (such as surface albedo and cloud top height) is also shown, as it sharply affects the structural uncertainty.

2.5 Recalculation of OMI HCHO

The AMF is needed to transform the slant column (SC) viewed by the satellite into a vertical column (Ω):

$$AMF = \frac{SC}{\Omega} \quad (2.3)$$

A slant or vertical column is expressed in molecules cm⁻². OMI HCHO vertical columns are calculated using modelled a priori HCHO profiles (see section 2.2.4). When comparing satellite measurements against models it is important to recognise the impact of this a priori on the total column values. This is due to how the sensitivity of OMI (to HCHO and other trace gases) varies vertically throughout the atmosphere.

When comparing OMI vertical columns (Ω_O) to GEOS-Chem (Ω_G), the satellite AMF needs to be recalculated using GEOS-Chem modelled vertical gas profiles as the a prioris. Without performing this step a bias between modelled and measured total column values may be due to the a priori rather than chemistry or measurements (Palmer et al. 2001; Lamsal et al. 2014).

Here, two new AMFs are calculated, both using GEOS-Chem HCHO profiles as the new a priori. The first AMF (AMF_{GC}) uses the original satellite scattering weights while recalculating the shape factor, whereas the second AMF (AMF_{PP}) also recalculates scattering weights. AMF_{PP} is created using code initially written by Professor P. Palmer (see sections 2.5.3 and 2.5.4 for more details). A reference sector correction is determined using the method described in Abad et al. (2016), (see section 2.5.5). This correction is unique for each of the 60 *measurement tracks* used by OMI. Finally the correction is applied to each pixel to create the corrected vertical column. The end product is three sets of corrected vertical columns (VCC): the original (Ω_{OC}), one using GEOS-Chem shape factors (Ω_{GC}), and one from Palmer's code (Ω_{PC}).

2.5.1 Outline

Here is an outline in computational order of what takes place when recalculating the Ω from OMI.

1. GEOS-Chem satellite overpass output (see section 2.3.8) is used to create new shape factors (S_z and S_σ).
 - (a) Pressure edges and geometric midpoints are determined, along with altitudes (z), and box heights (H).
 - (b) Number density and mixing ratio of HCHO (n_{HCHO} , C_{HCHO} respectively) are taken or created from model outputs HCHO(ppb) and air density (molec/cm³)
 - (c) Total column HCHO from GEOS-Chem (Ω_G) is calculated $\Omega_G = \sum_z (n_{\text{HCHO}}(z) \times H(z))$, along with total column air (Ω_A , calculated similarly)
 - (d) The shape factor $S_z(z)$ is calculated on each altitude $S_z(z) = n_{\text{HCHO}}/\Omega_{\text{HCHO}}$.
 - (e) Pressures (p) are used to create sigma coordinates $\sigma(z) = (p(z) - p_{TOA})/(p(0) - p_{TOA})$
 - (f) $S_\sigma(z)$ is calculated on each altitude: $S_\sigma(z) = C_{\text{HCHO}}(z) \times \Omega_A/\Omega_{\text{HCHO}}$
2. Satellite pixels (SC, scattering weights ($\omega(z)$)), pressure levels, latitude and longitude) are read from the OMHCHO dataset.
3. For each pixel, a new AMF (AMF_{GC}) is created using the GEOS-Chem shape factors and satellite scattering weights.
 - (a) scattering weights (ω) are interpolated onto the same vertical dimensions (z and σ) as the shape factors.

- (b) Integration (approximated using rectangular method) is performed along the vertical dimension to calculate the new AMF on both coordinate systems.

$$AMF_z = \Sigma_z (\omega(z) \times S_z(z) \times H(z)) \quad (2.4)$$

$$AMF_s = \Sigma_\sigma (\omega(\sigma) \times S_\sigma(\sigma) \times d\sigma) \quad (2.5)$$

4. These two AMFs represent the same thing using different coordinates. Either one can be used as the AMF_{GC} .
5. The AMF_{PP} (created separately) and AMF_{GC} are used to determine the new vertical columns (Ω_P, Ω_G respectively): $\Omega = SC / AMF$.
6. A reference sector correction (RSC) is defined each day using these AMFs along with modelled HCHO over the remote pacific.
 - (a) GEOS-Chem satellite overpass output ($\Omega_{GEOS-Chem}$ from 140°W to 160°W) are averaged monthly and longitudinally to provide a modelled reference sector $\Omega_0[lat]$.
 - (b) The modelled reference slant columns (MRSC) are calculated using $MRSC = \Omega_0 \times AMF$ for each AMF.
 - (c) For each satellite pixel between 140°W and 160°W , the correction is calculated as the measured SC minus the MRSC at the nearest latitude.

$$corr[lat, track] = SC[lat, track] - MRSC[lat]$$

- (d) These corrections are binned by satellite detector (track: 1-60), and latitude (0.36° ; 500 latitudes from 90°S to 90°N).
- (e) The median entry of each bin is determined and this forms the $RSC[lat, track]$
7. VCC are determined using $VCC = (SC - RSC[lat, track]) / AMF$ for each measured SC and using each AMF, with the RSC linearly interpolated to the latitude of the satellite pixel.
8. The VCC (along with most of the pixel and GEOS-Chem data) are binned onto a 0.25° by 0.3125° grid along with how many pixels have been binned, and the average pixel uncertainty in product OMHCHORP.

Figure ?? shows an overview of how these profiles are used in this thesis. Output from GEOS-Chem is combined with OMHCHO swath data to produce a gridded HCHO file which contains HCHO vertical columns recalculated with GEOS-Chem a prioris. The output keeps the original AMF as well as those recalculated using GEOS-Chem (AMF_{OMI} , AMF_{GC} , and AMF_{PP}). Additionally, MOD14A1, OMAERUVd and OMNO2d data are used to create masks which are also stored in the OMHCHORP dataset. The creation of fire, smoke, and anthropogenic influence masks is described in section 2.6.

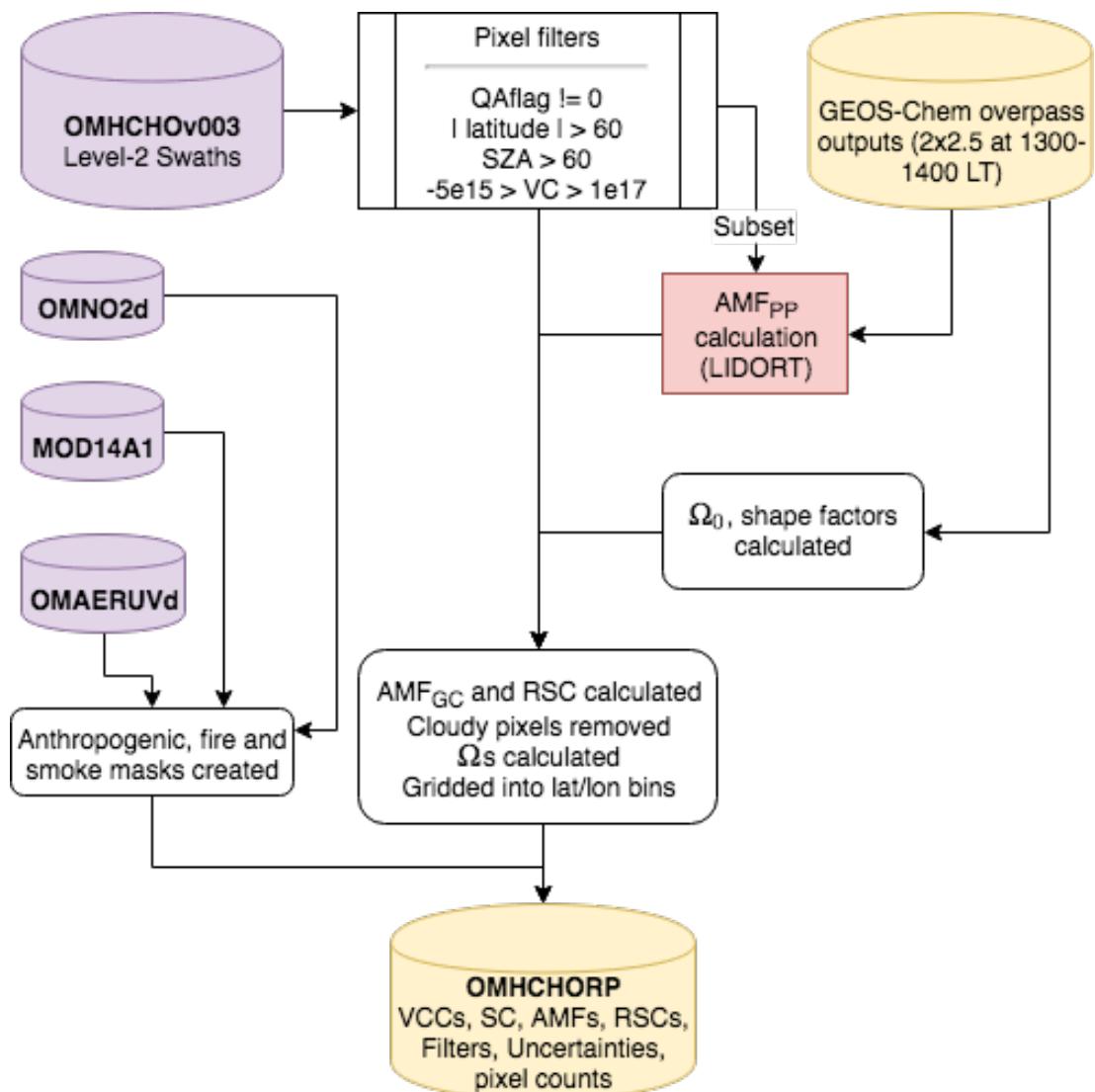


FIGURE 2.22: Flow diagram showing how OMHCHO level two swath data is read, processed, and gridded in this thesis

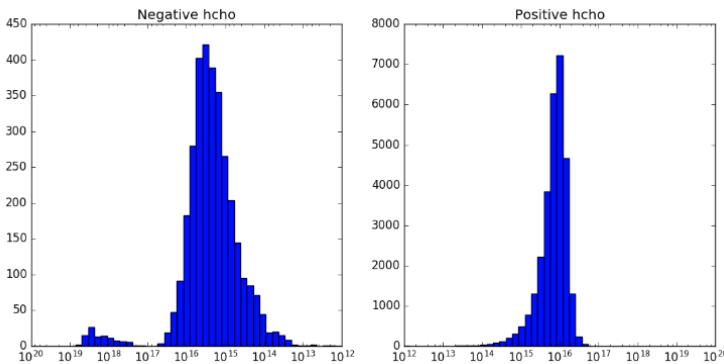


FIGURE 2.23: Column density histograms for a subset of OMI swaths over Australia on the 18th of March 2013. Negative entries are shown in the left panel, positive in the right, note the different scale between negative and positive panels.

2.5.2 Reading satellite data

First satellite slant columns of formaldehyde for the years January 1st, 2005 - April 1st, 2013 are downloaded (see section 2.2.4). The data set used is from the Ozone Monitoring Instrument (OMI) on board the Aura satellite, as it has data for the entire time line and sufficiently covers the southern hemisphere. When reading OMHCHO level 2 swath files, several factors are taken into account in order to filter uncertain and erroneous pixels. The process is outlined in figure 2.22 for a single day. First all *good* pixels (those with QA flag equal to 0) are read into a long list (roughly 1 million per day). These are then filtered by solar zenith angle (SZA) and latitude, similarly to other works (Marais et al. 2012; Barkley et al. 2013; Bauwens et al. 2016; Zhu et al. 2016, eg.). This filtering removes highly uncertain pixels, along with those for which instrument problems such as the OMI row anomaly (see section 2.2.4) may have affected.

Satellite measurements polewards of 60° north or south are removed as well as measurements with SZA greater than 60°. Pixels with a cloud fraction greater than 40% are filtered out after RSC is created. Further filtering is performed to remove the measurements which are most likely to be unrealistic: those with column density outside the range -0.5×10^{16} to 10^{17} molec cm^{-2} , as is performed by Zhu et al. (2016). This final filter is required due to currently unexplained large negative values which occur in the OMI HCHO product increasingly over time. Figure 2.23 shows how unfiltered HCHO columns are affected by a small set of highly negative values which heavily affect the mean column amount over any region. The histograms here show the negative (left) and positive (right) total column HCHO measurements from a subset of swaths over Australia, on the 18th of March 2013. The highly negative values can be seen around the -10^{19} molecules cm^{-2} region.

Each pixel and its relevant data are saved in a long list, around 1.1 million pixels per day. Additional information is added to each pixel, such as the new AMF calculated through replacing the a priori vertical profile with the newer GEOS-Chem simulated profile. The shape factors and scattering weights for each pixel lie along a z-axis which is vertically resolved to 47 layers.

Finally the pixels which may be too cloudy are dealt with after the pixel has been

used in determining the reference sector correction (see section 2.5.5), as is done in Gonzalez Abad et al. (2015) and De Smedt et al. (2015). Any pixel with a cloud fraction of greater than 40% is removed,

2.5.3 Scattering weight and the shape factor

In this section I describe how the shape factor can be used to recalculate OMI AMFs. The associated OMI per-pixel scattering weights are not changed in this calculation (unlike in section 2.5.4). The satellite shape factor is replaced by GEOS-Chem's over-pass time simulated HCHO profile, normalised and saved daily along with air density.

OMI's a priori shape factor is based on the GEOS-Chem (v9) model, which uses 47 layers between the earth's surface and the top of the atmosphere using a pressure-eta hybrid (the actual values are shown in table B.1). The equations and method of recalculating this a priori shape factor with the newer GEOS-Chem V10.01 code is layed out in this section. For each pixel the shape factor used in the AMF is recalculated using GEOS-Chem by following the method layed out initially in Palmer et al. (2001).

The Beer-Lambert law of extinction allows spectroscopic measurement of absorbing chemical species (absorbers) in the atmosphere:

$$I_B = I_{B_0} e^{-\tau_s} \quad (2.6)$$

where I_B , I_{B_0} is backscattered intensity with and without the absorber respectively, and τ_s is the optical thickness of the slant column absorber (see section 2.4.1).

The AMF is needed to transform the slant column, as viewed by the satellite, into a vertical column:

$$AMF = \frac{\Omega_s}{\Omega_v} = \frac{\tau_s}{\tau_v} \quad (2.7)$$

where s and v subscripts refer to slant and vertical values, while Ω represents a column of absorber in molecules cm^{-2} .

In the absence of atmospheric scattering a simple geometric AMF can be defined as a function of the solar zenith angle. The solar zenith angle (θ_s) and the satellite viewing angle (θ_v) are shown in image 2.24. However, in the UV-VIS region of the spectrum, Rayleigh and Mie scattering (see section ??) must be accounted for. Using equations 2.6 and 2.7 gives us an expression for the AMF which includes scattering:

$$AMF = \frac{\ln I_{B_0} - \ln I_B}{\tau_v} \quad (2.8)$$

We use $\nabla I = I_B - I_{B_0}$ to represent the change in intensity due to the absorber. Note that for optically thin absorption, $\nabla I / I_B \ll 1$, and we can use:

$$AMF = \frac{\ln \left(1 - \frac{\nabla I}{I_B} \right)}{\tau_v} \approx \frac{-\frac{\nabla I}{I_B}}{\tau_v} \quad (2.9)$$

∇I can also be expressed as the integral of the absorption slices over optical depth increments:

$$\nabla I = \int_0^{\tau_v} \frac{\partial I_B}{\partial \tau} d\tau$$

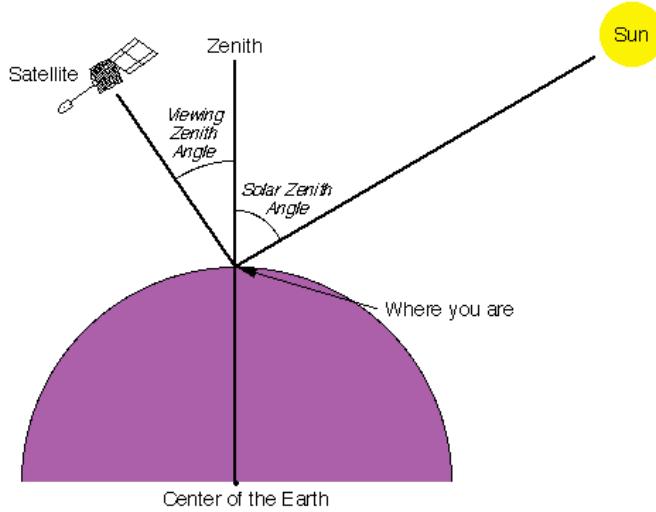


FIGURE 2.24: Solar and viewing zenith angles, image copied from Wikipedia (2016), originally from a NASA website.

which can be placed into equation 2.9:

$$AMF \approx \frac{-1}{\tau_v} \int_0^{\tau_v} \frac{\partial \ln I_B}{\partial \tau} d\tau$$

and rewritten as:

$$AMF = \frac{-1}{\tau_v} \int_0^{\infty} \frac{\partial \ln I_B}{\partial \tau} \alpha(z) \eta(z) dz \quad (2.10)$$

where $\alpha(z)$ and $\eta(z)$ represent absorption cross section in m^2 molecule $^{-1}$, and number density in molecules m^{-3} respectively. This uses the attenuation cross section relationship to optical depth (see section 2.4.1).

To represent an average cross section weighted by the absorbing species' vertical distribution, intended to account for temperature and pressure dependence of $\alpha(z)$, α_e is defined as:

$$\alpha_e = \frac{1}{\Omega_v} \int_0^{\infty} \alpha(z) \eta(z) dz$$

This is the same as $\tau_v = \Omega_v \alpha_e$, which we can place into equation 2.10 to obtain:

$$AMF = - \int_0^{\infty} \frac{\partial \ln I_B}{\partial \tau} \frac{\alpha(z)}{\alpha_e} \frac{\eta(z)}{\Omega_v} dz \quad (2.11)$$

Defining $\omega(z)$ as the scattering weights describing the sensitivity of the backscattered spectrum to the abundance of an absorber at altitude z :

$$\omega(z) = - \frac{1}{AMF_G} \frac{\alpha(z)}{\alpha_e} \frac{\partial \ln I_B}{\partial \tau} \quad (2.12)$$

It's worth noting that in the OMI satellite product, the provided $\omega(z)$ term does not include the $\frac{1}{AMF_G}$ term and the calculations which follow therefor do not include

this term when utilising the provided ω . This is not noted in any of the papers which recalculate the AMF from the OMI product, due to them recalculating the ω term themselves with a radiative transfer model such as LIDORT.

Plugging equation 2.12 into equation 2.11 gives us:

$$AMF = - \int_0^\infty \omega(z) \frac{\eta(z)}{\Omega_v} dz \quad (2.13)$$

which is best described as the scattering weights (ω) multiplied by the shape factor ($S_z(z) = \frac{\eta(z)}{\Omega_v}$) integrated over altitude z . Since we are using the ω provided by OMI, the AMF_G term is removed from this calculation when compared against Palmer et al. (2001).

The integration is done in Python using a simple rectangular method, which multiplies the integrand midpoints by the change in height, and then takes the sum. This assumes that the provided scattering weights and shape factors are linear between the 47 resolved values. I recalculate $S_z(z)$ using vertical columns of HCHO averaged over 1300-1400 LT from GEOS-Chem. As a sanity check I also recalculate S_σ and confirm that these shape factors are equivalent, for details see section 2.5.3.

All that remains for recalculating the total vertical column using our new a priori shape factor is to apply the new AMF to the slant columns and grid them onto our chosen resolution. Recalculation of the ω is done separately, as explained in section 2.5.4.

For example see figure TODO: sigma vs z shape factor plot. TODO add plot and describe here.

Normalised vertical shape factor

Additionally the vertical shape factor is normalised and the recalculated AMF is checked against that using the non-normalised version. The additional steps to calculate a normalised vertical shape factor S_σ are first layed out in Palmer et al. (2001).

The following equation converts model profile output from ppb into number densities:

$$\eta_{HCHO} = ppb_{HCHO} \times \eta_a \times 10^{-9} \quad (2.14)$$

where η_{HCHO} is the number density of a HCHO, η_a is the number density of air (from model output), and ppb_{HCHO} is the molecules of that species per billion molecules of air. In order to normalize these vertical density profiles over the globe, we divide by the modelled total vertical column Ω_{HCHO} which is determined by:

$$\Omega_{HCHO} = 2.12 \times 10^{13} \Sigma_z (ppb_{HCHO}(z)(P(z) - P(z + 1)))$$

where $P(z)$ is the pressure (hPa) at the bottom of altitude level z , the constant 2.12e13 is determined from equation (TODO: Add this equation). In effect this equation sums over the molecules per cm^2 in each altitude level.

The normalised vertical shape factor $S_z(z)$ is defined as:

$$S_z(z) = \frac{\eta(z)}{\Omega_v} \quad (2.15)$$

where $\eta(z)$ is the number density in molecules m⁻³.

Now the AMF can be expressed as

$$AMF = \int_0^\infty \omega(z) S_z(z) dz \quad (2.16)$$

We have $S_z(z)$ and $\omega(z)$ over the vertical pressure coordinate z at all latitude and longitude points on whatever grid we wish. A conversion to the sigma (σ) vertical coordinate is performed using $P = \sigma(P_S - P_T) + P_T$, where P_T is pressure at the top of the atmosphere and P_S is surface pressure. In the sigma coordinate system we calculated the shape factor as follows:

$$S_\sigma(\sigma) = \frac{\Omega_a}{\Omega_v} C_{HCHO}(\sigma) \quad (2.17)$$

where Ω_a is the vertical column of air from the surface to the top of the atmosphere and $C_{HCHO}(\sigma)$ is the mixing ratio of HCHO. This equation comes from Palmer et al. (2001), and is unitless since Ω_a/Ω_v is molecules of air per molecule of HCHO; the opposite of C_{HCHO} .

Sigma coordinate conversion

Using the vertical coordinate sigma (σ), which is related to pressure (P) by $P = \sigma(P_S - P_T) + P_T$, where subscripts S and T represent earth surface and top of the atmosphere respectively. The hydrostatic relation $P = -\rho_a g z$, with ρ_a , g , being density of air, gravity, respectively lets us switch to the sigma coordinate using:

$$\begin{aligned} \rho_a g z &= \sigma (P_S - P_T) + P_T \\ d\sigma &= -\frac{\rho_a g}{P_S - P_T} dz \end{aligned}$$

Substitution into 2.16 gives AMF using the sigma coordinates:

$$AMF = \int_0^1 w(\sigma) S_\sigma(\sigma) d\sigma \quad (2.18)$$

Where S_σ is defined as a vertical shape factor representing a normalised mixing ratio:

$$S_\sigma(\sigma) = \frac{\Omega_a}{\Omega_v} C(\sigma) \quad (2.19)$$

where Ω_a is the vertical column of air and $C(\sigma)$ is the mixing ratio of the absorber. This normalised shape factor is dimensionless. This can be useful when running global atmospheric models as the ground altitude is always at $\sigma = 1$ and we need not worry about topography.

2.5.4 Recalculating the AMF using PP code

Some of the pixels (those covering Australia and most of the zonal band) have their AMFs recalculated using Fortran code written by Paul Palmer, Randal Martin, and

updated by Luke Surl. I will refer to this as the PP code, and subscript the VCs and AMFs with PP when referring to those calculated through this method. This code is computationally expensive, and is only run on pixels within the region (50° - 10° S, 160° W- 160° E). The Fortran code uses a combination of GEOS-Chem a priori profile information and satellite measurement data to calculate the AMF after running the LIDORT radiative transfer calculations to determine scattering. The instrument sensitivity or scattering weights, and shape factors for each pixel are calculated which I then integrate vertically to get AMF_{PPS} .

Code for recalculating AMFs using satellite swaths and modelled aerosol optical depths and gas profiles can be found at http://fizz.phys.dal.ca/~atmos/martin/?page_id=129. The original method for HCHO is laid out in Palmer et al. (2001), with modifications for clouds and use of the LIDORT RTM (Spurr 2002) as described by Martin et al. (2003). This code does not work as is when using OMI satellite data, and requires modifications performed by Luke Surl at Edinburgh University. Additionally the tropopause heights averaged within satellite overpass times output by GEOS-Chem is required, which is achieved by modifying the ND51 diagnostic.

Mie scattering and clouds can complicate the calculation of $\omega(z)$, however tables of values for this function at various parameter inputs can be used with modeled vertical shape factors for local AMF calculations. This has been done in the PP code and the AMF look-up-table (LUT) can be found in the source code at TODO: add git repo with this code.

First special output is required from GEOS-Chem, averaged between 1300 and 1400 LT, including optical depths at several wavelengths (TODO: list), dust, and HCHO. I then pull out a subset of the OMI pixel information into a daily csv file, which can be read by the PP code as modified by Dr. Luke Surl, in conjunction with the GEOS-Chem outputs for each day. The PP code then produces a csv of recalculated AMFs which get read by my python code and associated with the corresponding pixel (outlined in ??).

2.5.5 Reference sector correction

Each satellite slant column measurement is corrected by some amount, based on the divergence from a modeled reference sector. HCHO products from OMI and SCIA-MACHY both use a median daily remote pacific ocean radiance reference spectrum, over 15° S- 15° N, 140° - 160° W where it is assumed that the only significant source of HCHO is methane oxidation (De Smedt et al. 2008; Barkley et al. 2013; Kurosu and Chance 2014).

The reference sector correction (RSC) for OMI satellite pixels in this thesis is calculated following Abad et al. (2016). The RSC is calculated daily and applied to each satellite pixel corresponding to each of the 60 measurement tracks used by OMI, depending also on latitude. The reference sector correction is based on the difference between the GEOS-Chem simulated slant columns and OMI slant columns over the meridional band from 140° W to 160° W. To get the modelled slant columns, each of the AMFs (calculated in prior sections) is applied to the modelled vertical columns using equation 2.3.

The RSC method corrects for several problems, however it introduces some a priori model influence. One of the problems removed through this correction method is

instrument degradation, which can introduce bias over time. Another is the possible influence of varying dead/hot pixel masks across 2-D detector arrays such as OMI (De Smedt et al. 2015). This method also corrects for the errors introduced through correlations between BrO and HCHO absorption cross sections, which are especially significant at high latitudes (Gonzalez Abad et al. 2015).

Vertical columns in OMI use this oceanic background instead of a solar irradiance spectrum. In order to recalculate the vertical columns using modelled data, a RSC needs to be applied. The correction uses the difference between the slant column (Ω_S) and the reference slant column (Ω_{S_0}) divided by the AMF, plus the modelled reference sector column (Ω_{V_B}):

$$\Omega_{VC} = \frac{(\Omega_S - \Omega_{S_0})}{AMF} + \Omega_{V_B}$$

This method is used in various papers, including De Smedt et al. (2008), De Smedt et al. (2012), De Smedt et al. (2015), Barkley et al. (2013), and Bauwens et al. (2016). Recently this correction was expanded (for OMI data) to include latitudinal and instrument track influence by Gonzalez Abad et al. (2015).

The correction used in this thesis is created based on the difference between the background HCHO measurements from OMI and the GEOS-Chem modelled HCHO columns within the reference sector. The longitudinal average is taken within the a priori reference sector, as corrections are assumed to be longitudinally invariant. The modeled reference sector is interpolated latitudinally to 500 equidistant bins. Figure 2.25 shows the simulated reference sector VCs as an example, calculated for January 1st 2005. In this figure the vertical resolution is increased from 2° to 0.36° , through linear interpolation, in order to form 500 vertical bins which are used in correcting the satellite data. Each day, good satellite measurements taken over the reference sector are used to determine a correction array. The correction is based on the difference between measured slant column and the modeled slant column within the reference sector. The model does not produce slant columns associated with each measurement, however one is created by multiplying the VC with the associated slant column's AMF.

For OMI swaths, each row of measured data contains 60 'Across track'(track) measurements. The track index (i) relates a the measurement to one of the 60 columns of data. Corrections for each measurement are calculated by taking the difference between the measured slant column and the a priori slant column as follows:

$$\text{Correction}(i, j) = SC_{HCHO}(i, j) - VC_{\text{GEOS-Chem}}(lat(j)) \times AMF(i, j) \quad (2.20)$$

where j represents a latitude index and $VC_{\text{GEOS-Chem}}(lat)$ represents the a priori reference sector vertical column HCHO at the latitude corresponding to j. Note that the correction is in molecules cm^{-2} . The RSC is independently calculated for each of the 60 tracks, at each latitude in the 500 0.36° bins. This provides a different RSC

Due incomplete latitudinal coverage, the correction for each track is interpolated linearly between measurements, with corrections outside of the highest measured latitudes being equal to the corrections at the highest measured latitudes. Figure 2.26 shows an example of the 60 track corrections for January 1st 2005, the points are satellite measurements and the lines are the interpolations for each track.

Another way to look at this correction is given in the OMI corrections panel of figure 2.25, which has the sensors along the x axis, and latitude on the y axis, and

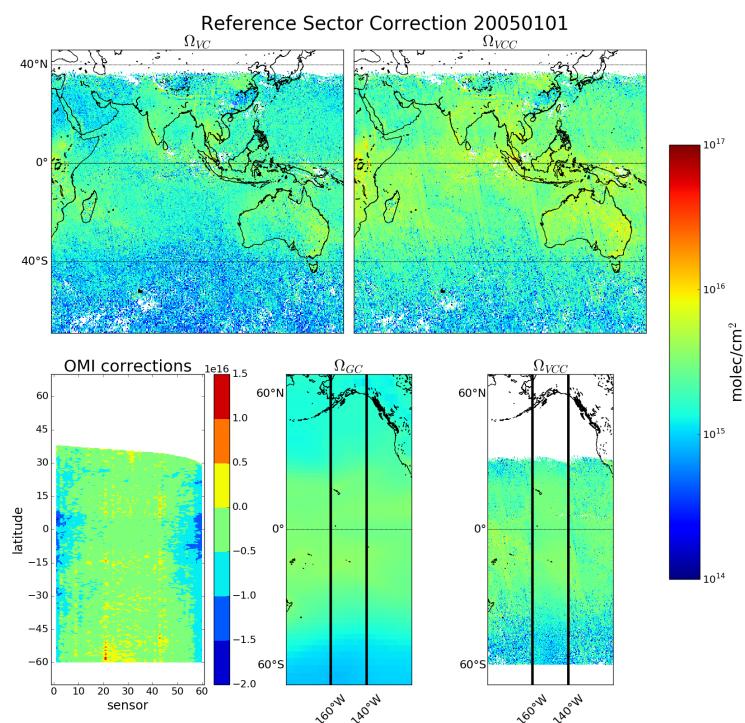


FIGURE 2.25: Example of remote pacific RSC using 8-day average measurements and one month modelled data. Ω_{VC} shows the uncorrected vertical columns, while Ω_{VCC} shows the corrected vertical columns. OMI corrections shows the correction applied globally based on latitude and OMI track number(sensor). Ω_{GC} shows the GEOS-Chem modelled HCHO VC over the RSC, with Ω_{VCC} showing the corrected VC over the same area.

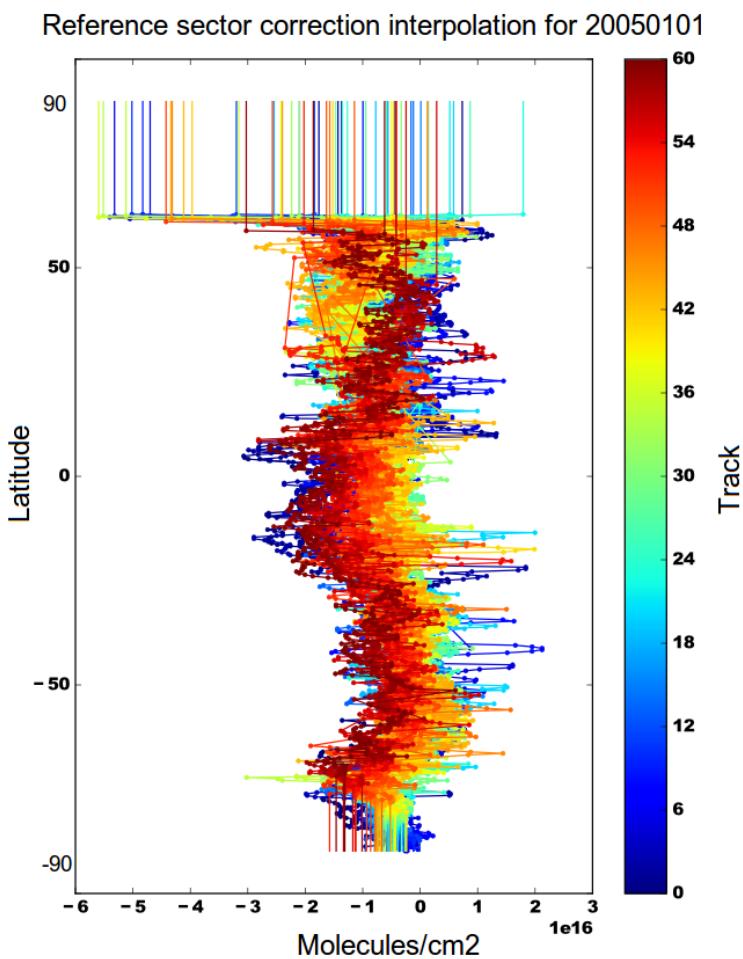


FIGURE 2.26: Example of track correction interpolations for January 1st 2005, points represent satellite slant column measurements, with lines interpolating and extrapolating along the latitudinal direction.

shows how for this example 8-day period, the corrections are distributed with more negative values towards the left or right sensors, especially in the tropics.

One correction is associated with every good satellite measurement which is used to create a reference sector corrected measurement (Vertical Column Corrected or VCC) through the following equation:

$$VCC(i,j) = \frac{SC_{HCHO}(i,j) - Correction(i, lat(j))}{AMF(i,j)} \quad (2.21)$$

Finally, for each day, the good satellite measurements are averaged into our own latitude longitude resolution bins along with the associated corrected SC, VC, VCC, AMF, and bin entry count. The bin entry count is used to create weighted averages from the daily binned data, which accounts for sparse entries due to filtering.

2.5.6 Binning the results daily

Finally the pixels are binned into a gridded dataset I've called OMHCHORP, as shown in figure 2.22. The resolution is chosen to match the native resolution of GEOS-Chem ($0.25 \times 0.3125^\circ$) and the GEOS met data fields. Data averaged into this dataset are as follows:

1. satellite SC
2. satellite AMF
3. satellite VC
4. satellite RSC VC
5. GEOS-Chem recalculated AMF
6. GEOS-Chem recalculated VC
7. GEOS-Chem recalculated RSC VC
8. GEOS-Chem AMF recalculated using Paul Palmer code (AMF_{PP})
9. GEOS-Chem RSC VC based on AMF_{PP}
10. Smoke AAOD from OMAERUVd (mapped into bins from $1 \times 1^\circ$ resolution)
11. satellite pixel counts (summed into bins)
12. fire counts (summed into bins)

TODO: time per regridding and reprocessing: This whole process requires some processing time as well as RAM and computer storage space, and has been performed on the National Computing Infrastructure (NCI) supercomputer cluster. In order to reprocess one year of swath files, X GB of daily data was downloaded and then transformed into Y GB of daily gridded data. This takes around 90 minutes per day, and is very parallelisable as each day is completely independent once the model has run in each required configuration. Initially parallelism was built into the python code,

however simply running sending separate ‘jobs’ to NCI’s process queue was simpler and more scalable. As much as possible, processing is done using the HDF-5 format, with some GEOS-Chem output being read and processed from bitpunch to HDF-EOS5 prior to reprocessing. The scripts to regrid and reprocess the swath data set are available in the supplementary (TODO).

2.5.7 Difference between new and old OMI HCHO columns

Two HCHO products are created, both using GEOS-Chem output at global 2 by 2.5°horizontal resolution. One uses the OMI product’s ω_z and equation ?? in order to calculate an AMF. While the other uses code provided by Dr. Paul Palmer, with alterations by Dr. Randal Martin, and Dr. Luke Surl to run LIDORT on the satellite slant columns and the GEOS-Chem output in order to calculate an AMF. These two calculations are compared over Australia in figure(s) TODO: Map comparison, regression, and time series once AMFpp is working properly. The effect of not recalculating the ω_z is can be seen in figure 2.27 which looks at the altered satellite vertical columns using each method.

Figure 2.27 shows vertical columns of HCHO for: column 1) the original satellite swaths, column 2) recalculated without changing the provided scattering weights, and column 3) fully recalculated vertical columns. Each grid square (at 0.25 by 0.3125°lat lon resolution) has been created by binning the recalculated satellite pixels within the month. The average pixels per land square is overlaid and changes due to how a fire filter is applied. Each row has a stricter fire filter applied from top to bottom, with no fire filter on the first row up to filtering pixels from squares with fires up to 8 days prior. This figure looks at March 2005 with biomass burning filtered differently in each row. Active fires over the last 0, 1, 2, 4, and 8 days are filtered as the row number increases.

Figure TODO shows an analysis of the differences between running the recalculation with and without updating the ω_z .

TODO: Ask luke if this is true: The AMF calculated using Dr. Palmer’s code uses a more strict series of filters, leading to fewer satellite based HCHO columns and reduced coverage over Australia. Stricter filtering must be balanced against both coverage and the sensitivity of the AMF determination to recalculating ω_z .

Figure TODO: shows global and Australian HCHO averaged total column maps for January 2005, along with the reduced major axis (RMA) regression correlation and percentage difference. This comparison shows how reprocessing with an updated model can have a systematic influence on the total column.

2.6 Filtering Data

In order to examine only biogenic processes, pyrogenic and anthropogenic influences need to be removed from modelled and measured data. As biomass burning can be a large local or transported source of HCHO, CHOCHO, glyoxal, and other compounds we would like to use to determine BVOC emissions, it is advantageous to filter out this source. Just filtering active fires does not account for transported smoke plumes, which can carry HCHO precursors. One complication when computing HCHO yield from VOC emissions is biomass burning interference, as smoke plumes can contribute

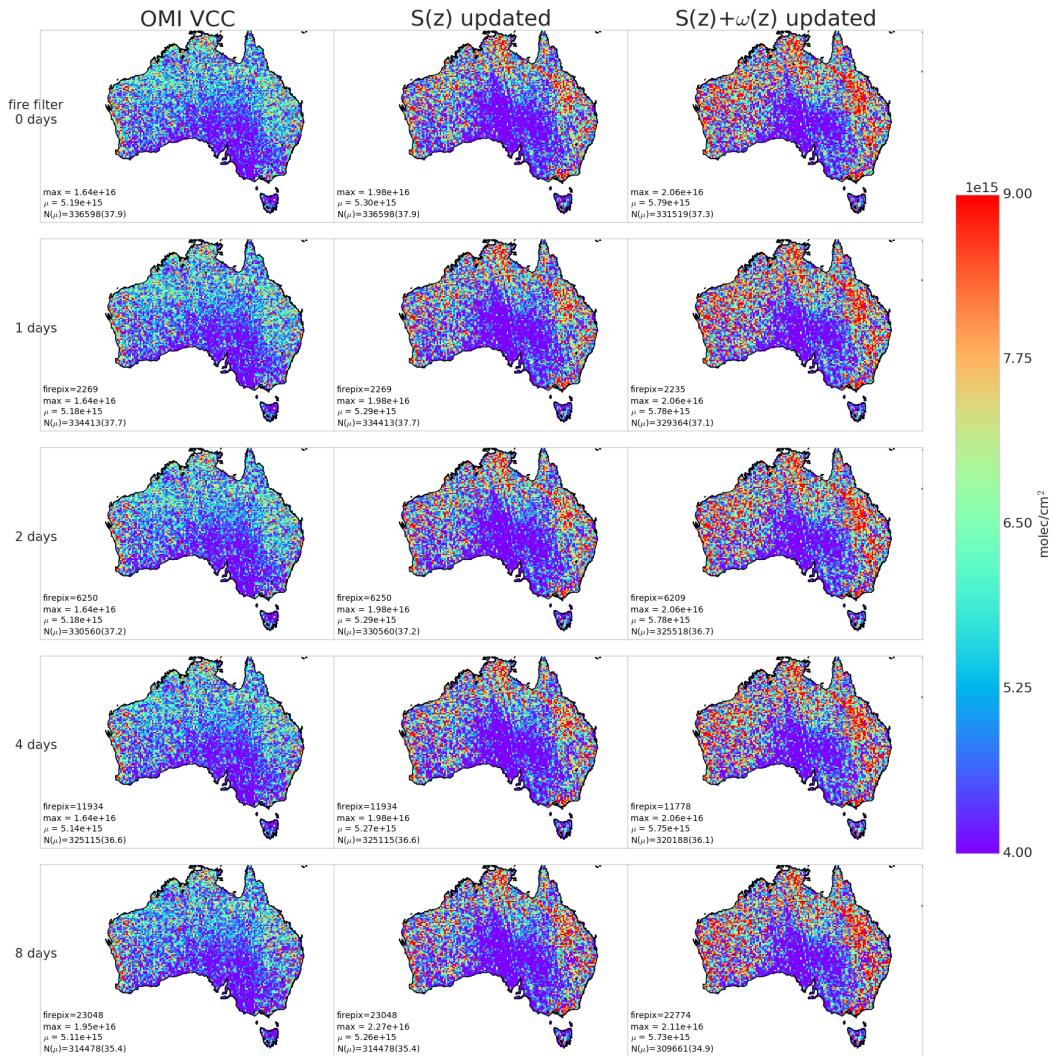


FIGURE 2.27: Column 1: Reference sector corrected HCHO vertical columns Ω from OMHCVOv003. Column 2: Ω with recalculated a priori shape factors using GEOS-Chem v10.01. Column 3: Ω with recalculated a priori shape factors and scattering weights using GEOS-Chem v10.01 and LIDORT. Row 1-5: increasing number of prior days which have active fires are included when masking fire influence.

to column HCHO. In GEOS-Chem we can simply turn off pyrogenic and anthropogenic emissions, however in satellite datasets we need to mask pixels affected by biomass burning.

Influence from biomass burning can be removed through measurements of acetonitrile and CO (eg: (Wolfe et al. 2016; Miller et al. 2016), or else removal of scenes coincident with satellite detected fire counts and aerosol absorption optical depth as done in Marais et al. (2014). Wolfe et al. (2016) disregard HCHO measurements when acetonitrile > 210 pptv and CO > 300 ppbv, while acetonitrile > 200 pptv is used to determine fire influence in Miller et al. (2016). TODO: look at yearly correlation, compare to exponential curve and look for fire outliers As seen in TODO: citation, HCHO concentrations scale exponentially with temperature. This allows another method for detecting the influence of non-biogenic HCHO emission/creation by looking for outliers above the curve at low temperature. Zhu et al. (2013) has a similar analysis over southeastern USA showing an exponential correlation of $HCHO = \exp(0.15 * T - 9.07)$.

Once the satellite data are quality filtered and gridded, I use additional data to account for anthropogenic and pyrogenic influences, which provides us an estimate of biogenic HCHO. MODIS fire counts are used in conjunction with smoke AAOD enhancements (also measured by satellite) to remove data points which may be affected by fires. This has a negative affect on uncertainty, as fewer measurements are available to be averaged. This section describes the creation and effects of filters used on satellite data.

2.6.1 Fire and smoke

The method used in this thesis follows that of Marais et al. (2012), and Barkley et al. (2013), with active fires filtered using fire counts, and smoke filtered out using smoke aerosol absorption optical depth (AAOD). Marais et al. (2012) remove pixels colocated with non zero fire counts in any of the prior eight days, within grid squares with $1 \times 1^\circ$ resolution. Barkley et al. (2013) use fires from the preceding and concurrent day, within local or adjacent grid squares, with grid resolution of $0.25 \times 0.3125^\circ$. We use the MODIS fire counts, detected from space using the combined product from Terra and Aqua (Terra at 10:30, 22:30 LT; Aqua at 13:30, 01:30 LT). Smoke plumes can be filtered using product OMAERUVd, although care needs to be taken when deciding the threshhold for smoke detection (Marais et al. 2012).

When analysing satellite OMHCHO vertical columns (Ω), the following steps are performed in order to mask influence from biomass burning:

1. MOD14A1 daily gridded Aqua/Terra combined fire counts are read at $1 \times 1 \text{ km}^2$ resolution, and binned into $0.25 \times 0.3125^\circ$ bins, matching the resolution of binned Ω .
2. A rolling mask is formed which removes Ω if one or more fires are detected in a grid square, or in the adjacent grid square, up to 2 days previously. This includes the 'current' day, making 3 days of fires in total being filtered out on each day.
3. AAOD at 500 nm is mapped from OMAERUVd $1 \times 1^\circ$ resolution onto the $0.25 \times 0.3125^\circ$ resolution.

4. An AAOD threshold of 0.03 is determined through visual analysis of AAOD distributions over several days, including days with and without influence from active fires, dust, and transported smoke plumes.
5. Grid squares with AAOD over this threshold are considered potentially affected by transported fire smoke.

Determining the AAOD due to smoke can be difficult since both smoke and dust absorb UV radiation (Ahn2008; Marais et al. 2012). AAOD is should be less sensitive to cloud contamination than AOD, and I use AAOD from the daily gridded level 3 satellite product OMAERUVd (Ahn2008) described in section 2.2.4 to provide a filter for smoke plumes. Although removing gridsquares with dust reduces how much data is available to analyse, it's considered a minor problem as dust in Australia is highly episodic and should not affect more than a few days per year, especially over regions with high tree coverage (Shao et al. 2007).

Filtering fire smoke using AAOD is done by removing OMHCHO gridsquares where the AAOD is above a 0.03, after the AAOD is mapped from $1 \times 1^\circ$ to the same $0.25 \times 0.3125^\circ$ resolution as our OMHCHO gridded product. The threshhold is determined through analysing AAOD over Australia in 4 scenarios: normal conditions, active local fires, during influence from transported fire smoke, and large scale dust storms. Figure 2.28 shows AAOD (columns 1 and 2), with AAOD distribution in column 3, along with satellite imagery on the same day in column 4 (from <https://worldview.earthdata.nasa.gov/>). The scenarios listed are shown from row 1 to 4, and AAOD = 0.03 is demarked by a horizontal line in the density plots in column 3.

Checking that fire masks are influencing pyrogenic HCHO

Looking at temperature can provide evidence of pyrogenic HCHO. HCHO precursors are heavily tied to temperature (TODO:cite), and model output shows how higher temperature leads to an increase in HCHO levels. Figures 2.29 - 2.31 show the relationship between temperature and HCHO, for January 2005, within subsets of Australia. A reduced major axis regression is used to determine the correlation between surface temperature (X axis) and HCHO (Y axis).. Using the natural log of HCHO we can take the linear regression and then exponentiate each side in the equation $\ln Y = mX + b$ to get $Y = \exp(mX + b)$. This gives us the exponential fit as shown, with the correlation coefficient between $\ln HCHO$ and temperature. The distributions of exponential correlation coefficients and exponential 'm' terms is shown in the embedded plot, with one datapoint available for each grid square where the regression is performed.

When applying the fire mask to the modelled HCHO, we can analyse how this affects the relationship to temperature. One cause of high HCHO at lower temperatures (above the exponential regression line) can be emissions from fires. If we see these datapoints affected by the firemask then this suggests that the mask is working correctly. TODO plot showing how fire mask affects HCHO - Temperature relationship

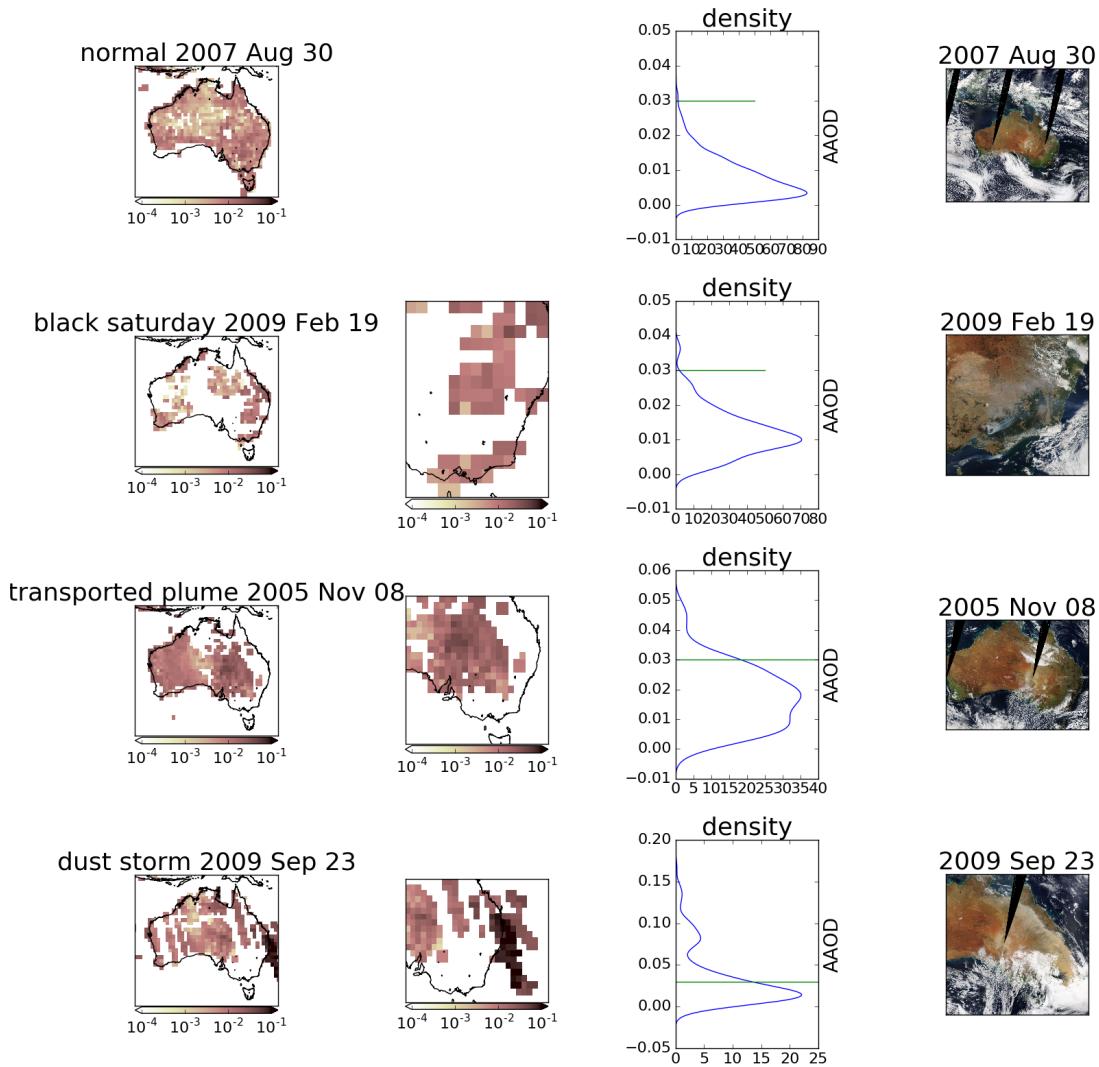


FIGURE 2.28: AAOD from OMAERUVd (columns 1, 2, 3) over Australia for four different scenarios (rows 1-4). Scenes from the same day are taken from the EOS Worldview website <https://worldview.earthdata.nasa.gov/>.

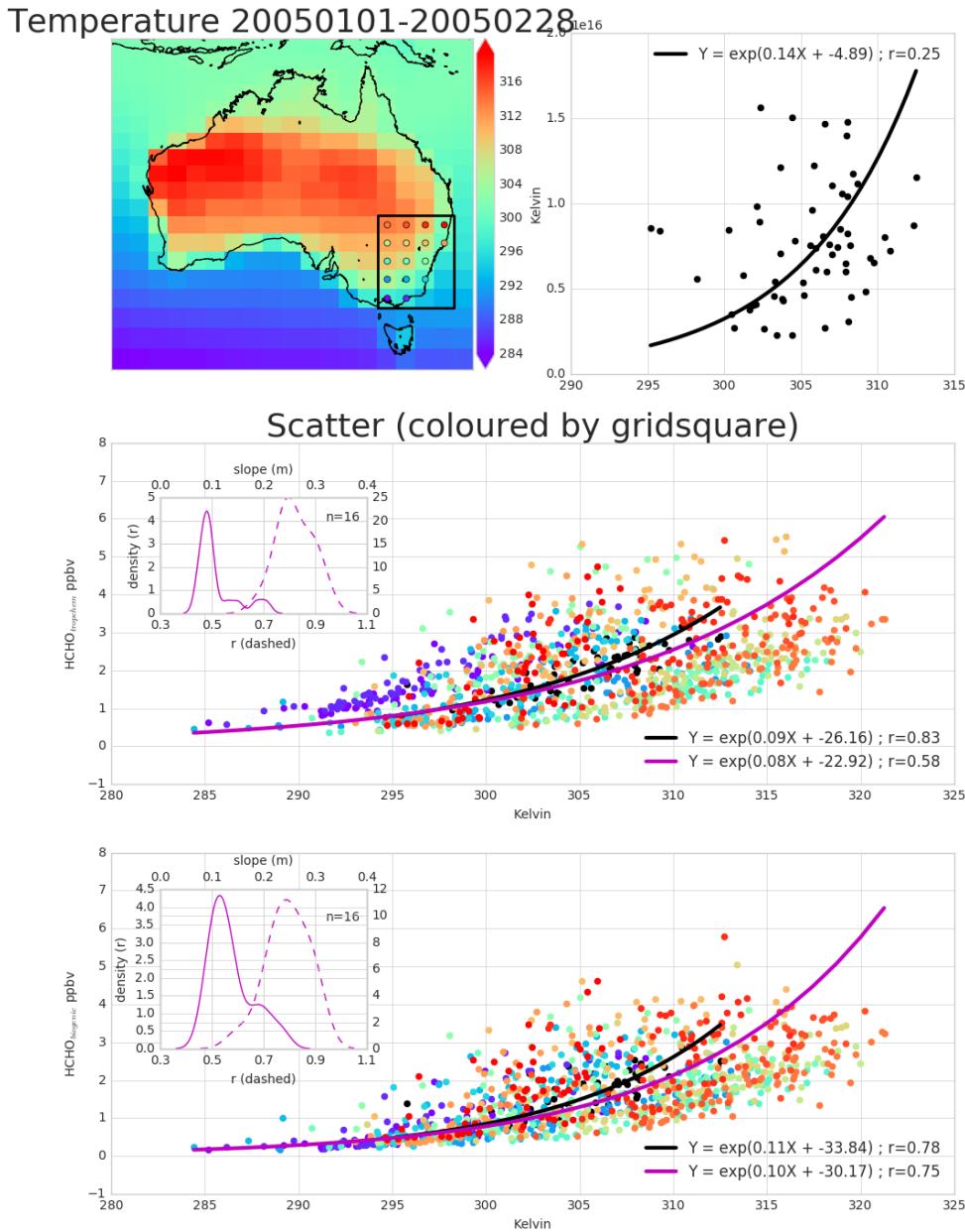


FIGURE 2.29: Top panel: surface temperature averaged over January and February 2005. Bottom panel: surface temperature correlated against temperature over, with different colours for each gridbox, and the combined correlation. A reduced major axis regression is used within each gridbox (shown in top panel) using daily overpass time surface temperature and HCHO amounts (ppbv). The distribution of slopes and regression correlation coefficients (one datapoint per gridbox) for the exponential regression is shown in the embedded plot.

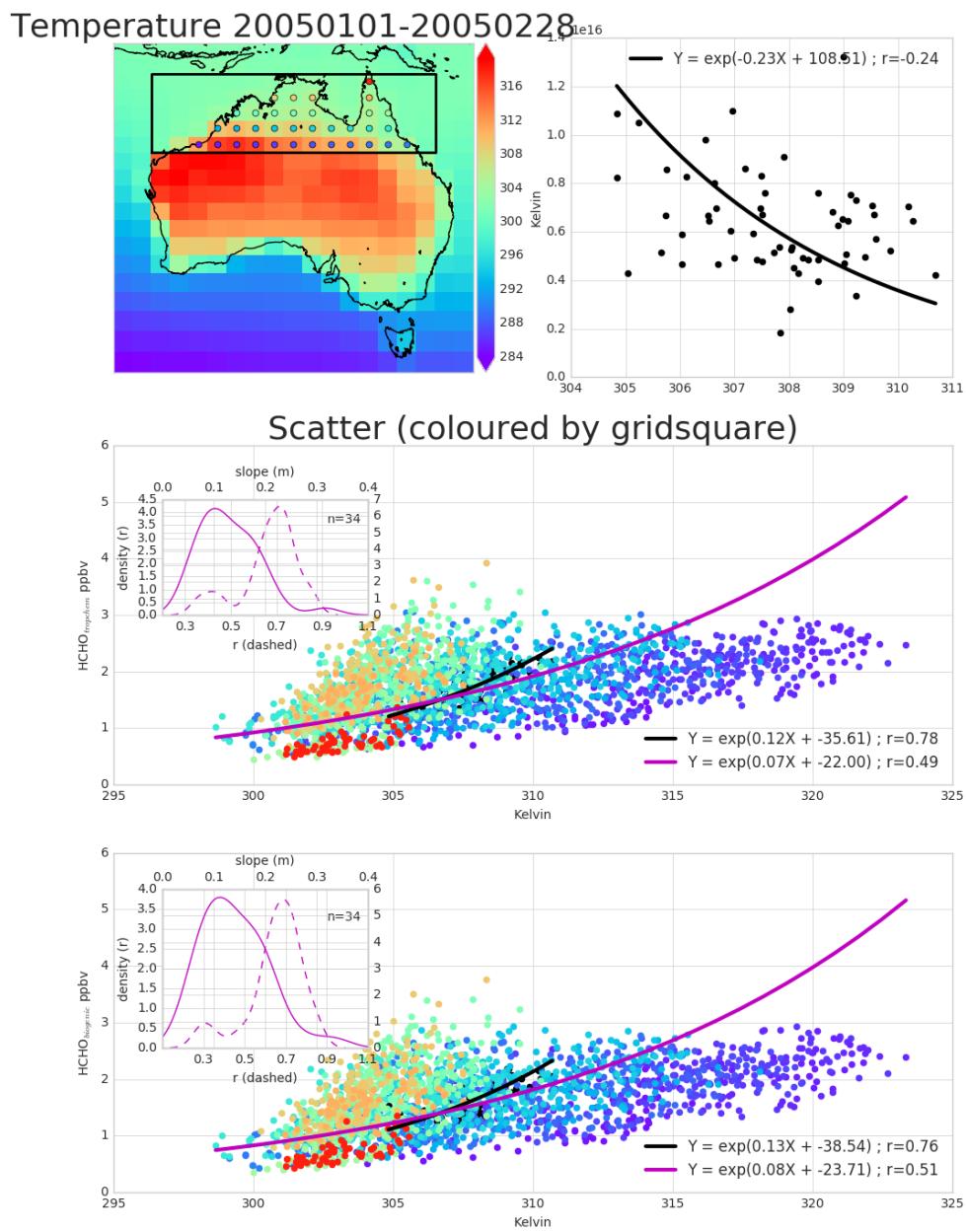


FIGURE 2.30: As figure 2.29 but for northern Australia.

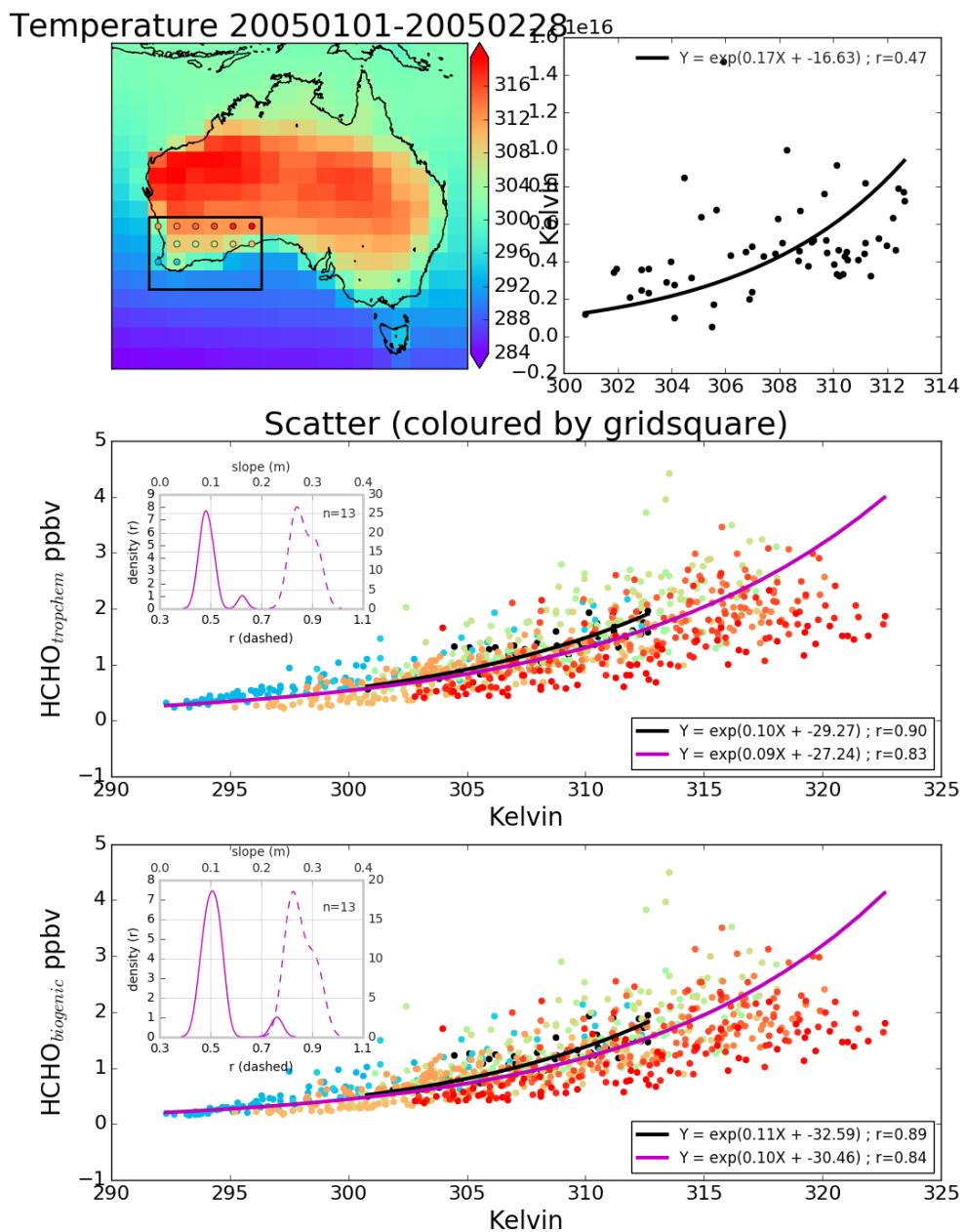


FIGURE 2.31: As figure 2.29 but for south-western Australia.

TABLE 2.2: NO₂ averages by region before and after filtering for anthropogenic emissions using 2005 data from the OMNO2d product.

Region	NO ₂	NO ₂ after filtering	% Data lost
Aus	1	2	3
BG	1	2	3
Syd	1	2	3
Melb	1	2	3
Adel	1	2	3

2.6.2 NOx

Enhanced NO₂ concentrations can indicate anthropogenic influence over Australia. In order to filter out these influences on satellite HCHO measurements, a filter is designed using the OMNO2d product which includes tropospheric NO₂ columns.

OMNO2d from 2005 is used to determine a suitable threshold for anthropogenic influence by looking at NO₂ columns near several major cities in the south eastern sector of Australia. The mean, standard deviation, and time series over Australia of tropospheric NO₂ seen by Aura is shown in figure 2.32. The average tropospheric NO₂ column for averaged within all of Australia and then each region shown in this figure is listed in table TODO 2.2.

Anthropogenic influences on the NO₂ columns are clearly visible near major cities in Australia. A filter is created each year from the OMNO2d product in two steps:

1. Daily gridsquares with NO₂ greater than 10^{15} molec cm⁻² are flagged as anthropogenic.
2. After taking the yearly average over Australia, any gridsquares greater than 1.5×10^{15} molec cm⁻² are flagged for the whole year.

This removes both the gridsquares close enough to cities to be affected by their emissions year round, as well as effects from transported pollution plumes. The affects of applying this filter to the OMNO2d product itself can be seen in figure 2.33

The same regions as in figure 2.32 are shown again in figure 2.34, with NO₂ pixels densities for each region shown, along with the threshold of 1×10^{15} molec cm⁻². This led to a reduction of TODO gridsquares from the total available measurement space over Australia. The removal of gridsquares which went above the yearly averaged limit of 1.5×10^{15} molec cm⁻² further reduced the available data by TODO gridsquares.

2.6.3 Summary of filters effects on HCHO

2.7 Data Access

TODO: ADD MORE HERE

OMNO2d Daily satellite NO₂ product downloaded from <https://search.earthdata.nasa.gov/search>, DOI:10.5067/Aura/OMI/DATA3007. See more information in section

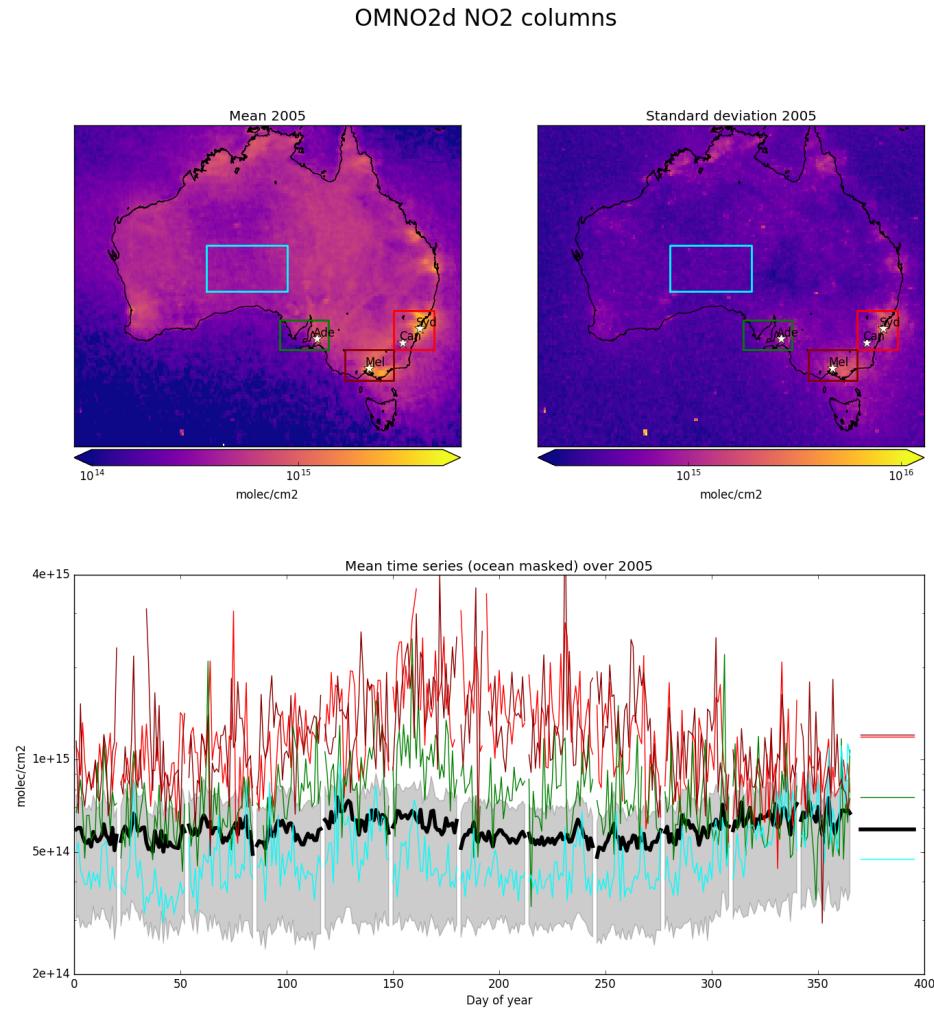


FIGURE 2.32: Mean (top left) and standard deviation (top right) of OMNO2d daily $0.25 \times 0.25^\circ$ tropospheric cloud filtered NO₂ columns. Time series for Australia, and each region (by colour) shown in the bottom panel, with mean for that region shown on the right. A grey shaded area depicts the 25th to 75th percentiles of Australia averaged NO₂ columns for each day in the time series, with a thicker black line showing the Australia-wide mean value.

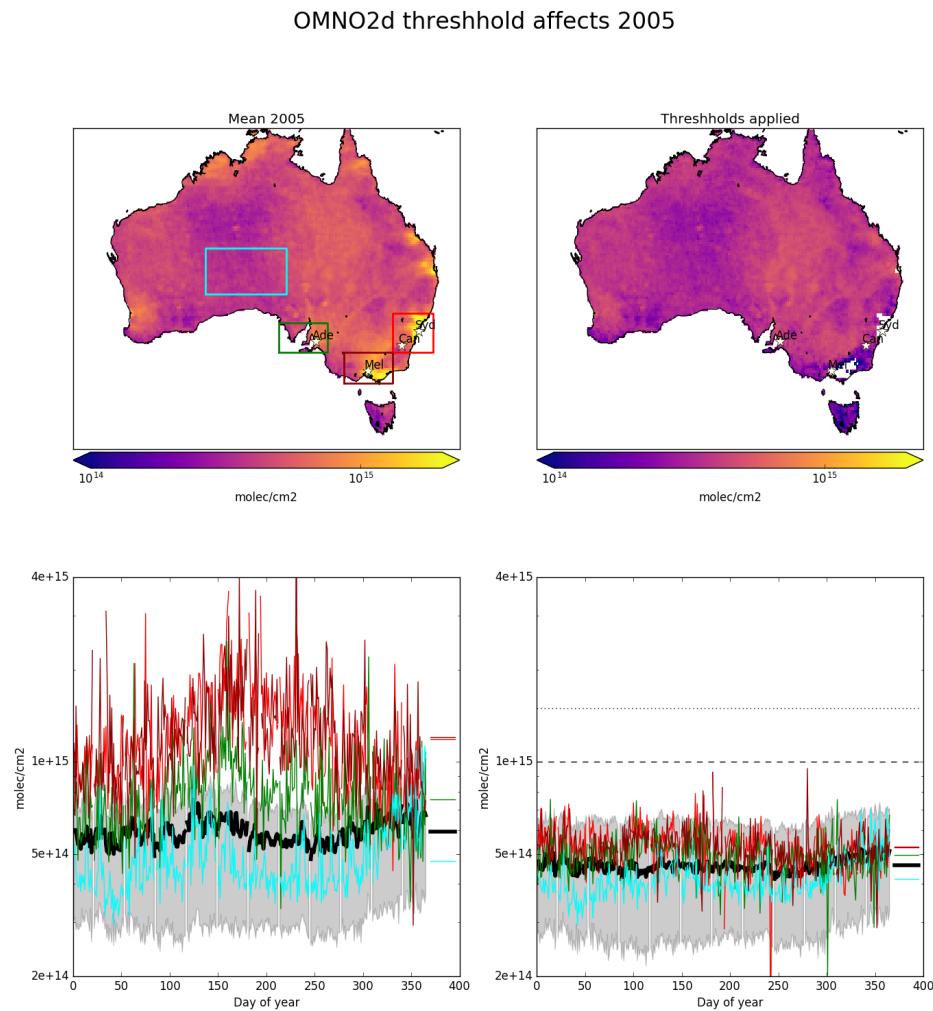


FIGURE 2.33: 2005 OMNO2d NO₂ column mean before (left) and after (right) applying the threshold filters as described in the text. Time series for Australia, and each region (by colour) shown in the bottom panel, with mean for that region shown on the right.

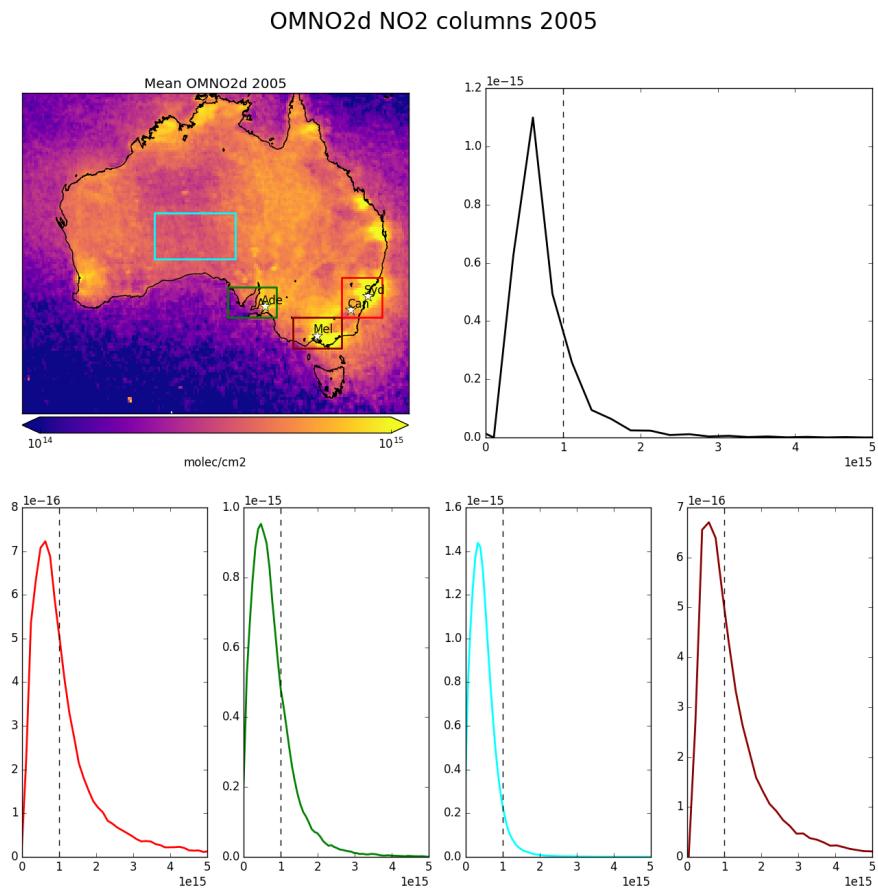


FIGURE 2.34: 2005 OMNO2d NO₂ column means (top left), along with column amount distributions for Australia (top right) and each region shown in the area map (by colour)

SPEI Monthly standardised precipitation evapotranspiration index (metric to determine drought stress) downloaded from <http://hdl.handle.net/10261/153475> with DOI:10.20350/digitalCSIC/8508. See more information in section

OMHCHO Satellite swaths of HCHO slant columns downloaded from TODO, with DOI TODO

Bibliography

- Young, P J et al. (2017). "Tropospheric Ozone Assessment Report (TOAR): Assessment of global-scale model performance for global and regional ozone distributions, variability, and trends". In: *Elementa: Science of the Anthropocene*, pp. 0–84. ISSN: 2325-1026. DOI: [10.1525/elementa.265](https://doi.org/10.1525/elementa.265). URL: http://eprints.lancs.ac.uk/88836/1/TOAR_Model_Performance_07062017.pdfhttp://www.igacproject.org/sites/default/files/2017-05/TOAR-Model_Performance_draft_for_open_comment.pdf.
- Dunne, Erin et al. (2017). "Comparison of VOC measurements made by PTR-MS, Adsorbent Tube/GC-FID-MS and DNPH-derivatization/HPLC during the Sydney Particle Study, 2012: a contribution to the assessment of uncertainty in current atmospheric VOC measurements". In: *Atmospheric Measurement Techniques Discussions*, pp. 1–24. ISSN: 1867-8610. DOI: [10.5194/amt-2016-349](https://doi.org/10.5194/amt-2016-349). URL: <https://www.atmos-meas-tech-discuss.net/amt-2016-349/>.
- Gonzalez Abad, G. et al. (2015). "Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval". In: *Atmospheric Measurement Techniques* 8.1, pp. 19–32. ISSN: 18678548. DOI: [10.5194/amt-8-19-2015](https://doi.org/10.5194/amt-8-19-2015).
- Kurosu, T and K Chance (2014). OMIReadme. URL: https://www.cfa.harvard.edu/atmosphere/Instruments/OMI/PGEReleases/READMEs/OMHCHO_README_v3.0.pdf.
- Wang, Yuxuan et al. (2017). "Adverse effects of increasing drought on air quality via natural processes". In: *Atmos. Chem. Phys.* 175194, pp. 12827–12843. DOI: [10.5194/acp-17-12827-2017](https://doi.org/10.5194/acp-17-12827-2017). URL: <https://www.atmos-chem-phys.net/17/12827/2017/acp-17-12827-2017.pdf>.
- Chen, D. et al. (2009). "Regional CO pollution in China simulated by the high-resolution nested-grid GEOS-Chem model". In: *Atmospheric Chemistry and Physics Discussions* 9.2, pp. 5853–5887. ISSN: 1680-7324. DOI: [10.5194/acpd-9-5853-2009](https://doi.org/10.5194/acpd-9-5853-2009). URL: [http://dx.doi.org/10.5194/acp-9-3825-2009](https://doi.org/10.5194/acp-9-3825-2009).
- Giglio, Louis, James T. Randerson, and Guido R. Van Der Werf (2013). "Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4)". In: *Journal of Geophysical Research* 118.1, pp. 317–328. ISSN: 21698961. DOI: [10.1002/jgrg.20042](https://doi.org/10.1002/jgrg.20042).
- Horowitz, Larry W. et al. (1998). "Export of reactive nitrogen from North America during summertime: Sensitivity to hydrocarbon chemistry". In: *Journal of Geophysical Research* 103.D11, pp. 13451–13476. ISSN: 0148-0227. DOI: [10.1029/97JD03142](https://doi.org/10.1029/97JD03142). URL: <http://doi.wiley.com/10.1029/97JD03142><http://www.agu.org/pubs/crossref/1998/97JD03142.shtml>.
- Mao, Jingqiu et al. (2013). "Ozone and organic nitrates over the eastern United States: Sensitivity to isoprene chemistry". In: *Journal of Geophysical Research Atmospheres* 118.19, pp. 11256–11268. ISSN: 21698996. DOI: [10.1002/jgrd.50817](https://doi.org/10.1002/jgrd.50817).

- Paulot, F. et al. (2009a). "Isoprene photooxidation: new insights into the production of acids and organic nitrates". In: *Atmospheric Chemistry and Physics* 9.4, pp. 1479–1501. ISSN: 1680-7324. DOI: [10.5194/acp-9-1479-2009](https://doi.org/10.5194/acp-9-1479-2009).
- Paulot, Fabien et al. (2009b). "Unexpected Epoxide Formation in the". In: *Science* 325.2009, pp. 730–733. ISSN: 0036-8075. DOI: [10.1126/science.1172910](https://doi.org/10.1126/science.1172910).
- Crounse, John D et al. (2011). "Peroxy radical isomerization in the oxidation of isoprene". In: *Physical Chemistry Chemical Physics* 13.30, pp. 13607–13613. ISSN: 1463-9076. DOI: [doi : 10 . 1039 / c1cp21330j](https://doi.org/10.1039/c1cp21330j). URL: [http : // dx . doi . org / 10 . 1039 / C1CP21330J](http://dx.doi.org/10.1039/C1CP21330J).
- Crounse, John D et al. (2012). "Atmospheric Fate of Methacrolein. 1. Peroxy Radical Isomerization Following Addition of OH and O₂". In: *Physical Chemistry* m.
- Peeters, Jozef and Jean-Francis Muller (2010). "HO_x radical regeneration in isoprene oxidation via peroxy radical isomerisations. II: experimental evidence and global impact". In: *Physical Chemistry Chemical Physics* 12.42, p. 14227. ISSN: 1463-9076. DOI: [10.1039/c0cp00811g](https://doi.org/10.1039/c0cp00811g). URL: <http://pubs.rsc.org/en/content/articlepdf/2010/cp/c0cp00811g>.
- Taraborrelli, D. et al. (2012). "Hydroxyl radical buffered by isoprene oxidation over tropical forests". In: *Nature Geoscience* 5.3, pp. 190–193. ISSN: 17520894. DOI: [10 . 1038 / ngeo1405](https://doi.org/10.1038/ngeo1405). URL: [http : // dx . doi . org / 10 . 1038 / ngeo1405](http://dx.doi.org/10.1038/ngeo1405)[https : // www . nature . com / articles / ngeo1405 . pdf](https://www.nature.com/articles/ngeo1405.pdf).
- Wolfe, Glenn M. et al. (2012). "Photolysis, OH reactivity and ozone reactivity of a proxy for isoprene-derived hydroperoxyenals (HPALDs)". In: *Physical Chemistry Chemical Physics* 14.20, p. 7276. ISSN: 1463-9076. DOI: [10 . 1039 / c2cp40388a](https://doi.org/10.1039/c2cp40388a). URL: <http://xlink.rsc.org/?DOI=c2cp40388a>.
- Jozef, Peeters et al. (2014). "Hydroxyl Radical Recycling in Isoprene Oxidation Driven by Hydrogen Bonding and Hydrogen Tunneling: The Upgraded LIM1 Mechanism". In: *Journal of Physical Chemistry*.
- Mao, J. et al. (2012). "Insights into hydroxyl measurements and atmospheric oxidation in a California forest". In: *Atmospheric Chemistry and Physics* 12.17, pp. 8009–8020. ISSN: 16807316. DOI: [10.5194/acp-12-8009-2012](https://doi.org/10.5194/acp-12-8009-2012).
- Kefauver, Shawn C., Iolanda Filella, and Josep Peñuelas (2014). "Remote sensing of atmospheric biogenic volatile organic compounds (BVOCs) via satellite-based formaldehyde vertical column assessments". en. In: *International Journal of Remote Sensing*. URL: [http://www.tandfonline.com/doi/abs/10.1080/01431161.2014.968690# .VqxEubNM61M](http://www.tandfonline.com/doi/abs/10.1080/01431161.2014.968690#.VqxEubNM61M).
- Guenther, Alex (2016). MEGAN. URL: <http://lar.wsu.edu/megan/>.
- Guenther, A. B. et al. (2012). "The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions". In: *Geoscientific Model Development* 5.6, pp. 1471–1492. ISSN: 1991959X. DOI: [10.5194/gmd-5-1471-2012](https://doi.org/10.5194/gmd-5-1471-2012).
- Guenther, A et al. (2006). "Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature)". In: *Atmospheric Chemistry and Physics* 6.11, pp. 3181–3210. DOI: [10 . 5194 / acp - 6 - 3181 - 2006](https://doi.org/10.5194/acp-6-3181-2006). URL: [http : // dx . doi . org / 10 . 5194 / acp - 6 - 3181 - 2006](http://dx.doi.org/10.5194/acp-6-3181-2006).
- Niinemets, U. et al. (1999). "A model of isoprene emission based on energetic requirements for isoprene synthesis and leaf photosynthetic properties for Liquidambar

- and Quercus". In: *Plant, Cell and Environment* 22.11, pp. 1319–1335. ISSN: 01407791. DOI: [10.1046/j.1365-3040.1999.00505.x](https://doi.org/10.1046/j.1365-3040.1999.00505.x).
- Guenther, Alex et al. (2000). "Natural emissions of non-methane volatile organic compounds, carbon monoxide, and oxides of nitrogen from North America". In: *Atmospheric Environment* 34.12-14, pp. 2205–2230. ISSN: 13522310. DOI: [10.1016/S1352-2310\(99\)00465-3](https://doi.org/10.1016/S1352-2310(99)00465-3).
- Sindelarova, K. et al. (2014). "Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years". In: *Atmospheric Chemistry and Physics* 14.17, pp. 9317–9341. ISSN: 16807324. DOI: [10.5194/acp-14-9317-2014](https://doi.org/10.5194/acp-14-9317-2014). arXiv: [arXiv:1011.1669v3](https://arxiv.org/abs/1011.1669v3).
- Jin, Xiaomeng et al. (2017). "Evaluating a Space-Based Indicator of Surface Ozone-NO_x-VOC Sensitivity Over Midlatitude Source Regions and Application to Decadal Trends". In: *Journal of Geophysical Research: Atmospheres* 122.19, pp. 10439–10461. ISSN: 21698996. DOI: [10.1002/2017JD026720](https://doi.org/10.1002/2017JD026720).
- Martin, Randall V et al. (2002b). "Interpretation of TOMS observations of tropical tropospheric ozone with a global model and in situ observations". In: 107. DOI: [10.1029/2001JD001480](https://doi.org/10.1029/2001JD001480).
- Leue, C et al. (2001). "Quantitative analysis of NO x emissions from Global Ozone Monitoring Experiment satellite image sequences". In: *J. Geophys. Res.* 106.D6, p. 5493. DOI: [10.1029/2000jd900572](https://doi.org/10.1029/2000jd900572). URL: <http://dx.doi.org/10.1029/2000jd900572>.
- Kwon, Hyeong-Ahn et al. (2017). "Sensitivity of formaldehyde (HCHO) column measurements from a geostationary satellite to temporal variation of the air mass factor in East Asia". In: *Atmospheric Chemistry and Physics* 17.7, pp. 4673–4686. ISSN: 1680-7324. DOI: [10.5194/acp-17-4673-2017](https://doi.org/10.5194/acp-17-4673-2017). URL: <http://www.atmos-chem-phys.net/17/4673/2017/>.
- Palmer, Paul I et al. (2001). "Air mass factor formulation for spectroscopic measurements from satellites' Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment". In: *Journal of Geophysical Research* 106.D13.
- Martin, Randall V. et al. (2002a). "An improved retrieval of tropospheric nitrogen dioxide from GOME". In: *Journal of Geophysical Research D: Atmospheres* 107.20. ISSN: 01480227. DOI: [10.1029/2001JD001027](https://doi.org/10.1029/2001JD001027).
- De Smedt, I. et al. (2015). "Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations". In: *Atmospheric Chemistry and Physics* 15.21, pp. 12519–12545. ISSN: 16807324. DOI: [10.5194/acp-15-12519-2015](https://doi.org/10.5194/acp-15-12519-2015). URL: <http://www.atmos-chem-phys-discuss.net/15/12241/2015/5Cnhttp://www.atmos-chem-phys-discuss.net/15/12241/2015/acpd-15-12241-2015.pdf>.
- Schenkeveld, V. M. Erik et al. (2017). "In-flight performance of the Ozone Monitoring Instrument". In: *Atmospheric Measurement Techniques* 10.5, pp. 1957–1986. ISSN: 1867-8548. DOI: [10.5194/amt-10-1957-2017](https://doi.org/10.5194/amt-10-1957-2017). URL: <http://www.atmos-meas-tech-discuss.net/amt-2016-420/https://www.atmos-meas-tech.net/10/1957/2017/>.
- Huang, Guanyu et al. (2017). "Validation of 10-year SAO OMI Ozone Profile (PROFOZ) Product Using Aura MLS Measurements". In: *Atmospheric Measurement Techniques Discussions*, pp. 1–25. ISSN: 1867-8610. DOI: [10.5194/amt-2017-92](https://doi.org/10.5194/amt-2017-92). URL: <https://www.atmos-meas-tech-discuss.net/amt-2017-92/>.

- Lorente, Alba et al. (2017). "Structural uncertainty in air mass factor calculation for NO₂ and HCHO satellite retrievals". In: *Atmospheric Measurement Techniques* 2, pp. 1–35. ISSN: 1867-8610. DOI: [10.5194/amt-10-759-2017](https://doi.org/10.5194/amt-10-759-2017). URL: <https://www.atmos-meas-tech.net/10/759/2017/amt-10-759-2017.html> <http://www.atmos-meas-tech-discuss.net/amt-2016-306/>.
- Instrument, O M I (2002). "OMI Algorithm Theoretical Basis Document Volume I". In: I.August, pp. 1–50.
- Millet, Dylan B et al. (2006). "Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission". In: *J. Geophys. Res.* 111.D24. DOI: [10.1029/2005jd006853](https://doi.org/10.1029/2005jd006853). URL: [TODO](#).
- Eskes, H J and K F Boersma (2003). "Averaging kernels for DOAS total-column satellite retrievals". In: *Atmospheric Chemistry and Physics* 3.1, pp. 1285–1291. ISSN: 1680-7324. DOI: [10.5194/acp-3-1285-2003](https://doi.org/10.5194/acp-3-1285-2003). URL: <http://dx.doi.org/10.5194/acpd-3-895-2003>.
- Lamsal, L N et al. (2014). "Evaluation of OMI operational standard NO₂ column retrievals using in situ and surface-based NO₂ observations". In: *Atmos. Chem. Phys* 14, pp. 11587–11609. DOI: [10.5194/acp-14-11587-2014](https://doi.org/10.5194/acp-14-11587-2014). URL: www.atmos-chem-phys.net/14/11587/2014/.
- Palmer, Paul I (2003). "Mapping isoprene emissions over North America using formaldehyde column observations from space". In: *J. Geophys. Res.* 108.D6. DOI: [10.1029/2002jd002153](https://doi.org/10.1029/2002jd002153). URL: <http://dx.doi.org/10.1029/2002jd002153>.
- Marais, E A et al. (2012). "Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns". In: *Atmospheric Chemistry and Physics* 12.3, pp. 7475–7520. DOI: [10.5194/acp-12-6219-2012](https://doi.org/10.5194/acp-12-6219-2012). URL: <http://dx.doi.org/10.5194/acp-12-6219-2012>.
- Vasilkov, A et al. (2017). "Accounting for the effects of surface BRDF on satellite cloud and trace-gas retrievals: a new approach based on geometry-dependent Lambertian equivalent reflectivity applied to OMI algorithms". In: *Atmospheric Measurement Techniques* 10.1, pp. 333–349. DOI: [10.5194/amt-10-333-2017](https://doi.org/10.5194/amt-10-333-2017). URL: <http://www.atmos-meas-tech.net/10/333/2017/>.
- Vigouroux, C. et al. (2009). "Ground-based FTIR and MAX-DOAS observations of formaldehyde at Réunion Island and comparisons with satellite and model data". In: *Atmospheric Chemistry and Physics Discussions* 9, pp. 15891–15957. ISSN: 1680-7316. DOI: [10.5194/acpd-9-15891-2009](https://doi.org/10.5194/acpd-9-15891-2009).
- Dufour, G. et al. (2008). "SCIAMACHY formaldehyde observations: constraint for isoprene emissions over Europe?" In: *Atmospheric Chemistry and Physics* 8.6, pp. 19273–19312. ISSN: 1680-7324. DOI: [10.5194/acpd-8-19273-2008](https://doi.org/10.5194/acpd-8-19273-2008).
- Millet, Dylan B. et al. (2008). "Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor". In: *Journal of Geophysical Research Atmospheres* 113.2, pp. 1–18. ISSN: 01480227. DOI: [10.1029/2007JD008950](https://doi.org/10.1029/2007JD008950).
- Zhu, Lei et al. (2014). "Anthropogenic emissions of highly reactive volatile organic compounds in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns". In: *Environmental Research Letters* 9.11, p. 114004. ISSN: 1748-9326. DOI: [10.1088/1748-9326/9/11/114004](https://doi.org/10.1088/1748-9326/9/11/114004). URL: <http://stacks.iop.org/1748-9326/9/i=11/a=114004?key=crossref.3d2869ee02fd4f0792f831ac8cbe117>.

- Marais, E A et al. (2014). "Improved model of isoprene emissions in Africa using Ozone Monitoring Instrument (OMI) satellite observations of formaldehyde: implications for oxidants and particulate matter". In: *Atmospheric Chemistry and Physics* 14.15, pp. 7693–7703. DOI: [10.5194/acp-14-7693-2014](https://doi.org/10.5194/acp-14-7693-2014). URL: <http://dx.doi.org/10.5194/acp-14-7693-2014>.
- Surl, Luke, Paul I Palmer, and Gonzalo González Abad (2018). "Which processes drive observed variations of HCHO columns over India?" In: *Atmospheric Chemistry and Physics* 18.March, pp. 4549–4566. DOI: [10.5194/acp-18-4549-2018](https://doi.org/10.5194/acp-18-4549-2018). URL: <https://doi.org/10.5194/acp-18-4549-2018>.
- Shim, Changsub et al. (2005). "Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements". In: *Journal of Geophysical Research Atmospheres* 110.24, pp. 1–14. ISSN: 01480227. DOI: [10.1029/2004JD005629](https://doi.org/10.1029/2004JD005629).
- Abad, Gonzalo González et al. (2016). "Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval". In: *Atmospheric Measurement Techniques* 9.7, pp. 2797–2812. ISSN: 18678548. DOI: [10.5194/amt-9-2797-2016](https://doi.org/10.5194/amt-9-2797-2016).
- De Smedt, I. et al. (2012). "Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues". In: *Atmospheric Measurement Techniques* 5.11, pp. 2933–2949. ISSN: 18671381. DOI: [10.5194/amt-5-2933-2012](https://doi.org/10.5194/amt-5-2933-2012).
- Zhu, Lei et al. (2016). "Observing atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC4RS aircraft observations over the Southeast US". In: *Atmospheric Chemistry and Physics* 0, pp. 1–24. ISSN: 1680-7375. DOI: [10.5194/acp-2016-162](https://doi.org/10.5194/acp-2016-162). URL: <http://www.atmos-chem-phys.net/16/13477/2016/acp-16-13477-2016.pdf>.
- Barkley, Michael P. et al. (2013). "Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns". In: *Journal of Geophysical Research Atmospheres* 118.12, pp. 6849–6868. ISSN: 21698996. DOI: [10.1002/jgrd.50552](https://doi.org/10.1002/jgrd.50552). URL: <http://dx.doi.org/10.1002/jgrd.50552>.
- Bauwens, Maite et al. (2016). "Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations". In: *Atmospheric Chemistry and Physics Discussions* March, pp. 1–45. ISSN: 1680-7375. DOI: [10.5194/acp-2016-221](https://doi.org/10.5194/acp-2016-221). URL: <http://www.atmos-chem-phys-discuss.net/acp-2016-221/>.
- Wikipedia (2016). *Solar zenith angle*. DOI: [10.1016/B978-012369407-2/50005-X](https://doi.org/10.1016/B978-012369407-2/50005-X). URL: <http://sacs.aeronomie.be/info/sza.php>.
- Spurr, R. J D (2002). "Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical discrete ordinate radiative transfer treatment". In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 75.2, pp. 129–175. ISSN: 00224073. DOI: [10.1016/S0022-4073\(01\)00245-X](https://doi.org/10.1016/S0022-4073(01)00245-X).
- Martin, Randall V et al. (2003). "Global inventory of nitrogen oxide emissions constrained by space-based observations of NO₂ columns". In: 108.2, pp. 1–12. DOI: [10.1029/2003JD003453](https://doi.org/10.1029/2003JD003453).
- De Smedt, I et al. (2008). "Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors". In: *Atmos. Chem. Phys.*

- 8.16, pp. 4947–4963. ISSN: 1680-7324. DOI: [10.5194/acp-8-4947-2008](https://doi.org/10.5194/acp-8-4947-2008). URL: <http://www.atmos-chem-phys.net/8/4947/2008/>.
- Wolfe, G. M. et al. (2016). "Formaldehyde production from isoprene oxidation across NO_x regimes". In: *Atmospheric Chemistry and Physics* 16.x, pp. 2597–2610. ISSN: 16807324. DOI: [10.5194/acp-16-2597-2016](https://doi.org/10.5194/acp-16-2597-2016). URL: www.atmos-chem-phys.net/16/2597/2016/.
- Miller, Christopher Chan et al. (2016). "Glyoxal yield from isoprene oxidation and relation to formaldehyde: chemical mechanism, constraints from SENEX aircraft observations, and interpretation of OMI satellite data". In: *Atmospheric Chemistry and Physics Discussions* x, pp. 1–25. ISSN: 1680-7375. DOI: [10.5194/acp-2016-1042](https://doi.org/10.5194/acp-2016-1042). URL: <http://www.atmos-chem-phys-discuss.net/acp-2016-1042/>.
- Zhu, Lei et al. (2013). "Variability of HCHO over the Southeastern United States observed from space : Implications for VOC emissions". In: vol. 1.
- Shao, Yaping et al. (2007). "Numerical simulation of the October 2002 dust event in Australia". In: *J. Geophys. Res.* 112.D8. DOI: [10.1029/2006jd007767](https://doi.org/10.1029/2006jd007767). URL: <http://dx.doi.org/10.1029/2006jd007767>.