

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 22, 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



UNIVERSITY OF WOLLONGONG

*Abstract*

Science Medicine And Health  
Chemistry Department

Doctor of Philosophy

**Thesis Title**

by Jesse GREENSLADE

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...



## *Acknowledgements*

Thanks to my supervisor Jenny Fisher, the whole Atmospheric Chemistry team at Wollongong who made me feel at home: Clare Murphy, Dagmar, Elise, Joel, Kaitlyn, Max, Nick Deutscher, Nick Jones, Stephen Wilson,

Thanks also to my personal saviour and her noodly appendages.



# Contents

<b>Declaration of Authorship</b>	<b>iii</b>
<b>Abstract</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>ix</b>
<b>1 Introduction and Literature Review</b>	<b>1</b>
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
1.3.3.1 Oxidation . . . . .	15
1.3.3.2 High NO <sub>x</sub> pathway . . . . .	17
1.3.3.3 Low NO <sub>x</sub> pathway . . . . .	17
1.3.3.4 Night time processes . . . . .	18
1.4 Formaldehyde . . . . .	19
1.4.1 Sources and sinks . . . . .	19
1.4.2 Measurement techniques . . . . .	20
1.4.2.1 Satellite measurements . . . . .	22
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 . . . . .	26
1.5.4.1 Emissions Inventories . . . . .	26
1.5.4.2 Resolution . . . . .	26
1.5.4.3 Chemistry mechanisms . . . . .	26
1.5.4.4 Clouds . . . . .	27
1.5.4.5 Soil Moisture . . . . .	27
1.6 Australia and the southern hemisphere . . . . .	28

1.6.1	Ozone	28
1.6.2	VOCs	30
1.6.3	Measurements	31
1.7	Aims	31
<b>2</b>	<b>Data and Modelling</b>	<b>33</b>
2.1	Introduction	33
2.2	Measurement Techniques	33
2.2.1	DOAS	34
2.2.2	Satellites	35
2.2.2.1	LIDORT	36
2.2.2.2	OMI	36
2.2.2.3	Air mass factor (AMF)	37
2.2.2.4	Uncertainties	39
2.3	Datasets	41
2.3.1	Satellite datasets	42
2.3.1.1	OMHCHO	42
2.3.1.2	OMNO <sub>2</sub>	44
2.3.1.3	OMAERUVd	44
2.3.1.4	MOD14A1	46
2.3.1.5	AIRS CO columns	46
2.3.2	Model datasets	46
2.3.2.1	Meteorological reanalysis	46
2.3.3	Campaign datasets	46
2.3.3.1	Daintree	46
2.3.3.2	Marine and Urban MBA ? (MUMBA)	48
2.3.3.3	Sydney Particle Studies (SPS1, SPS2)	48
2.3.3.4	Ozonesondes	48
2.4	GEOS-Chem	48
2.4.1	Overview	48
2.4.2	Running GEOS-Chem (before isop?)	51
2.4.2.1	Installation and requirements	51
2.4.3	GEOS-Chem isoprene modelling	51
2.4.4	Chemical Mechanisms	53
2.4.5	Emissions from MEGAN	53
2.4.6	Rescaling NO <sub>x</sub>	55
2.4.7	GEOS-Chem outputs	66
2.4.8	GEOS-Chem simulations	66
2.4.8.1	Run comparisons	67
2.5	Calculating an AMF	70
2.6	Recalculation of OMI HCHO	73
2.6.1	Outline	73
2.6.2	Creating new shape factors	75
2.6.3	Reading satellite data	77
2.6.4	Creating the new AMF	77
2.6.5	Recalculating the AMF using PP code	78
2.6.6	Vertical columns from AMF	79

2.6.7	Reference sector correction . . . . .	79
2.6.8	Binning the results daily . . . . .	82
2.6.9	Difference between new and old OMI HCHO columns . . . . .	83
2.7	Filtering Data . . . . .	83
2.7.1	Fire and smoke . . . . .	85
2.7.1.1	Checking that fire masks are influencing pyrogenic HCHO	86
2.7.2	NO <sub>x</sub> . . . . .	86
2.7.3	Summary of filters effects on HCHO . . . . .	91
2.8	Data Access . . . . .	91
<b>3</b>	<b>Biogenic Isoprene emissions in Australia</b>	<b>95</b>
3.1	Introduction . . . . .	95
3.1.1	Aims . . . . .	95
3.1.2	satellite inversions . . . . .	96
3.1.3	MEGAN emission model . . . . .	96
3.1.4	Top-down emissions estimates . . . . .	97
3.1.4.1	Linear . . . . .	97
3.1.4.2	Bayesian . . . . .	97
3.2	Methods . . . . .	98
3.2.1	Outline . . . . .	98
3.2.2	Calculating modelled slope . . . . .	99
3.2.3	Satellite inversion . . . . .	101
3.2.4	Calculation of Emissions . . . . .	103
3.2.5	Emissions drivers . . . . .	107
3.2.6	HCHO Products and yield . . . . .	107
3.2.7	Accounting for smearing . . . . .	108
3.3	Results . . . . .	110
3.3.1	Emissions comparisons . . . . .	110
3.3.2	Emissions affect on GEOS-Chem . . . . .	111
3.3.3	Comparison with in-situ measurements . . . . .	111
3.4	Uncertainty . . . . .	111
3.4.1	Model Uncertainty . . . . .	112
3.4.2	Satellite Uncertainty . . . . .	112
3.4.3	Fire Filtering . . . . .	113
3.4.4	MEGAN . . . . .	113
<b>4</b>	<b>Stratospheric ozone intrusions</b>	<b>117</b>
4.1	Introduction . . . . .	117
4.2	Data and Methods . . . . .	119
4.2.1	Ozonesonde record in the Southern Ocean . . . . .	119
4.2.2	Model description . . . . .	122
4.2.3	Characterisation of STT events and associated fluxes . . . . .	123
4.2.4	Biomass burning influence . . . . .	125
4.2.5	Classifying synoptic conditions during STT events . . . . .	126
4.3	STT event climatologies . . . . .	126
4.4	Simulated ozone columns . . . . .	130
4.5	Stratosphere-to-troposphere ozone flux from STT events . . . . .	134

4.5.1	Method . . . . .	134
4.5.2	Results . . . . .	136
4.5.3	Comparison to literature . . . . .	136
4.6	Sensitivities and limitations . . . . .	141
4.6.1	Event detection . . . . .	141
4.6.2	Flux calculations . . . . .	142
4.7	Conclusions . . . . .	143
4.8	Contributions and Acknowledgements . . . . .	144
<b>A</b>	<b>Supplementary Notes</b>	<b>145</b>
A.1	Measurement Techniques . . . . .	145
A.1.1	MAX-DOAS . . . . .	145
A.2	Data sets . . . . .	145
A.2.1	SPEI . . . . .	145
A.2.2	GOME . . . . .	147
A.2.3	NPI . . . . .	147
A.3	Chemistry . . . . .	147
A.3.0.1	SOA . . . . .	147
A.3.1	Relationship to Glyoxyl TODO: remove if never used . . . . .	151
A.4	CAABA/MECCA . . . . .	151
A.4.0.1	CAABA/MECCA outputs . . . . .	152
A.4.1	CAABA/MECCA Box model: isoprene source classifications . . . . .	152
A.5	Satellite Stuff . . . . .	154
A.5.1	OMI Algorithm BOAS . . . . .	154
A.5.2	AMF recaulculation using 72 level output . . . . .	155
A.5.3	Old Fire Product MYD14C8H . . . . .	157
<b>B</b>	<b>Appendix A</b>	<b>161</b>
B.1	Grid Resolution . . . . .	161
<b>C</b>	<b>Frequently Asked Questions</b>	<b>163</b>
C.1	How do I change the colors of links? . . . . .	163
<b>Bibliography</b>		<b>165</b>

# List of Abbreviations

<b>CTM</b>	Chemical Transport Model
<b>DOAS</b>	Differential Optical Absorption Spectroscopy
<b>HCHO</b>	Formaldehyde
<b>(S,P)OA</b>	(Secondary, Primary) Organic Aerosols
<b>PAN</b>	PeroxyAcetyl Nitrate
<b>PM</b>	Particulate Matter
<b>STT</b>	Stratosphere to Troposphere Transport
<b>(NM)VOC</b>	(Non-Methane) Volatile Organic Compounds



# Physical Constants

Speed of Light  $c = 2.997\,924\,58 \times 10^8 \text{ m s}^{-1}$  (exact)



# List of Symbols

$C_x$	mixing ratio or mole fraction of gas x	molecules of x per molecule of air
$n_x$	Number density of gas x	molecules of x per unit volume of air
$\Omega$	Vertical column	molec cm <sup>-2</sup>



## Chapter 1

# Introduction and Literature Review

### 1.1 The atmosphere

The atmosphere is made up of gases held to the earth's surface by gravity. These gases undergo transport on all scales, from barbecue smoke being blown about the garden, to smoke plumes from forest fires travelling across the world and depositing in the Antarctic snow. They take part in innumerable chemical reactions along the way, largely driven by solar input and interactions with each other. Many gases are lofted into the atmosphere by soil, trees, factories, cars, seas and oceans. They are also deposited back to the surface both directly and in rainfall.

The atmosphere is made up of nitrogen ( $N_2$ :  $\sim 78\%$ ), oxygen ( $O_2$ :  $\sim 21\%$ ), and argon ( $Ar$ :  $\sim 1\%$ ), along with water ( $H_2O$ ) and *trace gases* (those that make up less than 1% of the atmosphere). Water ( $H_2O$ ) ranges from 0.001 to 1% depending on evaporation and precipitation. Beyond these major constituents the atmosphere has a vast number of trace gases, including carbon dioxide ( $CO_2$ :  $\sim 0.4\%$ ), ozone ( $O_3$ : .000001 to 0.001%), and methane ( $CH_4$ :  $\sim 0.4\%$ ) (Brasseur and Jacob 2017, Ch. 2). Trace gases in the atmosphere can have a large impact on living conditions. They react in complex ways with other elements (anthropogenic and natural), affecting all surface ecosystems upon which life depends.

One important trace gas is ozone ( $O_3$ ), which affects climate, human health, and ecosystem productivity. This thesis will focus on ozone in the troposphere, which is relatively uncertain over Australia.

Ozone in the lower atmosphere is a serious hazard that causes health problems (Hsieh and Liao 2013), damages agricultural crops worth billions of dollars (Avnery et al. 2013; Yue et al. 2017), and increases the rate of climate warming (Myhre and Shindell 2013). Around 5 to 20 percent of all air pollution related deaths are due to ozone (Monks et al. 2015), roughly .8 million deaths per year (Lelieveld et al. 2013). In the short term, ozone concentrations of  $\sim 50\text{-}60$  ppbv over eight hours or  $\sim 80$  ppbv over one hour are agreed to constitute a human health hazard (Ayers and Simpson 2006; Lelieveld et al. 2009). Long term exposure causes problems with crop loss and ecosystem damage (Ashmore, Emberson, and Murray Frank 2003), and concentrations may get worse in the future (Lelieveld et al. 2009; Stevenson et al. 2013). Further tropospheric ozone enhancements are projected to drive reductions in global crop yields equivalent to losses of up to \$USD<sub>2000</sub> 35 billion per year by 2030 (Avnery et al. 2013), along with detrimental health outcomes equivalent to  $\sim$ \$USD<sub>2000</sub> 11.8 billion per year



FIGURE 1.1: Pressure (red) logarithmically decreasing, shown with percentage of atmosphere below at several points. Temperature (green) changes throughout the atmosphere. Figure edited from <https://climate.ncsu.edu/Structure>.

by 2050 (Selin et al. 2009). Recently Yue et al. (2017) showed that the net effect of near-surface ozone on is a  $\sim 14\%$  decrease in net primary productivity in China. They state that reducing this decrease by  $\sim 70\%$  before 2030 would require drastic measures.

### 1.1.1 Structure

Most of the atmosphere ( $\sim 85\%$ ) is within 10 km of the earth's surface. This is due to air pressure, which decreases logarithmically with altitude. Any entity is subjected to the weight of all the air above it, and the density of the atmosphere is driven by this pressure.

The atmosphere extends above the earth's surface to the edges of space. This is split into various layers, defined by the *lapse rate*: the decrease in temperature ( $T$ ) with increasing altitude ( $z$ ), or  $\frac{-dT}{dz}$ . Figure 1.1 shows the pressure and temperature profiles against altitude through the atmosphere. The first layer is the troposphere, which extends to roughly 10 km and is characterised by positive lapse rate (or decreasing temperature with altitude). At the top of the troposphere (the tropopause) the temperature stops decreasing, and then the stratosphere is defined by a negative lapse rate. This is due to UV radiation being absorbed by ozone, and leads to a very vertically stable environment.

In addition to these atmospheric layers, the troposphere can be subset into the *boundary layer* and the *free troposphere*. The *boundary layer* is the lowest layer and involves increased atmospheric mixing due to ground heating and friction effects. It generally extends anywhere from 200 - 1000 m, above which the ground effects have

fewer direct impacts. The *free troposphere* is the remainder of the troposphere and is more affected by transport, both horizontally and from the stratosphere.

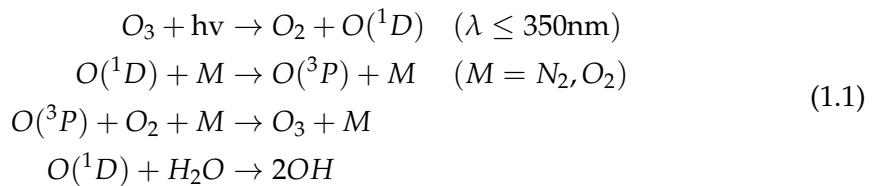
### 1.1.2 Composition and chemistry

There are a myriad of trace gases in the atmosphere, emitted by plants, animals, earth, and water. These gases react with one another and over time they either deposit back onto the earth or form more stable compounds such as CO<sub>2</sub>. Oxidation and photolysis (the process of being broken apart by photons) are the two main processes whereby compounds are broken down in the atmosphere.

OH and HO<sub>2</sub> concentrations largely determine the oxidative capacity of the atmosphere. Concentration of the OH radical drives many processes in the atmosphere, especially during the day when photolysis of ozone produces OH (Atkinson 2000). OH is a key species which reacts with nearly all the organic compounds in the troposphere, with only a few exceptions (Atkinson 2000). Over land, isoprene (C<sub>5</sub>H<sub>8</sub>) and monoterpenes (C<sub>10</sub>H<sub>16</sub>) account for 50% and 30% of the OH reactivity respectively (Fuentes et al. 2000).

Since radicals are involved in all oxidative chemistry in the atmosphere it is important for models to accurately represent them (eg. Travis2014). This is difficult as they are coupled with so many other species and measurements of OH are not readily available on a global scale. In the late 90's it was thought that OH radicals were formed exclusively from photolysis of O<sub>3</sub>, HONO, HCHO, and other carbonyls (R<sub>2</sub>C=O) (Atkinson 2000). It has been shown since that OH is recycled in various processes. Isoprene (C<sub>5</sub>H<sub>8</sub>) was thought to be a sink of OH until it was shown by Paulot et al. (2009b) that the radicals are recycled. This recycling process is discussed in more detail in section 1.3.3.

Ozone is an important precursor to OH, as excited oxygen atoms (O(<sup>1</sup>D)) are created through its photolysis, which then go on to react with water to form OH, as shown in this reaction sequence (Atkinson 2000; Atkinson and Arey 2003):



Where  $h\nu$  represents radiation and M is an inert molecule. This shows that some of the O(<sup>1</sup>D) recycles back to ozone, while some forms OH.

### 1.1.3 Radiative Forcing

One of the larger uncertainties in atmospheric modelling is how particles in the atmosphere affect radiative forcing. For 12 years it has been understood that aerosols overall cool the atmosphere, with smaller particles having a larger effect as they matching the wavelengths of visible light (Kanakidou et al. 2005). Aerosol products from gas phase emissions (or the children thereof) play an indirect and complex role in cloud properties, with a net cooling effect (Kanakidou et al. (2005), Stocker et al. (IPCC, 2013:

*Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Chapter 7,8)).

Transport and indirect effects complicate matters further, with cloud creation and modification of cloud properties being quite difficult to accurately predict. In the third IPCC report (*Intergovernmental Panel on Climate Change (IPCC): Climate Change: The Scientific Basis 2001*), the uncertainty involved if OA forcing was a factor of 3 times the estimated effect. This has since been improved however OA and cloud formation still remains a large uncertainty in more recent IPCC reports (Forster et al. 2007). Figure 1.2 shows the radiative forcing (RF) of various atmospheric constituents, it is clear that OA uncertainty dominates. Figure 1.3 shows the same summary updated in chapter 8 of the fifth report, where the SOA uncertainty remains quite large.

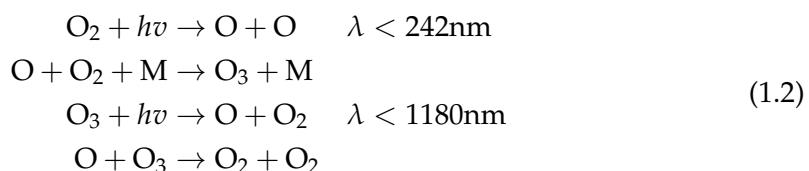
In order to improve understanding of processes involved in radiative forcing, Kanakidou et al. (2005) highlight the need for improving VOC emissions and flux measurements. They also advocate utilising satellite data in models as a means of improving the emissions inventories. VOCs can lead to changes in cloud formation, as nucleation can arise from the subsequent SOA. Kanakidou et al. (2005) concluded that it is very likely that organics contribute to particle growth and formation rates.

## 1.2 Ozone

Ozone ( $O_3$ ) is an important greenhouse gas and oxidant. It is mostly located in the stratosphere and prevents much of the shorter wavelength (UV) solar radiation from reaching the earth's surface. Ozone in the troposphere is less beneficial, leading to health issues, radiative forcing (Stevenson et al. 2013), and crop death. Understanding and accurately portraying ozone concentrations in the troposphere is important to allow accurate predictions of future climate. This will become even more important as projections of future climate changes suggest altered vertical mixing rates, ultra violet index (UVI) and ozone RF (Hegglin and Shepherd 2009).

### 1.2.1 Stratospheric ozone

In the stratosphere ozone production is driven by the Chapman mechanism, as high energy radiation (with wavelengths  $\lambda < 242 \text{ nm}$ ) photolyses the molecular oxygen ( $O_2$ ) in the atmosphere (Brasseur and Jacob 2017, Chapter 3, section 2). The Chapman mechanism involves several reactions which lead to rough equilibrium of  $O$ ,  $O_2$ ,  $O_3$  and pressure, as follows:



The high energy photons ( $\lambda < 242 \text{ nm}$ ) are present from the top of the atmosphere but are mostly removed before reaching the troposphere as their energy is used to split the  $O_2$  molecules. The lifetime of  $O$  against loss by  $O_2$  is less than a second in the troposphere, and produced  $O_3$  quickly returns to  $O$  and  $O_2$ , as low energy ( $\lambda < 1180 \text{ nm}$ )



FIGURE 1.2: The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from Forster et al. (2007), found at [https://www.ipcc.ch/publications\\_and\\_data/ar4/wg1/en/faq-2-1.html](https://www.ipcc.ch/publications_and_data/ar4/wg1/en/faq-2-1.html).



FIGURE 1.3: The overall radiative forcings and uncertainties of several atmospheric constituents. This is an image taken from Stocker et al. (*IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*), chapter 8.

photons and M are abundant. The reduced light penetration towards the surface, in addition to the logarithmic increase in atmospheric pressure (which affects M abundance) drives the vertical profile of ozone into what is called the *ozone layer*. This is a layer of relative ozone abundance within the stratosphere. The Chapman mechanism requires radiation so only takes place during the daytime, during the night this process slows to a halt, and the ozone concentrations remain stable unless pollution intrudes (Jacob 1999, Chapter 10).

Since the Montreal Protocol on Substances that Deplete the Ozone Layer was established in August 1987, and ratified in August 1989, several satellites and many measurement stations were set up to monitor ozone in the stratosphere. However, in the southern hemisphere there are relatively few records of ozone (Huang et al. 2017). This affects our ability to accurately determine sources of ozone in the troposphere.

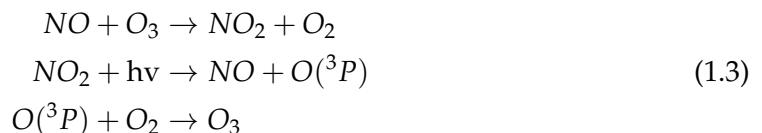
### 1.2.2 Tropospheric ozone

Figure 1.4, copied from Young et al. (2017), shows summary of the major processes and emissions affecting tropospheric ozone. This thesis involves improving the highly uncertain natural emissions of volatile organic compounds (VOCs) from Australia, and estimating impacts from STEs.

Generally there are two main drivers of tropospheric ozone concentrations; transport from the stratosphere and chemical production due to emissions of precursors. Tropospheric ozone is regulated by NO and NO<sub>2</sub> concentrations, which form an equilibrium (Cape 2008; Young et al. 2017). At small to medium scales, pyrogenic (fire) and anthropogenic (man-made) emissions can be important. Smoke plumes from biomass burning can carry ozone precursors, creating higher ozone concentrations downwind of the plume's source. Emissions of precursors from large cities (primarily traffic and power production) can impact ozone concentrations. These impacts are not always straightforwards due to the nonlinear relationship between ozone and its precursors.

NO<sub>x</sub> ( $\equiv$  NO<sub>2</sub> + NO) is another important chemical family in the atmosphere which interacts with ozone and regulates the atmospheric oxidative capacity. NO<sub>x</sub> or VOC emissions affect the tropospheric ozone equilibrium and can lead to enhanced ozone formation, shown in figure 1.4. NO<sub>x</sub> compounds are short lived, with emissions (Power generation and combustion transport) being the main driver of concentrations (Delmas, Serca, and Jambert 1997). NO<sub>x</sub> is removed primarily by conversion to nitric acid (HNO<sub>3</sub>) followed by wet or dry deposition (Ayers and Simpson 2006).

NO<sub>x</sub> and O<sub>3</sub> relative concentrations during the day are regulated by the following reactions (Sillman 1999; Atkinson 2000):



This process with and without the influence of VOCs (panel A and B respectively) is summarised in figure 1.5.



FIGURE 1.4: Tropospheric ozone processes, Figure 1 in Young et al. (2017). DOI: <https://doi.org/10.1525/elementa.265.f1>

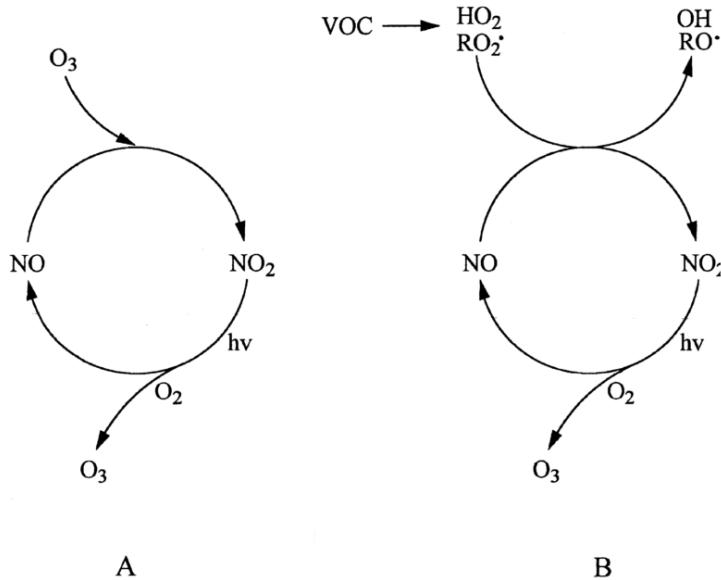


FIGURE 1.5: Figure showing NO, NO<sub>2</sub>, and Ozone photoequilibrium cycle with and without (B, A respectively) influence from VOCs. Figure copied from Atkinson (2000).

### 1.2.3 Stratosphere to troposphere transport

Historically, ozone transported down from the stratosphere was thought to contribute 10-40 ppb to tropospheric ozone levels, matching tropospheric production (Atkinson 2000; Stohl et al. 2003). The proportion of tropospheric ozone due to transport from the stratosphere was revised down to around 10% over the years as measurement and modelling campaigns improved our understanding of global scale transport, mixing, and chemistry (Guenther et al. 2006; Monks et al. 2015). Intrusions of stratospheric air into the troposphere are often called Stratosphere to Troposphere Transport (STT) events. Although most tropospheric ozone comes from production, STT enhancements of ozone are measurable and can be regionally important (eg. Jacobson and Hansson 2000; Lelieveld et al. 2009; Kuang et al. 2017), and upper tropospheric ozone can be transported long distances (Cooper et al. 2004). An analysis of the Atmospheric Chemistry and Climate Model Inter-comparison Project (ACCMIP) simulations by Young et al. (2013) found STT is responsible for  $540 \pm 140 \text{ Tg yr}^{-1}$ , equivalent to  $\sim 11\%$  of the tropospheric ozone column (Monks et al. 2015).

Ozone transported to the troposphere from the stratosphere can occur through diffusion (relatively slowly) or direct mixing (as STT). STT often occurs as tongues of stratospheric air descend and get disconnected from the stratosphere, potentially due to low pressure systems and jet streams (Sprenger, Croci Maspeli, and Wernli 2003). Recently global chemical transport models (CTMs) have been used to trace how much ozone is being transported to the troposphere in this manner. There are a few methods of doing this, such as modeling ozone formed in (and transported from) the stratosphere (Ojha et al. 2016). Model based estimates require validation against

actual measurements, such as those from ozonesondes or satellites. Hegglin and Shepherd (2009) estimate that climate change will lead to increased STT. They posit that this is due to an acceleration in the Brewer Dobson circulation; which is the global scale model of transport of air in the troposphere and stratosphere. They estimate  $\sim 30$ , and  $\sim 121 \text{ Tg yr}^{-1}$  increases by 2095 (relative to 1965) in the southern and northern hemispheres respectively, up by 23% globally.

Liu et al. (2017) examine southern hemispheric ozone and the processes which control its inter-annual variability (IAV). IAV is the standard deviation of ozone anomalies from the monthly mean. They show that ozone transported from the stratosphere plays a major role in the upper troposphere, especially over the southern Indian ocean during austral winter. STT mostly impacts the upper troposphere, although some areas are impacted right down to the surface. Kuang et al. (2017) found a measurable impact of STT ozone enhancement in the south east US using several different instruments. They also show how ozone depends on both the local topography, weather systems, and trace gases emitted and transported into the region. Liu et al. (2017) examined modelled tropospheric ozone sensitivity to various meteorological parameters. They found tropospheric ozone sensitivity to emissions from South America ( $0\text{--}20^\circ\text{S}$ ,  $72.5\text{--}37.5^\circ\text{W}$ ), southern Africa ( $5\text{--}10^\circ\text{S}$ ,  $12\text{--}38^\circ\text{E}$ ), and South to South east Asia ( $70\text{--}125^\circ\text{E}$ ,  $10^\circ\text{S}\text{--}40^\circ\text{N}$ ). In the US recent work by Lin et al. (2015) suggests that intrusions during spring are increasing surface ozone levels. They recommend improvements to understanding of the frequency and cause of STT are needed effectively implement air quality standards. Recently, modelled ozone concentrations has been shown to be most sensitive to  $\text{NO}_x$  sources (such as lightning and car exhaust emissions) and isoprene emissions (Christian2018).

### 1.2.4 Chemical production

Ozone produced in the troposphere from precursors and radiation drive ozone levels, especially in the lower (near-surface) troposphere. The main processes involved are shown in figure 1.4, with ozone regulated by reactions 1.3. As discussed above STTs source  $\sim 11\%$  of the tropospheric column of ozone, with the remainder produced photochemically (Monks et al. 2015). A recent summary by Young et al. (2017) estimates ozone production and loss in the troposphere to be  $\sim 4900 \text{ Tg yr}^{-1}$ , and  $\sim 4500 \text{ Tg yr}^{-1}$  respectively. These numbers are at the global scale, and it should be noted that meteorology and topography can play massive roles due to large spatial variability in ozone (eg. Kuang et al. 2017).

Tropospheric ozone concentrations require climate and ozone precursor emissions; including  $\text{NO}$ ,  $\text{NO}_2$ ,  $\text{CO}$ , and VOCs such as  $\text{HCHO}$  (Atkinson 2000; Young et al. 2013; Marvin et al. 2017). Ozone predictions are uncertain and changing climate affects transport, deposition, destruction, and plant based precursor emissions. All of these processes are tightly coupled and difficult to accurately model, as they depend on uncertain assumptions such as  $\text{CO}_2$  dependency (Young et al. 2013). Even with all the work done over the prior decades there remain large uncertainties about ozone precursors in the troposphere (Mazzuca et al. 2016).

Ozone is formed in the troposphere through oxidation of VOCs (described in Section 1.3) in the presence of  $\text{NO}_x$ . Net formation or loss of  $\text{O}_3$  is determined by interactions between VOCs,  $\text{NO}_x$ , and  $\text{HO}_x$ , and is a complicated system of positive and



FIGURE 1.6: Ozone production rate dependent on  $\text{NO}_X$  and VOC concentrations (Mazzuca et al. 2016).

negative feedbacks (Atkinson 2000). Figure 1.6 shows an example of this non-linear relationship between  $\text{NO}_X$ , VOCs, and ozone production as modelled in Mazzuca et al. (2016). This non-linear relationship is examined in more detail in the following section (1.3). Recently the relationship has been examined on the intradiel timescale showing that ozone production can be more or less sensitive to VOCs at different hours (Mazzuca et al. 2016). This shows how important it is to correctly determine the precursors concentrations in order to estimate ozone levels and production.

Tropospheric ozone is lost via chemical destruction and dry deposition, estimated to be  $4700 \pm 700 \text{ Tg yr}^{-1}$  and  $1000 \pm 200 \text{ Tg yr}^{-1}$ , respectively (Stevenson et al. 2006; Young et al. 2017). The main loss channel is through equation 1.1, where photolysis and collisions (increasing with pressure) create OH from the  $\text{O}_3$ .

### 1.3 Volatile Organic Compounds

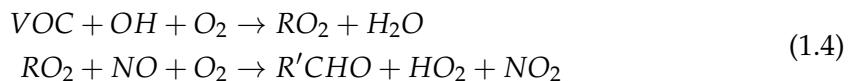
The least well understood precursors to tropospheric ozone production belong to a class of organic compounds. Organic compounds are members of a large class of chemicals whose molecules contain carbon, with the exception of a few compounds such as carbides, carbonates, and simple oxides of carbon and cyanide. Organic compounds can be categorised based on their vapour pressure, which is the tendency of a liquid or solid to vaporise. Compounds with high vapour pressures at standard temperature are classed as volatile, and have a facility to evaporate at low temperatures. Plants contain tens of thousands of organic compounds, with fewer than 40 having high enough volatility to be emitted (Guenther et al. 2000). Gas phase emissions with higher vapour pressures can be oxidised into lower vapour pressure products which will partition between gas and particle phase, often called semi or non-volatile.

Atmospheric organic compounds are legion and differ by orders of magnitude with respect to their fundamental properties, such as volatility, reactivity, and cloud droplet formation propensity, etc. Volatile organic compounds (VOCs) have vapour pressure greater than  $10^{-5}$  atm, and are mostly generated naturally by plants, which emit around 1000 Tg yr<sup>-1</sup> (Guenther et al. 1995; Glasius and Goldstein 2016). Due to their high volatility these compounds generally exist in the gas phase. Organic compounds with a lower volatility are classed as semi-volatile (SVOCs: vapour pressure between  $10^{-5}$  and  $10^{-11}$  atm) are found in both gas and particle phase depending on temperature and pressure. Organic compounds with even lower vapour pressure are generally found in the particle phase in aerosol particulate matter (Glasius and Goldstein 2016). Understanding the drivers of trends in biogenic VOC emissions (BVOCs) is required in order to estimate future carbon fluxes, changes in the water cycle, ozone production, air quality, and other climate responses (Yue, Unger, and Zheng 2015). In the last 20 years anthropogenic emissions of VOCs have been increasing while biogenic VOC emissions have decreased, due to rapid economic growth and lower annual temperatures (Stavrakou et al. 2014; Kwon et al. 2017).

Methane (CH<sub>4</sub>) is one of the more abundant VOCs, however it is often classified separately and compared against non-methane VOCs (NMVOCs). NMVOCs include alkanes, alkenes, and aromatic hydrocarbons, with isoprene (an alkene) being the most abundant (Guenther et al. 1995). Methane is relatively long lived (years) and is well mixed in the atmosphere while other VOC levels are spatially diverse due to their shorter lifetimes. In this thesis I work towards a better understanding of the isoprene emissions coming from Australia.

VOCs are an important driver of atmospheric processes, especially near forests. VOCs are broken down into HCHO, O<sub>3</sub>, CO<sub>2</sub> and many other species, mainly through oxidation by OH. VOC emissions result in radical cycling, acid deposition, production of tropospheric ozone, and secondary organic aerosols (SOAs) (Atkinson 2000; Kanakidou et al. 2005). VOC emissions affect surface pollution levels, potentially enhancing particulate matter (PM) and ozone levels. A regional-model study in Europe (Aksoyoglu et al. 2017) has also shown VOCs impact secondary inorganic aerosol concentrations. These have impacts on climate (through radiative forcing) and air quality (from ozone and SOA enhancements), affecting both human health and crop yields (Forster et al. 2007; Avnery et al. 2013; Lelieveld et al. 2015).

Ozone in rural areas is often higher than in populous cities, due to titration of ozone by NO in polluted areas (Cooper, Gilge, and Shindell 2014; Monks et al. 2015). In areas with high VOC concentrations, ozone production may be enhanced through the following reaction sequence (Sillman 1999):



with R and R' representing organic species. The reactions of VOCs or CO with OH convert NO to NO<sub>2</sub>, which leads to ozone formation as NO<sub>2</sub> production in reaction 1 of 1.3 is bypassed.

One aspect associated with VOC emissions is the production of aerosols. Aerosols are suspended particulates and liquid compounds in the atmosphere, often called particulate matter (PM). PM in the atmosphere is a major problem, causing an estimated

2-3 million deaths annually (Hoek et al. 2013; Krewski et al. 2009; Silva et al. 2013; Lelieveld et al. 2015). Fine particulate matter ( $\text{PM}_{2.5}$ ) penetrates deep into the lungs and is detrimental to human health. Some PM comes from small organic aerosols (OA) emitted in the particulate phase and referred to as primary OA (POA).

A substantial amount of PM is due to gaseous organic compounds transforming in the troposphere leading to what is known as secondary OA (SOA) (Kroll and Seinfeld 2008). Formation of SOA is generally due to VOC oxidation and subsequent reactions, while removal from the atmosphere is largely due to wet or dry deposition, and cloud scavenging (Kanakidou et al. 2005). It can be difficult to attribute the formation of SOA, in part due to the complex relationship between  $\text{NO}_x$ , OH,  $\text{O}_3$ , and the uncertainty surrounding precursor emissions. Most of the tropospheric SOA comes from biogenic precursors, the evidence for this has grown over the last two decades (Guenther et al. 1995; Kanakidou et al. 2005; Guenther et al. 2012). Improved concentration estimates of these precursors requires a better understanding of their emissions, which is one of the foci in this thesis.

Photolysis and oxidation of many VOCs initially form alkyl radicals ( $R\cdot$ ). VOCs are removed mainly by photolysis and oxidation, but also by wet and dry deposition, reaction with  $\text{NO}_3^-$ , and ozonolysis (at night time or in polluted areas) (Atkinson and Arey 2003; Brown et al. 2009). The process of deposition only accounts for a small fraction of the VOC loss, with the possible exception of the long lived methane compound (Atkinson and Arey 2003).

### 1.3.1 Emissions

VOC emissions are often classified as either anthropogenic, biogenic, or pyrogenic. There are ten times the mass of NMVOCs from natural sources as there are from anthropogenic sources (Guenther et al. 2006; Kanakidou et al. 2005; Millet et al. 2006). Methane and isoprene each comprise around a third of the global total emissions of VOCs (Guenther et al. 2006). Major emitters are broadleafs (notably eucalypts), and shrubs (Guenther et al. 2006; Arneth et al. 2008; Niinemets et al. 2010; Monks et al. 2015). TODO: why do plants emit? increased canopy light penetration? Emissions are affected by many factors including temperature, sunlight, soil moisture, drought stress, etc. The ocean also plays a role in VOC emissions, with the Oceanic Niño Index (ONI) showing positive VOC emission anomalies associated with neighbouring countries (Stavrakou et al. 2014).

Globally around  $710 - 1150 \text{ Tg C yr}^{-1}$  of BVOCs are emitted (Lathiere 2016; Guenther et al. 1995; Lathière et al. 2006; Guenther et al. 2012). 90% of these emissions come from plants and trees, with the most dominant species being isoprene ( $\text{C}_5\text{H}_8$ ) (~50%), monoterpenes ( $\text{C}_{10}\text{H}_{16}$ ), methanol ( $\text{CH}_3\text{OH}$ ), ethanol ( $\text{C}_2\text{H}_6\text{O}$ ), acetaldehyde ( $\text{CH}_3\text{CHO}$ ), acetone ( $(\text{CH}_3)_2\text{CO}$ ), ethene ( $\text{C}_2\text{H}_4$ ) and propene ( $\text{C}_3\text{H}_6$ ) (together making up ~30%) (Guenther et al. 2012). Many of these estimates come from MEGAN, a bottom-up biogenic emissions model which is highly sensitive to several parameters including soil moisture and plant functional type. MEGAN has recently been analysed using 30 years of meteorological reanalysis information by Sindelarova et al. (2014). They estimate emissions of Biogenic VOCs (BVOCs) to be  $760 \text{ Tg C yr}^{-1}$ , 70% ( $532 \text{ Tg C yr}^{-1}$ ) of which is isoprene. This is similar to isoprene emission estimates

from MEGAN itself, of 400-600 Tg C yr<sup>-1</sup> (Guenther et al. 2006). Another model (OR-CHIDEE, with inputs similar to MEGAN) estimates  $752 \pm 16$  Tg C yr<sup>-1</sup>, sensitive to terrestrial vegetation variations (Lathière et al. 2006). MEGAN emissions estimates are termed bottom-up, as opposed to top-down which are derived from satellite measurements of the products of various VOCs. Using GOME satellite HCHO and a Bayesian inversion technique to derive isoprene emissions, Shim et al. (2005) estimated global isoprene emissions to be  $\sim 566$  Tg C yr<sup>-1</sup>. This estimate decreases simulated OH concentrations by  $\sim 10\%$ , to  $9.5e5$  molec cm<sup>-3</sup>.

Global VOC levels are estimated at 85 %, 13 %, and 3 % from biogenic, anthropogenic, and pyrogenic sources respectively (Kanakidou et al. 2005; Kefauver, Filella, and Peñuelas 2014). Methane makes up a third of atmospheric VOCs and is relatively ubiquitous due to its longer lifetime. The main non-methane BVOC emissions are isoprene (44%) and monoterpenes (11%) (Guenther et al. 2000; Kefauver, Filella, and Peñuelas 2014). Land use changes can drastically affect isoprene sources, for instance in the tropics where large scale deforestation has converted forest into crop lands (Kanakidou et al. 2005). Due to the lack of in-situ ground based measurements, estimates of VOC emissions are uncertain, with large scale extrapolation required (Millet et al. 2006). In this thesis I'll be focusing on emissions of isoprene.

### 1.3.2 Isoprene

Isoprene, or 2-methylbuta-1,3-diene, is a VOC with the chemical formula C<sub>5</sub>H<sub>8</sub>. It is of major importance to the atmosphere, as it is involved in various processes which alter the oxidative capacity of the atmosphere. Isoprene affects NO<sub>X</sub> and HO<sub>X</sub> cycling, see for example formulae 1.1, 1.3. In the presence of NO<sub>X</sub>, isoprene forms tropospheric ozone and SOAs (Wagner 2002; Millet et al. 2006). It has a short lifetime during the day, roughly an hour due to OH oxidation (Atkinson and Arey 2003)).

Guenther et al. (1995), and subsequent updates (Guenther et al. 2000; Guenther et al. 2006; Guenther et al. 2012), have been used ubiquitously by the atmospheric community as a global estimate of isoprene emissions, at roughly 500-600 Tg yr<sup>-1</sup>, emitted mostly during the day. Recently an estimate of global isoprene emissions, of around 465 Tg C yr<sup>-1</sup>, has been made using a completely different model (Messina et al. 2016). The global emission factors used to derive both these estimates are based on modelling emissions from different plant species (phenotypes), and relatively few Australian species are used when forming in these estimates.

Measurements of isoprene are often uncertain or difficult to make accurately. Chamber experiments are used to determine how isoprene behaves once it is emitted into the atmosphere, however reaction rates may be unsuitable to the natural atmosphere which is often very different (Kanakidou et al. 2005; Nguyen et al. 2014). Improving chamber study methods could improve understanding of ambient atmospheric oxidation mechanisms of biogenic hydrocarbons (such as isoprene), which could reduce some of the high uncertainties involved with VOC chemistry (Nguyen et al. 2014). Uncertainties in emissions however require different techniques and measurements in order to be improved.

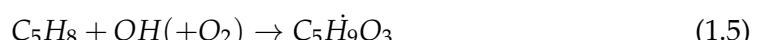
Isoprene emissions estimates are still fairly uncertain, as global measurements are difficult and regional emissions can be very different. The global uncertainty of isoprene emission was estimated to be a factor of 2 to 5 (250-750 Tg yr<sup>-1</sup>) (Kanakidou et

al. 2005). Improvements over the years have been incremental, and generally localised to regions of particular interest for air quality such as China and the USA TODO: find recent uncertainty estimate improvements examples. The lack of accuracy in BVOC emissions estimates prevents accurate determinations of the sources and distribution of pollutants including ozone and organic aerosols. Accuracy in VOC measurements is important: it has been shown that even the diurnal pattern of isoprene emissions has an effect on modelling ground level ozone (Hewitt et al. 2011; Fan and Zhang 2004).

### 1.3.3 Isoprene chemistry

Isoprene forms many products with various lifetimes, here I will present an overview of some important mechanisms and products. Isoprene is emitted and enters the atmosphere in the gas phase, where it reacts quickly with OH and other radicals. One common compound which is produced by these reactions is HCHO, which is easier to measure and often used to estimate how much isoprene is being emitted. Alkenes (VOCs with double bonded carbon, such as isoprene) react with OH, ozone, or NO<sub>3</sub>, leading to organic peroxy radicals (ROO'). These go on to form many products and lead to (amongst other things) aerosol, formaldehyde, and ozone formation, depending on sunlight and NO<sub>x</sub> concentrations (Atkinson 2000). Reactions with NO can lead to ozone production within environments rich in isoprene or other NMVOCs (PATCHEN et al. 2007; Atkinson and Arey 2003).

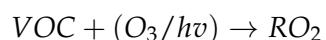
Figure 1.7 shows the first stages of oxidation of isoprene by OH. Isoprene reactions are important to understand due to their impacts on air quality, ozone, and physical properties in the lower troposphere. The primary first step for atmospheric isoprene is photooxidation, reacting with OH to form isoprene hydroxyperoxy radicals (ISOPOO - a subset of ROO') (Patchen2017; Wolfe et al. 2016; Marvin et al. 2017). This is largely split into two types of ISOPOO, based on which carbon the OH adducts to (see figure 1.7).



The many children processes and products which begin with isoprene oxidation are often called the isoprene (photochemical) cascade (Paulot2012; eg. Crounse et al. 2012; Wolfe et al. 2016).

#### 1.3.3.1 Oxidation

The primary sink for isoprene is oxidation by OH. First isoprene has its double bond replaced by OH, as summarised by the equation from PATCHEN et al. (2007): R-CH=CH-R' + OH → R-CH(OH)CH-R' where R and R' represent hydrocarbons. Ozonolysis and photolysis are lesser oxidation pathways for volatile alkenes, involving the splitting of carbon chains by ozone molecules or photons respectively:

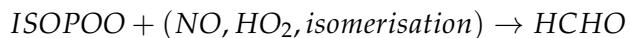


(Nguyen et al. 2016; Wolfe et al. 2016). Ozonolysis also leads to HCHO, with yields depending on subsequent reactions.



FIGURE 1.7: Isoprene products following oxidation by OH, figure from Mao et al. (2013)

After oxidation by OH, the adducted OH then reacts with O<sub>2</sub> to produce ISOPOO, which can be any of six different isomers (PATCHEN et al. 2007). ISOPOO reacts with HO<sub>2</sub> or NO, producing stable products (often called oxidised VOCs or OVOCs). One important product produced (with varying yields) through each oxidation pathway is HCHO:



During the day HCHO has a lifetime of 1-2 hrs, while ROO' lasts  $\sim$  100 s, making reaction 1.5 a rate limiting factor in HCHO production (Wolfe et al. 2016). ISOPOO also can isomerise and produce HPALDS (see figure 1.7), which also leads to HCHO.

There is uncertainty about which pathways are most important following ISOPOO production, affecting predictions by atmospheric models (Nguyen et al. 2014). This limits understanding of the relative importance of some chemical processes, such as auto-oxidation (of ISOPOO and other ROO') (Crounse et al. 2013). The reaction pathways depend on local concentrations of NO<sub>x</sub>: the high and low NO<sub>x</sub> pathways are dominated by NO and HO<sub>2</sub> reactions respectively. HO<sub>2</sub> reactions predominantly produce hydroxyhydroperoxides (ISOPOOH), while NO reactions produce isoprene nitrates (ISOPN) (Crounse et al. 2006). If measured, first generation ISOPN and ISOPOOH products can be used to determine the portion of isoprene oxidation following each pathway (eg. Yu et al. 2016). Globally around one third of ISOPOO react with HO<sub>2</sub>, and two thirds react with NO (Paulot et al. 2009b). Most of these reaction pathways produce HCHO, however this along with methyl vinyl ketone (MVK), and methacrolein (MACR) are formed at different yields between the two pathways (Marais et al. 2012; Liu et al. 2016b; Wolfe et al. 2016).

### 1.3.3.2 High NO<sub>x</sub> pathway

In the presence of NO<sub>x</sub>, ISOPOO reacts with NO and forms ISOPN, which affect levels of both HO<sub>x</sub> (H, OH, peroxy radicals) and NO<sub>x</sub>. ISOPN generally act as a sink of HO<sub>x</sub>, and can be a sink or reservoir for NO<sub>x</sub> (Mao et al. 2013). A portion of the ISOPN are recycled back to NO<sub>x</sub>, serving as a reservoir of nitrogen and allow its transport to the boundary layer of remote regions (PATCHEN et al. 2007; Paulot et al. 2009a; Yu et al. 2016). The nitrates can also build up in the winter, when removal processes are not as dominant (Lelieveld et al. 2009). Reactions of OH with NO<sub>2</sub> are the main radical sink in high-NO<sub>x</sub> systems (Wolfe et al. 2012).

First generation ISOPN produce MVK( $\sim$  40%), MACR( $\sim$  26%), and HCHO( $\sim$  60%) at higher yields than is produced by ISOPOOH (Liu et al. 2013; Mao et al. 2013). The MVK and MACR products form additional HCHO within a few hours due to oxidation by OH (Palmer et al. 2006). Under high NO<sub>x</sub> conditions there is a higher and faster yield of HCHO, with most of the ultimate HCHO production occurring within one day (Palmer et al. 2006).

### 1.3.3.3 Low NO<sub>x</sub> pathway

In low NO<sub>x</sub> environments, ISOPOOH is formed in yields  $>$  70%, while MACR, MVK, and HCHO are formed at  $\sim$  5%,  $\sim$  7%, and  $\sim$  12% respectively (Paulot et al. 2009b; Mao et al. 2013). This ISOPOOH mostly reacts with OH to form IEPOX while regenerating OH (Mao et al. 2013). This pathway has lower and slower ultimate yields of

HCHO from isoprene emissions when compared to the high- $\text{NO}_x$  pathway (Palmer et al. 2006).

Isoprene oxidation and subsequent reactions are less well understood when lower concentrations of NO are present in the atmosphere. It was thought that in low NO environments, like those far from anthropogenic pollution and fires, oxidation of isoprene would create ISOPOOH and reduce local concentrations of OH and  $\text{HO}_2$  (Guenther et al. 2000; Paulot et al. 2009b). However this reduction was not seen in measurements and  $\text{HO}_x$  levels have been shown to be largely unaffected by isoprene concentrations (Paulot et al. 2009b).  $\text{HO}_x$  is recycled through dihydroxyperoxides (IEPOX), formed from ISOPOOH oxidation, and some  $\text{HO}_x$  is produced in the formation of MACR and MVK (Paulot et al. 2009b). Paulot et al. (2009b) estimated that  $95 \pm 45 \text{ Tg yr}^{-1}$  of IEPOX was being created in the atmosphere, which (at the time) was not modelled by CTMs. Peeters and Muller (2010) suggested that the work of Paulot et al. (2009b) only partially bridges the gap between clean air OH concentration measurements and models. They suggested four new mechanisms for OH recycling in these pristine conditions. These can be summarised as OH regenerating reactions which occur during photolysis of hydroperoxy-methyl-butenals (HPALDs), and resulting photolabile peroxy-acid-aldehydes (PACALDs). These reactions are highly non-linear and subject to large uncertainty, however they were shown to improve modeled  $\text{HO}_x$  concentrations against several campaigns. Peeters and Muller (2010) showed that  $\text{HO}_2$  is produced at near unity yields following isoprene oxidation initiated by OH. Their results were backed up by observations of OH recycling observed in low NO conditions (Crounse et al. 2012).

Uncertainties and bias from measurements have made it more difficult to understand what happens in low  $\text{NO}_x$  conditions as many observations of OH were still quite under-predicted in models (Mao et al. 2012). Due to OVOC interference, measurements in low  $\text{NO}_x$  environments can lead to massively overestimated MVK and MACR yields (Nguyen et al. 2014). Nguyen et al. (2014) show preliminary estimates of low-NO yields of MVK and MACR to be  $6 \pm 3\%$  and  $4 \pm 2\%$  respectively, consistent with TODO:Liu et al. (2013), but only when cold-trapping methods are employed. Mao et al. (2012) showed that many instruments were generating OH internally. Model and measurement corrections have helped closed the gap between model predictions and detected concentrations of VOCs and OH (Mao et al. 2012).

Even with the recent boom in isoprene analysis, uncertainties remain in the isoprene oxidation mechanisms. Examples (taken from Nguyen et al. (2014)) include isoprene nitrate yields, which range from 4-15% (Paulot et al. 2009a), 90% disagreements in MACR and MVK yields (Liu et al. 2013), various possible sources for SOA (Chan et al. 2010; Surratt et al. 2010; Lin et al. 2013), unknown HPALD fates, incomplete  $\text{O}_2$  incorporation (Peeters, Nguyen, and Vereecken 2009; Crounse et al. 2013), and under-characterised RO<sub>2</sub> lifetime impacts (Wolfe et al. 2012).

### 1.3.3.4 Night time processes

At night when OH concentrations have dropped, isoprene can remain in the atmosphere. Typically less than half of this night time isoprene is removed through ozonolysis (Atkinson and Arey 2003), however, in polluted areas where high levels of  $\text{NO}_x$

exist, isoprene is consumed by nitrate radicals ( $\text{NO}_3$ ), which joins to one of the double bonds and produces organic nitrates in high yield (65% to 85%) (Mao et al. 2013). (todo: read mao2013 para 3 cites)  $\text{NO}_3$  are largely formed through ozone reactions, as in equation 1.3. A build up of  $\text{NO}_3$  radicals can be seen at night, when photolysis is not removing them (Atkinson 2000; Brown et al. 2009).

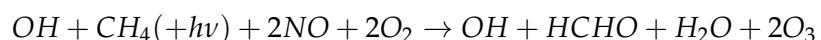
In areas with high  $\text{NO}_x$  levels, greater than 20% of the isoprene emitted late in the day ends up being oxidised by the  $\text{NO}_3$  radical overnight (Brown et al. 2009). At night isoprene affects on both  $\text{NO}_x$  concentrations and ozone levels, and can form harmful organic nitrates and SOAs (Brown et al. 2009; Mao et al. 2013). These nitrates go on to produce further SOAs, largely due to  $\text{NO}_3$  reacting with first generation isoprene oxidation products (Rollins et al. 2009). The night-time concentrations of OH and ozone also have a complex effect on  $\text{NO}_x$  removal in high latitude winters, when photolysis and NO reactions are reduced (Ayers and Simpson 2006).

## 1.4 Formaldehyde

formaldehyde (HCHO), aka methanal, methyl aldehyde, or methylene oxide, is of the aldehyde family. HCHO is an OVOC which is toxic, allergenic, and a potential carcinogen. In this thesis HCHO is used to estimate isoprene emissions over Australia. One of the major products of isoprene chemistry is HCHO. HCHO is important both for its own atmospheric impacts, and as a proxy for determination of isoprene emissions. Given a modelled yield of HCHO from isoprene, it is possible to work backwards from measured HCHO concentrations to determine the isoprene emissions. HCHO production does depend on  $\text{NO}_x$  concentrations, as it affects the yield from isoprene oxidation. HCHO yield is higher in the high- $\text{NO}_x$  pathway (compared to the low- $\text{NO}_x$  pathway) from isoprene reactions (Marais et al. 2012).

### 1.4.1 Sources and sinks

Background levels of HCHO in the atmosphere are driven by the oxidation of methane ( $\text{CH}_4$ ) by the hydroxyl radical ( $\text{OH}$ ), which produces  $\sim 970 \text{ Tg yr}^{-1}$  (Fortems-Cheiney et al. 2012). Atkinson (2000) summarised the background formation of HCHO with the following reaction:



which shows that photolysis and oxidation of methane forms HCHO and ozone in a process that regenerates the OH radicals.  $\text{CH}_4$  concentrations are relatively well constrained in models, with the ACCMIP comparison showing only  $\sim 3\%$  inter-quartile range (Young et al. 2013). There is a complex relationship between VOCs,  $\text{HO}_x$ , and  $\text{NO}_x$ : with higher levels of  $\text{NO}_x$  increasing the rate at which VOCs are converted into HCHO (Wolfe et al. 2016).

Within the continental boundary layer (CBL), HCHO is enhanced above background HCHO levels, due to NMVOC emissions reacting with OH radicals in the presence of  $\text{NO}_x$  (Wagner 2002; Millet et al. 2006; Kefauver, Filella, and Peñuelas 2014). The total contribution from NMVOC oxidation is  $\sim 358 \text{ Tg yr}^{-1}$  (Fortems-Cheiney et al. 2012). Enhancements to regional and continental HCHO are largely

driven by isoprene emissions (Guenther et al. 1995; Palmer 2003; Shim et al. 2005; Kefauver, Filella, and Peñuelas 2014). This is true except near fires or anthropogenic sources of HCHO and precursors (Guenther et al. 1995; Kefauver, Filella, and Peñuelas 2014; Wolfe et al. 2016). Biomass burning (BB) can be a source of HCHO, and various other pollutants, precursors, and aerosols (Guenther et al. 1995; Andreae 2001). Additionally HCHO is emitted into the atmosphere directly through fossil fuel combustion, natural gas flaring, ethanol refining, and agricultural activity (Wolfe et al. 2016).

Other terpenoids (monoterpenes, sesquiterpenes, etc.) can also produce HCHO, although generally to a lesser extent than isoprene, methane and biomass burning (Guenther et al. 2012). Many of the HCHO yields from terpenoids are estimated through chamber studies which examine molecular mass and charge after mixing the compound of choice into a known volume of air (eg. Nguyen et al. 2014). These conditions generally don't match those of the real world, where ambient air will have a cocktail of these compounds and other reactants. One issue with chamber studies is the difficulty they have trying to accurately reproduce ambient outside air, which limits the scope to which the studies may be applied (Nguyen et al. 2014).

Anthropogenic sources of HCHO are largely negligible, however in very large cities or by using oversampling techniques an anthropogenic signal can be found (Millet et al. 2008; Zhu et al. 2014). If the population centres and industrial districts are large enough they can emit huge amounts of VOCs into the atmosphere (Fu et al. 2007), leading to increased surface ozone levels (Zhu et al. 2014). In Australia this is not yet a major issue, however anthropogenic sources of pollution can be detected (see section 2.7.2).

In the past, HCHO levels were underestimated by models, often with large discrepancies, due to the poor understanding of methyl peroxy radical ( $\text{CH}_3\text{OO}$ ) chemistry (Wagner 2002). Nowadays HCHO concentrations are better understood, however precursor emissions are one of the main unknowns (eg. Emmerson et al. 2016; Marvin et al. 2017). Marvin et al. (2017) found that discrepancies in modelled HCHO concentrations are primarily due to second and later generation isoprene oxidation chemistry.

HCHO has two major sinks, reactions with OH (oxidation), and photolysis (adding up to  $\sim 1210 \text{ Tg yr}^{-1}$ ) (Levy 1972; CRUTZEN, LAWRENCE, and PÖSCHL 1999; Wagner 2002; Fortems-Cheiney et al. 2012; Kefauver, Filella, and Peñuelas 2014). These reactions lead to a daytime lifetime of a few hours (Atkinson 2000; Millet et al. 2006). Both these loss processes (photolysis, oxidation) form CO and hydroperoxyl radicals ( $\text{HO}_2$ ), and have global significance to radiative forcing and oxidative capacity (Franco et al. 2015). The other major sinks are wet and dry deposition, although these are not as significant ( $\sim 32 \text{ Tg yr}^{-1}$ ) (Atkinson 2000; Fortems-Cheiney et al. 2012).

#### 1.4.2 Measurement techniques

There are a few ways to measure HCHO, including Fourier Transform Infra-Red (FTIR) Spectrometry and Differential Optical Absorption Spectroscopy (DOAS). FTIR examines the Fourier transform of a measured spectrum in order to detect things which affect that spectrum. DOAS methods are based on light interference and absorption through air masses.

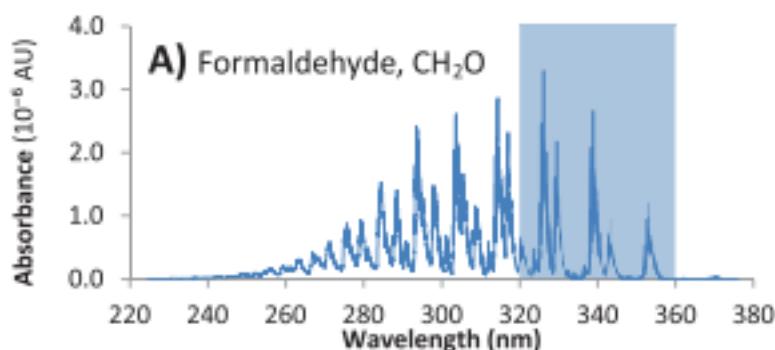


FIGURE 1.8: HCHO spectrum, with a typical band of wavelengths used for DOAS path measurements. This is a portion of an image from Davenport et al. (2015).

The DOAS technique takes advantage of the optically thin nature of HCHO in order to linearise the radiance differential through air masses with and without HCHO, using the Beer-Lambert intensity law. This method is used both on the ground, and from space, globally for HCHO detection (Guenther et al. 1995; Gonzalez Abad et al. 2015; Davenport et al. 2015). As a trace gas HCHO interferes with light over a few wavelength bands, which allows instruments to detect concentrations between a known light source and a detector. Figure 1.8 shows the interference spectrum of HCHO along with a typical band used to examine interference in the DOAS technique. One difficulty is that this interference is relatively small (HCHO is optically thin) and other compounds absorb light at similar wavelengths (Davenport et al. 2015).

FTIR and DOAS measurements have a range of uncertainties, including systematic and random measurement errors and uncertain apriori shape factors and water profiles (eg: Franco et al. (2015)). Other types of measurement involve directly measuring the air, and determining chemical compounds through their physical properties such as by mass spectrometry analysis of mass to charge ratios ( $m/z$ ) of ionised air masses. Two examples of this include proton transfer reaction mass spectrometers (PTR-MS), and gas chromatography mass spectrometers (GC-MS). These instruments can be used to determine gas phase evolution of isoprene and monoterpene products such as HCHO (eg. Lee et al. 2006a; Nguyen et al. 2014; Wolfe et al. 2016; Lerner et al. 2017).

Other measurement techniques include chromatographic and fluorimetric methods, both of which differ widely from each other and the spectroscopic methods (Hak et al. 2005). Hak et al. (2005) examine a single air mass with 8 instruments using the four techniques (MAX-DOAS, FTIR, chromatographic, and fluorimetric), and show that reasonable agreements can be achieved. Generally the measurements were close, the five Hantzsch instruments agreeing to within 11% (after removing two potentially faulty measurements), although different calibration standards were used. Titration for the different calibration solutions could not be resolved, which may account for absolute offsets up to 30%. These differences and non-uniformities between measurements (even among identical instruments) are part of the reason HCHO does not have a consistent network for global measurements like those for greenhouse gases or ozone (Fortems-Cheiney et al. 2012).

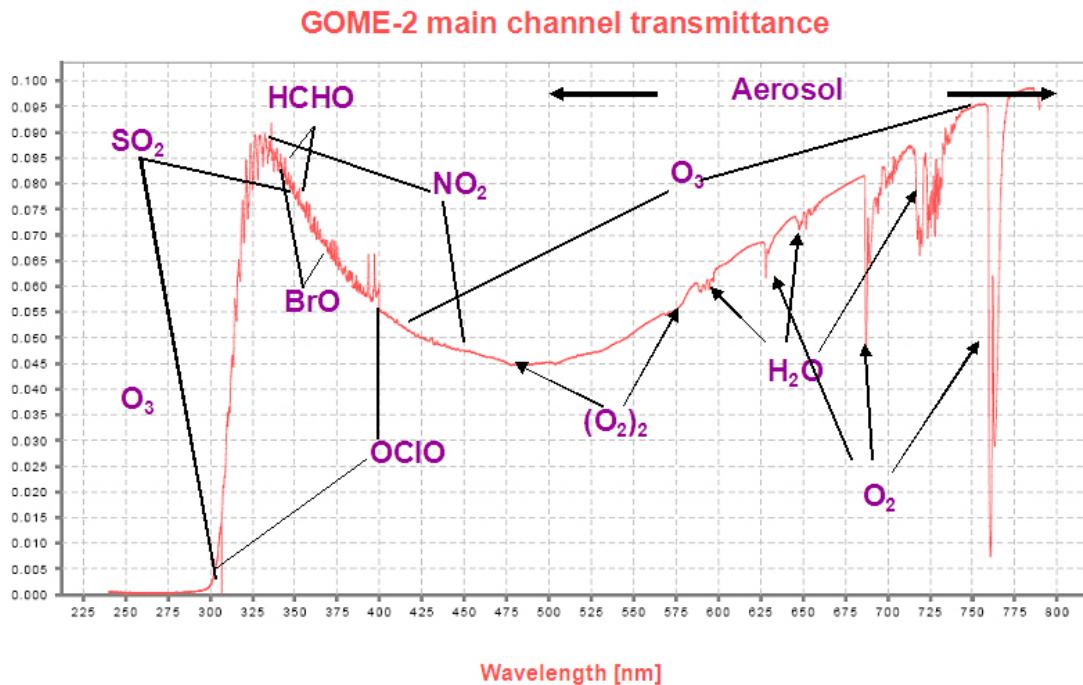


FIGURE 1.9: An example spectrum showing interferences used for species concentration measurements by GOME-2. Image by EUMETSAT and ESA (EUMETSAT 2015).

#### 1.4.2.1 Satellite measurements

Satellites remotely sense atmospheric HCHO through irradiance measurements of solar light which has reflected off the earth's surface. These irradiances are affected by gases which exist along the reflected path of light between the detector, earth, and sun. The irradiance is then used to estimate how much of a particular gas exists along this path, which gives us an estimate which is called a slant column (SC). The retrieved SC of a particular gas (or species) can be transformed into a vertical column (VC) by scaling the path length in conjunction with accounting for the trace gas' light scattering properties. The scaling coefficient created to transform from SC to VC is called the Air Mass Factor (AMF).

Several satellites provide long term trace gas observations with near complete global coverage, including the ERS-2 launched in April 1995 which houses the GOME ultraviolet and visible (UV-Vis) spectrometer, the AURA launched in July 2004 which houses the OMI UV-Vis spectrometer, the MetOp-A and B launched in October 2006 and September 2012 respectively both housing a GOME-2 UV-Vis spectrometer. These satellites are on Low Earth Orbit (LEO) trajectories and overpass any area up to once per day. Satellites use DOAS techniques with radiative transfer calculations on solar radiation absorption spectra to measure column HCHO . An example of a spectrum retrieved from the GOME-2 instrument is given in figure 1.9.

In conjunction with atmospheric chemistry and radiative models, satellite measurements quantify the abundance of HCHO in the atmosphere. Isoprene is hard to measure directly due to its short lifetime and weak spectral absorption, instead HCHO

is often used as a proxy (Millet et al. 2006; Fu et al. 2007; Dufour et al. 2008; Marais et al. 2012; Bauwens et al. 2013; Kefauver, Filella, and Peñuelas 2014; Bauwens et al. 2016; Surl, Palmer, and González Abad 2018). This leads to a method of isoprene emissions estimation termed top-down (as opposed to bottom-up estimates). The existence of satellite data covering remote areas provides an opportunity to improve VOC emissions estimates leading to more robust models of global climate and chemistry. Satellite data gives us another way to estimate large scale isoprene emissions, and their subsequent chemistry. This method is described in detail in section 3.2.

## 1.5 Atmospheric Chemistry Modelling

Models can fill the gaps (both spatial and temporal) in measurement records, and can help us improve our understanding of the natural world. They are used to examine future outcomes resulting from changing our emissions, from small to large scales. They can be used to increase measurement accuracy (for instance in satellite measurements) and determine where we lack information, while also checking the performance of new instruments. Precisely representing various chemicals and reactions in the atmosphere allows efficient mitigation of pollution, since we can compare scenarios against one another. Models can always be expanded to include new compounds or processes, however validation is always necessary. Currently they require improved isoprene emissions and subsequent chemistry understanding for effective air quality determination (Marvin et al. 2017).

### 1.5.1 Box models

Box models are much smaller scale than global CTMs, examining one uniform environment with many parametrisations such as transport and emissions. Box models can be used to check chemical mechanisms in specific scenarios, such as high or low NO<sub>x</sub> environments. For example: Marvin et al. (2017) use a box model matching conditions in southeast USA to evaluate isoprene mechanisms from several models. A box model involves modelling chemistry in a singular set of conditions without transport or any spatial gradients.

By allowing for interactions between boxes this concept can be extended to multiple-box models. These are simply multiple instances of single boxes with the addition of transport between them, which requires meteorological fields such as wind velocities and turbulence. The meteorology fields can be modelled, and/or input as parameters.

### 1.5.2 Chemical transport models

Chemical transport models (CTMs) provide a simulation of chemical densities and transport over time, through the atmosphere. They require many inputs (such as wind velocities) in order to accurately represent scenarios or regions on earth. Models of emissions are often used as drivers for atmospheric chemistry models, which require initial and boundary conditions in order to run. Chemistry in the atmosphere is a complex system of coupled reactions and dynamics, which can be solved using numerical partial differential equation solvers.

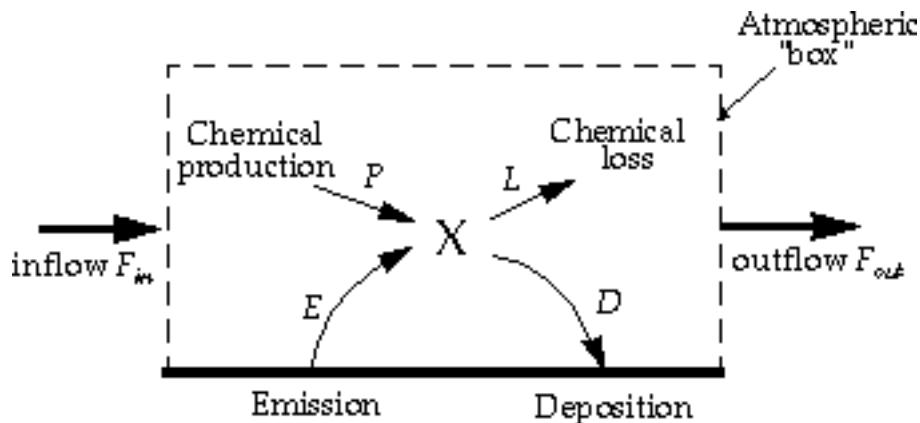


FIGURE 1.10: Standard box model parameters, image taken from Jacob (1999).

CTMs simulate production, loss, and transport of chemical species. This is generally calculated using one or both of the Eulerian (box) or Lagrangian (puff) frames of reference. CTMs normally solve continuity equations simultaneously for many coupled species. The continuity equations describe transport of a conserved quantity such as mass or energy, which, solved together with production and loss of a chemical can provide detailed simulations of natural processes.

The general continuity equation links a quantity of a substance ( $q$ ) to the field in which it flows and can be described by the formula:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot j = \sigma$$

where  $\rho$  is density of  $q$  in the field,  $t$  is time,  $\nabla$  is divergence,  $j$  is the flux ( $q$  per unit area per unit time entering or leaving the field), and  $\sigma$  is the generation or loss of  $q$  per unit volume per unit time.

The type of model best suited to modelling the entire earth uses the Eulerian frame of reference, where the atmosphere is broken up into 3-D boxes with densities and transport calculated and stored for sequential steps in time at each location. The mass balance equation must be satisfied in any realistic long term model and is as follows:

$$\begin{aligned}\frac{dm}{dt} &= \sum \text{sources} - \sum \text{sinks} \\ &= F_{in} + E + P - F_{out} - L - D\end{aligned}$$

where  $m$  is mass of a chemical,  $E$  and  $D$  are emission and deposition,  $P$  and  $L$  are production and loss, and  $F$  is chemical transport in and out, as shown in figure 1.10. Many chemical species interact with each other through production and loss. Any large chemical model will solve this mass balance equation over highly coupled arrays of partial differential equations, which becomes computation time expensive as complexity increases.

Contemporary models generally use mathematical differential solving tools of various complexity (often called chemical mechanisms) to solve chemical equations in order to predict chemical species evolutions over time. Different solvers may be slower or faster and more suited to particular situations based on the stability of the equations and systems involved, and chemical mechanisms may vary in how many reactions and chemicals are listed and grouped together. For example: Since  $[O] \ll [O_3]$  the chemical family  $O_x$  ( $O_x \equiv O + O_3$ ) can be used to simplify chemistry simulations and approximate  $O_3$  concentrations (Brasseur and Jacob 2017, Chapter 3). Different chemical mechanisms may find different solutions to the same problems, due to how the numerical solvers are implemented, which can affect model output (Zhang et al. 2012).

### 1.5.3 Emissions

There are two commonly used ways of estimating isoprene emissions, top-down or bottom-up. Bottom-up emission estimates generally model the flora which emit isoprene, along with the rates of emissions and things which affect these rates. Isoprene is emitted by trees or shrubs, depending on several parameters such as leaf area index (LAI), emission factors (EF), plant functional type (PFT), and light density fraction (LDF). Models use these properties of the emitters in order to estimate how much isoprene is being produced (eg. Guenther et al. 1995; Guenther et al. 2006). Understanding how much isoprene is emitted, when and by what, is complicated. One frequently used bottom up emissions model is the Model of Emissions of Gases and Aerosols from Nature (MEGAN, Guenther et al. (1995)). Since little data exists with which to verify many of these bottom-up emission inventories, they can be uncertain on a large scale.

Bottom up models of VOC emissions are sensitive to parameters. For example Stavrakou et al. (2014) examined modelled Asian emissions and altered model parameters for temperature, plant type emission factors, incoming solar radiation (insolation) intensity, land use changes, and palm tree forest expansion. Changes were constrained by a network of radiation measurements and some experiments with south east Asian forest emissions - and led to reduction in apriori isoprene emissions by a factor of two over the region in 2005. Sensitivity to these factors is pervasive in bottom up emissions models (eg. Marais et al. 2014; Miller et al. 2014; Messina et al. 2016). One of the important uncertainties seen in MEGAN is the isoprene emissions due to PFTs. If one plant species is emitting heavily near a measuring instrument, possible overestimations may occur due to extrapolation over the entire forest. Global emissions inventories like MEGAN often have large areas based on extrapolations which introduces uncertainties (Miller et al. 2014). Current emissions estimates require more validation against observations, and recently a comparison of two major VOC models (MEGAN and ORCHIDEE) was undertaken by Messina et al. (2016) reiterating this requirement. In their work they examine model sensitivities and show that the most important parameters are LAI, EF, PFT, and LDF. There is high uncertainty in LAI and EF, which require more or improved measurements at the global scale, as well as more PFTs and improved LDF parameterisation (Messina et al. 2016).

### 1.5.4 Uncertainties

Here I will attempt to list and partially explain the major uncertainties models have in relation to VOCs, and ozone. TODO: Move these uncertainties to the sections where they are applicable.

#### 1.5.4.1 Emissions Inventories

Using different emissions inventories in a CTM can have large impacts on the simulation. Natural (biogenic or pyrogenic) and human driven (anthropogenic) emissions often drive a large fraction of atmospheric oxidation and radical chemistry, especially in the continental boundary layer. Emissions inventories have been found to be generally OK at larger (regional to global) scales, as long as they are derived from accurate input measurements (Zeng et al. 2015). Modelled ozone concentrations has been found to be most sensitive to isoprene emissions and NO<sub>x</sub> sources, both of which have uncertainty factors of ~ 2 (Christian, Brune, and Mao 2017).

It is important to note that many estimates of isoprene emission are based on a few algorithms which can depend greatly on input parameters (Arneth et al. 2008; Niinemets et al. 2010). Arneth et al. (2008) argue that this monopoly of emissions estimates may be leading us to an incorrect understanding of isoprene chemistry. Yue, Unger, and Zheng (2015) has shown that this is still a problem by looking at land carbon fluxes and modelling the sensitivity to VOC emissions estimates using two independent models of VOC emission. One model is photosynthesis based and estimates isoprene emissions using electron transfer energies and leaf physiology (Niinemets et al. 1999), while the other (MEGAN) uses the light and canopy temperature ((Guenther et al. 1995; Arneth et al. 2007) TODO: Read Arneth et al., 2007; Unger et al., 2013). Both are sensitive to light and temperature parameterisations.

#### 1.5.4.2 Resolution

todo: Yu2016 paper.

GEOS-Chem simulations are somewhat sensitive to the gridbox resolution. For example: Wild and Prather (2006) show that reduced resolution increases OH concentrations and ozone production rates. Christian, Brune, and Mao (2017) find small changes in OH (< 10%) in OH, HO<sub>2</sub> and ozone concentrations local to the north american arctic, when changing from 4 by 5 to 2 by 2.5°resolution. For many global scale analyses, errors from resolution are less important than those from chemistry, meteorology, and emissions (Christian2018; Christian, Brune, and Mao 2017). Many models lack in-situ measurements with which to verify their chemical mechanisms, leading to large discrepancies, as seen in Marvin2017a TODO: briefly talk about Marvin2017a takeaways.

#### 1.5.4.3 Chemistry mechanisms

There is still much work to be done in models to correctly simulate emissions and processes which lead to HCHO and ozone. Often HCHO is used as a way of checking if precursors are correctly modelled since HCHO measurements are more readily available (for instance from satellites). GEOS-Chem has recently been analysed for

ozone and oxidant ( $\text{OH}$  and  $\text{HO}_2$ ) sensitivity to the processes within the model along with inputs which drive it (Christian, Brune, and Mao 2017). Christian, Brune, and Mao (2017) found that GEOS-Chem ozone was most sensitive to  $\text{NO}_2$  photolysis, the  $\text{NO}_2 + \text{OH}$  reaction rate, and precursor emissions such as VOCs.

Marvin et al. (2017) suggest that isoprene mechanisms in several contemporary models (including GEOS-Chem) are inadequate. They show that for a specific measurement campaign, the HCHO concentrations are underestimated in a way that can not be easily fixed through rate constant changes. Recently Marvin et al. (2017) compared five global CTMs isoprene mechanisms by evaluating simulated HCHO mixing ratios compared to in situ measurements from the Southeast Nexus (SENEX) aircraft campaign (in southeastern USA). They compared five models (GEOS-Chem, CB05, CB6r2, MCMv3.2, and MCMv3.3.1) and found all of them underestimated the HCHO concentrations (by 15 – 30%).

Another important factor in determining the yield of HCHO and ozone from BVOCs is the local concentration of  $\text{NO}_x$ . Travis et al. (2016) show how modelled surface ozone is overestimated due to high estimates of  $\text{NO}_x$  emissions, which affect oxidative capacity and VOC reactions in the US.  $\text{NO}_x$  and isoprene emissions are shown to be the most significant sources of uncertainty for ozone concentrations near the surface in GOES-Chem over the US, while isoprene derived products and lightning  $\text{NO}_x$  drives uncertainty in the upper atmosphere (Christian, Brune, and Mao 2017).

#### 1.5.4.4 Clouds

One of the major uncertainties in chemical, climate, radiation, and weather models is cloud formation and dynamics. Clouds are remarkably complex at a much finer scale than can be accurately modelled by global chemistry models (with current processing power). Globally over half (50-60%) of the world is covered by clouds, with ~ 10% of them being rain-clouds (Kanakidou et al. 2005). Wet scavenging performed in clouds not only depends on large scale cloud processes, but also on the micro-physics of aerosols being scavenged, differing between aerosol sizes and hygroscopic properties. Cloud filtering is performed on both OMHCHO and OMNO2d products, as those measurements are too uncertain or do not capture enough information. This has been seen to introduce a clear-sky bias - as monthly averages do not include cloudy conditions (Surl, Palmer, and González Abad 2018).

#### 1.5.4.5 Soil Moisture

Modelled emissions are sensitive to soil moisture, especially near the wilting point, below which trees stop emitting isoprene and other VOCs completely as they can no longer draw water (Bauwens et al. 2016). MEGAN accounts for soil moisture through a parameterisation which drops plant emissions to zero below a prescribed soil moisture level (the wilting point). Recently an update to this has been shown to improve modelled isoprene emissions in drought conditions (Jiang2018). Jiang2018 found that improving the parameterisation of drought based on a measurement campaign in the U.S. would lower isoprene emissions globally by ~ 17%. Many environmental parameters are affected by soil moisture, which all play a role at fine scales to surface emissions (Rowntree and Bolton 1983; Chen and Dudhia 2001). Droughts effects can

be difficult to measure, as they are a multi-scale problem which affects various aspects of the land-air interface including plant emissions and dry deposition (Wang et al. 2017).

## 1.6 Australia and the southern hemisphere

Australia has a unique climate, along with soil moisture, clay content and other important properties which affect VOC emissions. These properties are only sparsely measured in Australia due to the spread out distribution of population centres, which make many areas very difficult or expensive to reach. In Australia most long term air quality or composition measurements are performed in or near large cities. Australia is dominated by areas with little anthropogenic influence and few ground based measurements of the natural emissions taking place (VanDerA et al. 2008). Since many Australian cities are on the edge of regions with rich VOC emissions, it is very important to clarify the quantity, type, and cause of VOC emissions. Understanding of emissions from these areas is necessary to inform national policy on air pollution levels.

The vegetation in Australia is diverse, a summary is provided by ABARES using the national forest inventory at <http://www.agriculture.gov.au/abares/forestsaustralia/australias-forests>. Figure 1.11 shows the different forest types and their locations within Australia, highlighting that much of our forested lands are near population centres along the east coast. 16% of Australia is covered by forest, most (75%) of which is Eucalyptus.

Fire emissions include a range of chemical compounds and particulates and each year the effects of fire or burning seasons blanket the northern and southern hemispheres independently. Biomass burning in southern Africa and South America has previously been shown to have a major influence on atmospheric composition in Australia (Oltmans et al. 2001; Gloudemans et al. 2007; Edwards et al. 2006), particularly from July to December (Pak et al. 2003; Liu et al. 2016a).

It has been estimated by MEGAN that the Australian outback is among the worlds' strongest isoprene emitters with forests in SE Australia having emission factors greater than  $16 \text{ mg m}^{-2} \text{ h}^{-1}$  (see figure 1.12) (Guenther et al. 2006; Guenther et al. 2012). Measurement campaigns in SE Australia have since cast doubt on the emission factors used by MEGAN, potentially due to poor characterisation of Eucalyptus trees and soil moisture (Emmerson et al. 2016). These emissions factor estimates are not well verified and measurements of isoprene (or other BVOC) emissions are sparse and infrequent in Australia (Sindelarova et al. 2014; Bauwens et al. 2016).. In addition, monoterpene emissions are 2-4 times too low, which may be due to underestimated emission rates for many Eucalypt species (Winters et al. 2009; Emmerson et al. 2016).

### 1.6.1 Ozone

Surface ozone levels over Australia are relatively low ( $\sim 20 \text{ ppb}$ ) (Young et al. 2017), however it remains unclear how much we would expect this to change in the future as relatively little is known about precursors and influx for the continent. Australian air quality is monitored independently within each state, using several metrics.

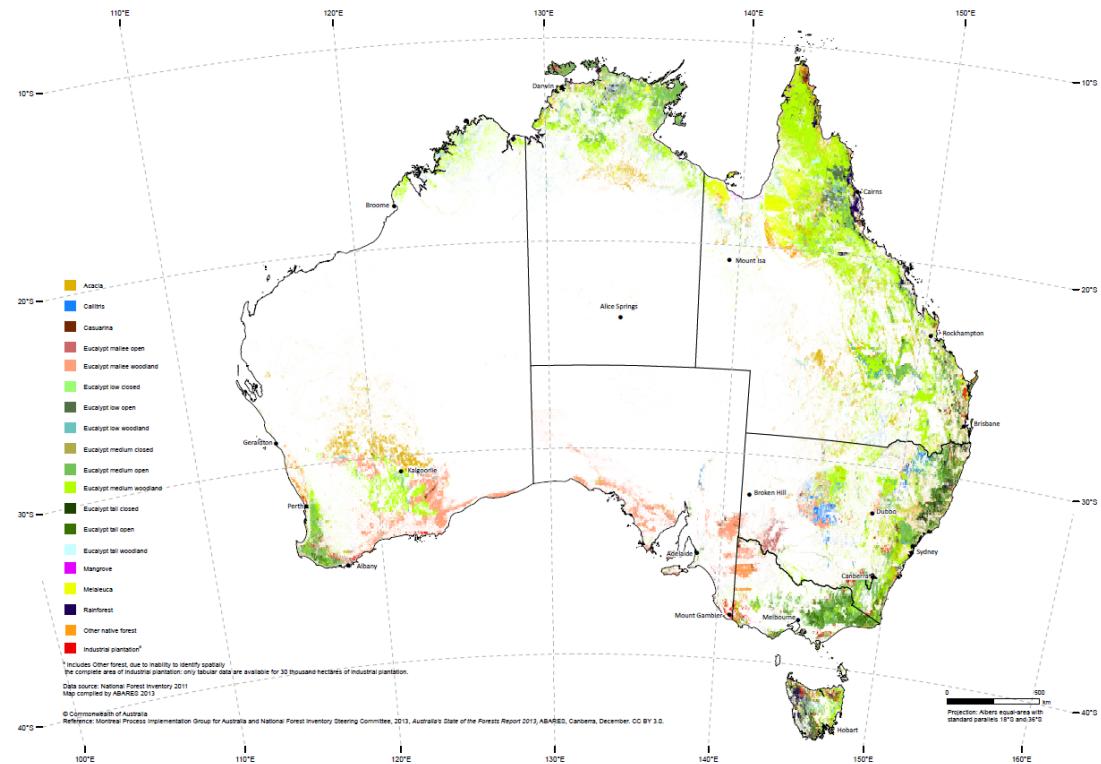


FIGURE 1.11: Forest types in Australia (<http://www.agriculture.gov.au/abares/forestsaustralia/australias-forests>)

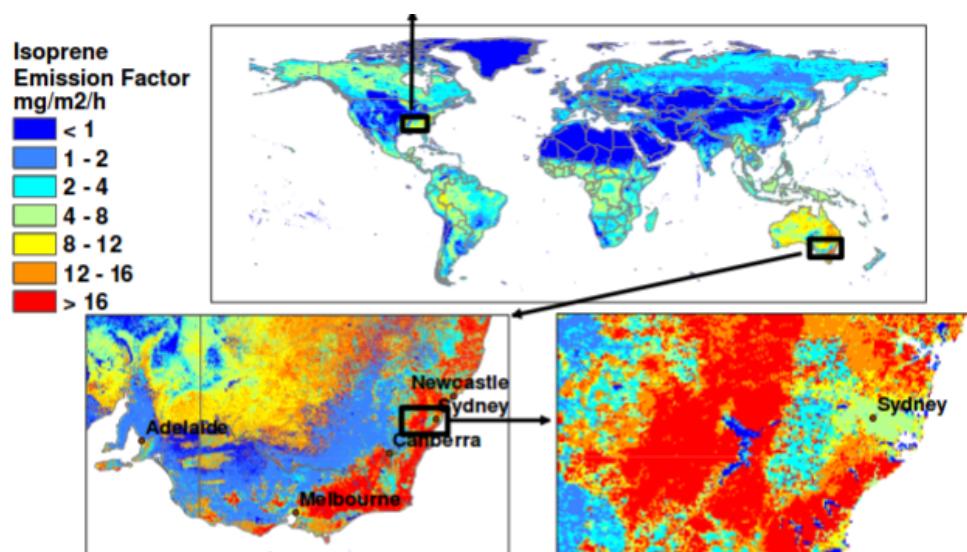


Fig. 2. Global distribution of landscape-average isoprene emission factors ( $\text{mg isoprene } \text{m}^{-2} \text{ h}^{-1}$ ). Spatial variability at the base resolution ( $\sim 1 \text{ km}$ ) is shown by regional images of the southeastern U.S. and southeastern Australia.

FIGURE 1.12: Part of a figure from Guenther et al. (2006) showing global isoprene emission factors.

These metrics are measured by varying numbers of monitoring stations in each state. Measurement stations are generally located in population centres, and don't regularly measure isoprenoid emissions. This is an important omission as these naturally emitted precursor gases often get transported into cities where they affect air quality through production of O<sub>3</sub> and other pollutants.

Generally STT over Australia affects the upper troposphere only, however ozone enhancements can reach quite low during heavy storms and cyclonic weather patterns (Alexander et al. 2013). TODO: More on Australian ozone STT Ozone enhancements over the southern ocean signify either transported pollution or stratospheric influx, due to lack of precursor emissions TODO:cite. Analysis over this region is relatively sparse TODO:cite, and quantification is difficult without large scale extrapolations.

### 1.6.2 VOCs

Bottom up inventories of VOCs remain largely uncertain due to extensive extrapolation over plant functional types, changing land cover, and parameterised environmental stressors (Guenther et al. 2000; Kanakidou et al. 2005; Millet et al. 2006). Müller et al. (2008) show how isoprene (a key VOC) is poorly captured by the MEGAN model and analyse the affect of changing the soil moisture parameter. Sindelarova et al. (2014) show reductions in modelled Australian isoprene emissions of 50% when incorporating soil moisture in MEGAN estimates. Uncertainties in isoprene emissions could explain why models of HCHO over Australia are poor at reproducing satellite measurements (Stavrakou et al. 2009). Improved parameterisation of the affect of drought on plant emissions could also lower modelled isoprene emissions (Jiang2018).

Australia suffers from poor characterisation of plant emissions, partly because emission factors are based on northern hemispheric data. Many plant emissions rates have not been published, such as those for any Australian acacias. Some Eucalypt emissions are based on samples from young trees, which may emit more isoprene than older trees (Emmerson et al. 2016). Additionally soil moisture is not well quantified which has a large effect on emissions. Soil type and moisture, along with drought threshholds, have poorly understood effects on plant emissions in Australia Changes in parameterisation of soil moisture in the MEGAN lead to massive changes in Australian isoprene emission estimates (Sindelarova et al. 2014). Over Australia MEGAN suffers from a lack of studied plant functional types and their emissions (eg. Müller et al. 2008). Emission rates from various species of Eucalyptus and other flora are highly complex, depending on current and recent weather, temperature, tree age, health, etc. (Guenther et al. 2012). With this complexity added to the diversity of tree species in Australia as well as sparse rural data collections it is hard to model and verify emissions.

Emmerson et al. (2016) analysed isoprene and monoterpene emissions sensitivities in a regional model of atmospheric chemistry over southeast Australia, using four campaigns which are also examined in this thesis. They show that modelled emissions require spatially and temporally resolved changes. Emmerson et al. (2016) suggest that monoterpenes may be emitted in similar quantities to isoprene, with more measurements required to determine if this is so. They compare emissions estimates from MEGAN against field campaign data and see overestimated isoprene emissions, as well as underestimated monoterpene emissions. Their work suggests that MEGAN

estimates of isoprene emissions may be 2-6 times too high, and monoterpene emissions  $\sim$  3 times too low over southeast Australia.

Improvements to emissions models require improved understanding of regions and their behaviour. Satellite measurements can be used to improve understanding of Australian emissions. As HCHO is produced with relatively high yield after isoprene is emitted, we can use satellite measurements to estimate isoprene emissions (Palmer et al. 2001; Millet et al. 2006; Bauwens et al. 2016, e.g.).

### 1.6.3 Measurements

TODO: Brief overview of all the measurement campaigns, pointing to Modelling and Data chapter for more details. There are relatively few measurements of isoprene in the southern hemisphere, including MUMBA(TODO CITE), SPSS(cite), Tumbarumba (Emmerson et al. 2016), and that girl from Macquarie University with an instrument in the daintree rainforest(TODO CITE, DESCRIBE). For details on the MUMBA campaign see Section 2.3.3.2. An airlift campaign (HIPPO) measuring isoprene was also performed in 2009-2011? TODO: ask Jenny re this one.

A particulate and air quality measurement campaign took place in Sydney using PTR-MS and GC-FID, for details see Section 2.3.3.3.

One method of measuring ozone in the troposphere and stratosphere is by releasing weather balloons (with attached ozone detectors) which take readings as they rise up to around 30 km, giving a vertical profile of concentrations. Since 1986, Lauder, New Zealand ( $45^{\circ}$ S,  $170^{\circ}$ E) has released ozonesondes allowing a multi-decadal analysis of ozone concentrations over the city (Brinksma et al. 2002). Kerguelan Island ( $49.2^{\circ}$ S,  $70.1^{\circ}$ E), also has a record of ozonesonde profiles, which are directly in the path of biomass burning smoke plumes transported off shore from Africa (Baray et al. 2012). SHADOZ is the southern hemispheric additional ozone project, which have released sondes from 15 sites at different times <http://tropo.gsfc.nasa.gov/shadoz/>.

A smaller network of ozonesonde release sites is operated by TODO: get details for sondes I use in ozone chapter.

## 1.7 Aims

**In this thesis I aim to improve understanding of natural contributions to ozone over Australia and the southern ocean.** The two largest contributors to tropospheric ozone concentrations are chemical production (driven by precursor emissions) and stratospheric transport. I aim to improve understanding of both of these sources using existing satellite and ground-based datasets along with GEOS-Chem modelled outputs.

Calculation of isoprene to HCHO yields over Australia is required to create top-down estimates. This requires among other things an idea of which VOCs are present and their yields of HCHO. The technique of determining isoprene emissions from satellite detected HCHO is called satellite inversion. **I aim to determine isoprene to HCHO yield over Australia, and the importance of relevant parameters using GEOS-Chem.** This includes an examination of how well GEOS-Chem simulates several species such as NO<sub>x</sub>, isoprene, and HCHO compared to campaign and satellite

data that exists for Australia. TODO: summary of stuff in modelling chapter. Work to complete this aim is in Chapter 2.

**One of the aims in this thesis is to use available satellite measurements to improve the estimates of isoprene emissions in Australia.** Satellites which overpass daily record reflected solar (and emitted terrestrial) radiation, and give us measurements over all of Australia. Combining satellite data with model outcomes provides a platform for the understanding of natural processes which is especially useful over Australia. Due to the low availability of in-situ data over most of the Australian continent, a combination of the models with satellite can fill the gap of understanding of emissions from Australian landscapes. Improved emissions estimates will in turn improve the accuracy of CTMs, providing better predictions of atmospheric composition and its response to ongoing environmental change. The work done to complete this aim is in Chapter 3.

**To improve understanding of ozone transported to the troposphere from the stratosphere in Australia and the southern ocean.** Stratospheric transport is the second largest driver of tropospheric ozone concentrations, and an improved understanding of transported ozone can be determined from ozonesonde measurements. Ozonesondes provide a glimpse of the vertical ozone profile up to  $\sim 30$  km, and we use a Fourier filter to determine how often stratospheric transport is occurring at three sites: Melbourne, Macquarie Island, and Davis Station. Combining transport event frequency analysis with modelled ozone distributions is used to derive a new method of detection and quantification of transported ozone in Chapter 4.

TODO: conclusions chapter aims? **I aim to describe relative importance of sources of tropospheric ozone in Australia, as well as seasonality.**

## 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 data from 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 2.3 lists and describes the campaign datasets, section 2.4 describes the GEOS-Chem model, the creation and effects of filters are described in section 2.7 TODO: rest of section outlined.

### 2.2 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.

### 2.2.1 DOAS

The DOAS technique uses solar radiation absorption spectra to measure trace gases through paths of light. Beer's law states

$$T = I/I_0 = e^{-\tau} \quad (2.1)$$

with T being transmittance,  $\tau$  being optical depth, and I,  $I_0$  being radiant flux received at instrument and emitted at source respectively. The Beer-Lambert law of extinction allows spectroscopic measurement of absorbing chemical species (absorbers) in the atmosphere:

$$I_B = I_{B_0} e^{-\tau_s} \quad (2.2)$$

where  $I_B$ ,  $I_{B_0}$  is backscattered intensity with and without the absorber respectively, and  $\tau_s$  is the optical thickness of the absorber along the measured path between source and instrument.

$\tau$  can be described using the scattering and absorption cross section area ( $\alpha$ ,  $\text{cm}^2$ ) and density ( $\eta$ , molec  $\text{cm}^{-3}$ ) of an absorber as follows:

$$\tau = \int \alpha(s)\eta(s)ds \quad (2.3)$$

$\tau$  through a medium is the sum of optical thicknesses of each absorber within the measured path, and subbing 2.3 into 2.2 leads to

$$I = I_0 \exp \left( \sum_i \int \eta_i \alpha_i ds \right)$$

Where i represents a chemical species index, and the integral over ds represents integration over the path from light source to instrument.

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 (from equation 2.2). 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  $\phi_e^i$  is radiant flux seen at the earth surface,  $\phi_e^t$  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.

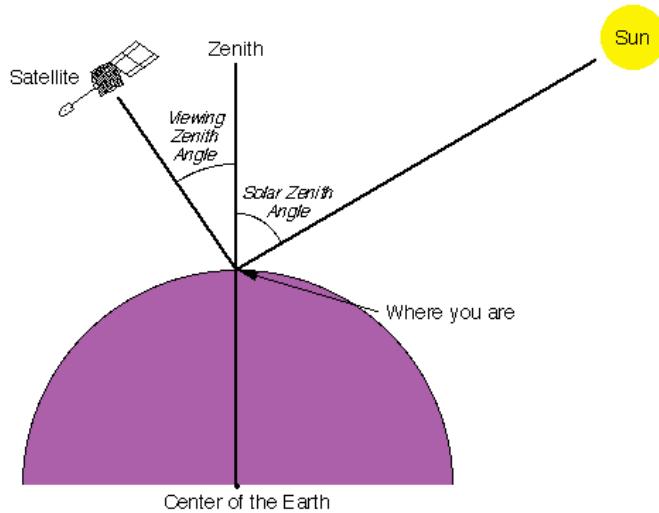


FIGURE 2.1: Solar and viewing zenith angles, image copied from Wikipedia (2016), originally from a NASA website.

### 2.2.2 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 2.2.2.1) 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.

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 ( $\theta_s$ ) and the satellite viewing angle ( $\theta_v$ ) are shown in image 2.1. However, in the UV-VIS region of the spectrum, Rayleigh and Mie scattering (see section 2.2.1) must be accounted for.

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 scattered light. Mie scattering is more generally involves 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 (e.g. Martin et al. 2002b).

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 O<sub>3</sub>, NO<sub>2</sub>, and HCHO (eg: Leue et al. (2001)). Satellite measurements are generally performed using spectral fitting followed by conversion to vertical column densities. Several public data servers are available which include products from satellites, including NASAs Earthdata portal (<https://earthdata.nasa.gov/>) and the Belgian Institute for Space Aeronomy (IASB-BIRA) Aeronomie site (<http://h2co.aeronomie.be/>).

Difficulties can arise when aerosols interfere with recorded spectra (eg. clouds, smoke, dust), however some of these can be detected and filtered out. Instruments including MODIS on board the AQUA and TERRA satellites are 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.

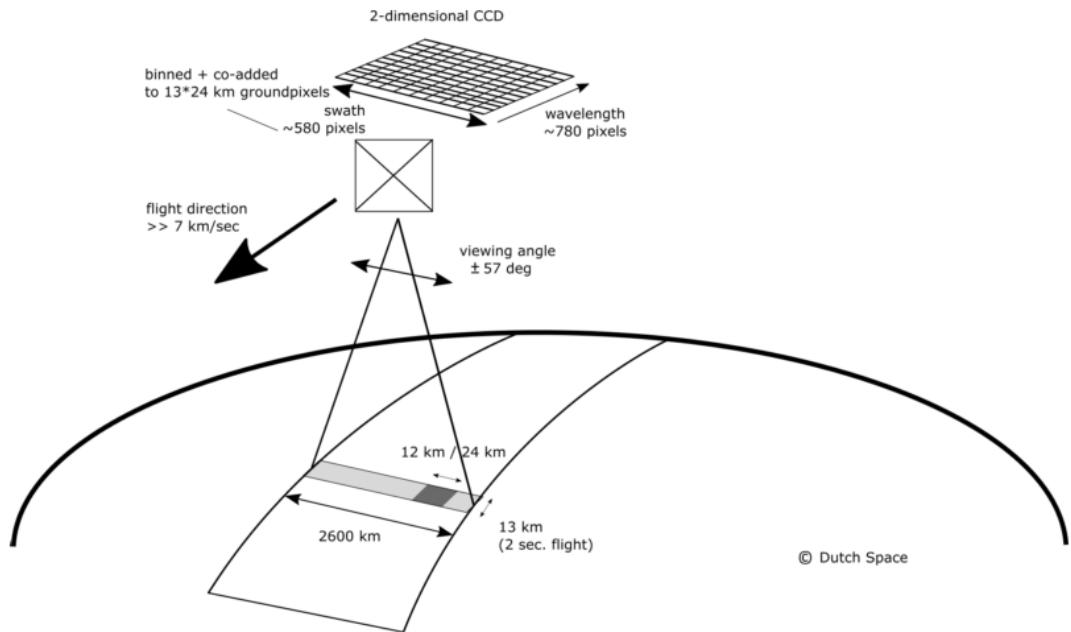
Soon even more HCHO data will be available in the form of geostationary satellite measurements (Kwon et al. (2017)). Kwon et al. (2017) examine simulated geostationary measurements against GEOS-Chem column simulations to determine the most important instrument sensitivities. Geostationary satellites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

### 2.2.2.1 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).

### 2.2.2.2 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 O<sub>3</sub>, NO<sub>2</sub>, SO<sub>2</sub>, 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. 2.2, as it shows the instruments 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



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

FIGURE 2.2: 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  $7 \text{ km s}^{-1}$  and an exposure time of 2 s lead to a ground pixel size of 13 km in the flight direction. The viewing angle of  $114^\circ$  leads to a swath width on the ground of 2600 km.” The table shows the optical properties for OMIs three channels.

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, 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)).

### 2.2.2.3 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 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](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:

$$AK(z) = \frac{\omega(z)}{AMF} \quad (2.4)$$

Note that this is an approximation for the OMI product, which does not include the AK but does include the  $\omega$  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  $2^\circ$  latitude by  $2.5^\circ$  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 2.6.

#### 2.2.2.4 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 km<sup>2</sup>) 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 2.3

### OMI HCHO and uncertainty for 200501

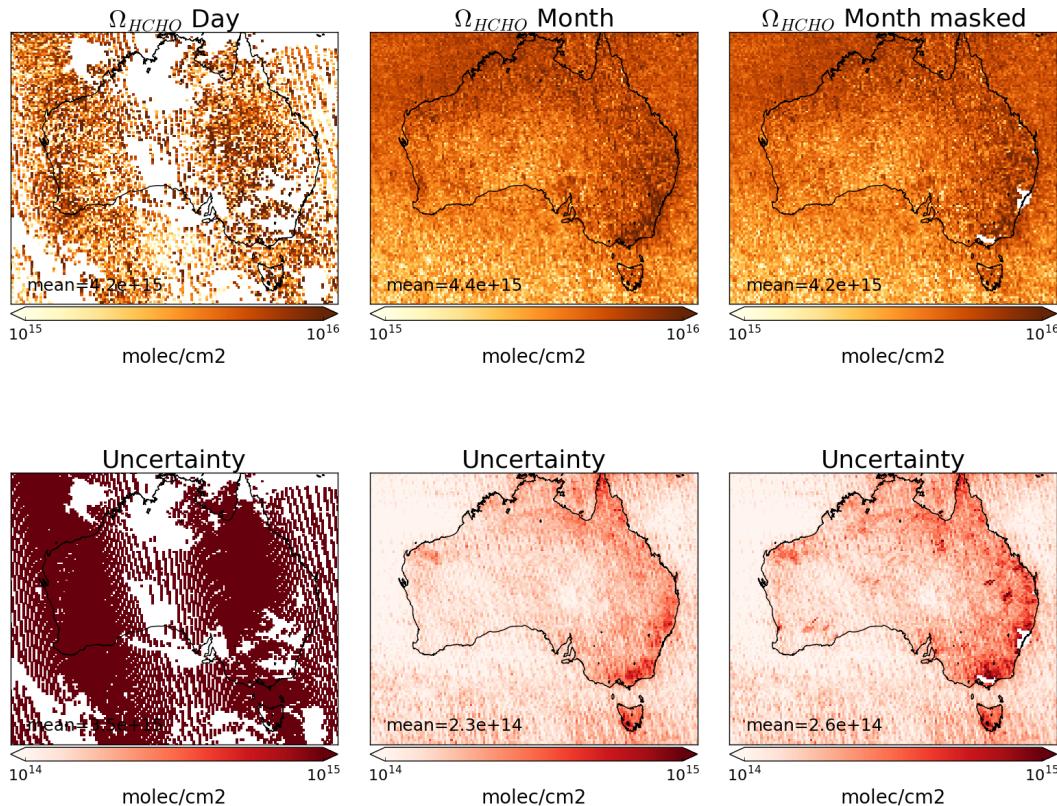


FIGURE 2.3

shows OMI HCHO columns binned to at  $0.25^\circ$ longitude by  $0.3125^\circ$ latitude averaged over one day, one month, and one month after filtering. Row two shows uncertainty 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 (SEAC<sup>4</sup>RS). 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.

AMF calculation uncertainty often dominates the total uncertainty in satellite retrievals, especially in polluted regions (Lorente et al. 2017). 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 (Lorente et al. 2017). Another impact often not included in uncertainty calculations is the structural uncertainty of retrieval methods. The structural uncertainty of AMF calculation approaches used by different retrieval groups is the uncertainty due to how the AMF is calculated, rather than uncertainty in the calculation components. The importance of a priori and ancillary data (such as surface albedo and cloud top height) sharply affects the structural uncertainty (Lorente et al. 2017).

## 2.3 Datasets

Here I will describe the various datasets used within this thesis, including modelled output, satellite measurements, and measurement campaigns. 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 for four purposes:

1. Model output validation in this chapter

2. Calculation of biogenic HCHO distribution over Australia in chapter 2
3. Determination of isoprene emissions in chapter 3
4. Extrapolation of ozone transport in chapter 4

### 2.3.1 Satellite datasets

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 (products). Aura orbits the earth in a polar sun-synchronous pattern, circling the earth on a plane coincident with the sun and the poles. 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 measures several atmospheric trace gases, including NO<sub>2</sub>, SO<sub>2</sub>, BrO, HCHO, O<sub>3</sub>, 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 km<sup>2</sup> at nadir to 26x135 km<sup>2</sup> 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. OMI spectra are used in several products used in this thesis, including OMNO2d, OMHCHO, and OMAERUVd.

#### 2.3.1.1 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 ( $\Omega_s$ ) are actually the difference with respect to the radiance reference column ( $\Omega_{S_0}$ ). The full method details for slant column retrieval by OMI are outlined in section A.5.1. Slant columns range from  $\sim 4 \times 10^{15}$  to  $\sim 6 \times 10^{16}$  molec cm<sup>-2</sup>, 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  $9 \times 10^4$  pixels, each of which contains various data including latitude, longitude vertical column HCHO, etc. The OMHCHO

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

<b>Value</b>	<b>Classification</b>	<b>Rational</b>
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"> <li>• Fit convergence flag is <math>&lt; 300</math> but <math>&gt; 0</math>: Convergence at noise level</li> <li>• Column <math>+2\sigma</math> uncertainty <math>&lt; 0 &lt;</math> Column <math>+3\sigma</math> uncertainty</li> <li>• Absolute column value <math>&gt;</math> Maximum column amount (<math>1e19</math> molec <math>cm^{-2}</math>)</li> </ul>
2	Bad	Avoid using as one of the following conditions are present: <ul style="list-style-type: none"> <li>• Fit convergence flag is <math>&lt; 0</math> : No convergence, abnormal termination</li> <li>• Column <math>+3\sigma</math> uncertainty <math>&lt; 0</math></li> </ul>
$< 0$	Missing	No column values have been computed; entries are missing

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. If greater than 40% of a pixel measurement is cloudy (ie. cloud fraction  $> 0.4$ ) then the pixel is removed from subsequent analysis. This removes around 30% of the pixels which remain after filtering out the bad or missing data.

Each  $\sim 90$  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 ( $\Omega$  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

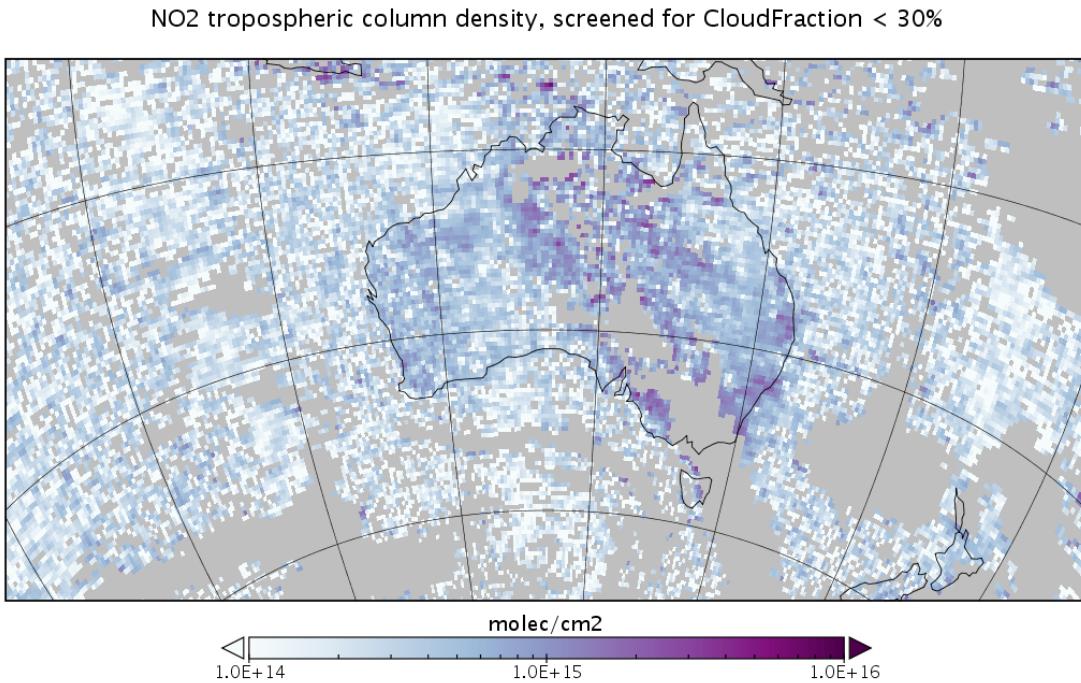


FIGURE 2.4: Example of NO<sub>2</sub> tropospheric columns taken from the OMNO2d product.

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.

### 2.3.1.2 OMNO2

NO<sub>2</sub> measured by OMI is used to check whether NO<sub>2</sub> 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.

### 2.3.1.3 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 2.2.1). 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](https://disc.gsfc.nasa.gov/datasets/OMAERUVd_V003/summary). The OMAERUVd

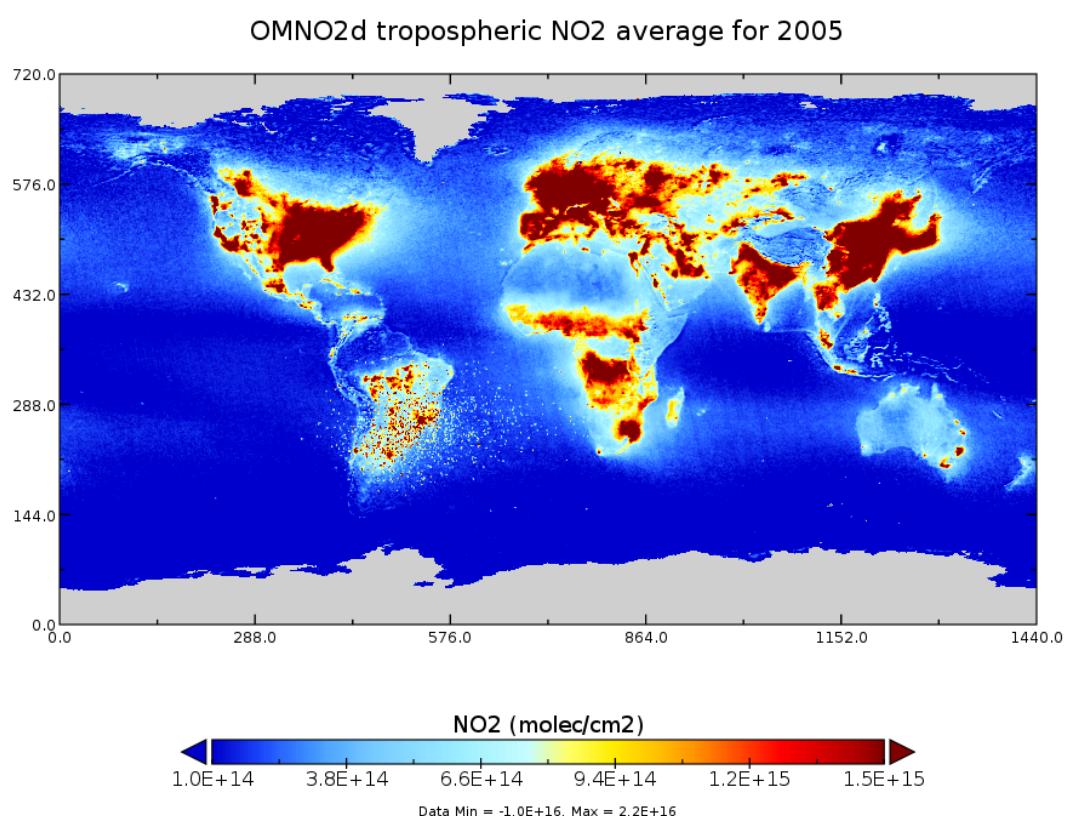


FIGURE 2.5: Average 2005 tropospheric NO<sub>2</sub> from OMNO2d with pixels screened for < 30% cloud cover.

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  $1 \times 1^\circ$  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  $> 0.03$ .

### **2.3.1.4 MOD14A1**

Daily gridded fire counts compiled from Terra and Aqua satellite into  $1 \times 1 \text{ km}^2$  resolution. Using this product after binning into a lower resolution allows an active fire influence mask (see section [2.7.1](#)).

### **2.3.1.5 AIRS CO columns**

In chapter [4](#), potential biomass burning plumes are identified using satellite observations of CO from the AIRS (Atmospheric Infra-red Sounder) instrument aboard the Aqua satellite ([Texeira 2013](#)).

## **2.3.2 Model datasets**

### **2.3.2.1 Meteorological reanalysis**

Synoptic scale weather patterns are taken from the European Centre for Medium-range Weather Forecasts (ECMWF) Interim Reanalysis (ERA-I) ([Dee et al. 2011](#)). These are used in chapter [4](#) to determine typical weather systems for stratospheric ozone intrusions.

### **2.3.3 Campaign datasets**

In this thesis data from several measurement campaigns are used to examine accuracy of modelled data at specific sites. Figure [2.6](#) shows the locations of each of the campaigns used in this work. 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.

#### **2.3.3.1 Daintree**

Daintree summary (P. Nelson)



FIGURE 2.6: Locations of Australian campaigns which are analysed within this thesis

### 2.3.3.2 Marine and Urban MBA ? (MUMBA)

### 2.3.3.3 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,  $\alpha$ -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 2.7 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 2.8 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.

### 2.3.3.4 Ozonesondes

Ozonesonde data comes from the World Ozone and Ultraviolet Data Centre (WOUDC).

## 2.4 GEOS-Chem

### 2.4.1 Overview

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). Chemistry, transport, and meteorology are simulated at 15 minute

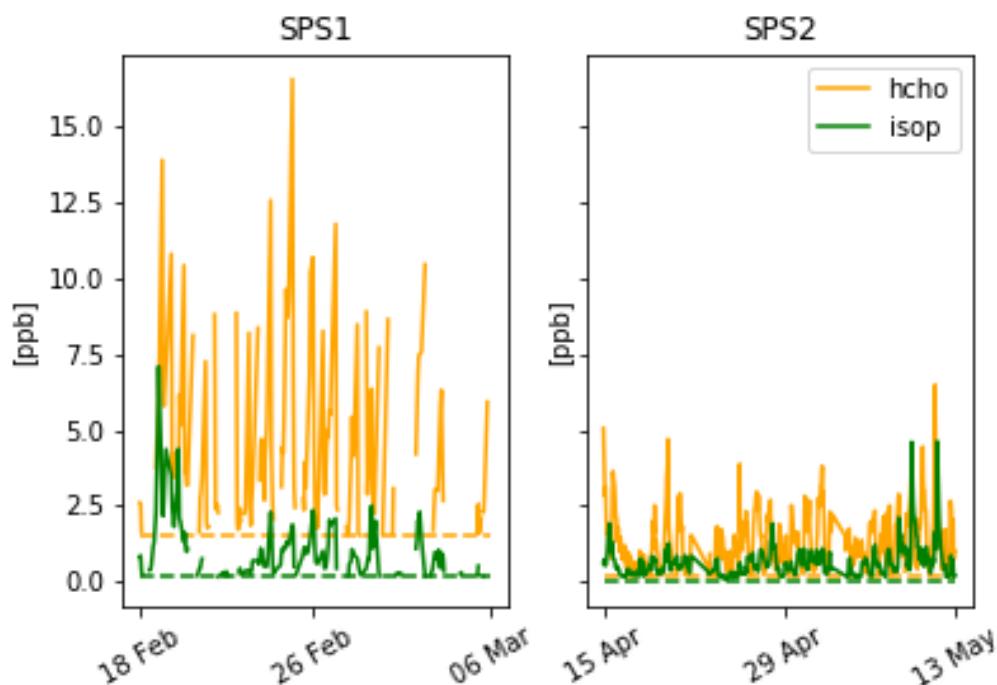


FIGURE 2.7: SPS HCHO (yellow) and isoprene (green) time series, along with detection limits (dashed). SPS 1 (left) took place in late summer 2011, while SPS 2 (right) occurred during Autumn.

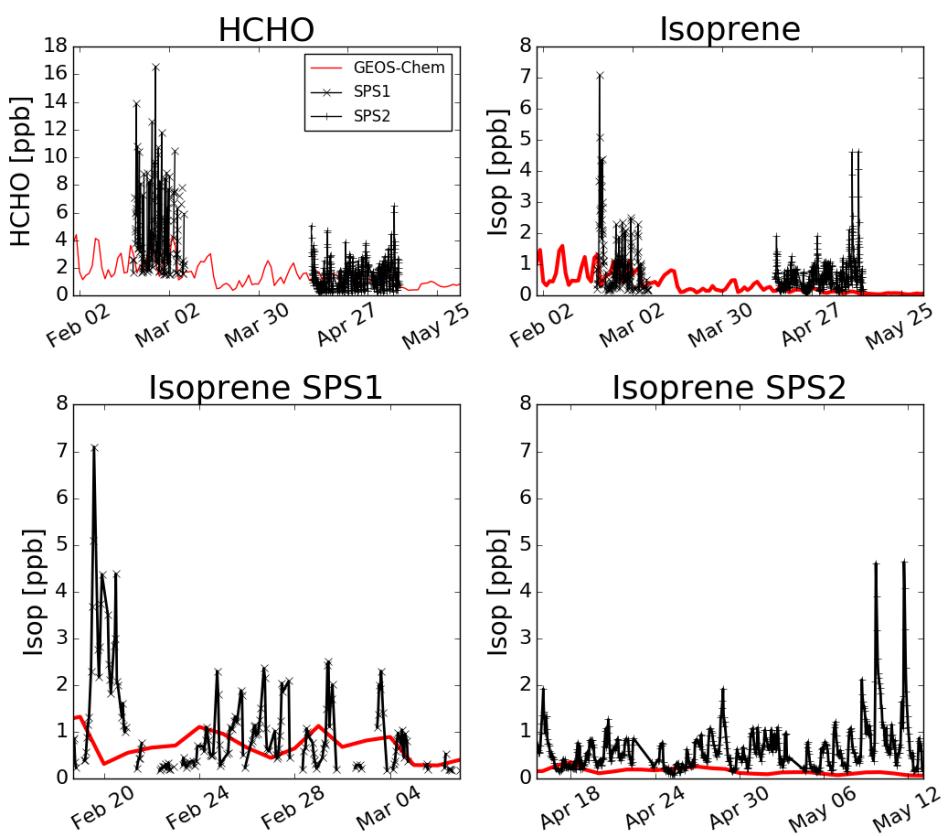


FIGURE 2.8: Comparison between GEOS-Chem HCHO concentrations in the gridsquare containing Sydney for the duration of the SPS 1 and 2 campaigns

time steps within a global set of 3-D boxes. Emissions are either prescribed by inventories or modelled (eg. biogenic emissions are created using the Model of Emissions of Gases and Aerosols from Nature (MEGAN)).

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 outputs up to 66 chemical species (tracers) in the standard run, at 2 by  $2.5^\circ$  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 used by GEOS-Chem are described here. Meteorological fields are taken from NASA's GEOS-5 dataset ( $0.5^\circ \times 0.666^\circ$ ) (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 (see section 2.4.5). MEGAN is used to determine biogenic emissions for our default GEOS-Chem simulation. The estimated biogenic VOC emissions are important for accurately simulating chemistry within models, as discussed in Section 1.1.2.

## 2.4.2 Running GEOS-Chem (before isop?)

### 2.4.2.1 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.

### 2.4.3 GEOS-Chem isoprene modelling

The isoprene reactions simulated by GEOS-Chem were originally based on Horowitz et al. (1998). This involved simulating NO<sub>x</sub>, O<sub>3</sub>, 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 C<sub>5</sub>-hydroperoxyaldehydes (HPALDs) by isomerisation. Prior to 2012 oxidation chamber studies were performed in high NO or HO<sub>2</sub> concentrations, giving peroxy lifetimes of less than 0.1 s (Crounse et al. 2012). In most environments NO and HO<sub>2</sub> concentrations are not so high, GEOS-Chem uses production rates for different NO concentrations and peroxy radical lifetimes determined by Crounse et al. (2012). OH regeneration through photolysis of hydroperoxy-methyl-butinals (HPALDs, produced by isoprene isomerisation) in areas with high isoprene emissions are included from Peeters and Muller (2010). Photolysis of photolabile peroxy-acid-aldehydes creates OH and improved model agreement with continental observations. 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 (ISOPN) have an effect on ozone levels through NO<sub>x</sub> 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<sub>x</sub> chemistry (150 min - 600 min), and finally HO<sub>x</sub> 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<sub>x</sub> ISOPOO reacts with HO<sub>2</sub> (70% yield of hydroxy hydroperoxides, ISOPOOH), RO<sub>2</sub> (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), recycling OH (Paulot et al. 2009b). HPALDs can photolise 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 low NO<sub>x</sub> conditions production of HCHO , MVK, and MACR is 4.7%, 7.3%, and 12% respectively.

Under high NO<sub>x</sub> conditions, isoprene undergoes OH addition at the 1 and 4 positions, becoming  $\beta$  (71%) or  $\delta$  (29%) hydroxyl peroxy radicals (ISOPO<sub>2</sub>). The  $\beta$ -hydroxyl reacts with NO<sub>x</sub> and produces HCHO (66%), methylvinylketone (40%) (MVK), methacrolein (26%), and  $\beta$ -hydroxyl nitrates (6.7%) (ISOPNB). The  $\delta$ -hydroxyl reacts with NO to form  $\delta$ -hydroxyl nitrates (24%) (ISOPND), and ISOPNB (6.7%). ISOPNB and ISOPND yield first generation isoprene at 4.7% and 7% respectively.

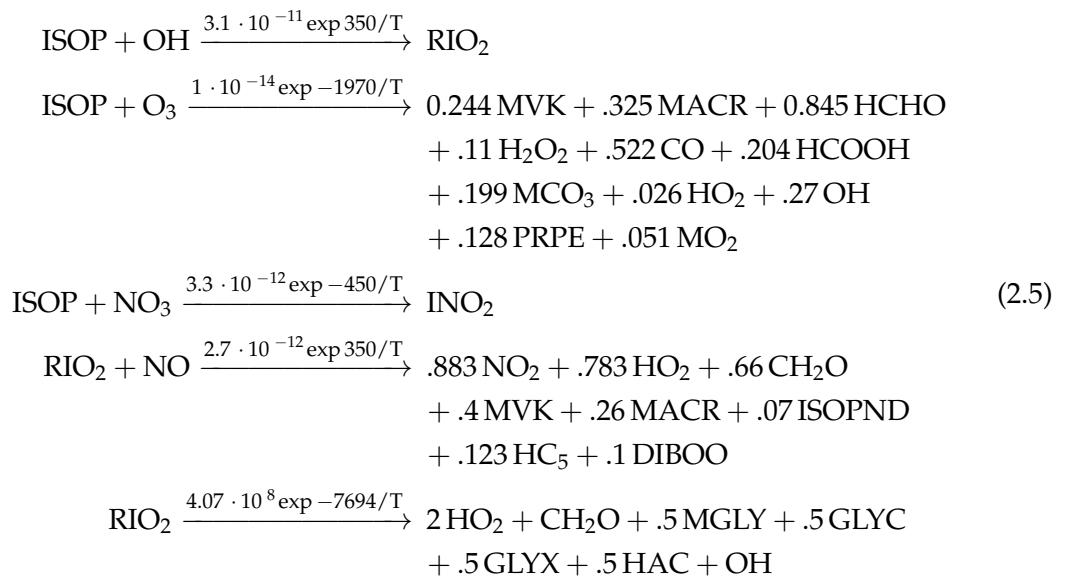
The isoprene mechanism in GEOS-Chem includes OH regeneration from oxidation of epoxydiols and slow isomerisation of ISOPO<sub>2</sub> (Mao et al. 2013). In older models isoprene produced ISOPOOH which then titrated OH, however, the loss of OH had not been seen in measurements (Paulot et al. 2009b; Mao et al. 2013). Mao et al. (2013) show that a lower (factor of 50) rate constant for ISOPO<sub>2</sub> isomerisation leads to better organic nitrate agreements with ICARTT. The chemistry updates have led to more accurate modelling of OH concentrations, especially in low NO<sub>x</sub> 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. 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](http://wiki.seas.harvard.edu/geos-chem/index.php/New_isoprene_scheme).

#### 2.4.4 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



In these reactions T is temperature.

#### 2.4.5 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 2.9.

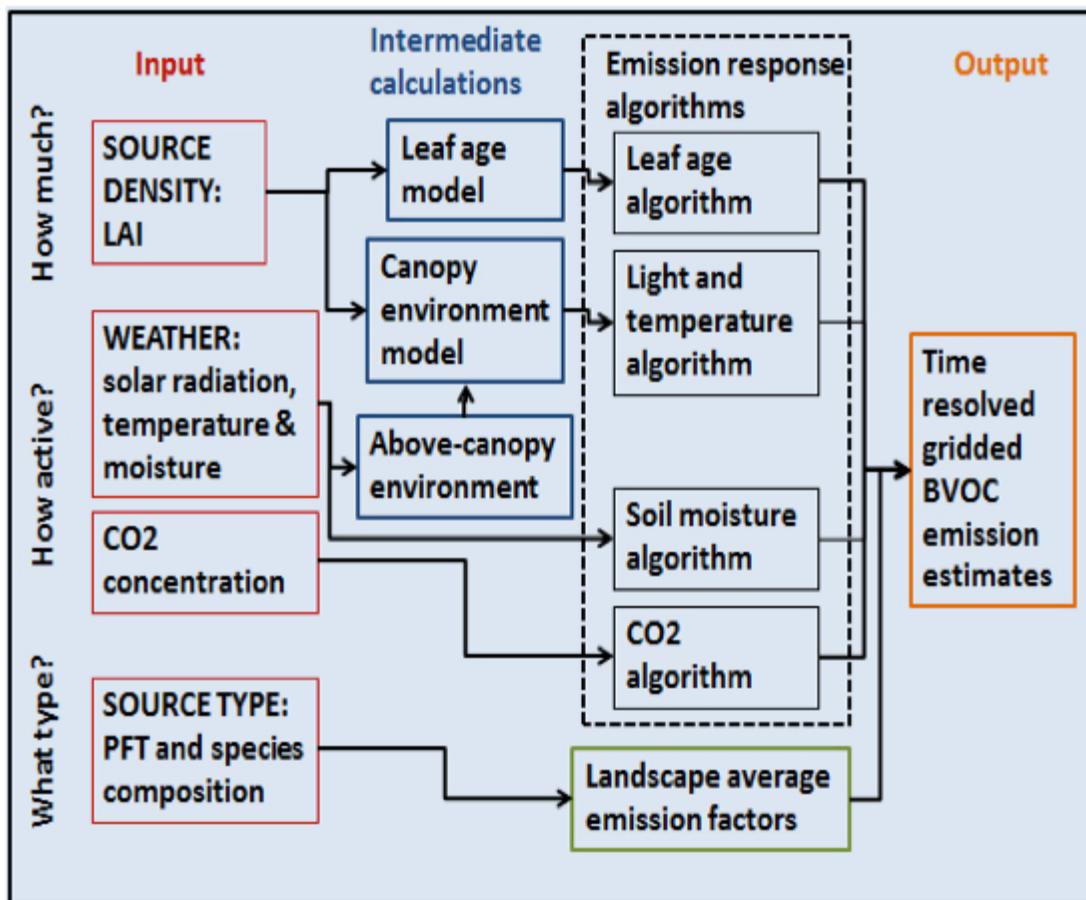


FIGURE 2.9: MEGAN schematic, copied from Guenther (2016)

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).

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) 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 affect estimates.

Instructions to run version 2.1 are available at [http://lar.wsu.edu/megan/docs/MEGAN2.1\\_User\\_GuideWSU.pdf](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.

#### 2.4.6 Rescaling NO<sub>x</sub>

NO<sub>x</sub> 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<sub>x</sub>, isoprene to HCHO yield and transport (see ??) may be poorly estimated. In order to determine if rescaling the NO emissions over Australia is necessary in GEOS-Chem, I looked at modelled NO<sub>2</sub> amounts compared to satellite data for most of 2005.

Simulated GEOS-Chem tropospheric NO<sub>2</sub> columns averaged from 1300-1400 LT are compared against OMNO2d data (Sec. 2.3.1.2). Figure 2.10 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 NO<sub>2</sub> hotspots, which are underestimated by GEOS-Chem due to averaging over the 2x2.5°horizontal resolution. Over much of the country GEOS-Chem overestimates NO<sub>2</sub> 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 2.10, and winter (JJA) of 2005 in Figure 2.10. The poor correlations for anthropogenic NO suggest that blanket alterations over Australia would not lead to improved NO<sub>2</sub> fit.

This comparison is expanded, including a comparison against modelled emissions, and repeated for autumn (MAM), winter (JJA), and spring (SON) in figures 2.12 to 2.19. These show an analysis of GEOS-Chem NO emissions and their correlations with the bias between GEOS-Chem NO<sub>2</sub> 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 NO<sub>2</sub> 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 2.12 to 2.19 show that the visible biases are not driven by modelled emissions of NO. While the correlation between column NO<sub>2</sub> 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 NO<sub>2</sub> columns away from satellite

## GC NO vs OMNO2d 20050101-20050228

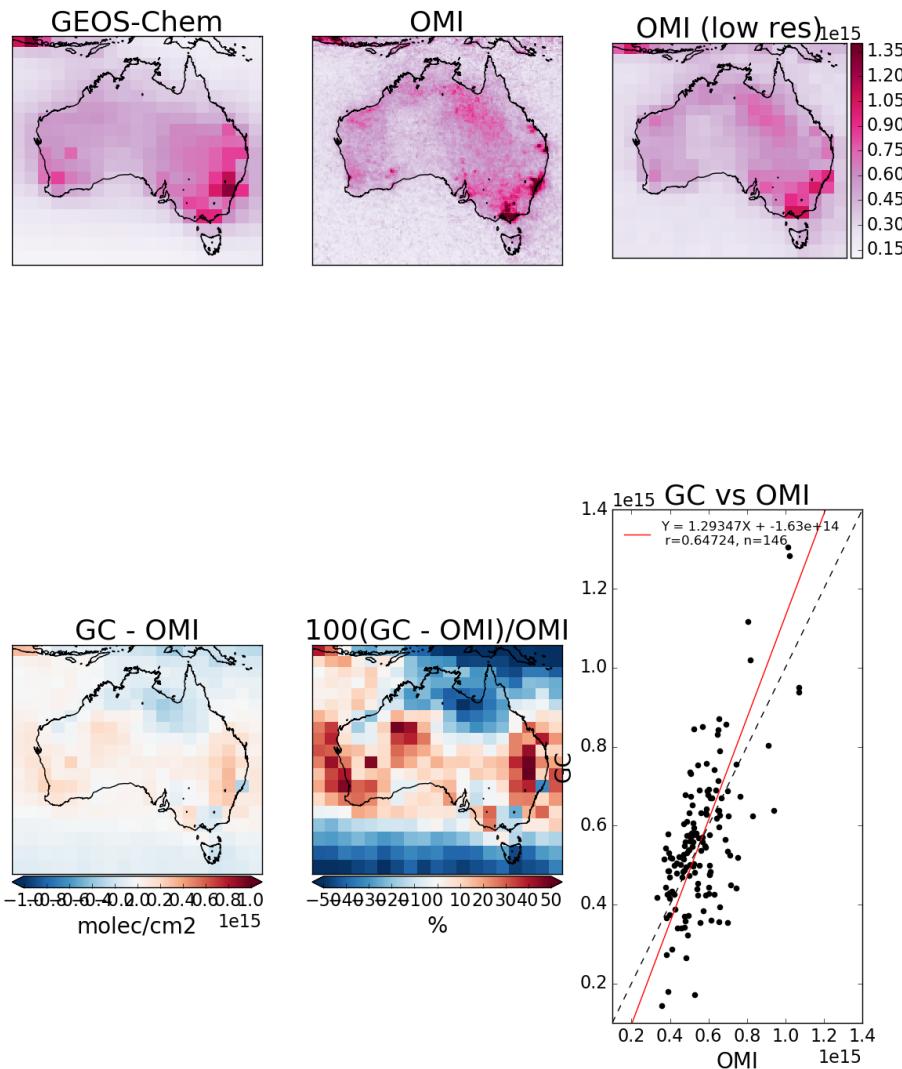


FIGURE 2.10: Row 1 shows the tropospheric columns in molec cm<sup>-2</sup>, GEOS-Chem, OMNO2d, and OMNO2d averaged onto the lower resolution of GEOS-Chem from left to right. Row 2 shows the correlations of GEOS-Chem (X axes) between daily anthropogenic emissions, and mid-day OMNO2d columns. Row 3 shows the differences with OMNO2d columns averaged into the lower resolution of GEOS-Chem.

### GC NO vs OMNO2d 20050101-20050228

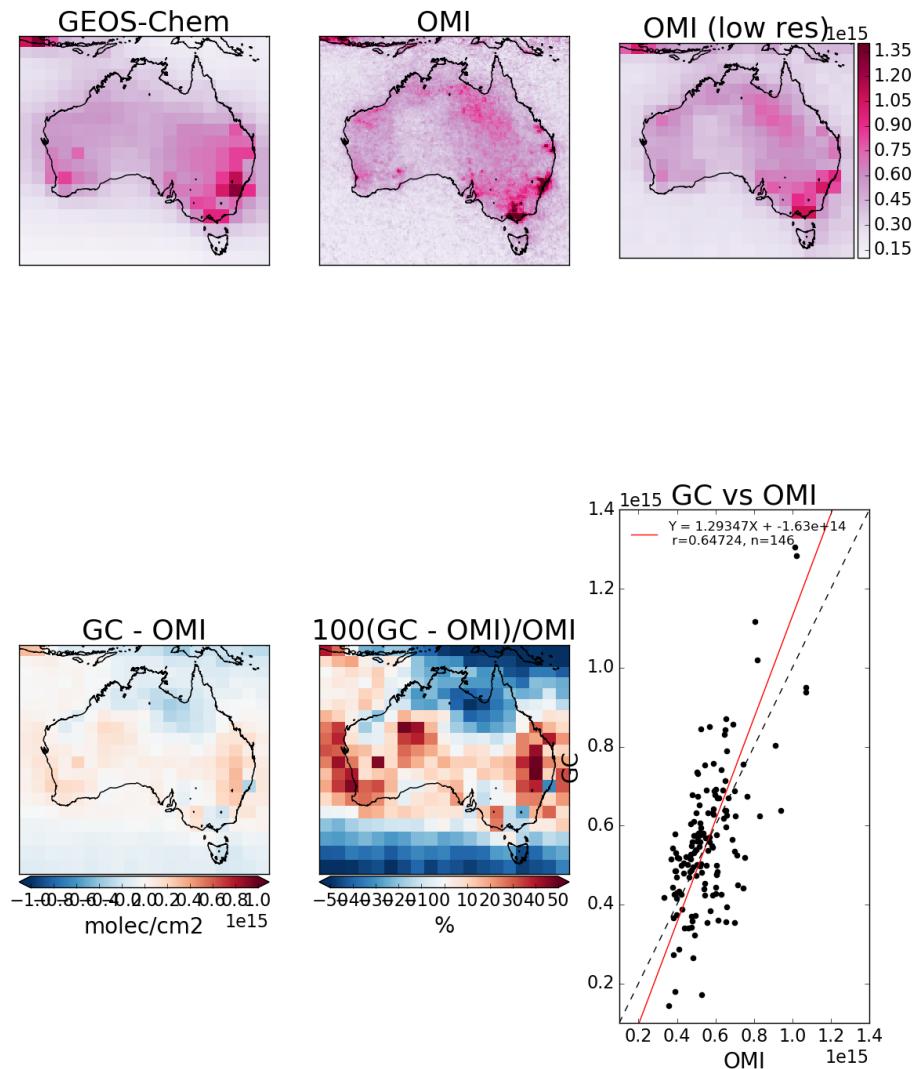


FIGURE 2.11: Row 1 shows the tropospheric columns in molec cm<sup>-2</sup>, GEOS-Chem, OMNO2d, and OMNO2d averaged onto the lower resolution of GEOS-Chem from left to right. Row 2 shows the correlations of GEOS-Chem (X axes) between daily anthropogenic emissions, and mid-day OMNO2d columns. Row 3 shows the differences with OMNO2d columns averaged into the lower resolution of GEOS-Chem.

### GEOS-Chem vs OMNO2d Jan-Feb 2005

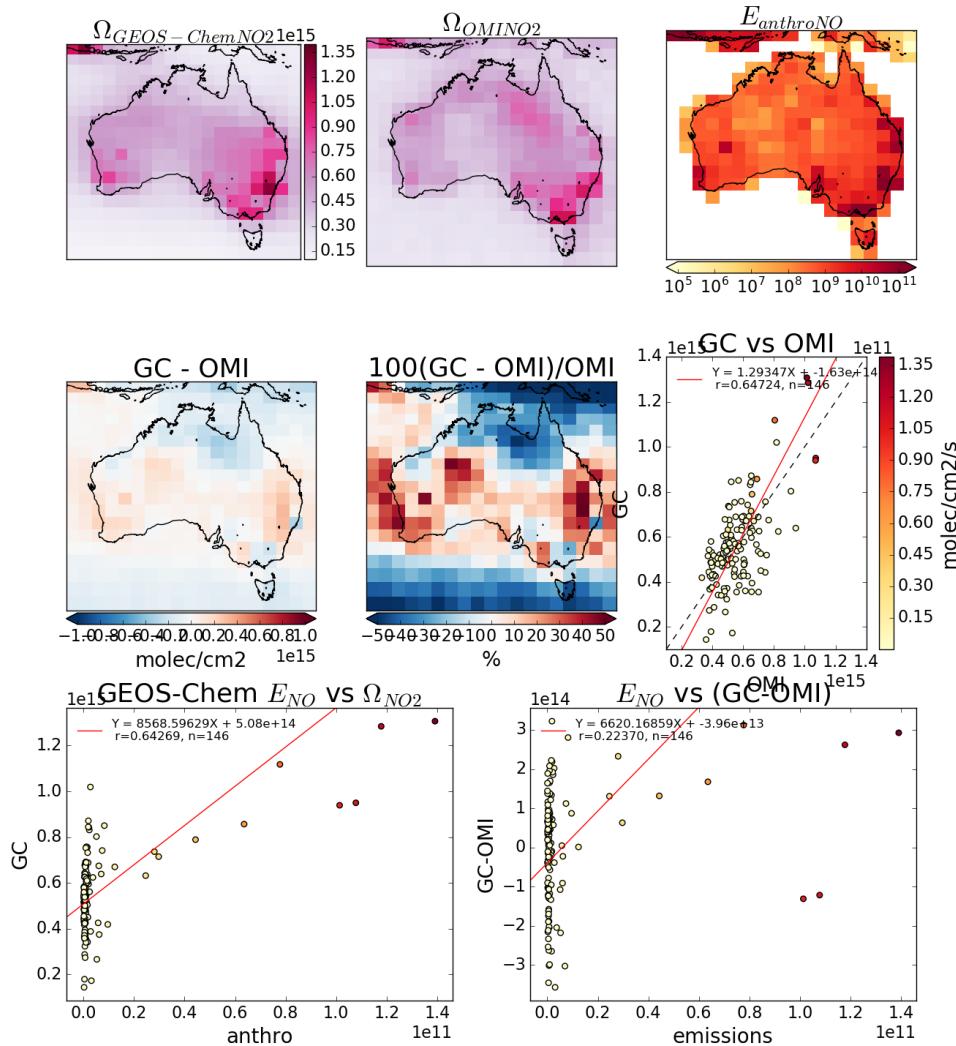


FIGURE 2.12: Top row (left to right): GEOS-Chem NO<sub>2</sub> mid-day tropospheric columns, OMNO2d NO<sub>2</sub> columns, modelled anthropogenic NO emissions. Second row: absolute and relative difference between GEOS-Chem and OMI NO<sub>2</sub> data, and the correlation. Third row: correlation between GEOS-Chem tropospheric column NO<sub>2</sub> and emitted NO, then between the model-satellite bias and the emissions. All correlation plots are coloured by emission rates.

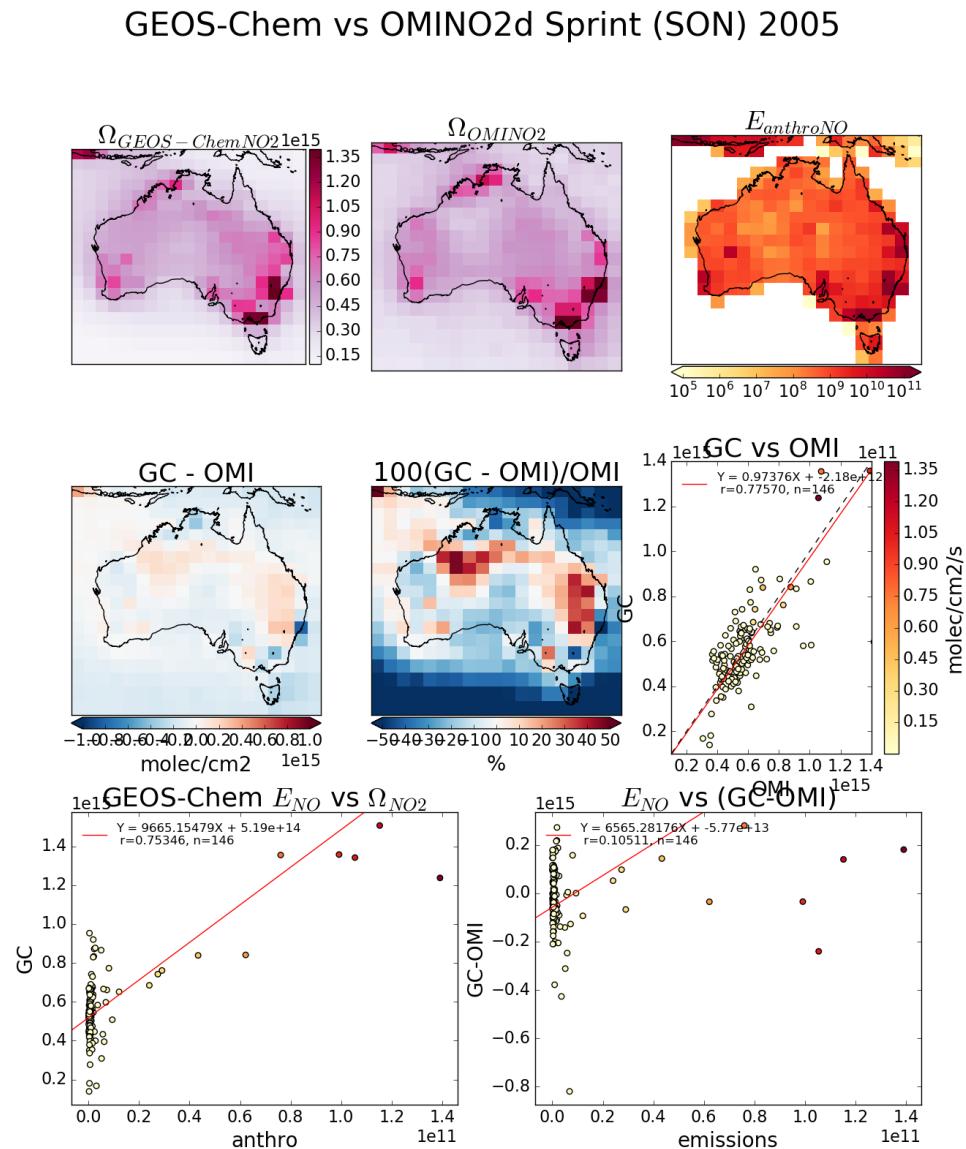


FIGURE 2.13: As figure 2.12, for Autumn 2005.

### GEOS-Chem vs OMINO2d Winter (JJA) 2005

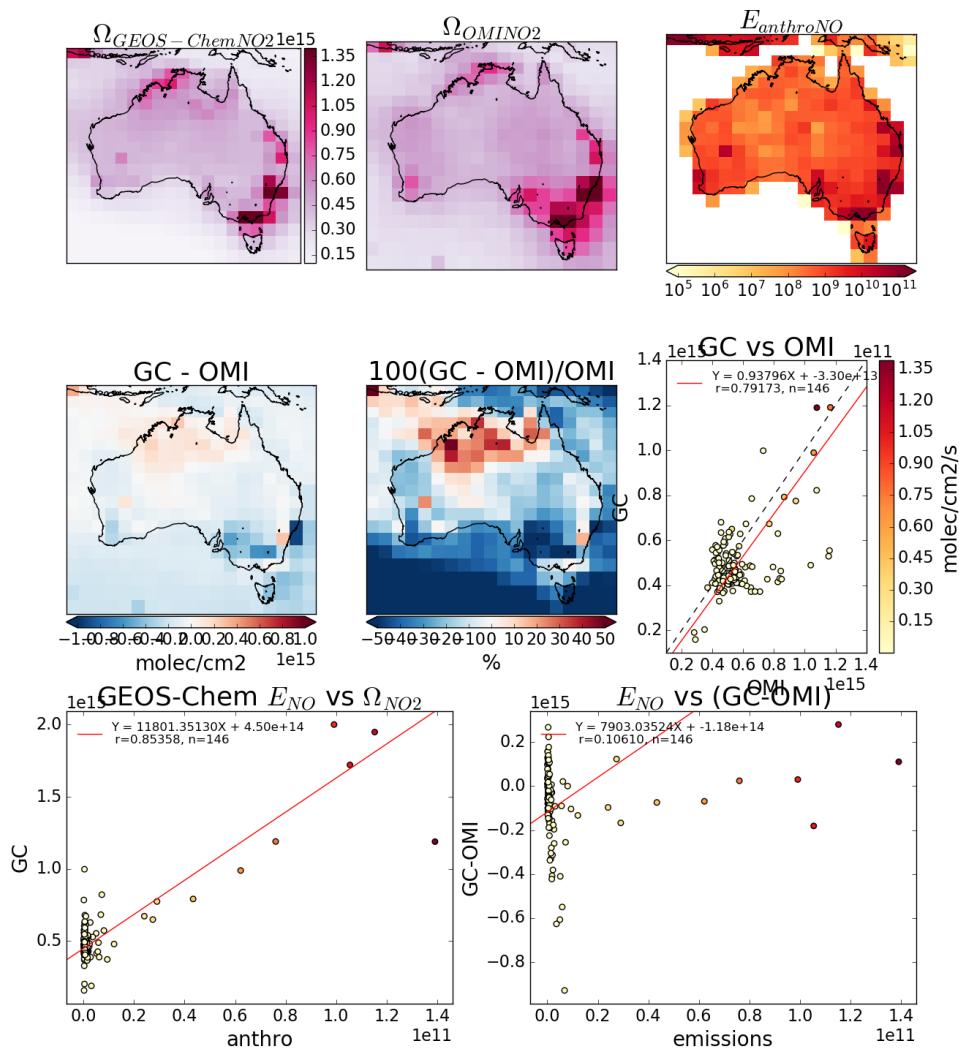


FIGURE 2.14: As figure 2.12, for Winter 2005.

GEOS-Chem vs OMINO2d Autumn (MAM) 2005

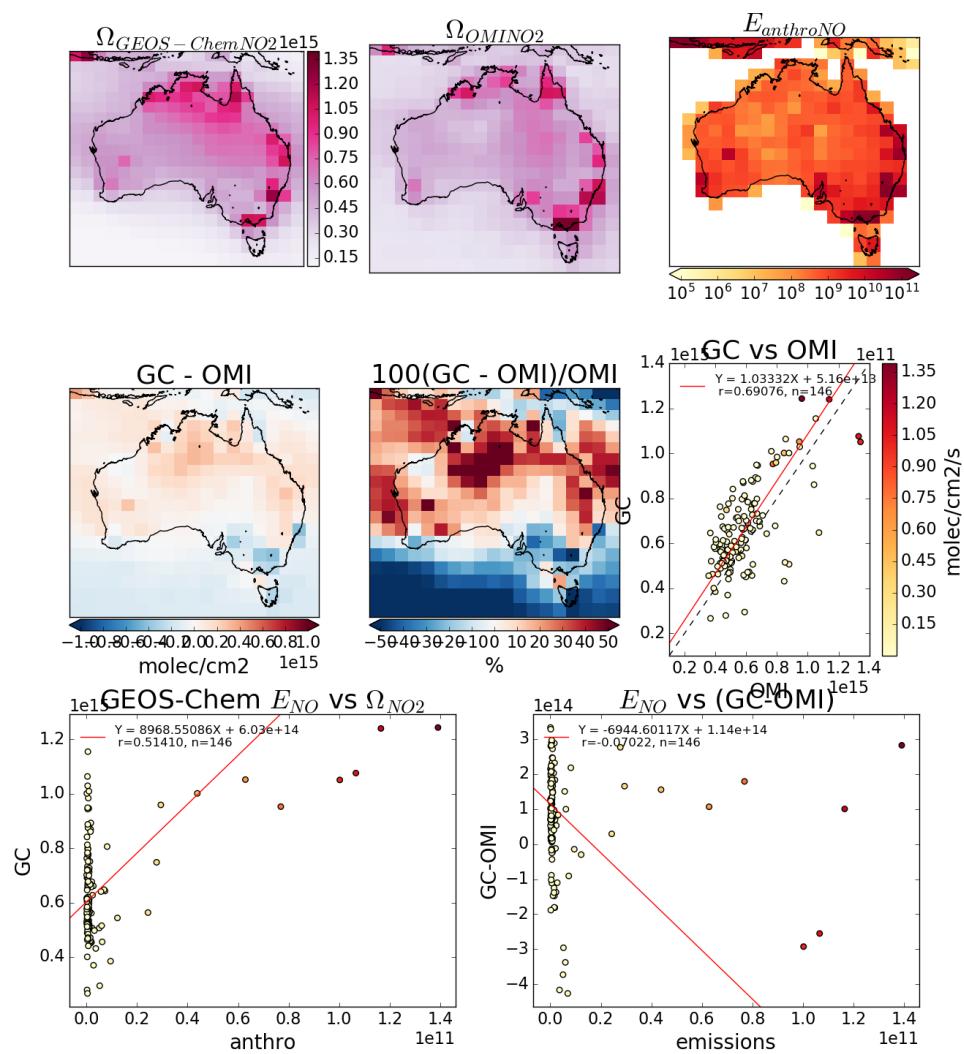


FIGURE 2.15: As figure 2.12, for Spring 2005.

### GEOS-Chem vs OMINO2d 20050101-20050228

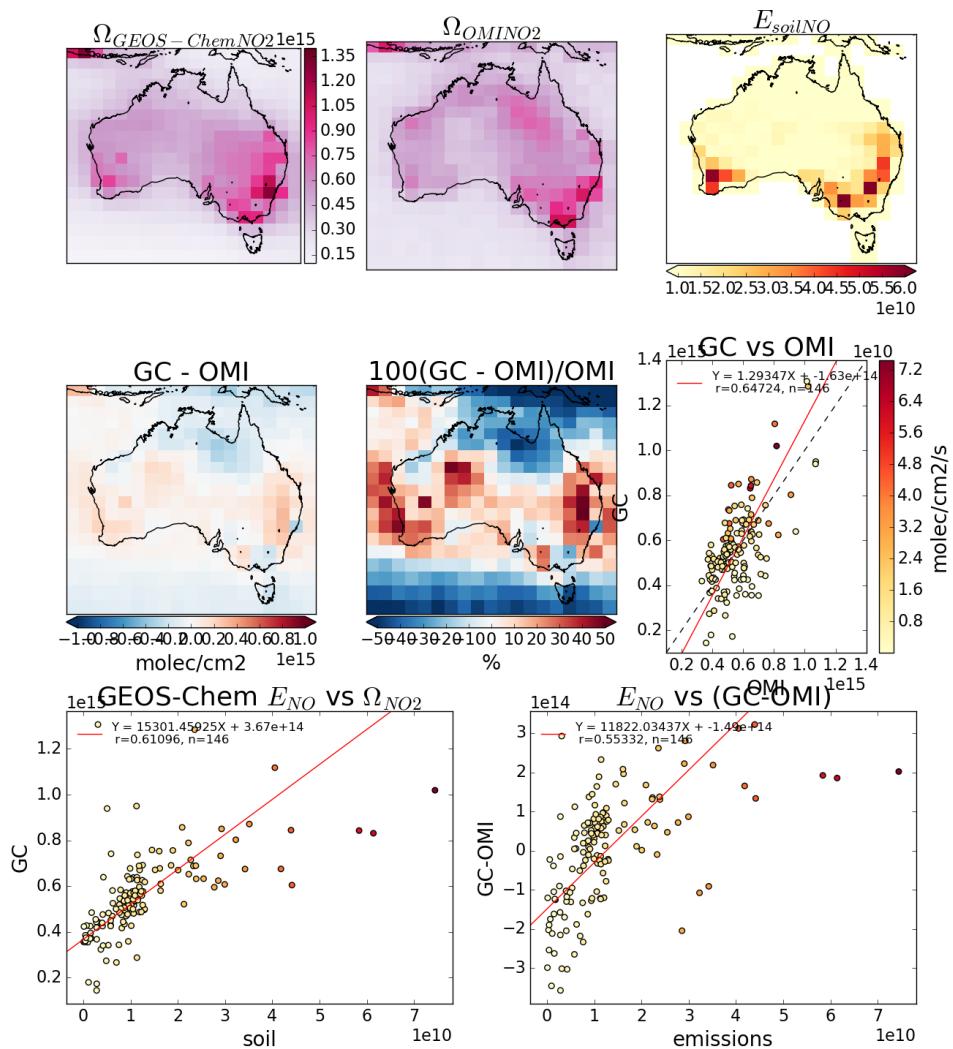


FIGURE 2.16: As figure 2.12, except anthropogenic NO emissions are replaced by soil NO emissions.

## GEOS-Chem vs OMINO2d 20050301-20050531

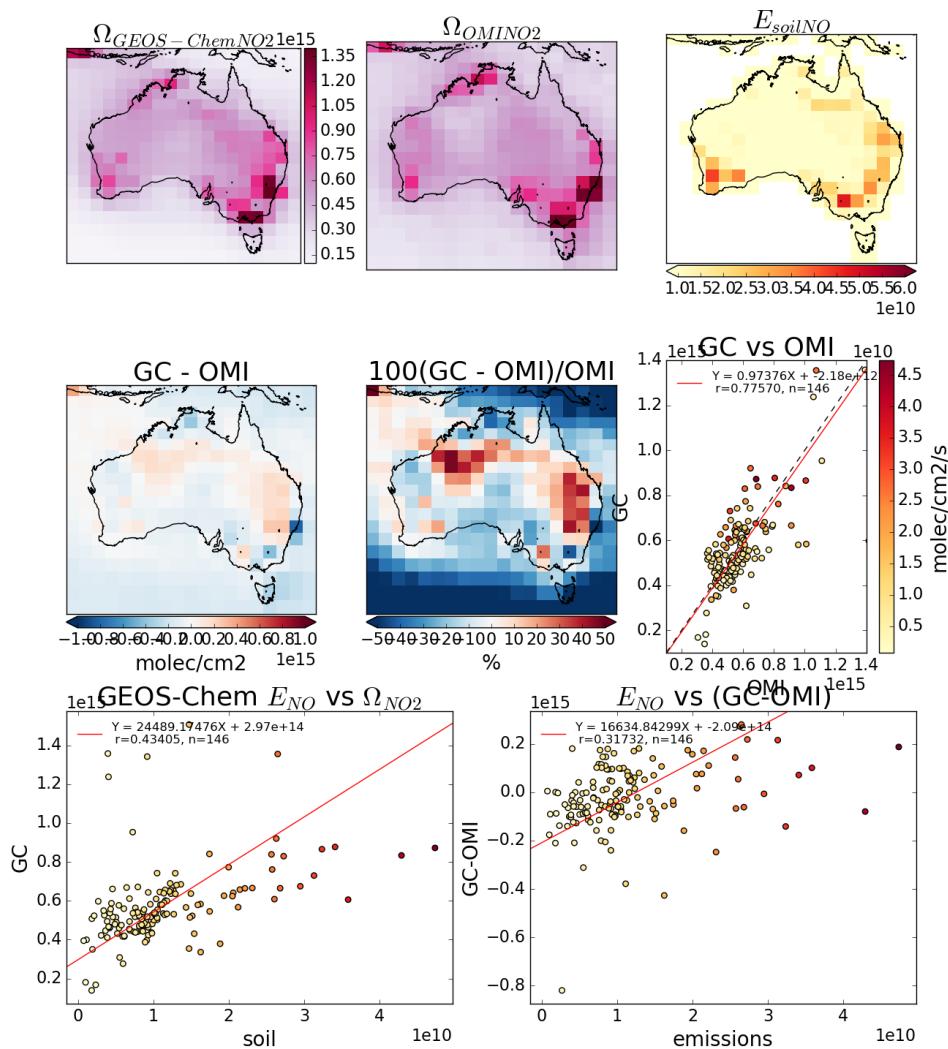


FIGURE 2.17: As figure 2.12, for Autumn 2005, with soil NO emissions replacing anthropogenic NO emissions.

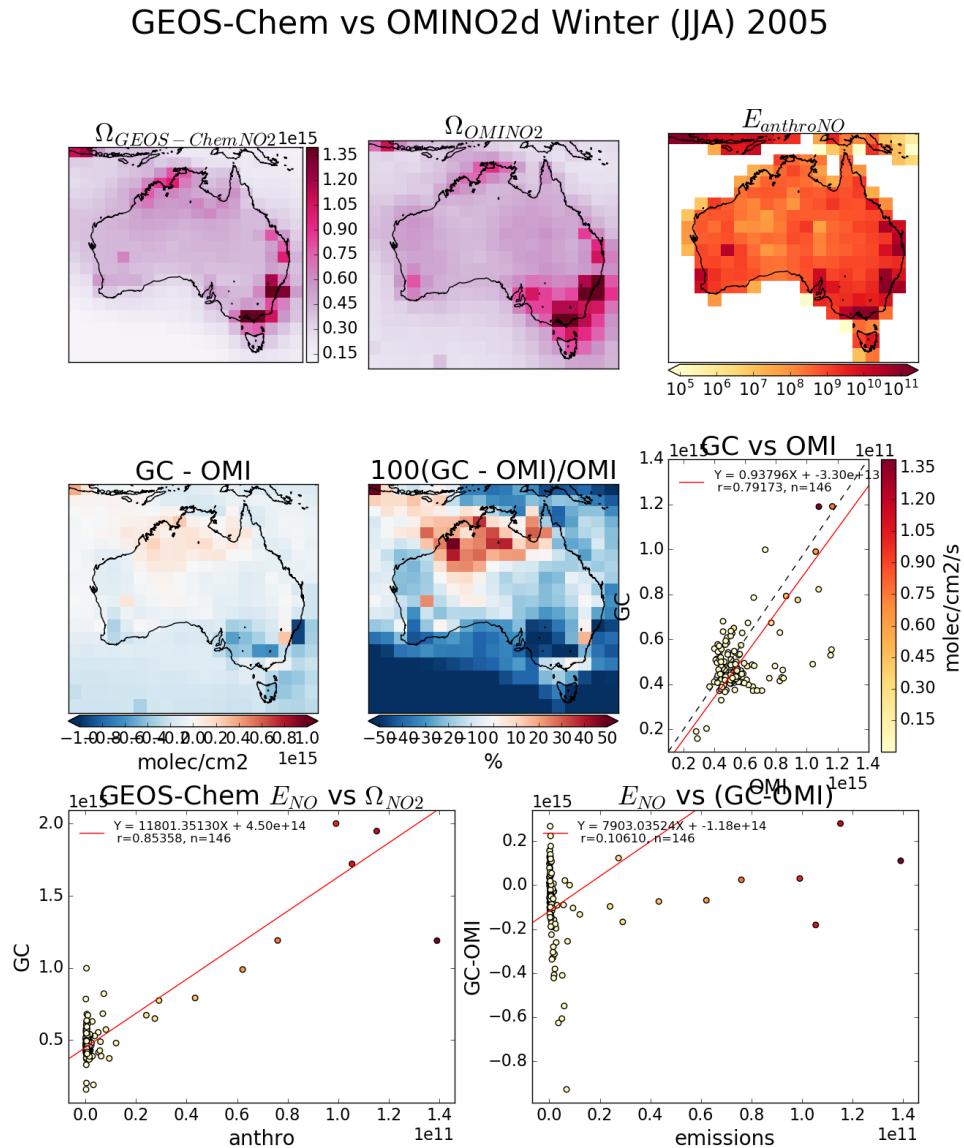


FIGURE 2.18: As figure 2.12, for Winter 2005, with soil NO emissions replacing anthropogenic NO emissions.

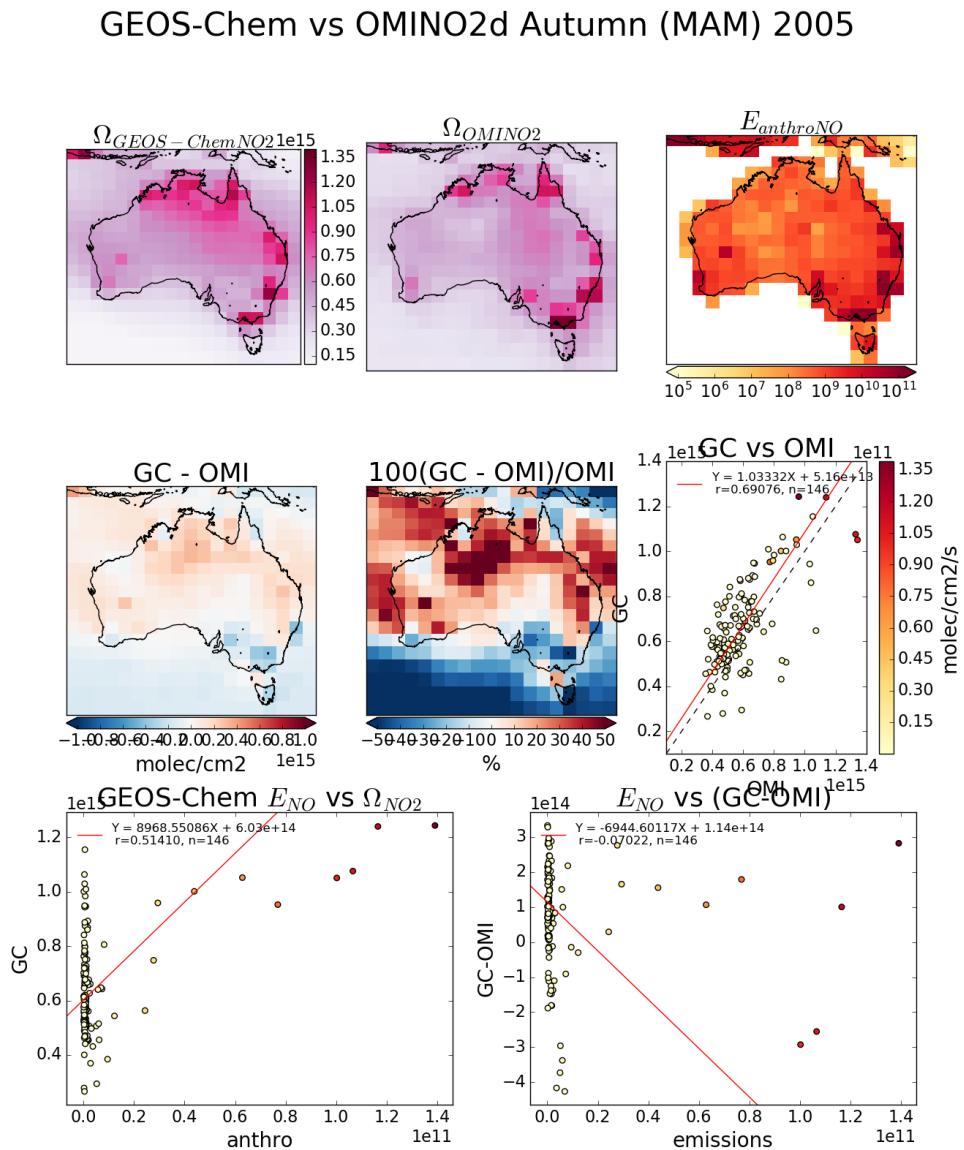


FIGURE 2.19: As figure 2.12, for Spring 2005, with soil NO emissions replacing anthropogenic NO emissions.

measurements over Australia. For this reason modelled NO emissions are not scaled in model runs in this thesis.

#### 2.4.7 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 4.

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

#### 2.4.8 GEOS-Chem simulations

GEOS-Chem is run four independent times in this thesis, with different outputs from each simulation used to determine specific information. Output averaged over 1300 - 1400 local time is saved for comparison with, and recalculation of, 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 2.6.7). 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

#### 2.4.8.1 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  $2^{\circ}$  latitude by  $2.5^{\circ}$  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

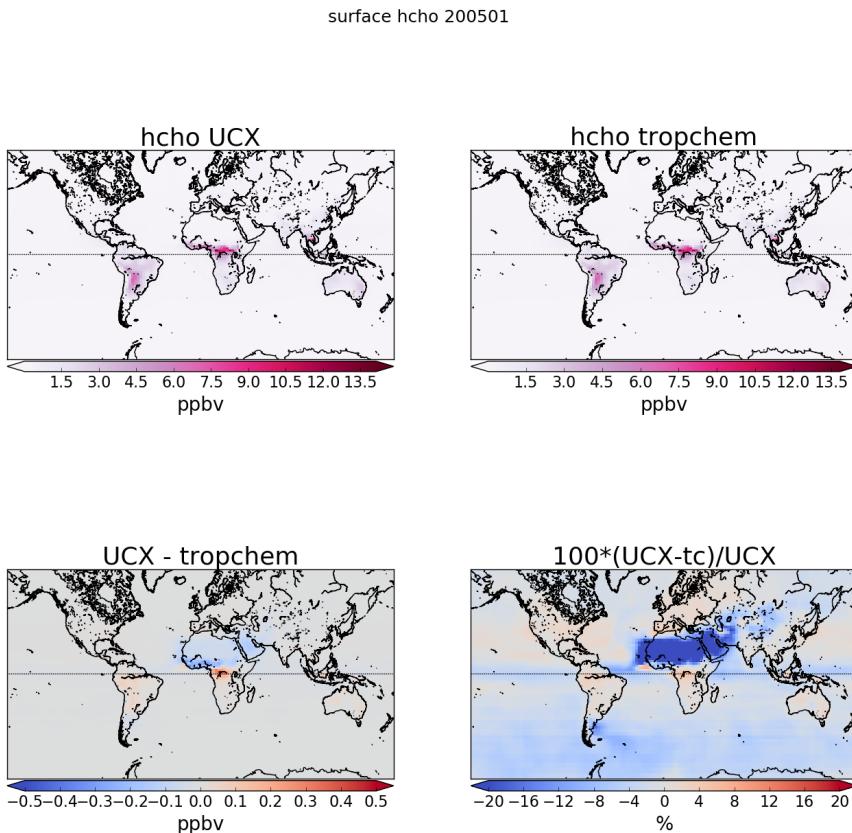


FIGURE 2.20: Surface HCHO simulated by GEOS-Chem with UCX (top left), and without UCX (top right), along with their absolute and relative differences (bottom left, right respectively). Amounts simulated by GEOS-Chem for the 1st of January, 2005.

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 2.20 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 2.21 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 ??.

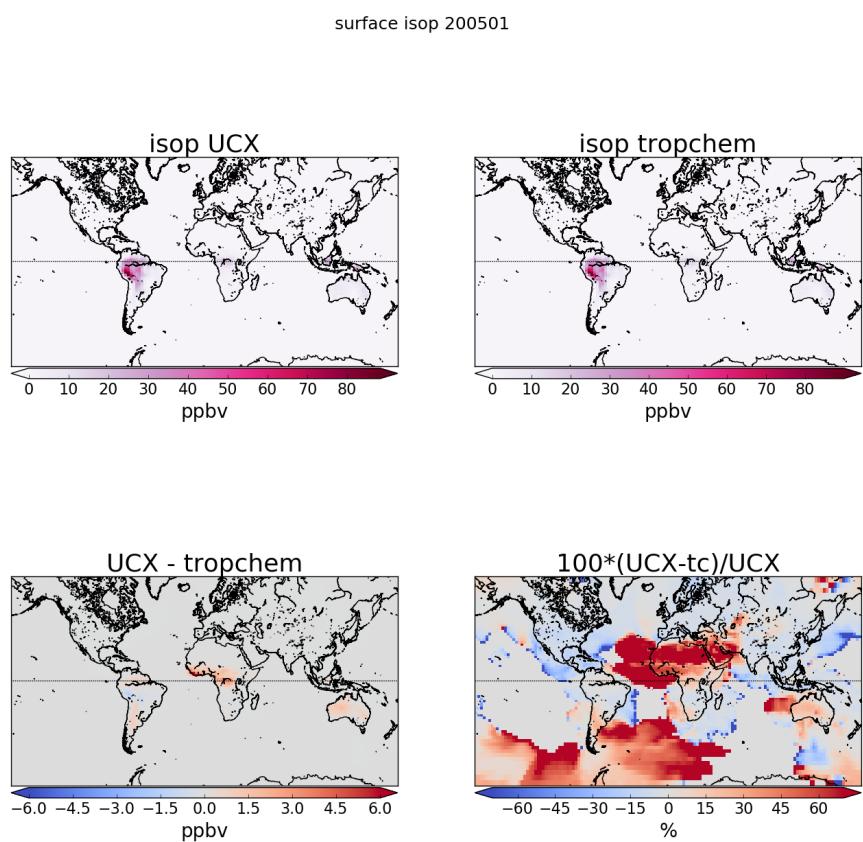


FIGURE 2.21: As figure 2.20, except looking at isoprene.

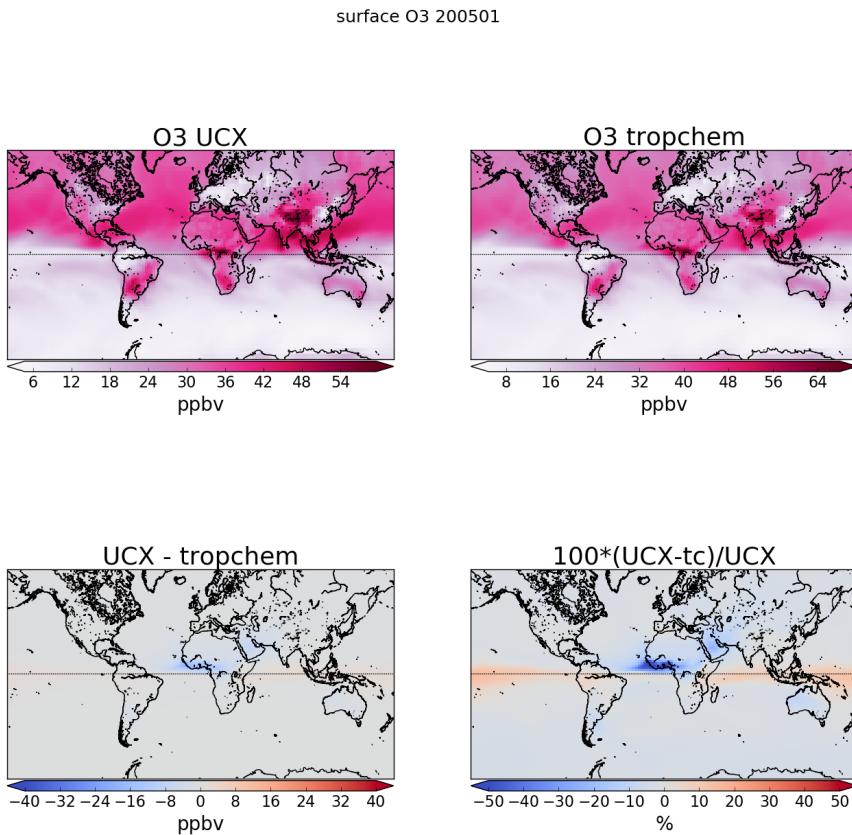


FIGURE 2.22: As figure 2.20, except looking at ozone.

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 2.22 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.

## 2.5 Calculating an AMF

The AMF is the ratio of the slant column ( $\Omega_s$ ) to the vertical column ( $\Omega_v$ )

$$AMF = \frac{\Omega_s}{\Omega_v} = \frac{\tau_s}{\tau_v} \quad (2.6)$$

with  $\tau$  being the optical depth or thickness of the absorber through the slant ( $s$ ) or vertical ( $v$ ) path of light

The OMI instrument records spectra of light which enters the viewing lens on board the satellite. The spectra provide backscattered intensity ( $I_B$ ) at various wavelengths (see section 2.3.1), with the light source ( $I_{B_0}$ ) being the sun. Using the log of beers law (equation 2.1) we get

$$\tau_s = \ln I_{B_0} - \ln I_B$$

which can be subbed into equation 2.6 to give an expression for the AMF which includes scattering:

$$AMF = \frac{\ln I_{B_0} - \ln I_B}{\tau_v} \quad (2.7)$$

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

$$AMF = \frac{\ln \left( 1 - \frac{\nabla I}{I_B} \right)}{\tau_v} \approx \frac{-\frac{\nabla I}{I_B}}{\tau_v} \quad (2.8)$$

This is due to the logarithmic property  $\ln(1-x) \approx -x$  for  $x \ll 1$ .  $\nabla I$  can also be expressed as the integral of the absorption slices over optical depth increments:

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

which can be placed into equation 2.8 leading to

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

We can then convert  $d\tau$  to our path using equation 2.3 leading to

$$AMF = \frac{-1}{\tau_v} \int_0^{\infty} \frac{\partial \ln I_B}{\partial \tau} \alpha(z) \eta(z) dz \quad (2.9)$$

where  $\alpha(z)$  and  $\eta(z)$  represent absorption cross section in  $\text{m}^2 \text{ molecule}^{-1}$ , and number density in molecules  $\text{m}^{-3}$  respectively. This uses the attenuation cross section relationship to optical depth (see section 2.2.1).

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

$$\begin{aligned} \alpha_e &= \frac{1}{\Omega_v} \int_0^{\infty} \alpha(z) \eta(z) dz \\ &= \frac{\tau_v}{\Omega_v} \end{aligned}$$

Then replacing the  $\tau_v$  in equation 2.9 we obtain:

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

Often the integrand of this AMF formula (equation 2.10) is broken apart into two defining terms: the scattering weights  $\omega(z)$  and the shape factor  $S(z)$ .

- $\omega$  The scattering weights describing the sensitivity of the backscattered spectrum to the abundance of an absorber at altitude  $z$ :

$$\omega(z) = - \frac{1}{AMF_G} \frac{\alpha(z)}{\alpha_e} \frac{\partial \ln I_B}{\partial \tau} \quad (2.11)$$

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

- $S$  the shape factor describes the profile of an absorber ( $\eta(z)$ ) normalised by its total vertical column amount ( $\Omega_v$ ):

$$S(z) = \frac{\eta(z)}{\Omega_v} \quad (2.12)$$

Plugging equations 2.11 and 2.12 into equation 2.10 gives us:

$$AMF = AMF_G \int_0^\infty \omega(z) S(z) dz \quad (2.13)$$

Since we are using the  $\omega$  provided by OMI, the  $AMF_G$  term is removed from this calculation as it is not part equation 2.11 leading to

$$AMF = \int_0^\infty \omega(z) S(z) dz \quad (2.14)$$

Additionally the AMF can be determined using the sigma ( $\sigma$ ) coordinate system. A conversion to the  $\sigma$  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. 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.  $S_\sigma$  is a dimensionless normalised shape factor on the  $\sigma$  coordinate system. In the sigma coordinate system we calculated the shape factor as defined in Palmer et al. (2001):

$$S_\sigma(\sigma) = \frac{\Omega_a}{\Omega_v} C_{HCHO}(\sigma) \quad (2.15)$$

where  $\Omega_a$  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. The hydrostatic relation  $P = -\rho_a g z$ , with  $\rho_a$ ,  $g$ , being density of air, gravity, respectively lets us integrate over the sigma

coordinates:

$$\rho_a g z = \sigma (P_S - P_T) + P_T$$

$$d\sigma = -\frac{\rho_a g}{P_S - P_T} dz$$

Substitution into 2.14 gives AMF using the sigma coordinates:

$$AMF = \int_0^1 w(\sigma) S_\sigma(\sigma) d\sigma \quad (2.16)$$

## 2.6 Recalculation of OMI HCHO

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

$$AMF = \frac{SC}{\Omega} \quad (2.17)$$

A slant or vertical column is expressed in molecules  $\text{cm}^{-2}$ . OMI HCHO vertical columns are calculated using modelled a priori HCHO profiles (see section 2.3.1.1). 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 ( $\Omega_O$ ) to GEOS-Chem ( $\Omega_G$ ), 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 2.6.4 and 2.6.5 for more details). A reference sector correction is determined using the method described in Abad et al. (2016), (see section 2.6.7). 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 ( $\Omega_{OC}$ ), one using GEOS-Chem shape factors ( $\Omega_{GC}$ ), and one from Palmer's code ( $\Omega_{PC}$ ).

### 2.6.1 Outline

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

1. GEOS-Chem satellite overpass output (see section 2.4.7) is used to create new shape factors ( $S_z$  and  $S_\sigma$ ).
  - (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/cm<sup>3</sup>)
  - (c) Total column HCHO from GEOS-Chem ( $\Omega_G$ ) is calculated  $\Omega_G = \Sigma_z (n_{HCHO}(z) \times H(z))$ , along with total column air ( $\Omega_A$ , calculated similarly)
  - (d) The shape factor  $S_z(z)$  is calculated on each altitude  $S_z(z) = n_{HCHO}/\Omega_{HCHO}$ .
  - (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:  $S_\sigma(z) = C_{HCHO}(z) \times \Omega_A/\Omega_{HCHO}$
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 ( $\omega$ ) are interpolated onto the same vertical dimensions ( $z$  and  $\sigma$ ) 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.

$$AMF_z = \Sigma_z (\omega(z) \times S_z(z) \times H(z)) \quad (2.18)$$

$$AMF_s = \Sigma_\sigma (\omega(\sigma) \times S_\sigma(\sigma) \times d\sigma) \quad (2.19)$$

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 ( $\Omega_P$ ,  $\Omega_G$  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] (eg. figure 2.26)

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  $0.25^\circ$ by  $0.3125^\circ$ grid along with how many pixels have been binned, and the average pixel uncertainty in product OMHCHORP.

Figure 2.23 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. PP code is run on a subset of the globe covering Australia and the pacific ocean, producing  $AMF_{PP}$ . 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 2.7.

### 2.6.2 Creating new shape factors

In order to visualise and analyse OMI HCHO columns, slant columns are transformed into vertical columns using the AMF. The shape factor ( $S$ ) is one of the key components in creation of the AMF (see section 2.5, equation 2.14). The shape factor is calculated using GEOS-Chem satellite output (see section 2.4.7) which provide simulated HCHO concentration profiles ( $\eta(z)$ ) and total columns ( $\Omega$ ) at  $2 \times 2.5^\circ$ horizontal resolution. Using equation 2.12 to determine the shape factor is straightforwards  $S(z) = \frac{\eta(z)}{\Omega}$ . The associated OMI per-pixel scattering weights are not changed in this calculation (unlike in section 2.6.5).

Model output is in molecules per billion molecules of air (ppb), and is converted before being used in the shape factor calculation. The following equation converts model profile output from ppb into number densities:

$$\eta_{HCHO} = ppb_{HCHO} \times \eta_a \times 10^{-9} \quad (2.20)$$

where  $\eta_{HCHO}$  is the number density of a HCHO,  $\eta_a$  is the number density of air (from model output), and  $ppb_{HCHO}$  is the molecules of that species per billion molecules of air. The modelled total vertical column  $\Omega_{HCHO}$  is determined by:

$$\Omega_{HCHO} = \sum_z (\eta_{HCHO} \times H(z))$$

where  $H(z)$  is the box height for level  $z$ . In effect this equation sums over the molecules per  $\text{cm}^2$  in each altitude level.

As a sanity check I also recalculate  $S_\sigma$  (through equation 2.15)and confirm that these shape factors are equivalent by comparing the resulting AMFs created by equation 2.16 and equation 2.14 for each pixel. For example see figure TODO: sigma vs z shape factor plot. TODO add plot and describe here.

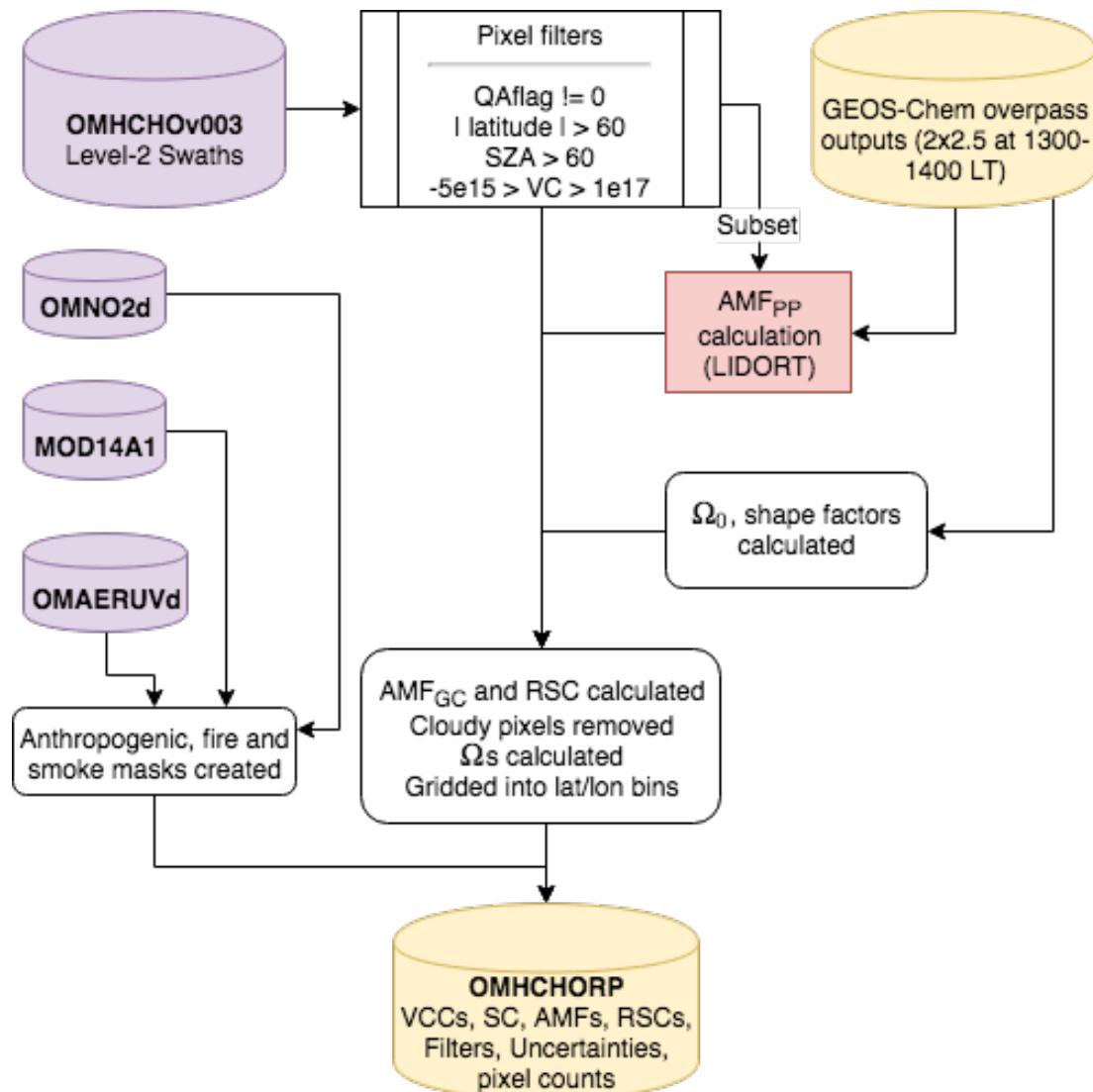


FIGURE 2.23: Flow diagram showing how OMHCHO level two swath data is read, processed, and gridded in this thesis

### 2.6.3 Reading satellite data

First satellite slant columns of formaldehyde for the years January 1st, 2005 - April 1st, 2013 are downloaded (see section 2.3.1.1). 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 2.23 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 2.3.1.1) may have affected.

Satellite measurements polewards of  $60^{\circ}$  north or south are removed as well as measurements with SZA greater than  $60^{\circ}$ . Pixels with cloud fraction greater than 40% are removed after being used in determining the reference sector correction (see section 2.6.7), as is done in Gonzalez Abad et al. (2015) and De Smedt et al. (2015). Further filtering is performed to remove the measurements which are most likely to be unrealistic: those with column density outside the range  $-0.5 \times 10^{16}$  to  $10^{17}$  molec  $\text{cm}^{-2}$ , as is performed by Zhu et al. (2016). This filter is required due to currently unexplained large negative values which occur in the OMI HCHO product increasingly over time. Figure 2.24 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  $\text{cm}^{-2}$  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.

### 2.6.4 Creating the new AMF

From equation 2.14 we have:

$$\text{AMF} = \int_0^{\infty} \omega(z) S(z) dz$$

then using the  $\omega(z)$  from satellite output, along with our calculated  $S_z$  interpolated linearly onto the same vertical grid as  $\omega(z)$ , the AMF can be determined through integration. 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.

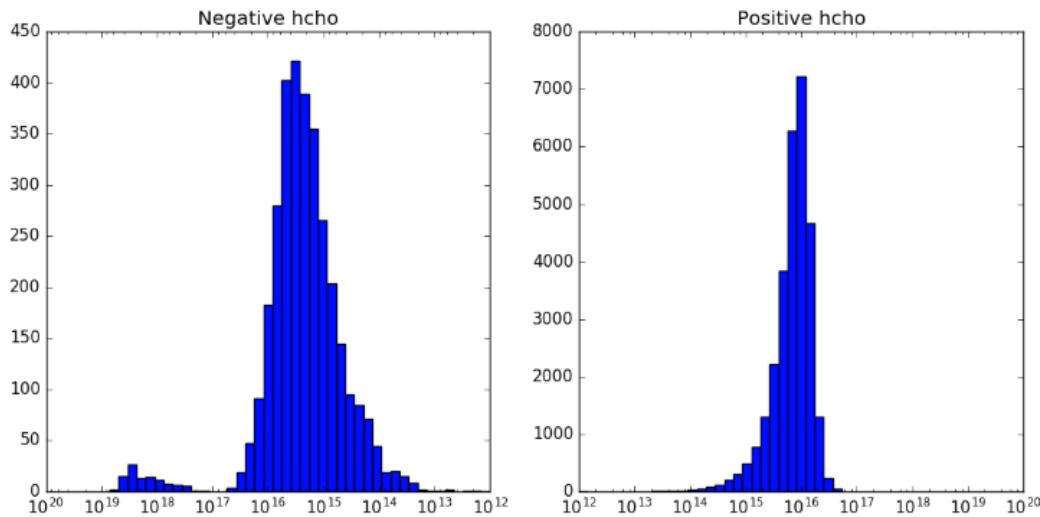


FIGURE 2.24: Column density histograms for a subset of OMI swaths over Australia on the 18th of March 2013. Negative entries are shown in the left panel, positive in the right, note the different scale between negative and positive panels.

### 2.6.5 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 within the PP code, which outputs  $\text{AMF}_{\text{PP}}$ .

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](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. In order to run, output is required from GEOS-Chem: the HCHO concentration profile averaged between 1300 and 1400 LT, including optical depths at

several wavelengths (TODO: list), and dust concentrations. A subset of the OMI pixel information is also required, which is coalesced from OMHCHO daily swath files into csv files (one per day), 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 2.23).

### 2.6.6 Vertical columns from AMF

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. Each satellite pixel at this stage has an associated SC along with three AMFs: the original ( $\text{AMF}_O$ ), one with recalculated shape factors ( $\text{AMF}_G$ ), and one completely recalculated using PP code ( $\text{AMF}_{PP}$ ). These are used to create vertical columns ( $\Omega$ ) through equation 2.6:  $\Omega = \text{SC} / \text{AMF}$ .

Figure TODO shows a comparison between the three  $\Omega$  over Australia for 2005. The first row shows the yearly average column amounts, as well as coloured rectangles describing regions which are averaged over time in row 3. Row two from left to right shows the relative differences between  $\Omega_P$  and  $\Omega_G$ ,  $\Omega_O$  and  $\Omega_G$ , and  $\Omega_O$  and  $\Omega_P$ . TODO: Add this plot here.

### 2.6.7 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  $15^\circ\text{S}$ - $15^\circ\text{N}$ ,  $140^\circ$ - $160^\circ\text{W}$  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 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 ( $\Omega_S$ ) and the reference slant column ( $\Omega_{S_0}$ ) divided by the AMF, plus the modelled reference sector column ( $\Omega_{V_B}$ ):

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

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).

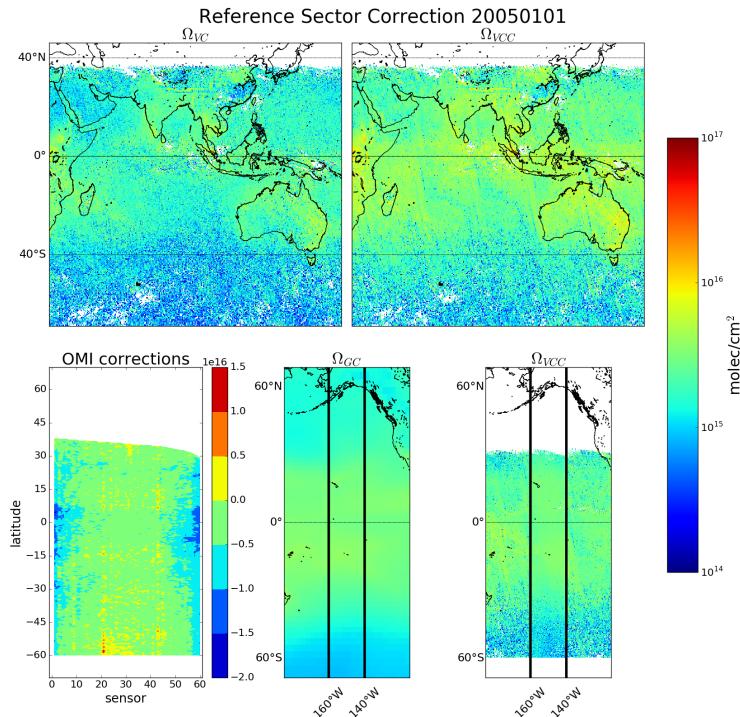


FIGURE 2.25: Example of remote pacific RSC using 8-day average measurements and one month modelled data.  $\Omega_{VC}$  shows the uncorrected vertical columns, while  $\Omega_{VCC}$  shows the corrected vertical columns. OMI corrections shows the correction applied globally based on latitude and OMI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

The reference sector correction (RSC) for OMI satellite pixels in this thesis is calculated following Abad et al. (2016). A correction is created based on the difference between the background HCHO measurements (OMI slant columns) and the GEOS-Chem modelled HCHO columns ( $\Omega_0$ ) within the reference sector ( $140^{\circ}\text{W}$  to  $160^{\circ}\text{W}$ ). To get the modelled slant columns, each of the AMFs (calculated in prior sections) is applied to the modelled vertical columns using equation 2.17. The longitudinal average is taken within the a priori reference sector, as corrections are assumed to be longitudinally invariant. The modeled reference sector is averaged over the month and 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  $2^{\circ}$  to  $0.36^{\circ}$ , 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 model does not produce slant columns associated with each measurement, however one is created by multiplying the vertical column 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

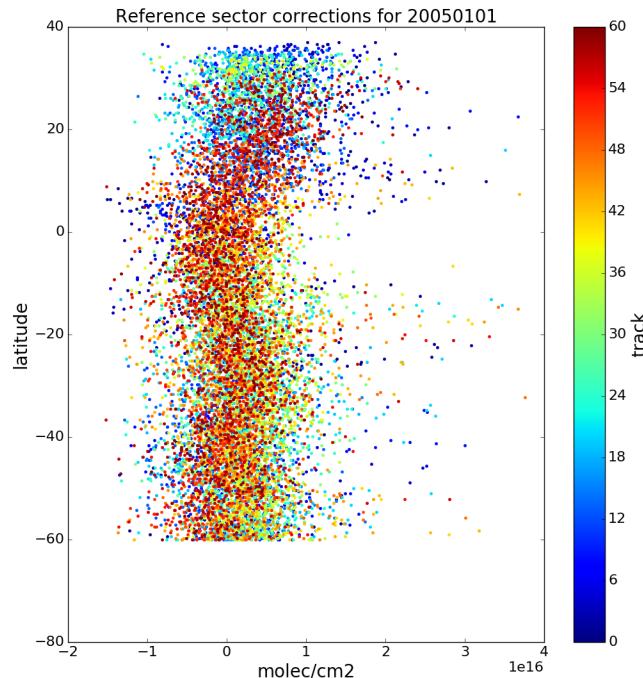


FIGURE 2.26: Example of track correction interpolations for January 1st 2005, points represent the difference between satellite slant column measurements and modelled slant columns over the remote pacific.

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:

$$\text{Correction}(i,j) = SC_{HCHO}(i,j) - VC_{GEOS-Chem}(\text{lat}(j)) \times AMF(i,j) \quad (2.21)$$

where  $j$  represents a latitude index and  $VC_{GEOS-Chem}(\text{lat})$  represents the a priori reference sector vertical column HCHO at the latitude corresponding to  $j$ . Note that the correction is in molecules  $\text{cm}^{-2}$ . 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 to 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 longitudinally averaged over the remote pacific, coloured by track number.

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 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:

$$VCC(i,j) = \frac{SC_{HCHO}(i,j) - Correction(i, lat(j))}{AMF(i,j)} \quad (2.22)$$

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.

### 2.6.8 Binning the results daily

Finally the pixels are binned into a gridded dataset I've called OMHCHORP, as shown in figure 2.23. 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 or NetCDF-4 formats, with some GEOS-Chem output being read from bitpunch. The scripts to regrid and reprocess the swath data set are available from github at TODO.

### 2.6.9 Difference between new and old OMI HCHO columns

New corrected vertical columns (VCC) of HCHO are created at global 0.25 by 0.3125°horizontal resolution. These calculations are compared over Australia in figure(s) TODO: Map comparison, regression, and time series. The effect of not recalculating the  $\omega_z$  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 recalculations with and without updating the  $\omega_z$ .

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  $\omega_z$ .

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.

## 2.7 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 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

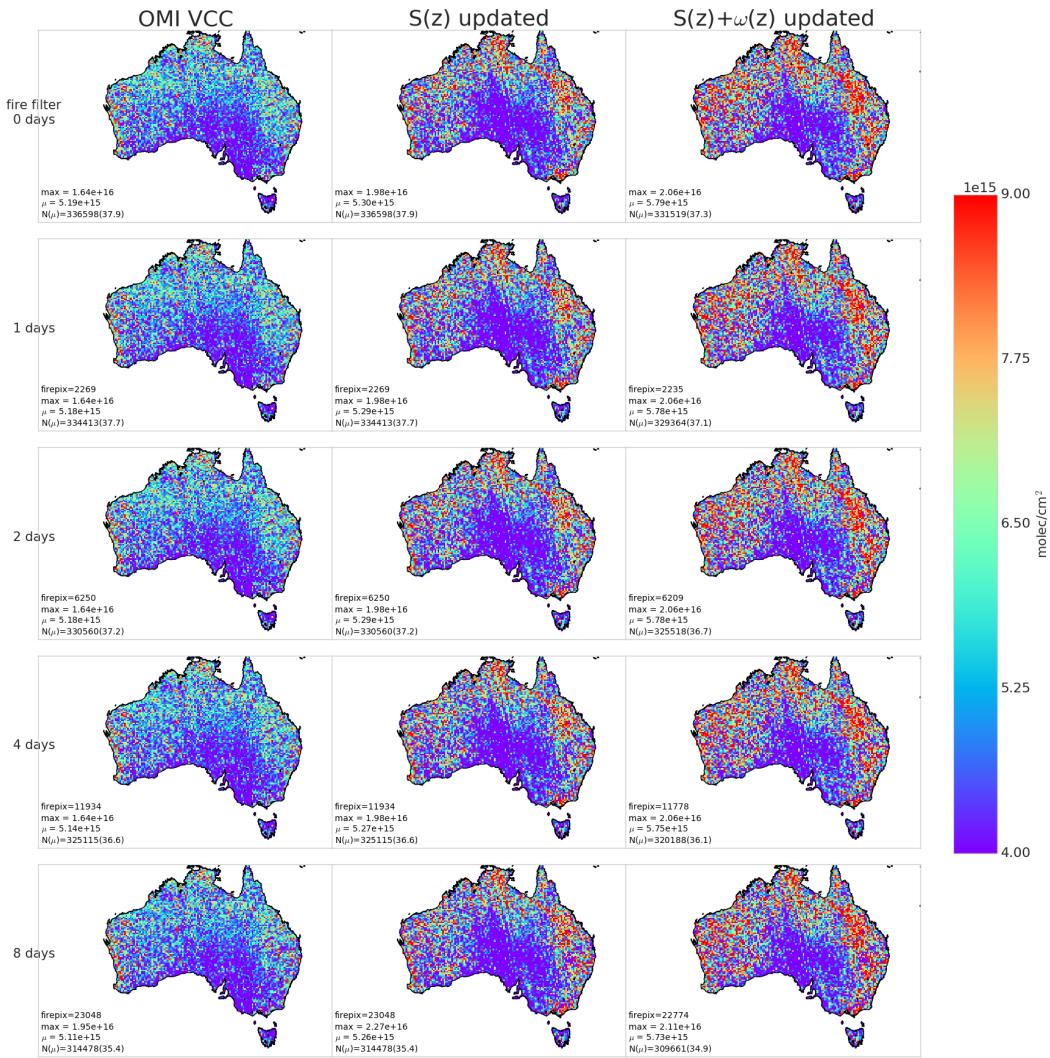


FIGURE 2.27: Column 1: Reference sector corrected HCHO vertical columns  $\Omega$  from OMHCHOv003. Column 2:  $\Omega$  with recalculated a priori shape factors using GEOS-Chem v10.01. Column 3:  $\Omega$  with recalculated a priori shape factors and scattering weights using GEOS-Chem v10.01 and LIDORT. Row 1-5: increasing number of prior days which have active fires are included when masking fire influence.

to 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 south-eastern 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.

### 2.7.1 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 threshold for smoke detection (Marais et al. 2012).

When analysing satellite OMHCHO vertical columns ( $\Omega$ ), 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  $\Omega$ .
2. A rolling mask is formed which removes  $\Omega$  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 2.3.1.3 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 threshold 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 demarcated by a horizontal line in the density plots in column 3.

### 2.7.1.1 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.

One cause of high HCHO at lower temperatures is direct or transported emissions and subsequent products from biomass burning. We test the fire mask by examining the relationship between modelled temperature and satellite HCHO with and without applying the filters for smoke and active fires. Figures TODO-TODO show the exponential fits for one month of datapoints (January 2005) in Northern, Southeastern, and Southwestern Australia respectively. Each grid square ( $2 \times 2.5^\circ$ ) provides one datapoint per day, with satellite HCHO initially averaged onto the lower resolution of the GEOS-Chem modelled surface (from 0 to  $\sim 100$  m) temperature. The scatter between HCHO and temperature is coloured by fire counts, and we see TODO. The TODO lines show the exponential fit before and after filtering fire and smoke. TODO plot showing how fire mask affects HCHO - Temperature relationship

## 2.7.2 NOx

Enhanced  $\text{NO}_2$  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  $\text{NO}_2$  columns.

OMNO2d from 2005 is used to determine a suitable threshold for anthropogenic influence by looking at  $\text{NO}_2$  columns near several major cities in the south eastern sector of Australia. The mean, standard deviation, and time series over Australia of

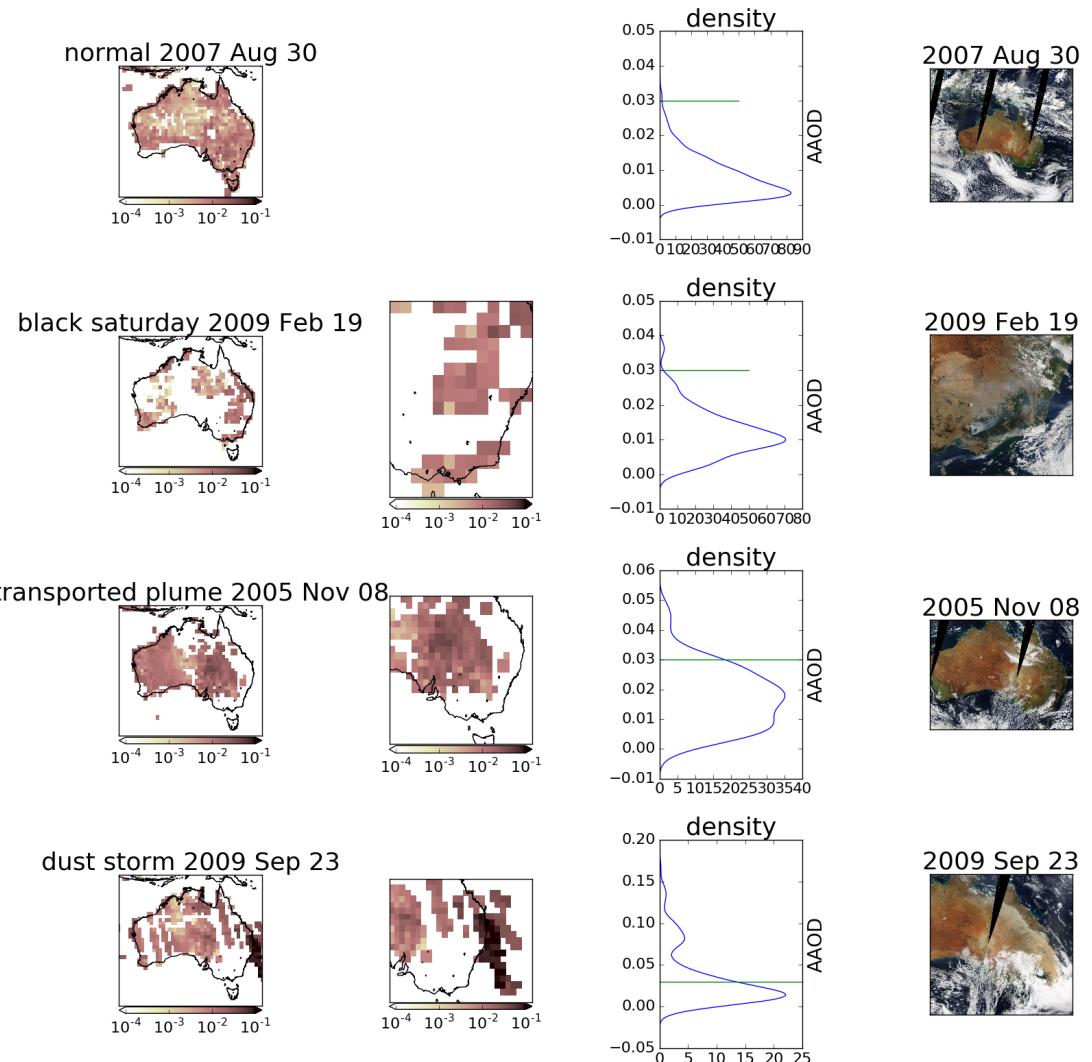


FIGURE 2.28: AAOD from OMAERUVd (columns 1, 2, 3) over Australia for four different scenarios (rows 1-4). Scenes from the same day are taken from the EOS Worldview website <https://worldview.earthdata.nasa.gov/>.

Temperature 20050101-20050228

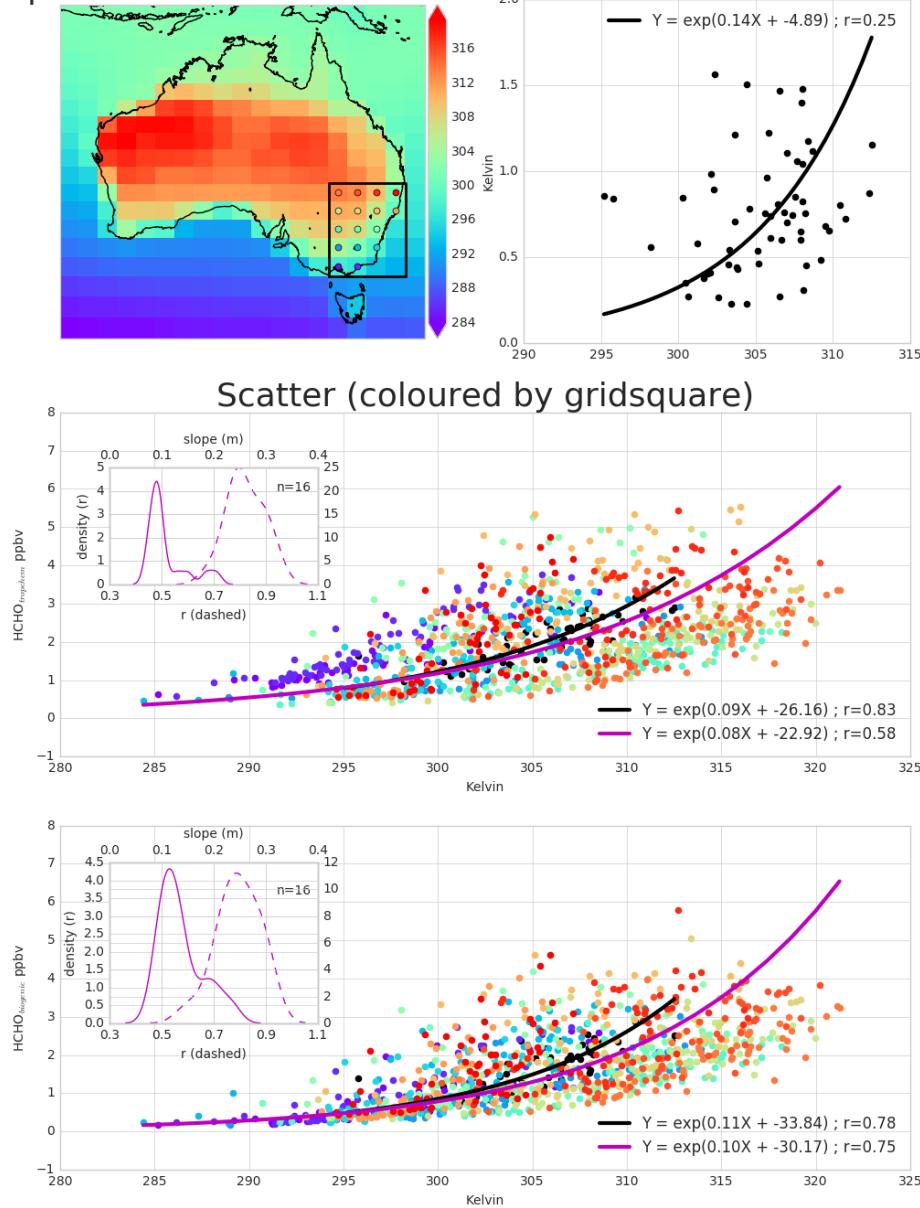


FIGURE 2.29: Top panel: surface temperature averaged over January and February 2005. Bottom panel: surface temperature correlated against temperature over, with different colours for each gridbox, and the combined correlation. A reduced major axis regression is used within each gridbox (shown in top 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.

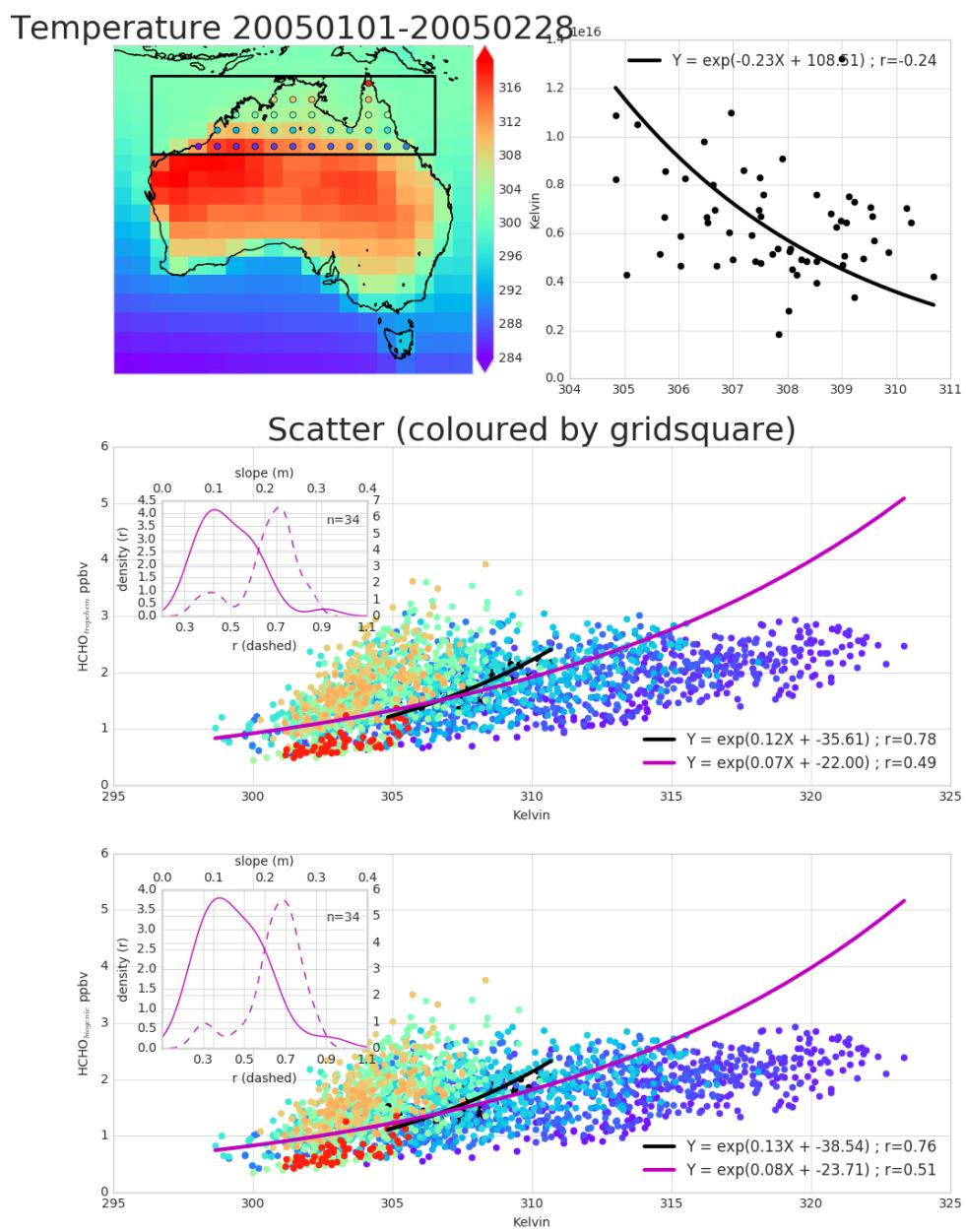


FIGURE 2.30: As figure 2.29 but for northern Australia.

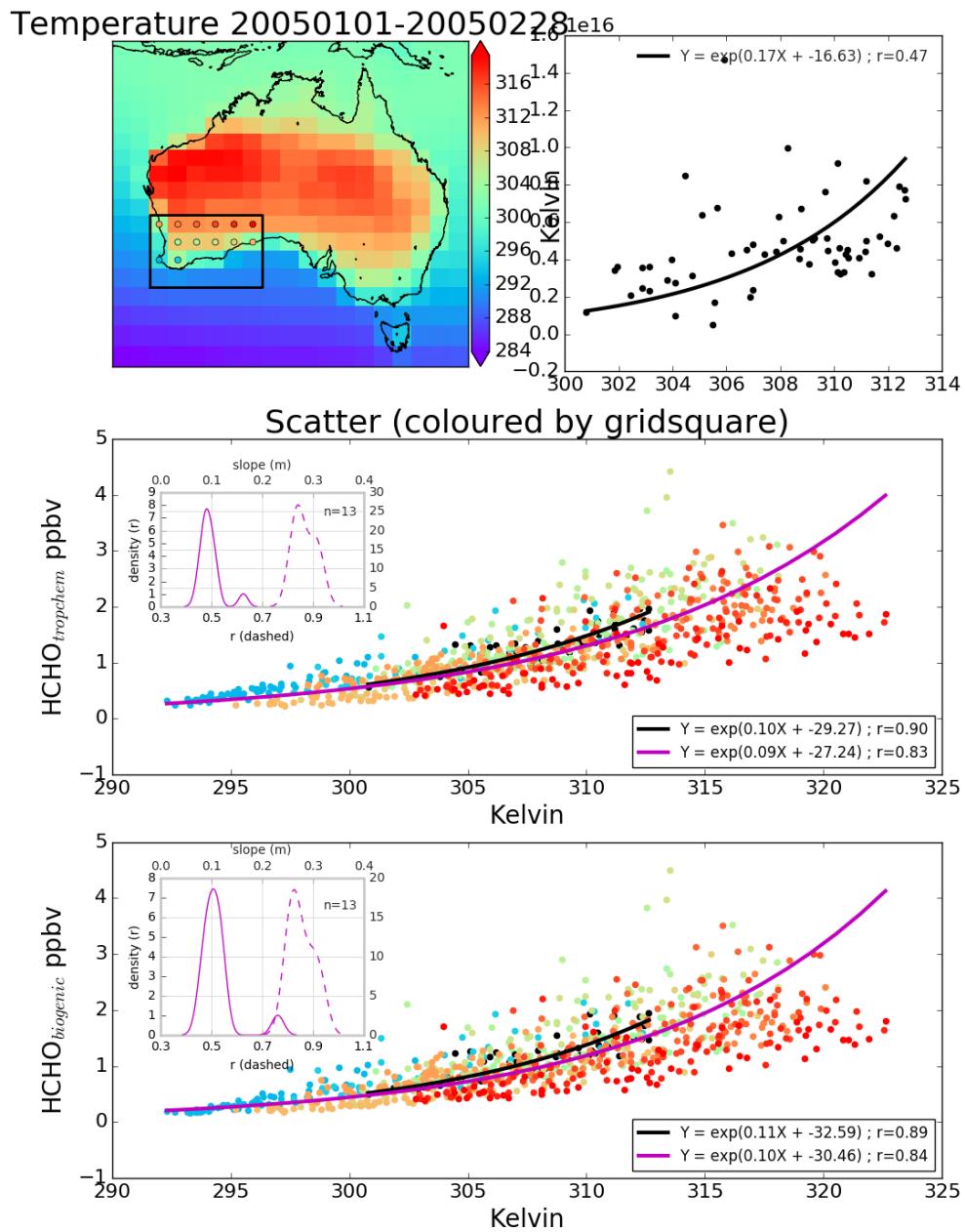


FIGURE 2.31: As figure 2.29 but for south-western Australia.

TABLE 2.2: NO<sub>2</sub> averages by region before and after filtering for anthropogenic emissions using 2005 data from the OMNO2d product.

Region	NO <sub>2</sub>	NO <sub>2</sub> after filtering	% Data lost
Aus	1	2	3
BG	1	2	3
Syd	1	2	3
Melb	1	2	3
Adel	1	2	3

tropospheric NO<sub>2</sub> seen by Aura is shown in figure 2.32. The average tropospheric NO<sub>2</sub> 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 NO<sub>2</sub> 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 NO<sub>2</sub> greater than  $10^{15}$  molec cm<sup>-2</sup> are flagged as anthropogenic.
2. After taking the yearly average over Australia, any gridsquares greater than  $1.5 \times 10^{15}$  molec cm<sup>-2</sup> 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 NO<sub>2</sub> pixels densities for each region shown, along with the threshold of  $1 \times 10^{15}$  molec cm<sup>-2</sup>. 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  $1.5 \times 10^{15}$  molec cm<sup>-2</sup> further reduced the available data by TODO gridsquares.

### 2.7.3 Summary of filters effects on HCHO

## 2.8 Data Access

TODO: ADD MORE HERE

**OMNO2d** Daily satellite NO<sub>2</sub> product downloaded from <https://search.earthdata.nasa.gov/search>, DOI:10.5067/Aura/OMI/DATA3007. See more information in section

**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

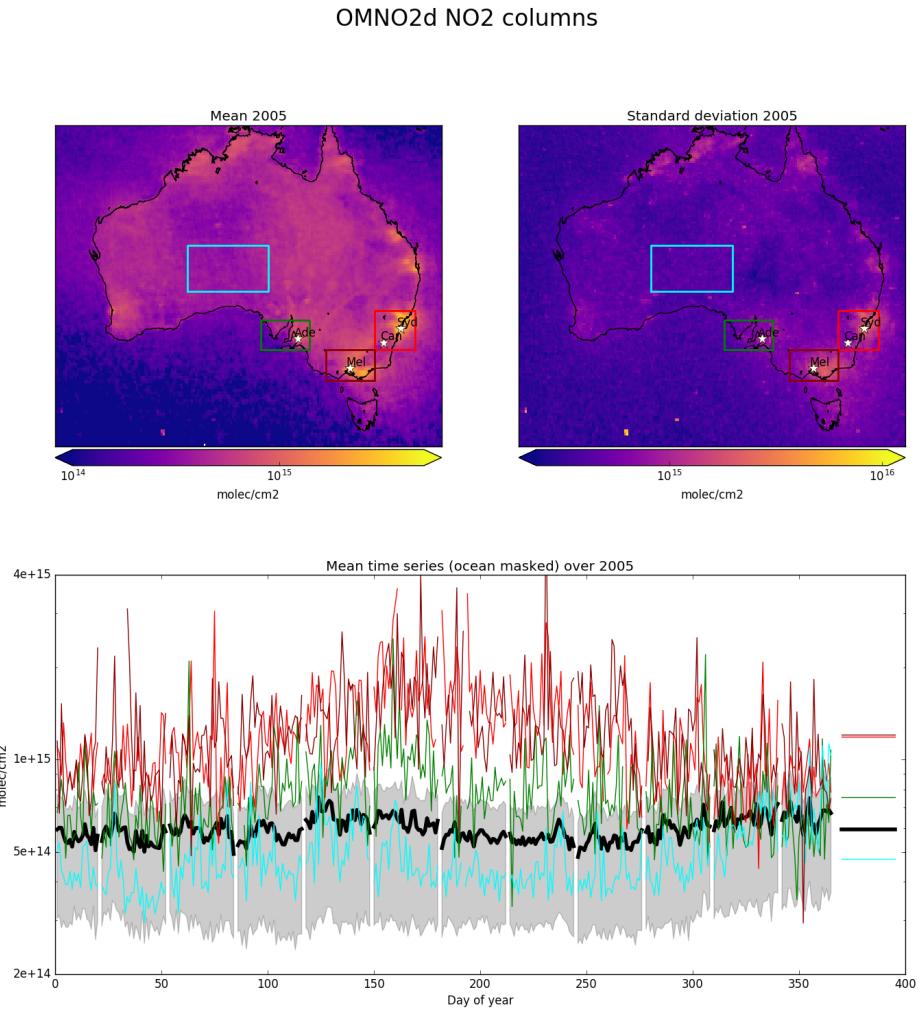


FIGURE 2.32: Mean (top left) and standard deviation (top right) of OMNO2d daily  $0.25 \times 0.25^\circ$  tropospheric cloud filtered NO<sub>2</sub> columns. Time series for Australia, and each region (by colour) shown in the bottom panel, with mean for that region shown on the right. A grey shaded area depicts the 25th to 75th percentiles of Australia averaged NO<sub>2</sub> columns for each day in the time series, with a thicker black line showing the Australia-wide mean value.

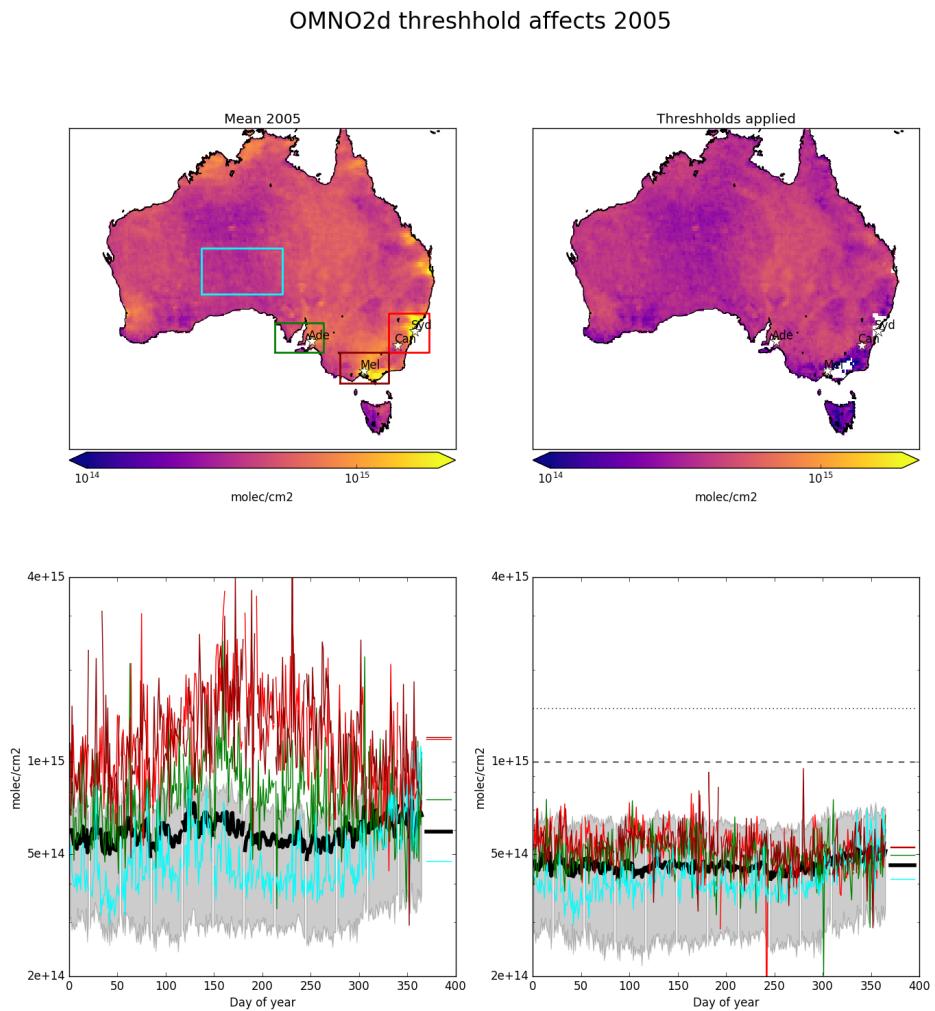


FIGURE 2.33: 2005 OMNO2d NO<sub>2</sub> column mean before (left) and after (right) applying the threshhold filters as described in the text. Time series for Australia, and each region (by colour) shown in the bottom panel, with mean for that region shown on the right.

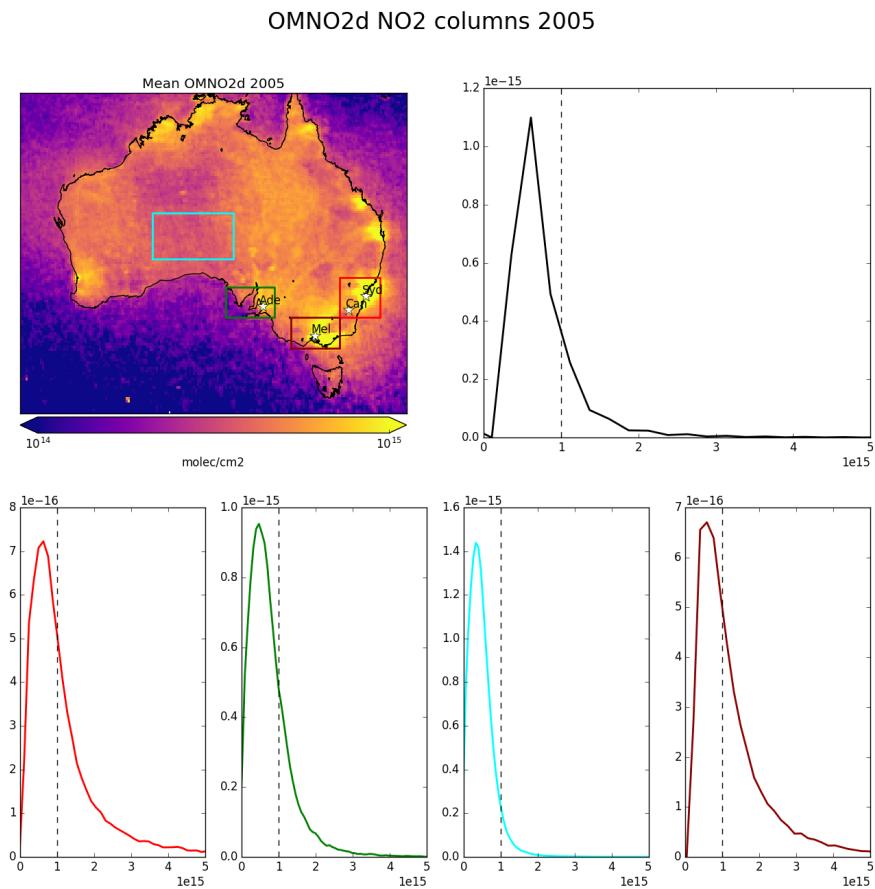


FIGURE 2.34: 2005 OMNO2d NO<sub>2</sub> column means (top left), along with column amount distributions for Australia (top right) and each region shown in the area map (by colour)

## Chapter 3

# Biogenic Isoprene emissions in Australia

### 3.1 Introduction

Isoprene has a large impact on the oxidative properties of the atmosphere, as it reacts quickly with the OH radical to form RO<sub>2</sub>. These react quickly with NO<sub>x</sub> to form OVOCs (such as HCHO), SOAs, and ozone. Australian isoprene emissions are poorly understood due to poor measurement coverage, and poor emission factor characterisation. The emissions of isoprene have been modelled at around 500 Tg C yr<sup>-1</sup> in Guenther et al. (1995) and Guenther et al. (2006) using MEGAN, and more recently around 465 Tg C yr<sup>-1</sup> in Messina et al. (2016) using ORCHIDEE. The global emission factors model used to derive these estimates is based on modelling emissions from different plant species (phenotypes), and very few are used to set the emission factors of Australian forests.

#### 3.1.1 Aims

This work aims to improve isoprene emission estimates, which are very uncertain in Australia. Here we introduce how uncertain isoprene emissions are over Australia, and discuss literature which shows how the estimates may be too high. We also describe how emissions estimates may be calculated using satellite datasets. Section ?? describes the model, satellite, and campaign data we use to determine and analyse isoprene emissions. Section 3.2 lays out how the isoprene emissions are estimated, and the results are examined in Section 3.3

We estimate isoprene emissions in Australia using top-down estimates based on OMI HCHO measurements and modelled isoprene to HCHO yields. The OMI HCHO total columns are recalculated using an updated estimate of HCHO profiles from GEOS-Chem v10.01 (see chapter 2). These estimates are compared to several campaigns (SPS1, SPS2, MUMBA, Daintree) Sensitivity to soil moisture, (maybe) LAI, and satellite AMF calculation is examined and quantified for some scenarios. The effect of using these new top down isoprene emissions as the boundary conditions for GEOS-Chem is studied Wellness of fit between in-situ (at Wollongong) HCHO, satellite (OMI), and modelled (GEOS-Chem) HCHO is determined with and without updated emissions estimates.

### 3.1.2 satellite inversions

Top-down estimates look at how much of a chemical is in the atmosphere and try to work out how much of its major precursors were emitted. This generally takes advantage of longer lived products which may reach a measurable equilibrium in the atmosphere. For isoprene this is done by looking at atmospheric HCHO enhancement, which can be largely attributed to isoprene emissions once transport and other factors are accounted for. Since 1997, when GOME first measured HCHO over Asia (Thomas et al. 1998), satellites have been able to provide a total column measurement of HCHO, one of the primary products of isoprene (Palmer et al. 2001; Bauwens et al. 2016).

Satellites recording reflected solar spectra use DOAS to measure various trace gases in the atmosphere, including formaldehyde. While satellite measurements can only be used during daytime hours, HCHO lifetimes are sufficiently short that any night-time chemistry will not affect midday observations (Wolfe et al. 2016). Satellites can be used to measure the seasonal and inter annual variability of HCHO over the globe. These records can be compared with modeled estimates of HCHO and used as a proxy to estimate isoprene emissions, as first shown by Palmer et al. (2001). This has been done in North America (Palmer 2003; Millet et al. 2006), South America, Africa, China, Europe (Dufour et al. 2008), India (Surl, Palmer, and González Abad 2018), and globally (Fortems-Cheiney et al. 2012; Bauwens et al. 2016). Often these works require both remote and in-situ measurements combined for validation (Marais et al. 2014). There is less information available from satellite measurements at higher latitudes due to increased errors (De Smedt et al. 2015).

In the remote troposphere HCHO production is dominated by methane oxidation, while in the continental boundary layer (CBL) production is largely due to NMVOCs (Abbot 2003; Kefauver, Filella, and Peñuelas 2014). This suggests that HCHO enhancement over continents can be used to determine NMVOC emissions. In the CBL, HCHO enhancement is generally driven by short lived (< 1 hr) precursors (most importantly isoprene). HCHO itself has a lifetime of a few hours (Kefauver, Filella, and Peñuelas 2014). Isoprene is emitted and enters the atmosphere in the gas phase, where it begins a complex series of reactions. Formaldehyde is produced with high yields in many of the isoprene reactions, which are discussed in more detail in Section 1.3.3. HCHO measurements are often used as a check on how well isoprene reactions are simulated, as model output can then be compared against them (Marvin et al. 2017).

### 3.1.3 MEGAN emission model

One of the most popular emissions inventories for biogenic isoprene, the Model of Emissions of Gases and Aerosols from Nature (MEGAN). Global atmospheric studies often use MEGAN along with a chemical transport model (CTM) to examine transport, deposition, and various chemical processes in the atmosphere. Emissions of Bio-generative Volatile Organic Compounds (BVOCs) including isoprene are often the subject of studies as they are still relatively uncertain, as well as being drivers for important oxidation and pollution events.

(MEGAN) is poorly calibrated for Australian conditions, their emissions of isoprene ( $C_5H_8$ ) may be overestimated, especially in the southeast. Müller et al. (2008) compared MEGAN against emissions calculated using top down estimates from the

GOME2 satellite measurements of formaldehyde. Stavrakou et al. (2015) showed that this overestimate may be a factor of 2-3 in January. Sindelarova et al. (2014) show how 50% of the isoprene emissions could be reduced by accounting for lower soil moisture. Emmerson et al. (2016) discuss the suitability of MEGAN's isoprene and monoterpene emission factors over southeast Australia, and suggest isoprene emissions are estimated 2-6 times too high. They also show that no blanket increase or decrease in emission factors is appropriate for the entire southeast of Australia.

### 3.1.4 Top-down emissions estimates

Using HCHO to determine emissions of isoprene was initially performed by Palmer et al. (2001) and Palmer (2003), who used in-situ summertime HCHO measurements over North America as model validation. Isoprene emissions fluxes were derived using the Global Ozone Monitoring Experiment (GOME) satellite instrument. Palmer's method improved biogenic isoprene emissions estimates (compared with in-situ measurements) over two available inventories: the U.S. EPA Biogenic Emissions Inventory System (BEIS2) and the Global Emissions Inventory Activity (GEIA). This showed an inversion technique which could be used to improve large scale emissions estimates without further expensive measurement campaigns.

There are now a few methods of estimating isoprene emissions using satellite measurements of emission products, here I describe them and briefly compare the pros and cons of each.

#### 3.1.4.1 Linear

This technique is the simplest, and is performed in this thesis. With the vertical columns of biogenic HCHO we can infer the local (grid space) isoprene emissions using effective molar formaldehyde yield (In other continents around 2-3, or 1 in low NO<sub>x</sub> conditions) (Palmer 2003; Marais et al. 2012; Bauwens et al. 2016). If we assume there is fast HCHO yield, so that the effect of chemical transport is minimal, and that HCHO and isoprene are at steady states, then we can calculate local yield from our CTM. Yield is calculated from the modelled slope between isoprene emissions and HCHO total column within each gridbox over Australia, as performed in Palmer (2003), using modelled values between 1200-1400 LT which is around the overpass time of the OMI. This modelled yield is then used in conjunction with the recalculated OMI measurements in order to estimate isoprene emissions. To calculate emissions we use a reduced major axis (RMA) regression between modelled average values of the loss rates and total columns, an example is shown in figure TODO: figure with RMA of these over whatever time and space I end up using.

#### 3.1.4.2 Bayesian

Satellite based emissions estimates may allow us to improve the models without requiring lots of hard work on calibrating MEGAN to the large data sparse continent of Australia. Emissions of monoterpenes (C<sub>10</sub>H<sub>16</sub>, two units of isoprene) may also be underestimated in southeastern Australia, which could lead to the unique scenario of

neither type of emission dominating VOC chemistry over the forests (Emmerson et al. 2016).

Another method of correcting isoprene emissions using observed HCHO total column involves a Bayesian inversion. Shim et al. (2005) work with GOME HCHO observations and GEOS-Chem, looking at areas with high signal to noise ratio (higher HCHO concentrations). They show that the model underestimates isoprene emissions and HCHO concentrations by 14-46%, with the corrected VOC emissions reducing the model biases to 3-25%.

The Bayesian inversion is also used in Curci et al. (2010), where a regional CTM (CHIMERE) simulates HCHO, which is compared against OMI observed HCHO and shown to be regionally biased. This bias is expected to be caused by errors in MEGANs isoprene emissions estimations. The CHIMERE model is used to derive yields of HCHO from the various local VOCs and these are then used in estimating local emissions. The model is run initially with emissions of BVOCs and reactive anthropogenic VOCs (RAVOCs) turned off in order to work out the background ( $b$ ) values of these compounds. The Bayesian inversion is used to correct regionally biased biogenic isoprene emissions by optimising these parameters in order to simulate HCHO closest to the observed HCHO levels. Curci et al. (2010) uses CHIMERE as the forward model to determine the relationship between HCHO ( $y$ ), isoprene and reactive anthropogenic VOCs ( $x$ ), using

$$y = \mathbf{K}x + b + \epsilon \quad (3.1)$$

where  $\epsilon$  are the (assumed) independent errors in measurements.  $K$  is the Jacobian matrix determined from CHIMERE representing the sensitivity of  $y$  to the state variable  $x$ . This  $K$  matrix is used in conjunction with error covariance in  $x$  to determine the Maximum A Posteriori (MAP) solution to calculate the optimal estimate of  $x$  ( $\hat{x}$ ).

## 3.2 Methods

### 3.2.1 Outline

Here is an overview of the steps involved in this thesis, which take satellite data and model output to estimate isoprene emissions. This process is summarised in figure 3.1.

1. Level three satellite data are used to make anthropogenic, fire, and smoke influence masks (see section 2.7)
2. Level two OMI HCHO satellite data (see section 2.3.1.1), along with GEOS-Chem model runs (see section 2.4.8) are used to recalculate corrected vertical columns (VCCs; saved in OMHCHORP dataset) at  $0.25^\circ$ by  $0.3125^\circ$ horizontal resolution (see section 2.6).
3. GEOS-Chem modelled biogenic emissions of isoprene along with total biogenic columns of HCHO, both averaged over  $2 \times 2.5^\circ$ horizontally and 1300-1400 LT temporally, are used to calculate a linear regression slope.
  - (a) Globally gridded biogenic emissions of isoprene output from GEOS-Chem are read in as  $E_{GC}$  atoms C  $\text{cm}^{-2} \text{ s}^{-1}$ .

- (b) Satellite overpass output of total HCHO columns from GEOS-Chem are read in as  $\Omega_{GC}$  molec cm $^{-2}$
  - (c) A reduced major axis regression slope is found between  $\Omega_{GC}$  and  $E_{GC}$  using a month of modelled output (one value per day) for each gridsquare (eg. see figure 3.2)
4. Fire and anthropogenic masks are applied to filter out satellite HCHO columns which may be influenced by non biogenic sources. For a full outline of these filters see section 2.7.
- (a) The fire mask is created daily from non-zero (MODIS) fire counts over the prior 2 days which occur in local or adjacent grid squares at 0.25°latitude by 0.3125°longitude
  - (b) A second filter to remove influence from transported smoke plumes is created where OMI aerosol absorption optical depth (AAOD, from OMAERUVd) greater than 0.03
  - (c) A filter for anthropogenic influence is created daily using OMNO2d NO<sub>2</sub> tropospheric column amounts; masking any grid squares with greater than  $1 \times 10^{15}$  molec cm $^{-2}$  on any particular day, along with grid squares where the yearly average is above  $1.5 \times 10^{15}$  molec cm $^{-2}$
5. For each of the corrected vertical columns ( $VCC_{OMI}$ ,  $VCC_{GC}$ , and  $VCC_{PP}$ ; see section 2.6), a top down estimate of biogenic isoprene emissions ( $E_{new}$ ; atoms C cm $^{-2}$  s $^{-1}$ ) is calculated
- (a) Background total columns ( $VCC_{BG}$ ) are calculated by averaging longitudinally the VCCs between 180 and 120°W
  - (b) The masks (fire, anthropogenic, smoke) created earlier are applied at this point to VCCs, in order to remove non biogenic influences
  - (c) Emissions are calculated using
- $$E_{new} = \frac{VCC - VCC_{BG}}{\text{slope}} \quad (3.2)$$
6. TODO: Biogenic yield masked when smearing term is above a threshhold (TODO: determine). This smearing value is determined as  $S = \Delta\Omega_{HCHO}/\Delta E_{isop}$ , the change in each gridsquare of the ratio between modelled HCHO columns and isoprene emissions. A full description and analysis of this term used to determine the smearing filter is given in section 3.2.7.
7. Compare against MEGAN, MUMBA, SPS, Daintree, Wollongong, and one set of airplane measurements (HIPPO).

### 3.2.2 Calculating modelled slope

Figure 3.2 (top left) shows the RMA regression slope between modelled HCHO columns and isoprene emissions calculated within each grid square over January 2005, using

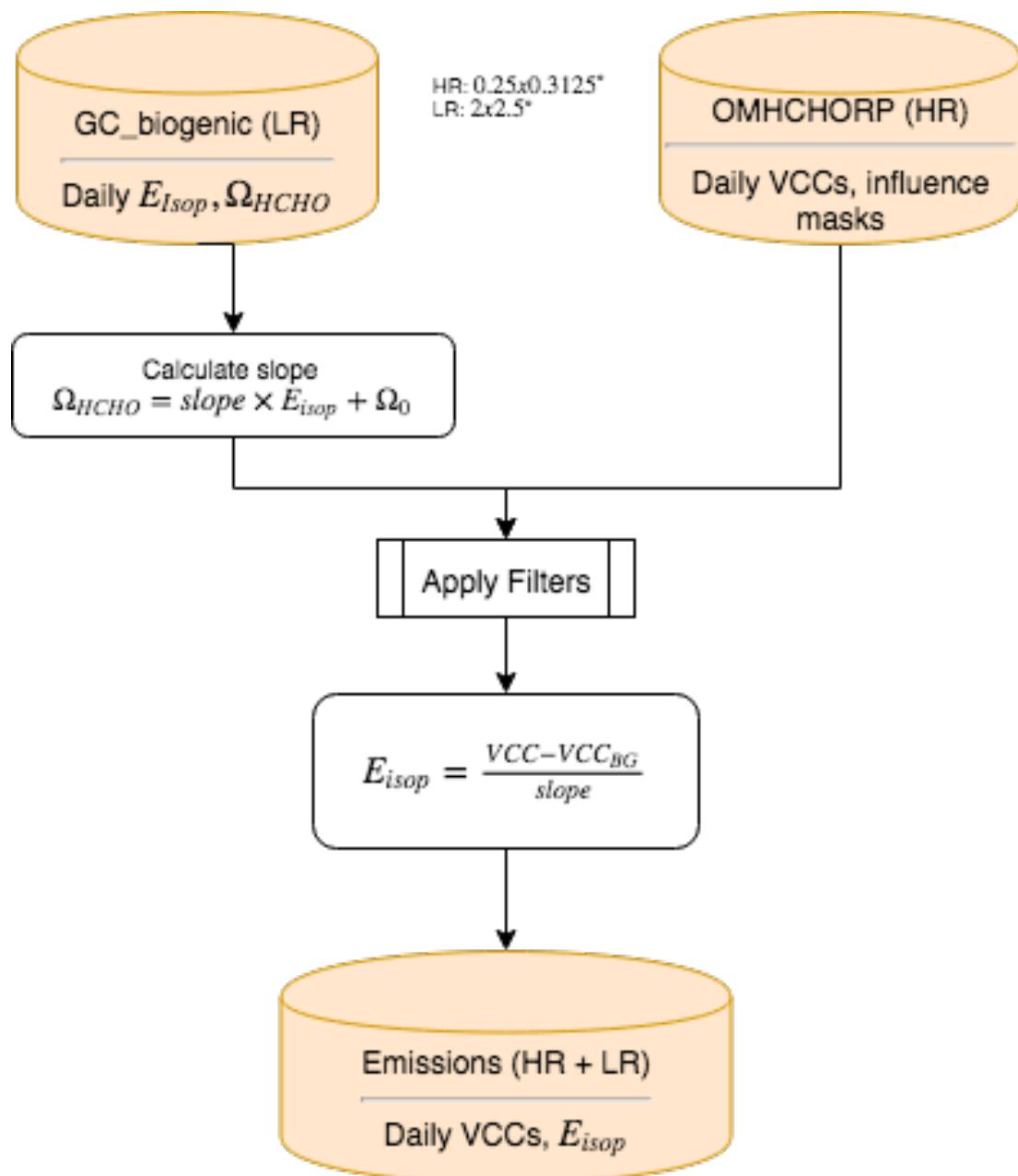


FIGURE 3.1: Creation of isoprene emissions dataset using OMH-CHORP and biogenic GEOS-Chem outputs.

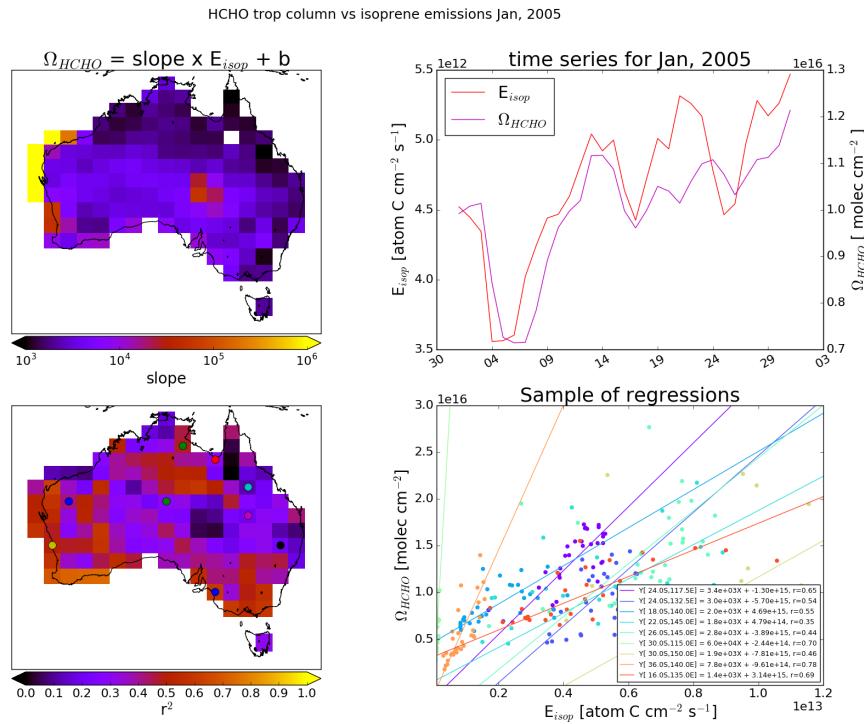


FIGURE 3.2: Top left panel: the RMA regression slope between modelled HCHO columns and isoprene emissions calculated within each grid box over January 2005. Bottom left panel: the square of the regression coefficients. Top right panel: Australia averaged time series Bottom right panel: a sample of the regressions coloured by latitude and longitude pairs, matching dots in the bottom left panel.

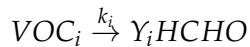
one datapoint for each day. Also shown are the regression coefficients (bottom left), Australia averaged time series (top right), and a sample of the regressions (bottom right). Each regression is for a single latitude and longitude, over the course of one month, and is coloured by location (matching dots shown in the bottom left panel).

### 3.2.3 Satellite inversion

TODO: Read through this list of sources on the hcho to isop process : taken from Wolfe2015 Such techniques have informed isoprene emission inventories in North America (Abbot et al., 2003; Millet et al., 2008 (Palmer 2003; Millet et al. 2006; Palmer et al. 2006)), South America ((Barkley et al. 2013), 2008), Europe (Curci et al. 2010; Du-four et al. 2008), Africa (Marais et al. 2012), Asia (Fu et al., 2007; Stavrakou et al., 2014), and globally (Fortems-Cheiney et al., 2012; (Shim et al. 2005); Stavrakou et al., 2009).

Initially studies assumed a simple linear steady-state relationship between HCHO and its precursors (Palmer 2003; Palmer et al. 2006; Millet et al. 2006). This allowed a simple calculation of isoprene using the measured HCHO, with estimated reaction rates and yields. The methodology for calculating VOCs from HCHO is laid out in Palmer (2003), and takes into account the expected lifetime and reaction rates of the

precursor VOCs and HCHO. Assuming HCHO is produced quickly from short-lived intermediates, and the column is at steady state:



Where  $Y_i$  is HCHO yield per C atom (a measure of how much HCHO will form per gram of C from a VOC within a system), and  $k_i$  is the reaction rate. Then assuming a steady state of atmospheric HCHO ( $\Omega$  molecules  $\text{cm}^{-2}$ ) produced by oxidation of VOCs ( $\text{VOC}_i$ ) and no horizontal transport:

$$\Omega = \frac{1}{k_{\text{HCHO}}} \sum_i Y_i E_i$$

Where  $i$  indexes a chemical species,  $k_{\text{HCHO}}$  is the HCHO loss rate due to OH and photolysis,  $Y_i$  is the molar HCHO yield from oxidation of  $i$ , and  $E_i$  is emission fluxes (C atoms  $\text{cm}^{-2}\text{s}^{-1}$ ).

Estimates of  $Y_i$  can be attained from a model as shown in Millet et al. (2006). This involves a reduced major axis (RMA) correlation calculation between modelled HCHO and isoprene columns, multiplied by their loss rates (to photolysis and oxidation) (as a normalising factor). In high NO<sub>x</sub> environments where HCHO has a lifetime on the order of 30 minutes, it can be used to map isoprene emissions with spatial resolution from 10-100 kms. Horizontal transport 'smears' the HCHO signal so that source location would need to be calculated using windspeeds and loss rates (Palmer et al. 2001; Palmer 2003). Smearing is explicitly handled in these studies due to the importance of transport and NO<sub>x</sub> on forming robust and accurate estimates. Over Australia NO<sub>x</sub> levels are generally not high enough to ensure quick HCHO formation and we must take extra care that we can account for the transport or 'smearing' caused by slower HCHO formation, details on this process can be found in Section 3.2.7.

More recently, full inversions that better account for transport, source attribution, and chemical schemes have been implemented (Fortems-Cheiney et al. 2012). TODO: full description of this better inversion technique going through FortemsCheiney2012.

Kefauver, Filella, and Peñuelas (2014) reviews remote sensing of BVOCs, which are on the rise, examining the last 20 years of data and analysis of the satellite products. Their review encompasses the latest reports up to 2014. The modelled isoprene and BVOC emissions from MEGAN (Guenther et al. 2000) of 500 and 1150 Tga<sup>-1</sup> respectively are still the global go to estimates. Their review reinforces the message that NMVOCs affect the oxidative capacity of the atmosphere and are largely driven by and sensitive to vegetation. The tropospheric affects from NMVOCs on the hydroxyl radical (OH), ozone (O<sub>3</sub>), SOAs, and methane longevity, all interconnect to form a very complex system which still suffers from relatively large uncertainties in both measurement and chemistry mechanisms. One focus of Kefauver, Filella, and Peñuelas (2014) is HCHO, which is the dominant product of most BVOCs which is measurable by remote sensing. The main datasets of HCHO are from four satellite instruments: GOME on ERS-2, SCIAMACHY on ENVI-SAT, OMI on EOS AURA, and GOME2 on MetOp-A. These satellites have slightly different spectral and spatial resolutions, as well as using varied processes to estimate HCHO from detected radiances. This can lead to different estimates between instruments or methodologies as

described in [Lorent2017](#) which means validation and comparison is more important when using these remotely sensed data.

Total HCHO is measured by satellite over the entire world, however the technique is not perfect and suffers from uncertainties and interferences. Satellite based chemical concentrations rely on ground-based measurements and modelled data for validation. They provide various readings with daily global coverage which is not otherwise feasible.

Validation is important due to the various uncertainties in the satellite remote sensing process, with apriori assumptions having the greatest effect on structural uncertainty between measurements techniques [Lorente et al. \(2017\)](#). [Zhu et al. \(2016\)](#) use SEAC<sup>4</sup>RS aircraft HCHO measurements over the southeastern US as model validation, and show a bias in the assumed OMI shape factor that leads to a bias between satellite and SEAC<sup>4</sup>RS measurements. [Marais et al. \(2014\)](#) compare OMI based isoprene emission estimates against relaxed eddy accumulation measurements from African field campaigns, as well as MEGAN and GEOS-5 inventories. [Dufour et al. \(2008\)](#) use HCHO from SCIAMACHY, and examine Europe using CHIMERE as the chemical model. In their work they show that satellite measurements can reduce source emission uncertainty by a factor of two, where emissions are relatively large.

### 3.2.4 Calculation of Emissions

As is done in [Palmer \(2003\)](#), [Millet et al. \(2006\)](#), and [Bauwens et al. \(2016\)](#), we assume that HCHO, and Isoprene columns are in a steady state, with no horizontal transport. We also assume that isoprene is the only compound enhancing the HCHO levels, which requires that we filter out influence from fires. Emissions of precursors are easy to calculate as long as we know the molar HCHO yields ( $Y_i$ ) and effective chemical loss rates ( $k_i$ ):

$$\Omega_{HCHO} = \frac{1}{k_{HCHO}} \sum_i k_i Y_i \Omega_i = \frac{1}{k_{HCHO}} \sum_i Y_i E_i \quad (3.3)$$

This works if there is fast HCHO yield, so that the effect of chemical transport is minimal. The background HCHO is calculated using measurements in the remote pacific at the same time and latitude. Table 3.1 shows the average yield calculated for Australia. (TODO: this table and some notes)

In order to approximate the isoprene to HCHO yields over Australia, GEOS-Chem is run and the slope ( $S$ ) between modelled tropospheric HCHO columns and emissions of isoprene within each grid box. Figure (TODO: example grid box and regression plot.) shows the regression between emitted isoprene and tropospheric column HCHO, averaged between 1200-1300 LT each day. We can infer the local (grid space) isoprene emissions ( $E_{isop}$ ) using effective formaldehyde yield from isoprene ( $Y_{isop}$ ).

$$\Omega_{HCHO} = S \times E_{isop} + B \quad (3.4)$$

Where  $B$  is the background HCHO, and  $S = Y_{isop}/k_{HCHO}$  is determined monthly as the regression between  $\Omega_{HCHO}$  and  $E_{isop}$  on daily saved outputs from GEOS-Chem over Australia using 2 by  $2.5^\circ$  horizontal resolution. Modeled background emissions can be ignored here as they do not affect the slope calculation. Once we have calculated

this slope, we use the same formula (Eqn. 3.4) to determine the isoprene emissions. By replacing  $\Omega_{HCHO}$  and  $B$  with OMI based values,  $E_{isop}$  is the only unknown.

In several studies OMI satellite HCHO columns are scaled up by up to 40% to match in-situ measurements TODO:citations. Since we don't have enough in-situ data to reasonably scale the satellite columns we instead re-run the calculation with them scaled up by 40% and consider this as a possible bound on satellite uncertainty. Yield and estimated top-down emissions are therefore given an upper and lower bound on satellite based uncertainty through this method.

There are a couple of ways to determine the modelled background HCHO concentration, one of which involves running the model with isoprene emissions turned off, which allows us to see exactly how much the modelled isoprene emissions alter each vertical column of HCHO. This is effective since we have assumed variation in HCHO columns only depends on isoprene emissions, so our background term is theoretically identical to the emission free simulated HCHO. The other way involves looking at HCHO over the remote pacific at matching latitudes and times, which emulates how the background is determined for the measured HCHO. These modelled background HCHO concentrations are mostly used for comparison with other datasets. TODO: show how these two numbers compare! figure with time series of both backgrounds for Aus, SEAus, and remote ocean.

The Background from OMI is determined using the mean column HCHO measured over the remote pacific ocean ( $180\text{--}120^\circ\text{W}$ ). TODO: Update the background term to do as follows: For this term we average each month of remote ocean measurements, as well as averaging longitudinally within  $180\text{--}120^\circ\text{W}$ , and finally  $B$  is estimated at each latitude using  $\pm 10^\circ$ . This is gives us a background which is appropriate for any latitude, and is shown in Figure TODO: figure with background region highlighted and a time series of background values. When calculating the  $E_{isop}$  from our modeled slope with OMI HCHO and background, we end up with negative emissions whenever the OMI HCHO column is less than the OMI background (as  $E_{isop} = \frac{\Omega_{HCHO}-B}{S}$ ). These are set to zero, which increases the average by around TODO: X%. The measured background HCHO is the average concentration measured in the remote pacific at the same time.

Figure 3.3 shows the modelled isoprene emissions and column HCHO concentrations along with the RMA regression line, sampled from grid boxes over Australia for January 2005. Some affects from the low emissions in grid boxes which are largely oceanic can be seen and are handled by TODO: handle these and document here. Due to the low horizontal resolution of GEOS-Chem (2 by  $2.5^\circ$ , latitude by longitude), calculations from grid boxes on the coast which are largely oceanic need to be discarded as the change in HCHO is not dominated by emissions of isoprene, as is assumed for Eqn 3.4. A nested version of GEOS-Chem allows a much better analysis of coastal regions, at  $0.25$  by  $0.3125^\circ$  resolution.

Using this modelled slope at  $2^\circ$  by  $2.5^\circ$  and applying it to equation 3.4 with  $B$  and  $\Omega_{HCHO}$  calculated using OMI satellite measurements provides a new estimate of isoprene emissions. Figure 3.4 shows the emissions calculated this way along with the Emissions output by GEOS-Chem averaged over January, 2005.

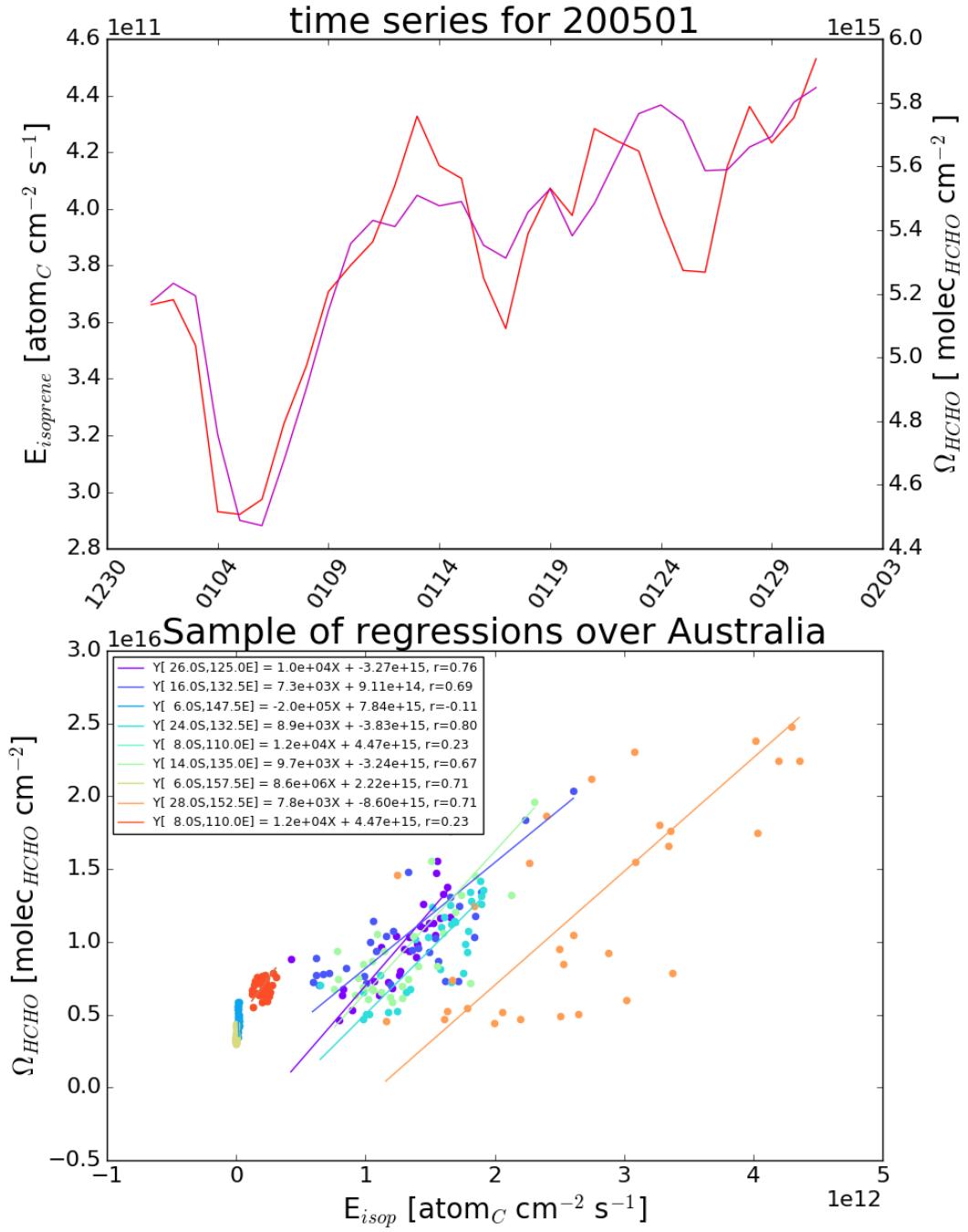


FIGURE 3.3: Top panel: isoprene emissions for January, 2005, shown in red, coplotted with tropospheric hcho columns, shown in magenta. Both series are daily averages over Australia. Bottom panel: (RMA) linear regressions from between emissions of isoprene and tropospheric hcho columns, sampled randomly from the  $2^\circ$  by  $2.5^\circ$  latitude longitude gridboxes over Australia for the month of January (2005).

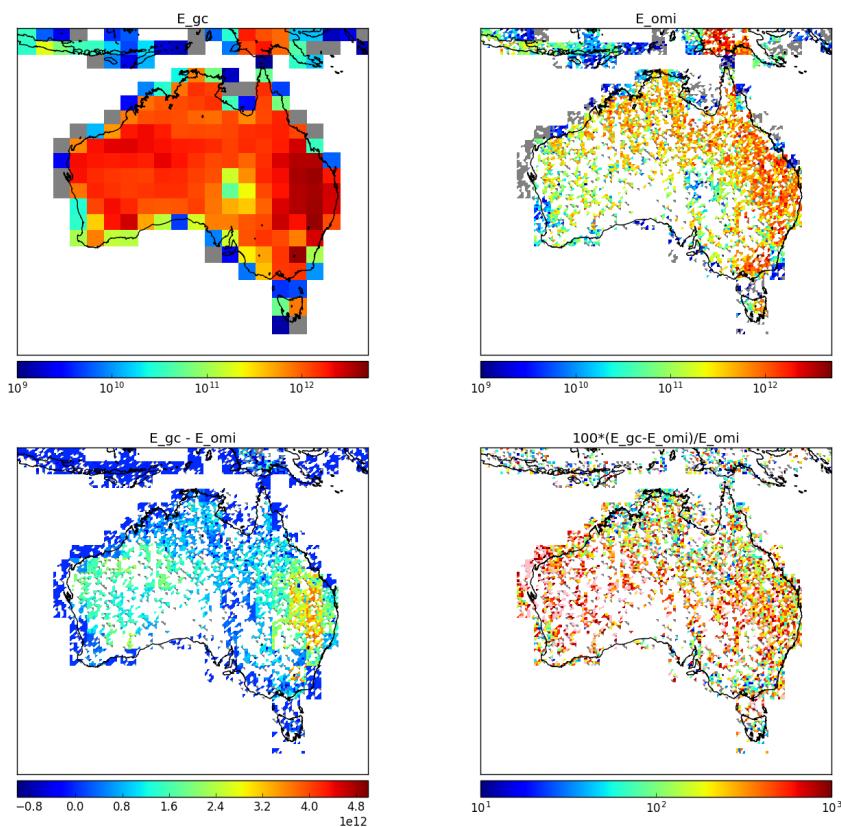


FIGURE 3.4: Top row is isoprene emissions for the month of January, in 2005, from GEOS-Chem and estimated from OMI respectively. Bottom row shows the absolute and relative differences between the two.

### 3.2.5 Emissions drivers

Calculated yields of HCHO can be classified using a box model which approximates specific environments, as described in Section ?? TODO: A table of different factors affecting emissions for three scenarios; urban, forest, shrublands is given in Table XX. The calculated yields for these scenarios is based on the CAABA/MECCA box model (described in Section ??) TODO: compare scenarios yields and show map of Australia with mapped closest scenario(one colour for each scenario, contourf).

### 3.2.6 HCHO Products and yield

Australian forests are strong emitters of both isoprene and monoterpenes, which go on to form various products including secondary organic aerosols, oxygenated VOCs (OVOCs), ozone, OH, and HO<sub>2</sub>. This production occurs over several steps, yields are often classed into at least two categories. First generation yield refers to the amount of HCHO produced per unit isoprene consumed by initial oxidation, total yield (sometimes molar yield) refers to time dependent yield of HCHO over multiple oxidation stages (Wolfe et al. 2016). Wolfe et al. (2016) define prompt yield as the change in formaldehyde measurement per unit change in initial isoprene emissions. Some argue that isoprene emissions are overestimated, due to the fact that they are based on relatively few measurements of isoprene emission factors (Winters et al. 2009; Fortems-Cheiney et al. 2012) TODO: read and cite paper mentioned in Fortems. Recently Emmerson2017 showed that MEGAN estimates 3-6 times too much isoprene emissions, and 4 times too little monoterpenes when compared against 4 (relatively small scale) measurement campaigns in southeastern Australia.

Isoprene production of HCHO depends on several factors, importantly NO<sub>x</sub> levels have a direct effect on the fate of VOCs in the atmosphere. At higher NO mixing ratios (at least a few hundred pptv), organic peroxy radicals (RO<sub>2</sub>) react mostly with NO. At low NO (less than 10's of pptv), reaction with HO<sub>2</sub>, other RO<sub>2</sub>, and isomerization dominate the fate of RO<sub>2</sub>. In low NO<sub>x</sub> environments, reported HCHO yields from isoprene are from XtoY%, while in high NO<sub>x</sub> environments this value is XtoY% TODO: these values from table. For monoterpenes the yields are around X, Y% for low, high NO<sub>x</sub> respectively. Emissions and yields for various species including some terpenes can be seen in table 3.1. Wolfe et al. (2016) determine that going from NO<sub>x</sub> = 0.1 to 2.0 ppbv triples the prompt yield of HCHO, from 0.3 to 0.9 ppbv ppbv<sup>-1</sup> due to isoprene, while the background HCHO doubles. They determine prompt yield as the change in HCHO per change in ISOP<sub>0</sub>, using  $[ISOP]_0 = [ISOP] \exp(k_1[OH]t)$ ; where  $k_1$  is first order loss rate. This effectively relates HCHO abundance with isoprene emission strength.

NO<sub>2</sub> measured by OMNO2d gives us a daily mid-day measurement which we can compare to output from GEOS-Chem to determine how well the model does at simulating NO<sub>2</sub>. This is also done in Travis et al. (2016), as a way to examine model bias in ozone (potentially due to NO<sub>2</sub> bias) over the USA.

Looking at Australian emissions from running GEOS-Chem and using yields provided by XYZ (TODO other table), we see that Australia may be more or less likely to do something TODO: this comparison sentence would be good to tie up tables and be copied to conclusions.

TABLE 3.1: HCHO yields from various species averaged over Australia during Summer.

Species	Emissions <sup>a</sup> (Tg C per month)	Lifetime <sup>b</sup>	HCHO Yield <sup>c</sup> (per C reacted)	HCHO production <sup>d</sup> %
Isoprene	Y	n minutes	0.x	10
$\alpha$ -Pinene	Y	n minutes	0.x	10
$\beta$ -Pinene	Y	n minutes	0.x	10
HCHO	Y	n minutes	1.0	10

a: Calculated using GEOS-Chem emissions over Australia in January 2005.

b:

c:

d: Production determined by dividing emission\*yield by the sum of all VOC emissions\*yields.

Conversions between HCHO per unit C yield and molar % yield from species X given by the equation  $Y_{molar\%} = 100 \times C_X \times Y_{HCHO \text{ per unit } C}$ , where  $C_X$  is how many Carbon are within species X (5 for isoprene, 10 for monoterpenes, etc...). For instance a 200% molar yield of HCHO from isoprene implies 1 Mole of  $C_5H_8$  becomes 2 Mole HCHO which is a 0.4 HCHO per unit C yield.

TODO: Fill out this table

Yields table looking at literature provided yields of HCHO.

### 3.2.7 Accounting for smearing

Accounting for transport of the precursors is important, especially in low  $NO_x$  conditions in which isoprene has a longer lifetime (days). When estimating emissions of isoprene using one of its products, it is often assumed that isoprene has a short lifetime, however when low  $NO_x$  environments (which are prevalent in the Australian outback) this assumption can be wrong. Smearing (or spatial smearing) is a measure of how much formaldehyde (the product) was created from isoprene (the precursor) emissions in a different grid box. Smearing has been measured in order to account for this uncertainty in various works (Martin et al. 2003; Palmer 2003; Millet et al. 2006; Marais et al. 2012; Barkley et al. 2013; Zhu et al. 2014; Wolfe et al. 2016; Surl, Palmer, and González Abad 2018), often implementing the method designed in Palmer (2003).

Horizontal transport complicates estimation of precursor emissions, as the smearing length scale which increases beyond our gridbox size. The smearing length scale; the distance travelled downwind ( $L_{d,i}$ ) by a precursor (i) before becoming HCHO can be estimated using:

$$L_{d,i} = \frac{U}{k_i - k_{HCHO}} \ln \left( \frac{k_i}{k_{HCHO}} \right)$$

where U is wind-speed. Palmer (2003) further define a smearing length scale:  $L_{s,i}$  as the distance downwind where a fraction ( $1 - 1/e$ ) of the precursor is completely transformed into HCHO. This equation uses the initial VOC column concentration

TABLE 3.2: HCHO yields from various species, and lifetime against oxidation by OH.

Species	HCHO Yield (molar %)	Life vs OH	NO <sub>x</sub> background	Source
Isoprene	315±50	35 min	High	a
	285±30		High	a
	225		High	b
	150		Low	b
	150		Low	d
	450		High	d
	235		1 ppbv	e
$\alpha$ -Pinene	150	1 hour	0.1 ppbv	e
	28±3		Low	c
	X±3		X	d
	230±90		High	a
	190±50		High	a
	19			b
	210		1 ppbv	e
$\beta$ -Pinene	70	40 min	0.1 ppbv	e
	65±6		Low	c
	X±3		X	d
	540±50		High	a
	450±80		High	a
	45			b
Methane	100		1 year	b
Ethane	180	10 days		b
Propane	60		2 days	b
Methylbutanol	.13(per C)	1 hour		b
HCHO	100	2 hour		b
Acetone	.67(per C)	10 days		b
Methanol	100	2 days		b

a Atkinson and Arey (2003): Table 2, Yield from Isoprene reaction with OH, two values are from two referenced papers therein.

b Palmer (2003): lifetimes assume [OH] is 1e15 mol cm<sup>-3</sup>.

c (Lee et al. 2006b): Calculated through change in concentration of parent and product linear least squares regression. Estimates assume 20° C conditions.

d Wolfe et al. (2016): “prompt yield”: change in HCHO per change in ISOP<sub>0</sub>.  $[ISOP]_0 = [ISOP] \exp(k_1[OH]t)$ ; where  $k_1$  is first order loss rate. Effectively relates HCHO abundance with isoprene emission strength

e Dufour et al. (2008): One-day yields from oxidation modelled by CHIMERE, using MCM reference scheme.

f Calculated using PTR-MS and iWAS on SENEX campaign data.

( $[VOC]_0$ ) at the point of emission and mass balance equations as follows:

$$\frac{1}{k_{HCHO} - k_i} \left( k_{HCHO} \exp \left[ \frac{-k_i L_{s,i}}{U} \right] - k_i \exp \left[ \frac{-k_{HCHO} L_{s,i}}{U} \right] \right) = \frac{1}{e} \quad (3.5)$$

with limiting values  $L_{s,i} \rightarrow U/k_i$  for  $k_i \ll k_{HCHO}$ , and  $L_{s,i} \rightarrow U/k_{HCHO}$  for  $k_{HCHO} \ll k_i$ .

TODO: calculation of smearing Smearing sensitive grid boxes within the model can be detected by running with and without isoprene emissions halved. Similarly to smearing sensitivity calculations in Marais et al. (2012), we run GEOS-Chem with isoprene emissions halved, then calculate  $\hat{S} = \frac{\Delta \Omega_{HCHO}}{\Delta E_{Isop}}$ , where  $\Delta$  represents the monthly mean departure over 1300-1400LT from default run values. This allows us to determine which gridboxes are disproportionately affected by emissions from non-local sources. Consider halving the isoprene emitted globally and rerunning the model, you would expect HCHO enhancement (above background levels) to be halved in isoprene emitting grid-squares. If the local grid-square HCHO enhancement is reduced by much more than half (factoring yield) then you can infer sensitivity to non-local isoprene emissions.

Smearing can be dependent on local or regional weather patterns, as greater wind speeds will reduce the time any emitted compound stays within the local grid box. As such smearing sensitivity is both spatially and temporally diverse, shown in figure TODO: is a picture of the smearing sensitivity over Australia. Large smearing values can be seen near many coastlines as only a fraction of the grid squares actually emit isoprene, which makes transported isoprene relatively more important in these gridboxes. Once the smearing sensitive grid squares are filtered out, application of equation ?? gives us an estimate of isoprene emissions across the nation.

TODO: Plots of S hat showing worst smearing affected areas per season.

Most recently a Bauwens et al. (2016) estimated isoprene emissions with this technique using the IMAGESv2 global CTM. They calculate emissions which create the closest match between model and satellite vertical columns, and compare these posteriori data with the a priori (satellite data) and independent data sets. They found TODO: what they found and how this analysis differs from my own. (TODO: simple outline of what they did and how my focus is different, this paper will also need to be summarised in the LitReview)

### 3.3 Results

#### 3.3.1 Emissions comparisons

Some global numbers (TODO: where do I throw these?) Guenther et al. (2012) Estimate global biogenic isoprene emissions at roughly  $535 \text{ Tg yr}^{-1}$ , using MEGAN. Sindelarova et al. (2014) Estimate around  $594 \text{ Tg yr}^{-1}$  using MEGAN with MACC, showing isoprene as 69.2% of the total BVOC emissions, with monoterpenes at  $10.9 \text{ Tg yr}^{-1}$  (10.9%). They show  $41 \text{ Tg yr}^{-1}$  decrease in Australia when introducing soil moisture parameterisation.

When comparing the GEOS-Chem (which runs MEGAN) emissions to those calculated using our top-down inversion, we see a decrease over TODO: locations and

seasons. TODO: table or figure showing summary of isoprene emissions changes over the whole of our time domain.

Satellite measured HCHO has been found to be biased low in several studies Zhu et al. (eg. 2016), De Smedt et al. (2015), and Barkley et al. (2013). These papers use in-situ data to scale up the satellite HCHO columns for their areas of interest, however Australia lacks sufficient HCHO measurements to do this. In these papers bias is seen as high as 40%, which we use as our upper bound. Scaling up the satellite columns by 40% gives us an upper bound on the uncertainty due to satellite bias. The 'Scaled Satellite' column refers to the calculations when using the 40% scaled up OMI HCHO columns. This can be considered as a boundary on satellite based HCHO column uncertainty.

One set of data from the Daintree rainforest in Queensland exists (TODO: summary from P. Nelson). Although the data set lies outside our run times, as it was measured in TODO(runtime), we compare against the seasonal average of our GEOS-Chem output for the matching months (TODO: name the months). This is done for both GEOS-Chem output and our recalculated isoprene emissions. When compared against GEOS-Chem output we see TODO. When compared against recalculated emissions we see TODO.

TODO: Figure showing campaign data against model and recalculated emissions over region for averaged months and eventually different resolutions.

We examine the affect of decreased isoprene emissions on the correlation between modelled and satellite based HCHO columns. Figure TODO: shows the regressions between GEOS-Chem tropospheric column amounts of HCHO and satellite columns for two runs of GEOS-Chem: a) using standard MEGAN emissions, b) using our updated emissions.

### 3.3.2 Emissions affect on GEOS-Chem

We interpolated or something (TODO) the emissions over Australia into the inventories used by GEOS-Chem which reduced the emissions by X% per year (over Australia). The resulting simulation output shows that HCHO was reduced by X%, although if we boost monoterpenes by X% where the isoprene emissions were lowered then

### 3.3.3 Comparison with in-situ measurements

TODO: Analyse comparison of gridbox with campaigns of measurements

## 3.4 Uncertainty

There are several factors which need to be considered when looking at the uncertainty in emissions estimates. Things with their own inherent uncertainty include the modelled a-priori, modelled relationship between HCHO and isoprene, and satellite measurements. Important factors which need to be analysed for confidence in results include the steady state assumptions, filtering techniques for fire and human influences, and the regression model for determining the isoprene to HCHO yield.

Model uncertainty is difficult to accurately ascertain, generally an analysis of the model compared to in-situ measurements is performed, however there are few of these measurements over Australia.

Yield calculations are performed at low resolution ( $2^{\circ}$  by  $2.5^{\circ}$ ) which may lead to an overestimate as seen in Yu et al. (2016). TODO: how could I check this or estimate sensitivity?

### **3.4.1 Model Uncertainty**

Uncertainty in modelled yield is estimated somehow (TODO:), and a upper and lower bound for the yield is determined using satellite scaling. Since OMI HCHO is scaled up by up to 40% in several papers, we consider HCHO scaled by 1 and 1.4 to be boundaries for modelled yield calculations. These prior works use flight campaigns and in-situ data to verify HCHO columns in various locations (TODO: redo cite list from lit review).

### **3.4.2 Satellite Uncertainty**

Uncertainty in satellite measurements is generally provided along with the data, although uncertainty introduced through AMF calculation needs to be determined to give a representation of the confidence in vertical column amounts.

Provided with the OMI product is the measurement of uncertainty in each pixel, calculated by SAO from the backscattered solar radiation fit (Gonzalez Abad et al. 2015; Abad et al. 2016). BIRA use another method, and calculate the standard deviation of HCHO over the remote pacific ocean as the uncertainty (De Smedt et al. 2012; De Smedt et al. 2015). In the remote pacific, it can be assumed that HCHO variations are weak, with concentrations remaining steady in the short term ( $\sim 1$  month). This means the standard deviation over this region can be used as a proxy for determination of the instrument error.

There are three main sources of error in the resulting HCHO columns:

**a** Fitting error from the OMI retrieval.

**b** Uncertainty in AMF calculations.

**c** Uncertainty of HCHO background.

a) is available in the OMI product and reduced through spatial and temporal averaging. Taking the eight day grided average with horizontal resolution of  $0.25^{\circ}$  by  $0.3125^{\circ}$  degrees (latitude by longitude) typically reduces uncertainty by a factor of 1.5 to 4. Another method for examining uncertainty of OMI is to analyse the standard deviation of the HCHO columns over the remote pacific. If we assume there is no HCHO variation from background levels over any 8-day period, then this method infers variations in the measuring instrument, and can be used as a metric for uncertainty as done in De Smedt et al. (2012). TODO: uncertainty calculation on remote pacific OMI. Millet et al. (2006) and Palmer et al. (2006) both examine OMI HCHO columns over North America and determine overall uncertainty to be 40%, with most of this coming from cloud interference.

b) is determined through an analysis of GEOS-Chem output, validated against the total column of HCHO at Wollongong using FTIR measurements from the (TODO: Nicholas Jones roof HCHO citation here). Palmer et al. (2006) calculate the error in AMF through combining estimates of error in the UV albedo database ( $\sim 8\%$ ), model error based on in-situ measurements, cloud error (20 – 30%) (Martin et al. 2003), and aerosol errors (< 20%), totalling AMF error of around  $\sim 30\%$ . It is worth noting here that independent error estimates are added in quadrature, which means total error equals the root of the sum of the independent errors each squared ( $e_{Total} = \sqrt{\sum_i e_i^2}$ ). TODO: Paul palmer calculation and combination for overall Satellite VC uncertainty per pixel and gridded. TODO: Millet2008?

c) is also determined through a study of GEOS-Chem output, in relation to in-situ measurements. TODO: calculate this uncertainty. Compare this error estimate with that of Curci et al. (2010), where the error in b) and c) are respectively found to be 30% and 15% based on their analysis of CHIMERE. Millet et al. (2008) also examine this uncertainty and determine an overall uncertainty ( $1\sigma$ ) of 25 – 27% in HCHO vertical columns with calculated AMFs where cloud fraction  $< 0.2$ .

### 3.4.3 Fire Filtering

Figure 3.5 shows emissions estimates for January 2005, using three different HCHO columns as the basis: the original OMI satellite HCHO columns ( $\Omega$ ), those with AMF recalculated using a new apriori ( $\Omega_{GC}$ ), and those with AMFs recalculated using PP code ( $\Omega_{PP}$ ). The first row shows emissions estimates calculated as shown in 3.2, while the second row runs the same calculations without applying any fire or smoke filter. The Third row is the absolute difference between them: fire filtered minus standard emissions.

### 3.4.4 MEGAN

One of the important parameters in Australia is the soil moisture activity factor ( $\gamma_{SM}$ ), which can have large regional effects on the isoprene emissions (Sindelarova et al. 2014; Bauwens et al. 2016). Generally if soil moisture is too low, isoprene emissions stop (Pegoraro et al. 2004; Niinemets et al. 2010), however in many Australian regions the plants may be more adapted to lower moisture levels. (TODO: Find cites for this - talk from K Emerson at Stanley indicated this) GEOS-Chem runs MEGANv2.1, which has three possible states for isoprene emissions based on the soil moisture ( $\theta$ ):

$$\begin{array}{ll} \gamma_{SM} = 1 & \theta > \theta_1 \\ \gamma_{SM} = (\theta - \theta_w) / \Delta\theta_1 & \theta_w < \theta < \theta_1 \\ \gamma_{SM} = 0 & \theta < \theta_w \end{array}$$

where  $\theta_w$  is the wilting point, and  $\theta_1$  determines when plants are near the wilting point. The wilting point is set by a land based database from Chen and Dudhia (2001), while  $\theta_1$  is set globally based on Pegoraro et al. (2004). Potentially importantly, these

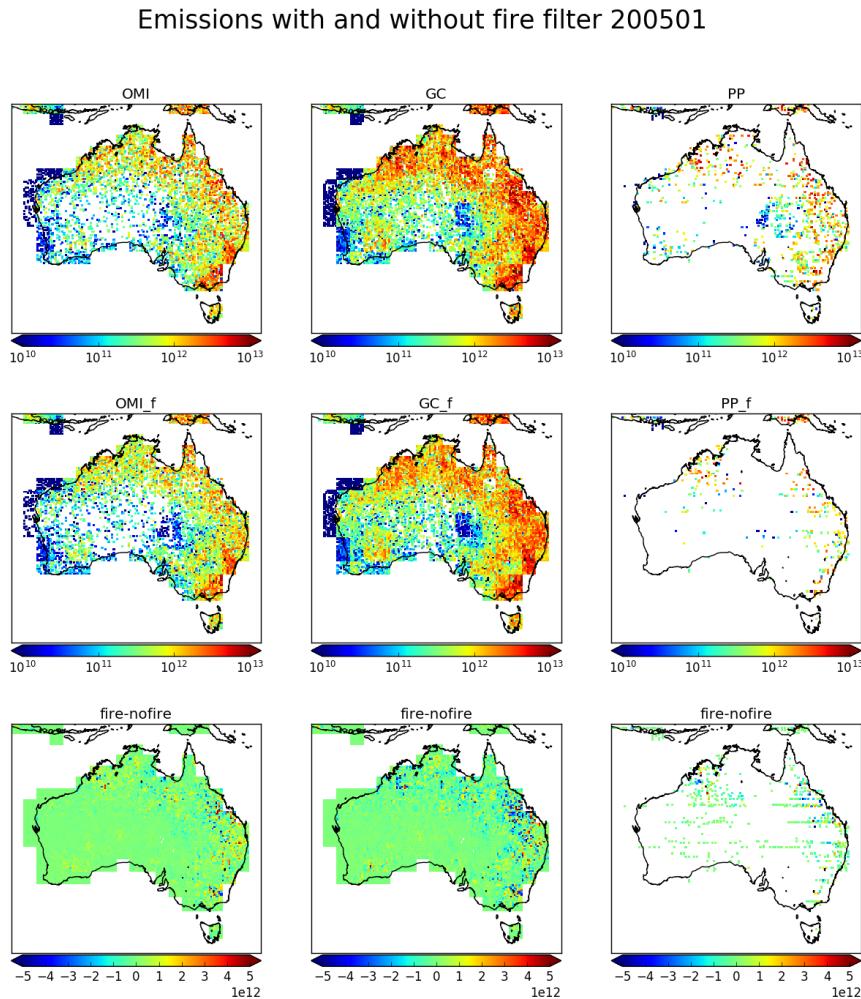


FIGURE 3.5: Emissions estimates using OMI satellite columns (column 1) recalculated with updated shape factor (column 2) and scattering weights (column 3). Turning off the fire and smoke filters gives emissions in row 2, while the difference between row 1 and row 2 is shown in row 3.

moisture states are disabled in GEOS-Chem V10.01, which is partly because accurate maps of soil moisture are not available.



## Chapter 4

# Stratospheric ozone intrusions

### 4.1 Introduction

Tropospheric ozone constitutes only 10% of the total ozone column but is an important oxidant and greenhouse gas which is toxic to life, harming natural ecosystems and reducing agricultural productivity. Over the industrial period, increasing tropospheric ozone has been estimated to exert a radiative forcing (RF) of  $365 \text{ mWm}^{-2}$  (Stevenson et al. 2013), equivalent to a quarter of the CO<sub>2</sub> forcing (Forster et al. 2007). While much tropospheric ozone is produced photochemically from anthropogenic and natural precursors, downward transport from the ozone-rich stratosphere provides an additional natural source of ozone that is particularly important in the upper troposphere (Jacobson and Hansson 2000, and references therein). The contribution of this source to overall tropospheric ozone budgets remains uncertain (Škerlav, Sprenger, and Wernli 2014), especially in the southern hemisphere (SH). Models show that stratospheric ozone depletion has propagated to the upper troposphere (Stevenson et al. 2013). However, work based on the Southern Hemisphere Additional OZonesonde (SHADOZ) network suggests stratospheric mixing may be increasing upper tropospheric ozone near southern Africa (Liu et al. 2015; Thompson et al. 2014). Uncertainties in the various processes which produce tropospheric ozone limit predictions of future ozone-induced radiative forcing. Here we use a multi-year record of ozonesonde observations from sites in the southern hemisphere extra-tropics, combined with a global model, to better characterise the impact of stratospheric ozone on the tropospheric ozone budget in the southern hemisphere.

Stratosphere-to-troposphere transport (STT) primarily impacts the ozone budget in the upper troposphere but can also increase regional surface ozone levels above the legal thresholds set by air quality standards (Danielsen 1968; Lelieveld et al. 2009; Lefohn et al. 2011; Langford et al. 2012; Zhang et al. 2014; Lin et al. 2015). In the western US, for example, deep STT events during spring can add 20-40 ppbv of ozone to the ground-level ozone concentration, which can provide over half the ozone needed to exceed the standard set by the U.S. Environmental Protection Agency (Lin et al. 2012; Lin et al. 2015). Another hotspot for STT is the Middle East, where surface ozone exceeds values of 80 ppbv in summer, with a stratospheric contribution of 10 ppb (Lelieveld et al. 2009). Estimates of the overall contribution of STT to tropospheric ozone vary widely (e.g. Galani 2003; Stohl et al. 2003; Stevenson et al. 2006; Lefohn et al. 2011). Early work based on two photochemical models showed that 25-50% of the tropospheric ozone column can be attributed to STT events globally, with most contribution in the upper troposphere (Stohl et al. 2003). In contrast, a more

recent analysis of the Atmospheric Chemistry and Climate Model Inter-comparison Project (ACCMIP) simulations by Young et al. (2013) found that STT is responsible for  $540 \pm 140 \text{ Tg yr}^{-1}$ , equivalent to  $\sim 11\%$  of the tropospheric ozone column, with the remainder produced photochemically (Monks et al. 2015). This wide range in model estimates exists in part because STT is challenging to accurately represent, and finer model resolution is necessary to simulate small scale turbulence. Observation-based process studies are therefore key in determining the relative frequency of STT events, with models then able to quantify STT impact over large regions. Ozonesondes are particularly valuable for this purpose as they provide multi-year datasets with high vertical resolution.

Lower stratospheric and upper tropospheric ozone concentrations are highly correlated, suggesting mixing across the tropopause mainly associated with the jet streams over the Atlantic and Pacific oceans (Terao et al. 2008). Extra-tropical STT events most commonly occur during synoptic-scale tropopause folds (Sprenger, Croci Maspoli, and Wernli 2003; Tang and Prather 2012; Frey et al. 2015) and are characterised by tongues of high potential vorticity (PV) air descending to lower altitudes. As these tongues become elongated, filaments disperse away from the tongue and mix irreversibly into the troposphere. STT can also be induced by deep overshooting convection (Frey et al. 2015), tropical cyclones (Das et al. 2016) and mid-latitude synoptic scale disturbances (e.g. Stohl et al. 2003; Mihalikova et al. 2012). STT events have been observed in tropopause folds around both the polar front jet (Vaughan, Price, and Howells 1993; Beekmann et al. 1997) and the subtropical jet (Baray et al. 2000). The summertime pool of high tropospheric ozone over the eastern Mediterranean (EM) is mainly attributed to the downward ozone transport, as a result of the enhanced subsidence (Zanis et al. 2014) and the tropopause fold activity (Akritidis et al. 2016) over the region. The EM exhibits a summer maximum of subsidence, which sits between  $20^\circ\text{E}$  and  $35^\circ\text{E}$  and  $31^\circ\text{N}$  to  $39^\circ\text{N}$ , while zonally most subtropical tropopause folds occur during winter (Tyrlis et al. 2014, and references therein). They are also observed near cut-off lows (Price and Vaughan 1993; Wirth 1995), so both regional weather patterns and stratospheric mixing are important to understand for STT analysis.

Stratospheric ozone intrusions undergo transport and mixing, with up to half of the ozone diffusing within 12 hours following descent from the upper troposphere (Trickl et al. 2014). The long range transport of enhanced ozone can be facilitated by upper tropospheric winds, with remarkably little convective mixing, as shown by Trickl et al. 2014 who measure STT air masses two days and thousands of kilometres from their source. Cooper et al. (2004) also shows how STT advection can transport stratospheric air over long distances, with a modelled STT event spreading from the northern Pacific to the East coast of the USA over a few days.

The strength (ozone enhancement above background levels), horizontal scale, vertical depth, and longevity of these intruding ozone tongues vary with wind direction and strength, topography, and season. While the frequency, seasonality, and impacts of STT events have been well described in the tropics and northern hemisphere (NH), observational estimates from the SH extra-tropics are noticeably absent in the literature. The role of STT in the SH remains highly uncertain due to the more limited data availability compared to the NH and the temporal sparsity of these datasets (Mze et al. 2010; Thompson et al. 2014; Liu et al. 2015).

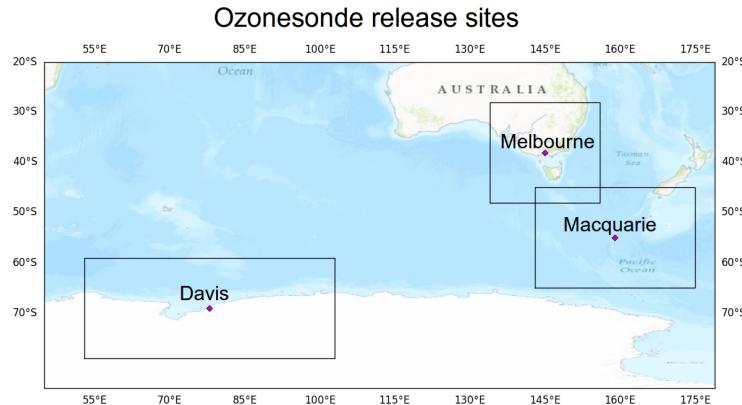


FIGURE 4.1: Ozonesonde release sites and the regions used to examine STT effect on tropospheric ozone levels.

Here, we characterise the seasonal cycle of STT events and quantify their contribution to the SH extra-tropical tropospheric ozone budget using nearly a decade of ozonesonde observations from three locations around the Southern Ocean spanning latitudes from 38°S-69°S. In Section 4.2 we describe the observations and methods used to identify STT events and to relate STT occurrence to meteorological events. We show how possible biomass burning smoke plume influence is detected and handled, and we introduce the GEOS-Chem model which is used for ozone flux estimation. Within Section 4.3 we show the seasonality, altitude, depth, and frequency of detected STT events, along with a comparison of our findings to other literature where possible. In Section 4.4 we analyse how well GEOS-Chem captures the tropospheric ozone seasonality and quantity near our three sites. In Section 4.5 an extrapolation of the STT detection frequencies along with GEOS-Chem monthly tropospheric ozone columns is used to estimate STT ozone flux near our three sites. We also compare and contrast our results against relevant literature. Finally, in Section 4.6 we examine in detail the uncertainties involved in our STT event detection technique and ozone flux estimations.

## 4.2 Data and Methods

### 4.2.1 Ozonesonde record in the Southern Ocean

Ozonesondes provide a high vertical resolution profile of ozone, temperature, pressure, and humidity from the surface and up to 35 km. In the troposphere, the ozonesondes generally perform 150-300 measurements. Ozone mixing ratio is quantified with an electrochemical concentration cell, using standardised procedures when constructing, transporting, and releasing the ozonesondes (<http://www.ndsc.ncep.noaa.gov/organize/protocols/appendix5/>). Ozonesondes are estimated to provide around 2% precision in the stratosphere, which improves at lower altitudes, and ozonesondes have been shown to be accurate to within 5% when the correct procedures are followed (Smit et al. 2007).

TABLE 4.1: Number of sonde releases at each site over the period of analysis.

Site	Total Releases	Monthly Releases (J, F, M, ...)	Date Range
Davis	240	11, 12, 13, 12, 17, 31, 29, 28, 32, 28, 15, 12	2006/04/13 - 2013/11/13
Macquarie Island	390	32, 31, 45, 28, 34, 33, 28, 35, 29, 33, 31, 31	2004/01/20 - 2013/01/09
Melbourne	456	31, 38, 40, 38, 41, 36, 38, 39, 46, 40, 38, 31	2004/01/08 - 2013/12/18

Ozonesondes are launched approximately weekly from Melbourne ( $38^{\circ}$  S,  $145^{\circ}$  E), Macquarie Island ( $55^{\circ}$  S,  $159^{\circ}$  E) and Davis ( $69^{\circ}$  S,  $78^{\circ}$  E), as shown in Fig. 4.1. Melbourne is a major city in the south east of Australia, and may be affected by anthropogenic pollution in the lower troposphere. Macquarie Island is isolated from the Australian mainland, situated in the remote Southern Ocean and unlikely to be affected by any local pollution events. Davis is on the coast of Antarctica and also unlikely to experience the effects of anthropogenic pollution.

For this study, we use the 2004-2013 data for Melbourne and Macquarie Island and the 2006-2013 data for Davis because both ozone and geopotential height (GPH) are available from the World Ozone and Ultraviolet Data Centre archived data in these periods. At Davis, ozonesondes are launched twice as frequently during the ozone hole season and preceding months (June-October) as at other times of year (Alexander et al. 2013). A summary of ozonesonde releases at each site can be seen in Table 4.1.

Characterisation of STT events requires a clear definition of the tropopause. Two common tropopause height definitions are the standard lapse rate tropopause (WMO 1957) and the ozone tropopause (Bethan, Vaughan, and Reid 1996). The lapse rate tropopause is defined as the lowest altitude where the lapse rate (vertical gradient of temperature) is less than  $2^{\circ}\text{C km}^{-1}$ , provided the lapse rate averaged between this altitude and 2 km above is also below  $2^{\circ}\text{C km}^{-1}$ . The ozone tropopause is defined as the lowest altitude satisfying the following three conditions for the ozone mixing ratio (OMR) (Bethan, Vaughan, and Reid 1996):

1. Vertical gradient of OMR is greater than  $60 \text{ ppb km}^{-1}$ ;
2. OMR is greater than 80 ppb; and
3. OMR exceeds 110 ppb between 500 m and 2000 m above the altitude under inspection (modified to between 500 m and 1500 m in the Antarctic, including the site at Davis).

The ozone tropopause may misdiagnose the real tropopause altitude during stratosphere-troposphere exchange; however, it is useful at polar latitudes in winter, where the lapse-rate definition may result in artificially high values for tropopause height (Bethan, Vaughan, and Reid 1996; Tomikawa, Nishimura, and Yamanouchi 2009; Alexander et al. 2013). We require lapse rate defined tropopauses to be at a minimum of 4 km altitude. Another commonly used tropopause definition is determined with the use of PV (dynamical tropopause). In the extra-tropics the isosurface where  $\text{PV} = 2 \text{ PVU}$

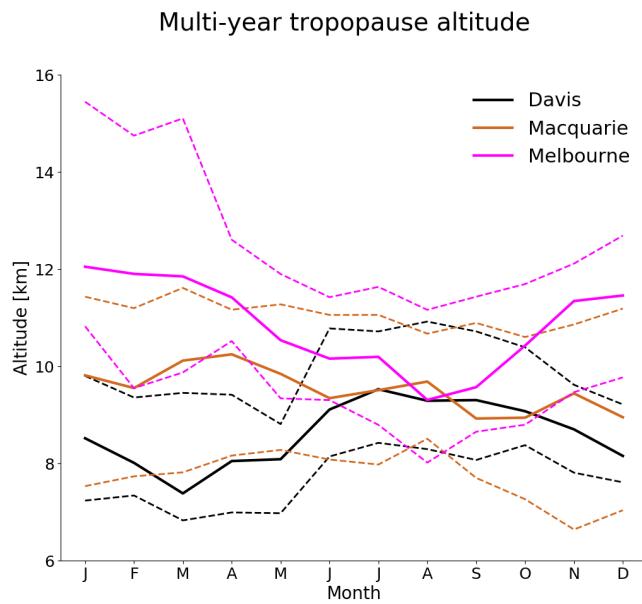


FIGURE 4.2: Multi-year monthly median tropopause altitude (using the ozone defined tropopause) determined from ozonesondes measurements at Davis (2006-2013), Macquarie Island (2004-2013), and Melbourne (2004-2013) (solid lines). Dashed lines show the 10th to the 90th percentile of tropopause altitude for each site.

(1 PVU =  $10^{-6} \text{ m}^2 \text{ s}^{-1} \text{ K kg}^{-1}$ ) is often used to define the tropopause, allowing the 3D representation of tropopause folds and other tropopause features in a sufficiently resolved model (Škerlak, Sprenger, and Wernli 2014; Tyrlis et al. 2014). The PV is not calculable using the ozonesonde measurements alone, so in this work the ozone tropopause is used when determining STT events or measured tropopause altitude.

Figure 4.2 shows the monthly median ozone tropopause altitudes at each location (solid lines). At Melbourne, the tropopause altitude displays a seasonal cycle with maximum in summer and minimum in winter. This seasonality is missing at Macquarie Island and almost reversed at Davis, which has a minimum during autumn and maximum from winter to spring. Tropopause altitude decreases with latitude from 9-14 km at Melbourne ( $38^\circ \text{ S}$ ) to 7-9 km at Davis ( $69^\circ \text{ S}$ ).

Figure 4.3 shows multi-year averaged ozone mixing ratios measured by ozonesonde over the three stations. Over Melbourne, increased ozone extending down through the troposphere is apparent from December to March and from September to November. The increased tropospheric ozone in these months is due to STT (in summer), and possible biomass burning influence (in spring), both discussed in more detail in the following sections. Over Davis and Macquarie Island, tropospheric ozone is higher between March and October, although the seasonal differences are small compared to those at Melbourne. The seasonality shown in Fig. 4.3 for Davis is consistent with remote free tropospheric photochemistry determined by solar radiation availability and temperature, resulting in higher ozone in winter (Lelieveld and Dentener 2000). NO<sub>2</sub> stratospheric observations have been conducted in the Southern hemisphere at

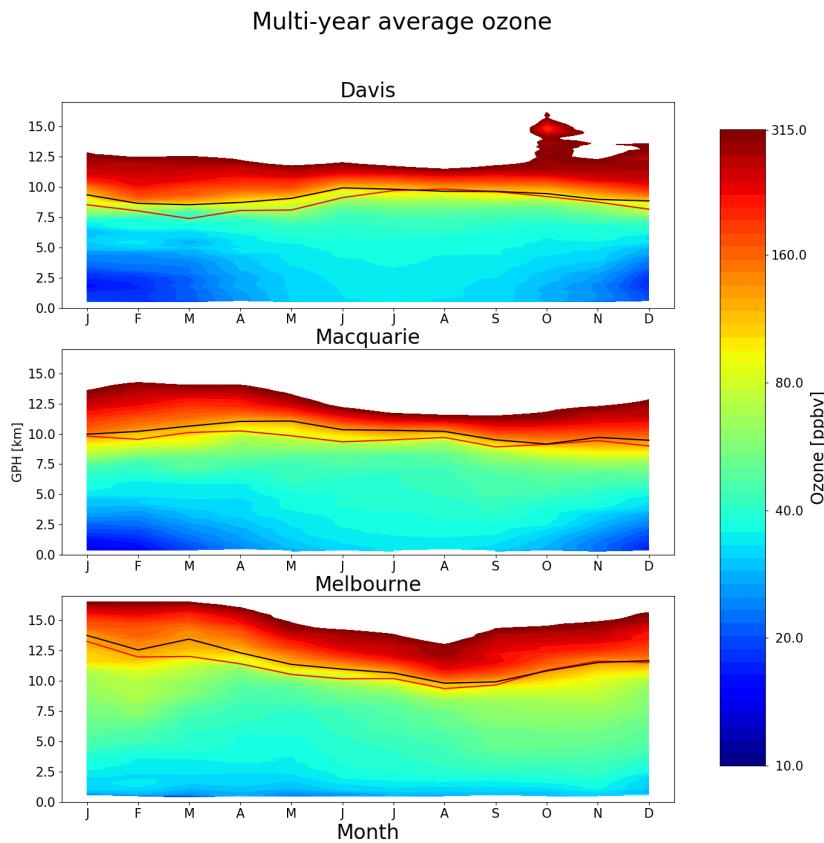


FIGURE 4.3: Multi-year mean seasonal cycle of ozone mixing ratio over Davis, Macquarie Island, and Melbourne as measured by ozonesondes. Measurements were interpolated to every 100 m and then binned monthly. Black and red solid lines show median ozone and lapse-rate defined tropopause altitudes (respectively), as defined in the text.

Lauder, Macquarie Island and Arrival Heights (i.e. Struthers et al. 2004) which displays a winter minima in seasonality consistent with an ozone maxima. Influence from the ozone hole can be seen over Davis in October, with relatively low ozone levels extending up 5-6 km into the stratosphere.

#### 4.2.2 Model description

To provide regional and global context to the ozonesonde observations, we use the GEOS-Chem version 10-01 global chemical transport model (Bey et al. 2001), which simulates ozone along with more than 100 other trace gases throughout the troposphere and stratosphere. Stratosphere-troposphere coupling is calculated using the stratospheric unified chemistry extension (UCX) (Eastham, Weisenstein, and Barrett 2014). Transport is driven by assimilated meteorological fields from the Goddard Earth Observing System (GEOS-5) maintained by the Global Modeling and Assimilation Office (GMAO) at NASA. Ozone precursor emissions are from the Model of

Emissions of Gases and Aerosols from Nature (MEGAN) version 2.1 (Guenther et al. 2012) for biogenic emissions, the Emissions Database for Global Atmospheric Research (EDGAR) version 4.2 for anthropogenic emissions, and the Global Fire Emissions Database (GFED4) inventory (Giglio, Randerson, and Van Der Werf 2013) for biomass burning emissions. Our simulation was modified from the standard v10-01 to fix an error in the treatment of ozone data from the Total Ozone Mapping Spectrometer (TOMS) satellite used to calculate photolysis (see [http://wiki.seas.harvard.edu/geos-chem/index.php/FAST-JX\\_v7.0\\_photolysis\\_mechanism#Fix\\_for\\_TOMS\\_to\\_address\\_strange\\_cycle\\_in\\_OH\\_output.](http://wiki.seas.harvard.edu/geos-chem/index.php/FAST-JX_v7.0_photolysis_mechanism#Fix_for_TOMS_to_address_strange_cycle_in_OH_output.)).

Our simulations span 2005-2012 (following a 1-year spin-up) with horizontal resolution of  $2^{\circ}$  latitude by  $2.5^{\circ}$  longitude and 72 vertical levels from the surface to 0.01 hPa. The vertical resolution is finer near the surface at  $\sim 60$  m between levels, spreading out to  $\sim 500$  m near 10 km altitude. For comparison to the ozonesonde observations, the model state was saved every 6 hours within the grid boxes containing each site. When comparing against ozonesondes, GEOS-Chem UTC+0 time samples are used for all sites. This means that the simulated ozone profiles are analysed at local times of 7AM for Davis, and 11AM for Macquarie Island and Melbourne. GEOS-Chem uses the tropopause height provided by GEOS-5 meteorological fields, which are calculated using a lapse-rate tropopause definition using the first minimum above the surface in the function  $0.03 \times T(p) - \log(p)$ , with p in hPa (Rienecker 2007).

#### 4.2.3 Characterisation of STT events and associated fluxes

We characterise STT events using the ozonesonde vertical profiles to identify tropospheric ozone enhancements above a local background (in moles per billion moles of dry air, referred to here as ppb). The process is illustrated in Figure 4.4 on an example ozone profile. First, the ozone vertical profiles are linearly interpolated to a regular grid with 20 m resolution from the surface to 14 km altitude. Small vertical-scale fluctuations in ozone, which are captured by the high-resolution ozonesondes, can be regarded as sinusoidal waves superimposed on the large vertical scale background tropospheric ozone. As such, the interpolated profiles are bandpass-filtered using a fast Fourier transform (Press et al. 1992) to retain these small vertical scales, between 0.5 km and 5 km (removing low and high frequency perturbations). The high frequency perturbations are removed as they may represent noise in the measurements. The perturbations with scales longer than 5 km represent the vertical gradient of ozone concentration from the surface to the stratosphere. In what follows, these filtered vertical profiles are referred to as perturbation profiles.

For an event to qualify as STT, a clear increase above the background ozone level is needed, as a bandpass filter leaves us with enhancements minus any noise or seasonal scale vertical profile effects. We next use all the perturbation profiles at each site to calculate the 95th percentile perturbation value for the site. The threshold is calculated from all the interpolated filtered values between 2 km above the surface and 1 km below the tropopause. This is our threshold for tropospheric ozone perturbations, and any profiles with perturbations exceeding this value in individual ozonesondes are classified as STT events. STT events at altitudes below 4 km are removed to avoid surface pollution, and events within 0.5 km of the tropopause are removed to avoid false positives induced by the sharp transition to stratospheric air. We note that this

### Ozone at Melbourne on 2004/01/08

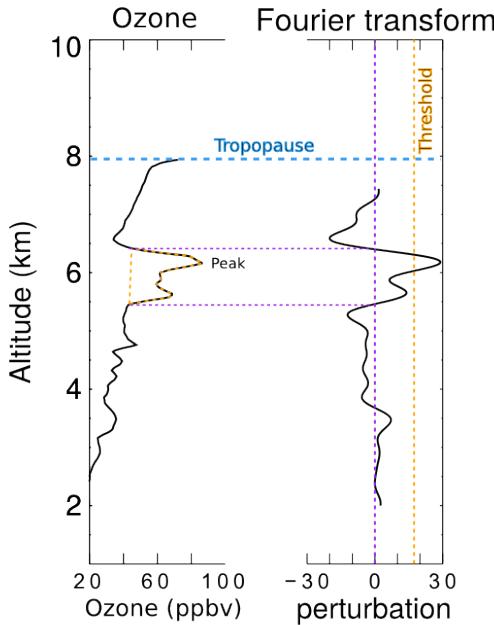


FIGURE 4.4: An example of the STT identification and flux estimation methods used in this work. The left panel shows an ozone profile from Melbourne on 8 January 2004 from 2 km to the tropopause (blue dashed horizontal line). The right panel shows the perturbation profile created from bandpass filtering of the mixing ratio profile. The STT occurrence threshold calculated from the 95th percentile of all perturbation profiles is shown as the orange dashed line, and the vertical extent of the event is shown with the purple dashed lines (see details in text). The ozone flux associated with the STT event is calculated using the area outlined with the orange dashed line in the left panel.

ozone detection methodology detailed above does not allow us to resolve STT events where the ozone flux is spread diffusely across the troposphere without a peak-like structure in the ozonesonde profile. In other words, STT events which might have occurred some distance and time away from the location of the ozonesonde profiles may not be readily detected using the high vertical resolution, but infrequent, ozonesonde launches.

We define the ozone peak as the altitude where the perturbation profile is greatest between 2 km from the surface and 0.5 km below the tropopause. The STT event is confirmed if the perturbation profile drops below zero between the ozone peak and the tropopause, as this represents a return to non-enhanced ozone concentrations. Alternatively, the STT event is also confirmed if the OMR between the ozone peak and the tropopause drops below 80 ppb and is at least 20 ppb lower than the OMR at the ozone peak. If neither of these conditions are met, the profile is rejected as a non-event. This final step removes near-tropopause anomalies for which there is insufficient evidence of detachment from the stratosphere. Vertical ozone profiles recorded by ozonesondes are highly dependent on the time of launch (Sprenger, Croci Maspoli, and Wernli

2003), and it cannot be guaranteed that detected ozone enhancements are fully separated from the stratosphere, although this method minimises that risk by removing detected events too near the tropopause.

We estimate the ozone flux into the troposphere associated with each event by integrating the ozone concentration enhancement vertically over the altitude range for which an STT event is identified (i.e. enhancement near the ozone peak over which the perturbation profile is greater than zero). This estimate is conservative because it does not take into account any ozone enhancements outside of the detected peak that may have been caused by the STT, and also ignores any enhanced ozone background amounts from synoptic-scale stratospheric mixing into the troposphere.

Our method differs somewhat from that used by Tang and Prather (2010) to detect STT events from ozonesonde measurements. Their definition is based on subjective analysis of sondes released from 20 stations ranging in latitude from 35° S to 40° N. They identify an STT event if, starting from 5 km altitude, ozone exceeds 80 ppb and then within 3 km decreases by 20 ppb or more to a value less than 120 ppb. Their technique would miss many events due to the lower ozone concentrations found in the cleaner Southern Hemisphere.

#### 4.2.4 Biomass burning influence

The STT detection algorithm described in Sect. 4.2.3 assumes all ozone enhancements are caused by stratospheric intrusions. In some cases, however, these perturbations may in fact reflect ozone production in lofted smoke plumes. Biomass burning in southern Africa and South America has previously been shown to have a major influence on atmospheric composition in the vicinity of our measurement sites (Oltmans et al. 2001; Gloudemans et al. 2007; Edwards et al. 2006), particularly from July to December (Pak et al. 2003; Liu et al. 2017). On occasion, smoke plumes from Australian and Indonesian fires can also reach the mid-high southern latitudes, as seen from satellite measurements of carbon monoxide (CO) discussed below.

Large biomass burning events emit substantial quantities of ozone precursors, some of which are capable of being transported over long distances and driving ozone production far from the fire source (Jaffe and Wigder 2012). Ozone production from biomass burning is complex and affected by photochemistry, fuel nitrogen load, and time since emission, among other factors. While ozone production occurs in some biomass burning plumes, this is not always the case; therefore ozone perturbations detected during transported smoke events may or may not be caused by the plume. For this reason all detected STT events which could be caused by smoke plumes are flagged, following the procedure outlined below. Calculations of seasonality, and ozone flux do not include flagged events, however they are included in summary plots in this work.

Possible biomass burning influence is identified using satellite observations of CO from the AIRS (Atmospheric Infra-red Sounder) instrument on board the Aqua satellite (Texeira 2013). CO is emitted during incomplete combustion and is an effective tracer of long-range transport due to its long lifetime (Edwards 2003; Edwards et al. 2006). In the Southern Hemisphere, biomass burning is the primary source of CO, making CO a good proxy for fire plumes (e.g. Sinha et al. 2004; Mari et al. 2008). To

identify possible biomass burning influence, AIRS vertical column CO is visually inspected for all dates with detected STT events. Smoke plumes are diagnosed over areas with elevated CO columns ( $\sim 2 \times 10^{18}$  molecules  $\text{cm}^{-2}$  or higher), and any sonde-detected STT event that occurs near (within  $\sim 150$  km of) a smoke plume is flagged. Removal of these detections reduces the yearly estimated ozone flux by  $\sim 15\%$  at Macquarie Island and  $\sim 20\%$  at Melbourne.

All days with detected STT events were screened, with the exception of one event during which there were no available AIRS data (January 2010). We find that biomass burning may have influenced 27 events over Melbourne and 21 events over Macquarie Island. These events are flagged in the following sections, and are not used in our calculation of total STT flux. All of the flagged events except for two occurred during the SH burning season (July to December). No events at Davis were seen to be influenced by smoke transport.

#### 4.2.5 Classifying synoptic conditions during STT events

Synoptic scale weather patterns are examined using data from the European Centre for Medium-range Weather Forecasts (ECMWF) Interim Reanalysis (ERA-I) (Dee et al. 2011). This is done using the ERA-I data products over the three sites on dates matching the detected STT events. We use the ERA-I 500 hPa data to subjectively classify the events based on their likely meteorological cause, by visually examining each date where an event was detected. During STT occurrence, the upper troposphere is typically characterised by nearby cyclones, cut-off lows, or cold fronts. Over Melbourne and Macquarie Island, we find that frontal and low pressure activity are prevalent during STT events (see Sect. 4.3). Over Davis, the weather systems are often less clear, however we see a higher portion of probable cut-off lows. The stratospheric polar vortex may create tropopause folds without other sources of upper tropospheric turbulence such as low pressure fronts or cyclones (e.g. Baray et al. 2000; Sprenger, Croci Maspoli, and Wernli 2003; Tyrlis et al. 2014). Cut-off low pressure systems can be seen clearly in synoptic scale weather maps as regions with lowered pressure and cyclonic winds. Low pressure fronts in the higher southern latitudes travel from west to east and lower the tropopause height. We examine two cases in detail to illustrate the relationship between synoptic-scale conditions and STT events over Melbourne. These are included in a supplementary document (Fig. S2 and S3) which show an archetypal cut-off low and low pressure front. To detect cut-off low pressure systems we look for cyclonic winds and a detached area of low pressure within  $\sim 500$  km of a site on days of event detection. For low pressure fronts we look for low pressure troughs within  $\sim 500$  km. Frontal passage is a known cause of STT as stratospheric air descends and streamers of ozone-rich air break off and mix into the troposphere (Sprenger, Croci Maspoli, and Wernli 2003).

### 4.3 STT event climatologies

Figure 4.5 shows the seasonal cycles of STT frequency at Davis, Macquarie Island, and Melbourne. Frequency is determined as detected event count divided by total launched ozonesondes for each month. STT events in Figures 4.5-4.8 are coloured

TABLE 4.2: Total number of ozonesonde detected STT events, along with the number of events in each category (see text).

Site	Events	Cut-offs	Frontals	Misc	Fire
Davis	80	44	19	17	0
Macquarie Island	105	19	31	34	21
Melbourne	127	28	31	41	27

based on the meteorological classification described in Sect. 4.2.5, with events classified as either low pressure fronts (“frontal”, dark blue), cut-off low pressure systems (“cutoff” teal), or indeterminate (“misc”, cyan). Events that may have been influenced by transported smoke plumes (Sect. 4.2.4) are shown in red. Ozonesonde releases are summarised in Table 4.1 and detected event counts are summarised in Table 4.2.

There is an annual cycle in the frequency of STT events (Fig. 4.5) with a summertime peak at all three sites. This summertime peak is due to a prevalence of summer low-pressure storms and fronts, which increase turbulence and lower the tropopause (Reutter et al. 2015). At Davis, there are more STT detections during winter relative to our other sites, which may be due to the polar vortex and its associated lowered tropopause and increased turbulence. STT events associated with cut-off low pressure systems are more prevalent during summer, while STT events associated with frontal passage occur throughout the year. The high frequency of STT ozone enhancements is comparable to the > 25% frequencies seen over Turkey and east of the Caspian sea in an ERA-I analysis performed by Tyrlis et al. (2014).

The SH summer maximum we see for STT ozone flux can also be seen in Fig. 16 of Škerlak, Sprenger, and Wernli 2014, which shows seasonal flux over the southern ocean, although this is less clear over Melbourne. This seasonality is not clear in the recent ERA-Interim tropopause fold analysis performed by Škerlak et al. (2015), where a winter maximum of tropopause fold frequency ( $\sim 0.5\%$  more folds in winter) over Australia can be seen to the north of Melbourne. Their work seems to show slightly higher fold frequencies over Melbourne in summer (Škerlak et al. 2015, Fig. 5), however not to the same extent that our summer peak suggests. Their winter maximum is in the subtropics only - from around  $20^\circ$  S to  $40^\circ$  S, which can be seen as the prevalent feature over Australia in their Fig. 5. Wauben, Fortuin, and Velthoven 1998 look at modelled (CTM driven by ECMWF output) and measured ozone distributions and find more SH ozone in the lower troposphere during austral winter, however they note that the ECMWF fields are uncertain here again due to lack of measurements. Their work shows a generally cleaner lower troposphere in the SH summer but this can not be construed to suggest more or less STT folds in either season. Sprenger, Croci Maspoli, and Wernli 2003 examine modelled STT folds using ECMWF output over March 2000 - April 2001, and show that for this year there is a clear austral winter maximum, again over the  $20^\circ$  S to  $40^\circ$  S band. The winter maximum does not include Melbourne, or the southern ocean, which explains why we see a seasonality not readily evident in these global-scale studies.

The measurement sites are not in the regions which have a clear winter maximum seen in Fig. 1 Sprenger, Croci Maspoli, and Wernli 2003, and the large scale winter maximum shown by all three studies seems to be dominated by the system in that

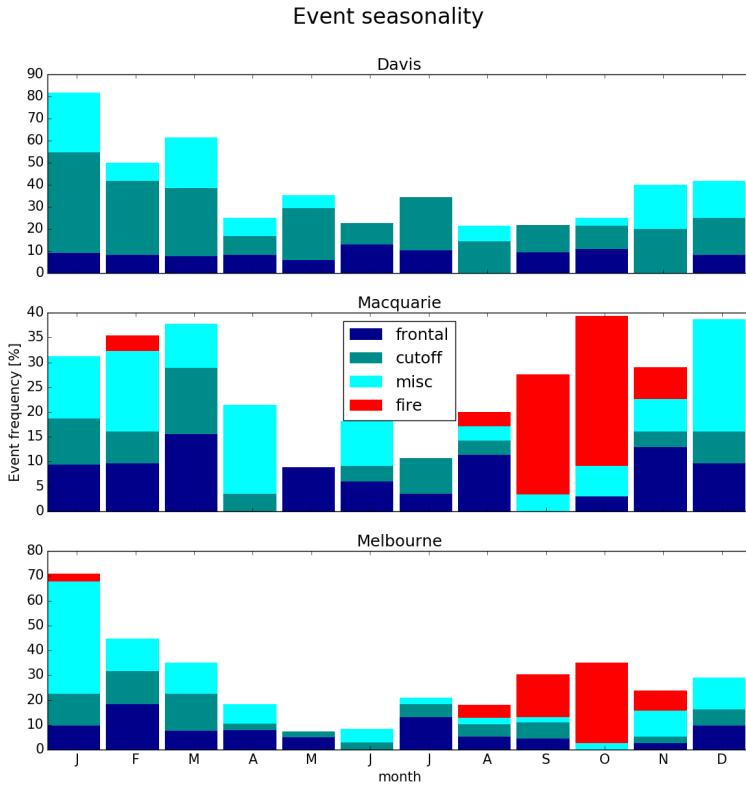


FIGURE 4.5: Seasonal cycle of STT event frequency at Davis (top), Macquarie Island (middle), and Melbourne (bottom). Events are categorised by associated meteorological conditions as described in the text, with low pressure fronts (“frontal”) in dark blue, cut-off low pressure systems (“cutoff”) in teal, and indeterminate meteorology (“misc”) in cyan. Events that may have been influenced by transported smoke plumes are shown in red (see text for details).

region. The seasonality of our three sites is not driven by the larger STT system seen over the southern Indian ocean and middle Australia which dominates prior analysis near or over Australia.

To examine the robustness of the distributions shown in Fig. 4.5, we developed an alternative assessment of the seasonal occurrence of STT events, with results shown in Fig. 4.6. Here STT occurrence is evaluated by consideration of the square of the dry Brunt-Väisälä frequency ( $N^2$ ) at the heights of the ozone tropopause ( $z_{OT}$ ) and lapse rate tropopause ( $z_{LRT}$ ) in each ozonesonde profile that has been binned to 100 m resolution. We use  $N^2$  to assess atmospheric stability, which is normally distinctly higher in the stratosphere than in the troposphere, and assume that the vertical temperature gradients within the intrusion respond most rapidly to transported heat, which is an additional characteristic of stratospheric air.  $N^2$  is evaluated using 250 m resolution data (to smooth variability in the vertical gradient of potential temperature that is due to small temperature fluctuations likely associated with gravity waves). The altitude binning chosen is a compromise between vertical resolution and the level of

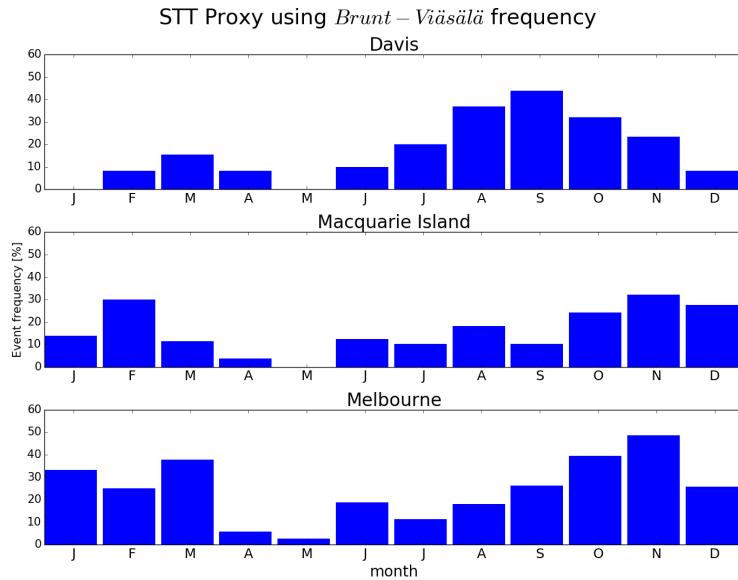


FIGURE 4.6: Seasonal distribution of STT events using the alternative STT proxy, obtained from consideration of the static stability at the ozone and lapse rate tropopauses, for Davis (2006-2013), Macquarie Island (2004-2013), and Melbourne (2004-2013).

variability in  $N^2$  introduced by temperature gradients associated with perturbations from gravity waves and changes near the lapse rate tropopause, and is the minimum that produces a robust seasonal distribution. We define STT as having taken place if  $N^2(z_{OT}) > N^2(z_{LRT})$  and  $z_{OT} < z_{LRT}$ ; in this way the characteristically higher static stability and ozone concentration of stratospheric intrusion is regarded as being retained as it penetrates below the lapse rate tropopause. The seasonal distributions shown for the three stations in Fig. 4.6 are generally similar to those shown in Fig. 4.5 (although detected events are less frequent), with the main exception that very few events are identified with the alternative method at Davis in the first half of the year. For our STT proxy, we only detect intrusions where the lowest altitude of the intrusion satisfies the ozone tropopause definition. During summer and autumn, the vertical ozone gradients at Davis are weaker compared with the other seasons, and the detected ozone tropopause tends to lie above the lapse rate tropopause potentially reducing the ability to identify STT events based on the definition of our proxy.

Figure 4.7 shows the altitudes of detected events, based on the altitude of peak tropospheric ozone (local maximum ozone within enhancement altitude) in the ozonesonde profile. STT event peaks most commonly occur at 6–11 km above Melbourne and anywhere from 4–9 km at Davis and Macquarie Island. There is no clear relationship between meteorological conditions and event altitude, which may reflect the fact that the ozonesondes observe a snapshot of an event at different stages of its life cycle.

Figure 4.8 shows the distance from the event peak to the ozone defined tropopause, referred to as event depth. The majority of STT events occur within 2.5 km of the tropopause at Davis and Macquarie Island. Over Melbourne, the event depth is more spread out, with peak ozone enhancement generally occurring up to 6 km below the

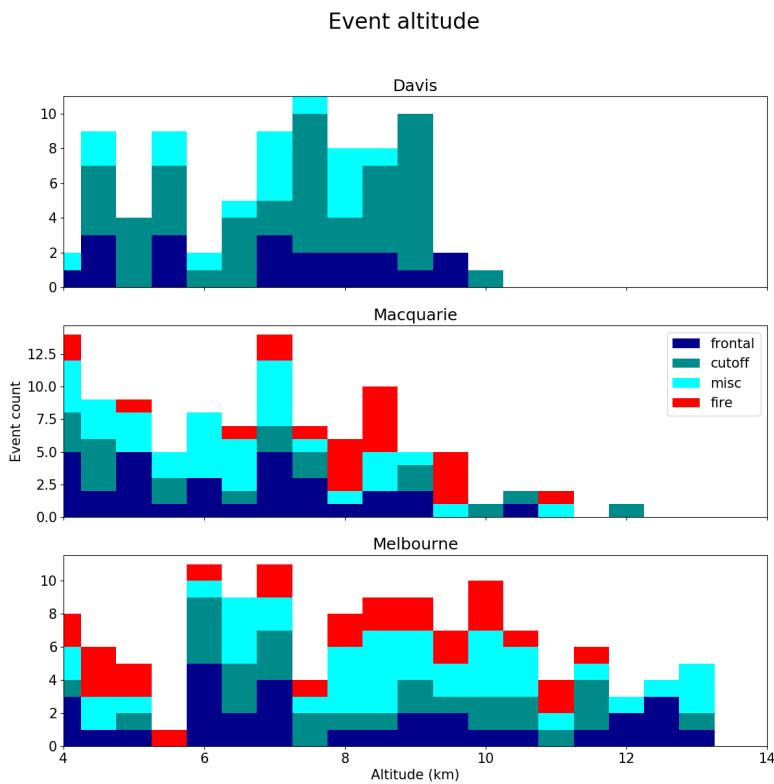


FIGURE 4.7: The distribution of STT events' altitudes at Davis (top), Macquarie Island (middle), and Melbourne (bottom), determined as described in the text. Events are coloured as described in Fig. 4.5.

tropopause. Again, there is no clear relationships between meteorological conditions and event depth.

#### 4.4 Simulated ozone columns

Figure 4.9 compares the time series of tropospheric ozone columns ( $\Omega_{O_3}$ ) in molecules  $\text{cm}^{-2}$  simulated by GEOS-Chem (red) to the measured tropospheric ozone columns (black). GEOS-Chem outputs ozone density ( $\text{molecules cm}^{-3}$ ), and height of each simulated box, as well as which level contains the tropopause, allowing modelled  $\Omega_{O_3}$  to be calculated as the product of density and height summed up to the box below the tropopause level. In both observations and model, the maximum ozone column at Melbourne occurs in austral summer, with a minimum in winter, while Macquarie Island and Davis show the opposite seasonality.

GEOS-Chem provides a reasonable simulation of the observed seasonality and magnitude of  $\Omega_{O_3}$ . Reduced major axis regression of observed versus simulated  $\Omega_{O_3}$  gives a line of best fit with slopes of 1.08 for Davis, 0.99 for Macquarie Island, and 1.34 for Melbourne. The model is only partially able to reproduce the variability in the observations, with paired  $r^2$  values of 0.38 for Davis, 0.18 for Macquarie Island, and

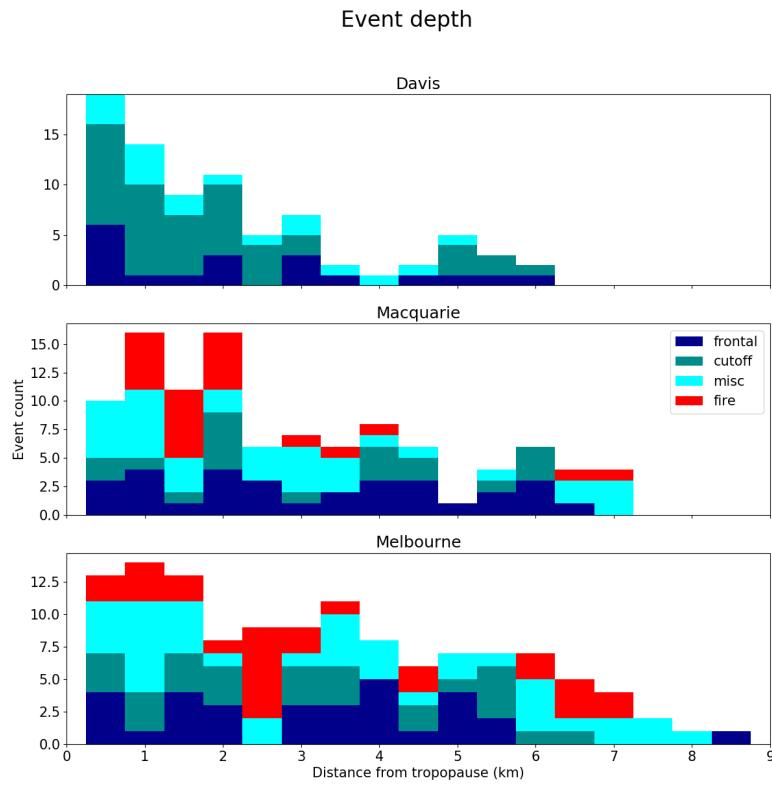


FIGURE 4.8: The distribution of STT events' depths, defined as the distance from the event to the tropopause, at Davis (top), Macquarie Island (middle), and Melbourne (bottom), determined as described in the text.

Events are coloured as described in Fig. 4.5.

0.37 for Melbourne. Much of the variability is driven by the seasonal cycle, and after removing this effect (by subtracting the multi-year monthly means), the  $r^2$  values decrease to 0.07, 0.11, and 0.30 respectively, although the slope improves at Melbourne to 1.08.

Figure 4.10 shows the observed and simulated ozone profiles at all sites, averaged seasonally. The model generally underestimates ozone in the lower troposphere (up to 6 km) over Davis, although this bias is less pronounced during summer. Over Melbourne, ozone in the lower troposphere is well represented, but the model overestimates ozone from around 4 km to the tropopause. Over Macquarie Island we see model overestimation of ozone above 4 km, as well as underestimated ozone in the lower troposphere, suggesting that this region is influenced by processes seen at both of our other sites. Also shown is the mean tropopause height simulated by the model (horizontal dashed red line), which is always higher than the observed average, although this difference is not statistically significant. The effect of local pollution over Melbourne during austral summer (DJF) can be seen from the increased mean mixing ratios and enhanced variance near the surface. The gradient of the  $O_3$  profiles is steeper in the measurements than the model, at all sites during all seasons. Recently

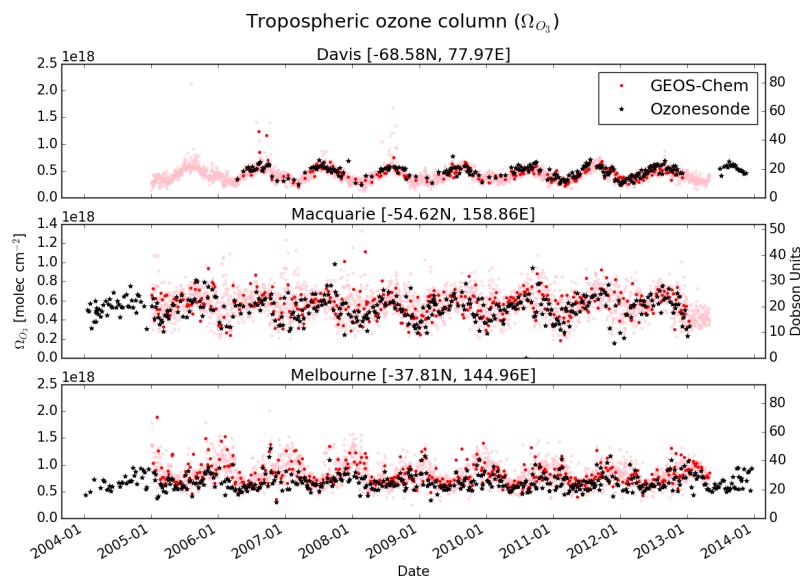


FIGURE 4.9: Comparison between observed (black) and simulated (pink, red) tropospheric ozone columns ( $\Omega_{O_3}$ , in molecules  $\text{cm}^{-2}$ ) from 1 January 2004 to 30 April 2013. For the model, daily output is shown in pink, while output from days with ozonesonde measurements are shown in red. For each site, the model has been sampled in the relevant grid square.

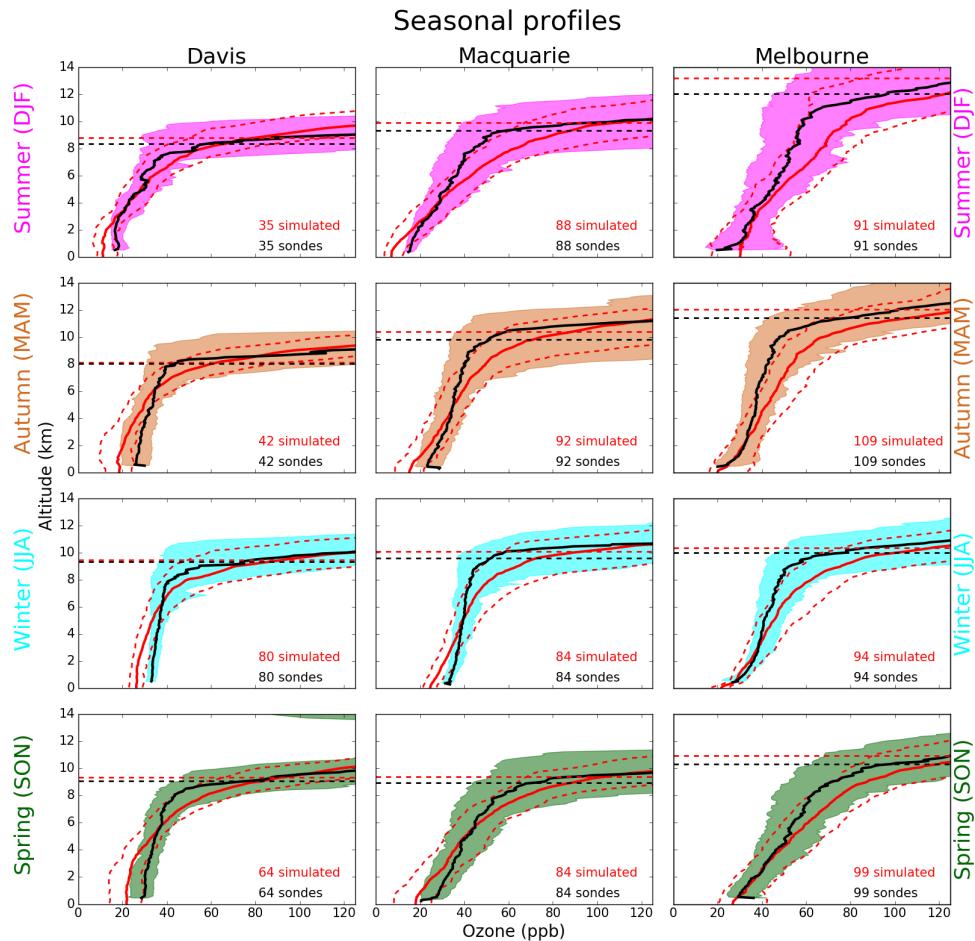


FIGURE 4.10: Observed and simulated tropospheric ozone profiles over Davis, Macquarie Island, and Melbourne, averaged seasonally. Model medians (2005-2013 average) are shown as red solid lines, with red dashed lines showing the 10th and 90th percentiles. Ozonesonde medians (over each season, for all years) are shown as black solid lines, with coloured shaded areas showing the 10th and 90th percentiles. The horizontal dashed lines show the median tropopause heights from the model (red) and the observations (black).

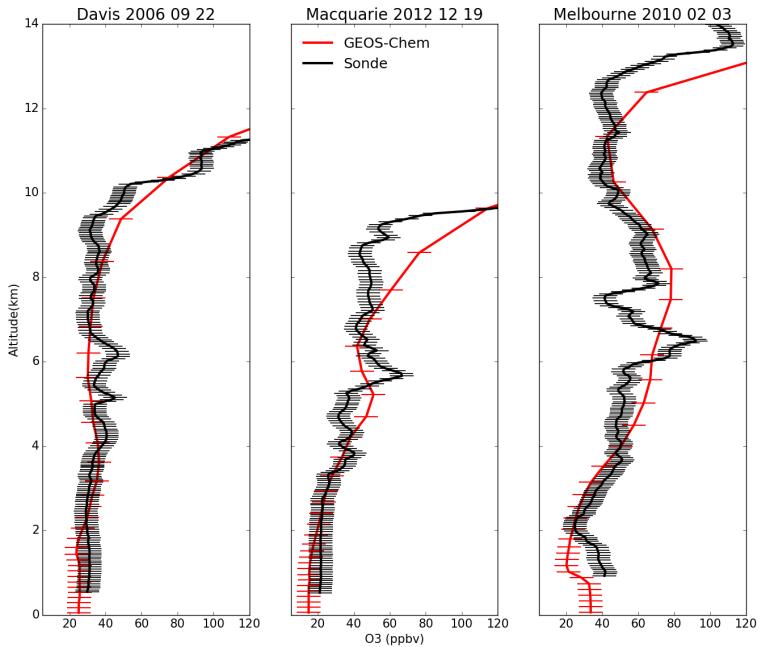


FIGURE 4.11: Example comparisons of ozone profiles from ozonesondes (black) and GEOS-Chem (red) from three different dates during which STT events were detected from the measurements. The dates were picked based on subjective visual analysis. The examples show the best match between model and observations for each site. GEOS-Chem and ozonesonde pressure levels are marked with red and black dashes respectively.

Hu et al. 2017 examined GEOS-Chem ozone simulations and found a similar overestimation of upper troposphere ozone in the mid southern latitudes when using the GEOS5 meteorological fields.

Figure 4.11 compares modeled (red) and observed (black) ozone profiles on three example days when STT events were detected using the ozonesondes. The figures show the profile for each site with the closest (qualitative) match between model and observations. The resolution (both vertical and horizontal) of GEOS-Chem in the upper troposphere is too low to consistently allow detection of STTs, although in a few cases (e.g., Melbourne in Fig. 4.11) it appears that the event was large enough to be visible in the model output.

## 4.5 Stratosphere-to-troposphere ozone flux from STT events

### 4.5.1 Method

We quantify the mean stratosphere-to-troposphere ozone flux due to STT events at each site based on the integrated ozone amount associated with each STT event (see Sect. 4.2.3). Events that may have been influenced by transported biomass burning

are excluded from this calculation. Our estimate provides a preliminary estimate of how much ozone is transported from the stratosphere by the events detected by our method. The estimate is conservative for several reasons: it ignores secondary ozone peaks which may also be transported from the stratosphere, it ignores potential ozone enhancements which may have dispersed and increased the local background mixing ratio, and any influence from STT events nearby which may also increase the local background ozone.

Observed tropospheric columns are calculated from the ozonesondes by calculating the ozone number density (molecules cm<sup>-3</sup>) using measured ozone partial pressure ( $P_{O_3}$ ) and integrating vertically up to the tropopause:

$$\Omega_{O_3} = \int_0^{TP} \frac{P_{O_3}(z)}{k_B \times T(z)} dz$$

where  $z$  is the altitude (GPH),  $TP$  is the altitude at the tropopause,  $T$  is the temperature, and  $k_B$  is the Boltzmann constant.

Three regions are used to examine possible STT flux over a larger area using modeled tropospheric ozone concentrations. The regions are shown in Fig. 4.1. The regions are centred at each site, plus or minus ten degrees latitude, and plus or minus 25, 16, and 11 degrees longitude for Davis, Macquarie Island, and Melbourne respectively. These boundaries approximate a rectangle centred at each site with  $\sim 2000$  km side lengths, covering  $\sim 4.4$ ,  $4.6$ , and  $4.8$  million square km, for Davis, Macquarie Island, and Melbourne respectively.

To determine the ozone column attributable to STT, we determine monthly averaged STT impact ( $I$ ; fraction of tropospheric ozone sourced from the stratosphere as shown above) and the monthly mean tropospheric ozone column (from the GEOS-Chem multi-year mean,  $\Omega_{O_3}$ ) over the regions described above. This can be expressed simply as the STT flux per event (flux<sub>*i*</sub> in each month:  $\text{flux}_i = \Omega_{O_3} \times I$ ). Next we determine how many events are occurring per month by assuming only one event can occur at one time, and that no event is measured twice. These assumptions allow a simple estimate of events per month from the relatively sparse dataset and should hold true as long as our regions of extrapolation are not too large. The ( $P$ )robability of any sonde launch detecting an event is calculated as the fraction of ozonesonde releases for which an STT event was detected, calculated for each month. We assume events last  $N$  days, then find how many events per month we expect by multiplying the days in a month by  $P$  and dividing by this assumed event lifetime. For example if we expect to see an event 25% of the time in a month, and events last one day, we expect one event every four days ( $\sim 7.5$  events in that month) whereas if we expect events to last a week then we would expect  $\sim$ one event in that month. This leads us to multiply our  $\text{flux}_i$  by  $P$ , and then by the term  $M$  ( $M = \frac{\text{days per month}}{N}$ ) determined by our assumed event lifetime in order to determine monthly STT ozone flux.

The longevity of ozone events is very difficult to determine, and we have chosen 2 days as a representative number based on several examples in Lin et al. 2012 where intrusions were seen to last from 1-3 days (occasionally longer) and an analysis of one large event by Cooper et al. 2004 showing that most of the ozone had dispersed after 48 hours. Worth noting is the recent work of Trickl et al. 2014, where intrusions are detected  $> 2$  days and thousands of kilometres away from their initial descent into

the troposphere over Greenland or the Arctic. In those regions with high wind shear, mixing appears to be slower, which allows ozone intrusions to be transported further without dissipating into the troposphere. Relative uncertainty in our  $M$  term is set to 50%, as we assume these synoptic events to generally last from 1-3 days.

### 4.5.2 Results

The top panel of Fig. 4.12 shows the STT ozone enhancements, based on a vertical integration of the ozone above baseline levels for each ozonesonde where an event was detected. The area considered to be ‘enhanced’ ozone is outlined with yellow dashes on the left panel of Fig. 4.4. We find that the mean ozone flux associated with STT events is  $\sim 0.5\text{--}2.0 \times 10^{16}$  molecules  $\text{cm}^{-2}$ . The bottom panel shows the mean fraction of total tropospheric column ozone (calculated from ozonesonde profiles) attributed to stratospheric ozone intrusions at each site for days when an STT event occurred. First the tropospheric ozone column is calculated, then the enhanced ozone column amount is used to determine the relative increase. At all sites, the mean fraction of tropospheric ozone attributed to STT events is  $\sim 1.0\text{--}3.5\%$ . On three separate days over Macquarie and Melbourne, this value exceeds 10%.

The upper panels in figures 4.13-4.15 show the factors  $I$ ,  $P$ , and  $\Omega_{O_3}$  which are used along with the assumed event lifetime to estimate the STT flux. The tropospheric ozone and area of our region is calculated using the output and surface area from GEOS-Chem over our three regions. The lower panel of these figures show the results of the calculation when we choose two days for our flux estimation, with dotted lines showing the range of flux calculated if we assume events last from one day to one week. The seasonal cycle of ozone flux is mostly driven by the  $P$  term, which peaks in the SH summer over all three sites. Total uncertainty (shaded) is on the order of 100% (see Sect. 4.6.2). We calculate the annual amount based on the sum of the monthly values. The regions over Davis, Macquarie Island, and Melbourne have estimated STT ozone contributions of  $\sim 5.7 \times 10^{17}$ ,  $\sim 5.7 \times 10^{17}$ , and  $\sim 8.7 \times 10^{17}$  molecules  $\text{cm}^{-2} \text{ a}^{-1}$  respectively, or equivalently  $\sim 2.0, 2.1$ , and  $3.3 \text{ Tg a}^{-1}$ .

### 4.5.3 Comparison to literature

Škerlak, Sprenger, and Wernli 2014 show an estimate of roughly 40 to 150  $\text{kg km}^{-2} \text{ month}^{-1}$  in these regions, over all seasons (see Fig. 16, 17 in their publication) while we estimate from 0 to 180  $\text{kg km}^{-2} \text{ month}^{-1}$  STT impact, following a seasonal cycle with the maximum in austral summer. We estimate higher maximum flux over Melbourne, (178, and 150  $\text{kg km}^{-2} \text{ month}^{-1}$  in January and February) than in either Davis (89  $\text{kg km}^{-2} \text{ month}^{-1}$  in March) or Macquarie Island (68  $\text{kg km}^{-2} \text{ month}^{-1}$  in January). Our calculated seasonal contributions, along with total uncertainty are shown in Table 4.3.

This result disagrees with several other studies which have found STT ozone fluxes in the SH extra-tropics are largest from autumn or winter to early spring. Roelofs and Lelieveld 1997 used a model carrying a tracer for stratospheric ozone to estimate STT impacts. They see higher SH tropospheric ozone concentrations, as well as STT flux, in the SH winter. Our model also shows ozone column amounts peaking in winter, however flux is maximised in summer due to our detected event frequencies. Elbern,

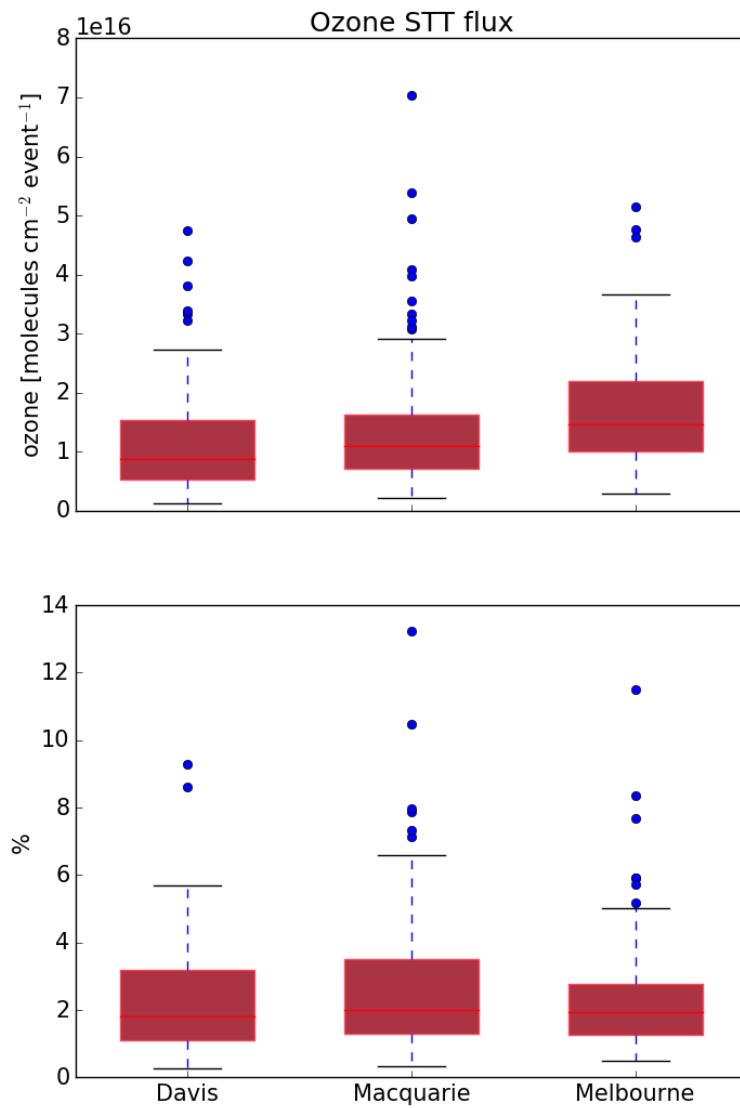


FIGURE 4.12: Top panel: tropospheric ozone attributed to STT events. Bottom panel: percent of total tropospheric column ozone attributed to STT events. Boxes show the inter-quartile range (IQR), with the centre line being the median, whiskers show the minimum and maximum, circles show values which lie more than 1.5 IQR from the median. Values calculated from ozonesonde measurements as described in the text.

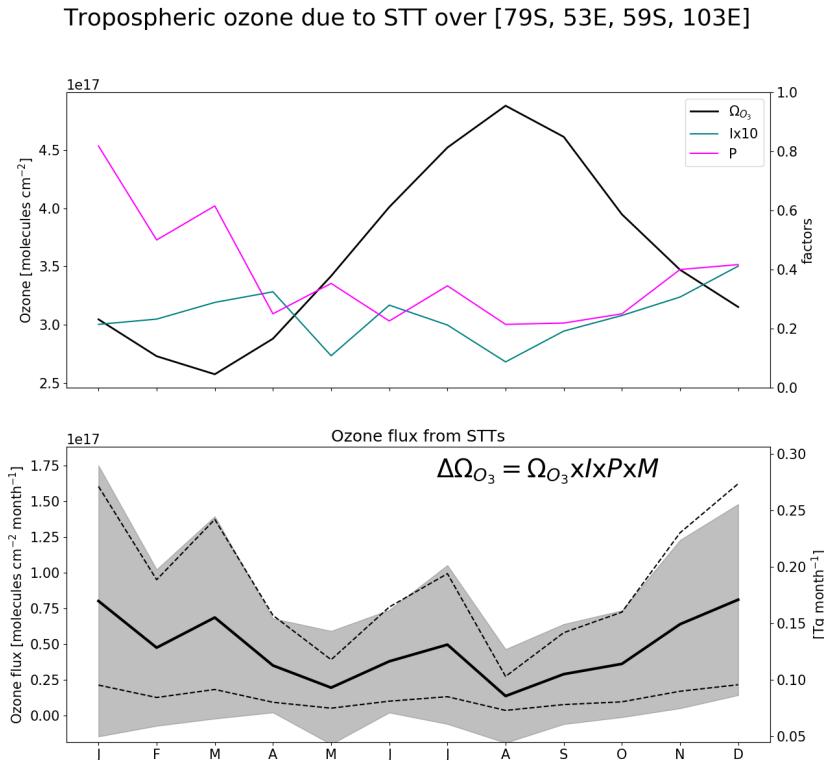


FIGURE 4.13: (Top) Tropospheric ozone, ( $I$ )impact per event, and ( $P$ )probability of event detection per sonde launch, averaged over the region above Davis. The tropospheric ozone column  $\Omega_{O_3}$  (black, left axis) is from GEOS-Chem, while the STT probability  $P$ (magenta, right axis) and impact  $I$  (teal, right axis) are from the ozonesonde measurements. The STT impact is multiplied by ten to better show the seasonality. (Bottom) Estimated contribution of STT to tropospheric ozone columns over the region, with uncertainty (shaded area) estimated as outlined in Sect. 4.6. The black line shows STT ozone flux if event lifetime is assumed to be two days, with dashed lines showing the range of flux estimation if we assumed events lasted from one day to one week.

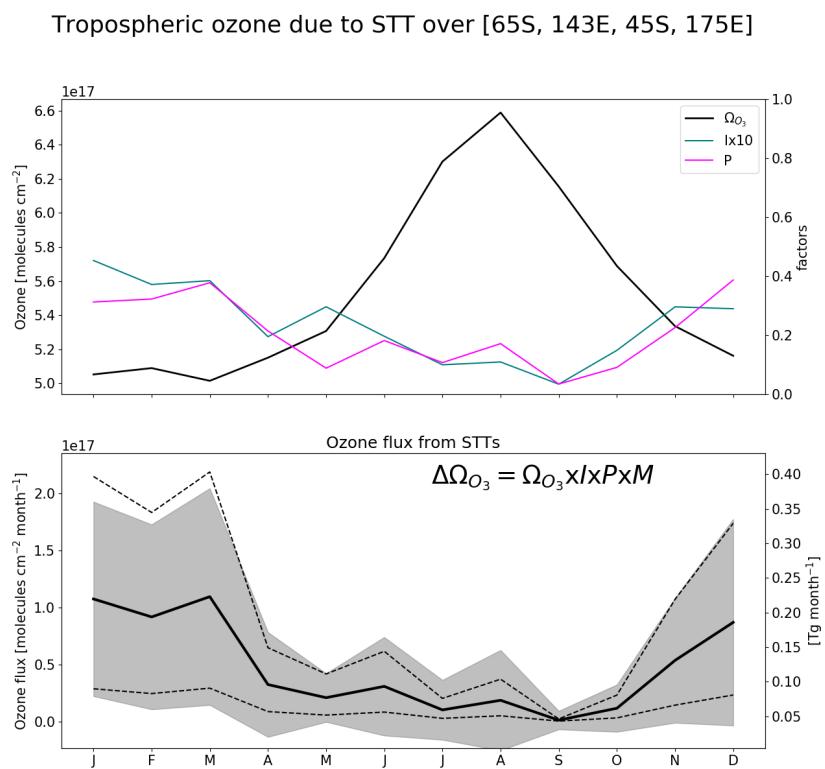


FIGURE 4.14: As described in 4.13, for the region containing Macquarie Island.

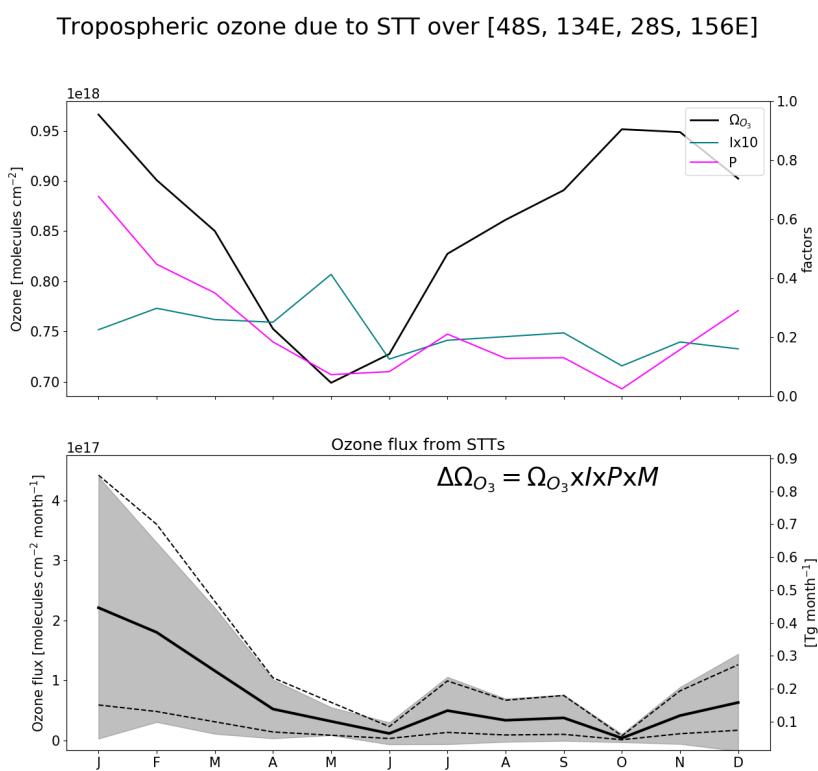


FIGURE 4.15: As described in 4.13, for the region containing Melbourne.

TABLE 4.3: Seasonal STT ozone contribution in the regions near each site, in  $\text{kg km}^{-2} \text{ month}^{-1}$ . In parentheses are the relative uncertainties.

Region	DJF	MAM	JJA	SON
Davis	54.5 (102%)	47.7 ( 97%)	30.7 (114%)	18.8 (127%)
Macquarie Island	61.3 ( 85%)	70.7 ( 91%)	17.9 (139%)	7.7 (229%)
Melbourne	96.7 (103%)	88.6 ( 89%)	26.7 (102%)	21.4 (109%)

Hendricks, and Ebel 1998 examine STT using ECMWF data for prior to 1996, using PV and Q-vectors to determine STT frequency and strength, and suggest fewer fold events in the SH occur from December to February. Olsen (2003) used PV and winds from the GEOS reanalysis combined with ozone measurements from the TOMS satellite to estimate that the ozone flux between  $30^\circ$  S and  $60^\circ$  S is  $210 \text{ Tg yr}^{-1}$ , with the maximum occurring over SH winter. Liu et al. (2017) model the upper tropospheric ozone and its source (emissions/lightning/stratospheric) over the Atlantic ocean between  $30^\circ$  S and  $45^\circ$  S, and suggest that most of this is transported from the stratosphere from March to September, which is when the subtropical jet system is strongest.

The disagreements largely reflect the difference between point source based estimates and zonally averaged estimates, as the meteorological behaviour at our three sites is not the same as the system that dominates the southern hemisphere in general. As detailed in Sect. 4.3, the maximum STT influx which occurs during SH winter is almost entirely due to the dominant STT system which occurs annually over the southern Indian ocean and middle of Australia. It is difficult to compare remote ozonesonde datasets with area averaged models or re-analyses based on non-co-located measurements (such as ERA).

## 4.6 Sensitivities and limitations

### 4.6.1 Event detection

Our method uses several subjectively-defined quantities in the process of STT event detection. Here we briefly discuss these quantities and the sensitivity of the method to each. Using the algorithm discussed in Sect. 4.2.3, we detect 80 events at Davis, 105 (21 fire influenced) events at Macquarie Island, and 127 (27 fire influenced) events at Melbourne.

The cut-off threshold (defined separately for each site) is determined from the 95th percentile of the ozone perturbation profiles between 2 km above the earth's surface and 1 km below the tropopause. We use the 95th percentile because at this point the filter locates clear events with fewer than 5% obvious false positive detections. Event detection is sensitive to this choice; for example, using the 96th, and 97th percentile instead decreased detected events by 2, 9 (2,10%) at Davis, 13, 31 (11, 28%) at Macquarie Island, and 8, 24 (6, 18%) at Melbourne. Event detection is therefore also sensitive to the range over which the percentile is calculated. This range was chosen to remove anomalous edge effects of the Fourier bandpass filter and to discount the highly variable ozone concentration which occurs near the tropopause.

Ozone enhancements are only considered STT events if they occur from 4 km altitude up to 500 m below the tropopause. This range removes possible ground pollution and events not sufficiently separated from the stratosphere, while still capturing many well-defined events that occur within 1 km of the tropopause. An example of a well-defined event that occurs within 1 km of the tropopause is shown in the supplementary (Fig. S2). However, STT events which reach below 4 km are physically possible and we may have some false negative detections due to the altitude restricted detections.

#### 4.6.2 Flux calculations

Flux is calculated as  $I \times P \times M \times \Omega_{O_3}$ , with each term calculated as described in Sect. 4.5.1. The uncertainty is determined using the standard deviation of the product, with variance calculated using the variance of a product formula, assuming that each of our terms is independent:

$$\text{var}(\Pi_i X_i) = \Pi_i (\text{var}(X_i) + E(X_i)^2) - (\Pi_i E(X_i))^2$$

The standard deviations for the  $I$  and  $\Omega_{O_3}$  terms are calculated over the entire dataset. These terms are considered to be homoskedastic (unchanging variance over time). Uncertainty in assumed event lifetime is set at 50%, as we believe it is reasonable to expect events to last 1-3 days.  $P$  is the probability of any ozonesonde detecting an event, and is assumed to be constant (for any month). The overall uncertainty as a percentage is shown in parentheses in Table 4.3, these values are on the order of 100%, largely due to relative uncertainty in the  $I$  factor which ranges from 50-120% for each month.

Small changes in the region don't have a large affect on the per area flux calculations: increasing or decreasing the regions by  $1^\circ$  on each side ( $\sim 10\%$  change in area) change the resulting flux by  $\sim 1\%$ . However due to the large portion of winter STT events being flagged due to potential smoke plume influence, a significant change in the yearly flux is seen when we don't remove these events. Without removing smoke flagged events we see an increase in estimated yearly flux of  $\sim 1.1, 2.1 \times 10^{17}$  molecules  $\text{cm}^{-2} \text{ yr}^{-1}$  (which is a change of  $\sim 15, 20\%$ ), over Macquarie Island and Melbourne respectively.

Considering the  $I$  factor, as discussed in here and in Sect. 4.6, there are several uncertainties in our method that are likely to lead to a low bias, such as the conservative estimate of flux within each event. Although there is little available data on SH ozone events for us to compare against, consider Terao et al. (2008), who estimated that up to 30–40% of the ozone at 500 hPa was transported from the stratosphere, in the northern hemisphere.

Our STT event impact estimates have some sensitivity to our biomass burning filter: including smoke-influenced days increases the mean per area flux by 15-20%. Although events which are detected near fire smoke plumes are removed, some portion of these could be actual STTs. The change in our  $P$  parameter when we include potentially smoke influenced events leads to a yearly estimated STT of  $11 \times 10^{17}$  molecules  $\text{cm}^{-2} \text{ yr}^{-1}$  over Melbourne, which suggests that up to  $2.1 \times 10^{17}$  molecules  $\text{cm}^{-2} \text{ yr}^{-1}$  ozone enhancement could be caused by smoke plume transported precursors. This is

a potential area for improvement, as a better method of determining smoke influenced columns would improve confidence in our estimate.

Other possibly important uncertainties in our calculation of STT flux which we don't cover are listed here. Filtering events which occur within 500 m of the tropopause may also lead to more false negatives. This could also cause lower impact estimates due to only measuring ozone enhancements which have descended and potentially slightly dissipated. On the other hand we have no measure of how often the detached ozone intrusion reascends into the stratosphere, which would lead to a reduced stratospheric impact. The estimated tropospheric ozone columns modelled by GEOS-Chem may be biased, for instance Hu et al. 2017 suggest that in general GEOS-Chem (with GEOS-5 met. fields) underestimates STT, with  $\sim 360 \text{ Tg a}^{-1}$  simulated globally, compared to  $\sim 550 \text{ Tg a}^{-1}$  observationally constrained. Transport uncertainty is very difficult to estimate with the disparate point measurements; it's possible that detected events are (at least partially) advected out of the analysis regions, which would mean we overestimate the influx into the region, and it is also possible that we are influenced by STT events outside the regions of analysis. Uncertainty in event longevity is set to 50%, however this implies a very simplistic model of event lifetimes. A great deal of work could be done to properly model the regional event lifetimes, however this is beyond the scope of our work.

Uncertainties in STT ozone flux detection are ( $\sim 100\%$ ), and could be directly improved with larger or longer datasets. Possible parameterisations and an improved model of event lifetime could also improve the confidence in our estimate of event impacts, as well as allowing fewer assumptions.

## 4.7 Conclusions

Stratosphere-to-troposphere transport (STT) can be a major source of ozone to the remote free troposphere, but the occurrence and influence of STT events remains poorly quantified in the southern extra-tropics. Ozonesonde observations in the SH provide a satellite-independent quantification of the frequency of STT events, as well as an estimate of their impact and source. Using almost ten years of ozonesonde profiles over the southern high latitudes, we have quantified the frequency, seasonality, and altitude distributions of STT events in the SH extra-tropics. By combining this information with ozone column estimates from a global chemical transport model, we provided a first, conservative estimate of the influence of STT events on tropospheric ozone over the Southern Ocean.

Our method involved applying a bandpass filter to the measured ozone profiles to determine STT event occurrence and strength. The filter removed seasonal influences and allowed clear detection of ozone-enhanced tongues of air in the troposphere. By setting empirically-derived thresholds, this method clearly distinguished tropospheric ozone enhancements that are separated from the stratosphere. Our method is sensitive to various parameters involved in the calculation; however, for our sites we saw few false positive detections of STT events.

Detected STT events at three sites spanning the SH extra-tropics ( $38^\circ\text{S}$ ,  $55^\circ\text{S}$ , and  $69^\circ\text{S}$ ) showed a distinct seasonal cycle. All three sites displayed a summer (DJF) maximum and an autumn to winter (AMJJA) minimum, although the seasonal amplitude

was less apparent at the Antarctic site (Davis) as events were also detected regularly in winter and spring (likely due to polar jet stream-caused turbulence). Analysis of ERA-Interim reanalysis data suggested the majority of events were caused by turbulent weather in the upper troposphere due to low pressure fronts, followed by cut-off low pressure systems. Comparison of ozonesonde-measured ozone profiles against those simulated by the GEOS-Chem global chemical transport model showed the model is able to reproduce seasonal features but does not have sufficient vertical resolution to distinguish STT events.

By combining the simulated tropospheric column ozone from GEOS-Chem with ozonesonde-derived STT estimates, we provide a first estimate of the total contribution of STT events to tropospheric ozone in these southern extra-tropical regions. We estimate that the ozone enhancement due to STT events near our three sites ranges from 300–570 kg km<sup>-2</sup> a<sup>-1</sup>, with seasonal maximum in SH summer.

Estimating STT flux using ozonesonde data alone remains challenging; however, the very high vertical resolution provided by ozonesondes suggests they are capable of detecting STT events that models, re-analyses, and satellites may not. Further work is needed to more accurately translate these ozonesonde measurements into STT ozone fluxes, particularly in the SH where data are sparse and STT is likely to be a major contributor to upper tropospheric ozone in some regions. More frequent ozonesonde releases at SH sites could facilitate development of better STT flux estimates for this region.

## 4.8 Contributions and Acknowledgements

JWG wrote the algorithms, ran the GEOS-Chem simulations, performed the analysis and led the writing of the paper under the supervision and guidance of SPA, RS, and JAF. AK contributed the Davis ozonesonde data and performed the analysis of the alternate STT proxy. All authors contributed to editing and revising the manuscript.

*Data availability.* All GEOS-Chem model output and the ozonesonde observational data are available from the authors upon request.

We thank Dr. Sandy Burden for help clarifying some of the uncertainties involved in methods within this work. We also thank Dr. Clare Paton-Walsh, who identified the need to account for smoke-influenced events, and provided discussions on how to go about doing such. Ozonesonde data comes from the World Ozone and Ultraviolet Data Centre (WOUDC). The ERA-Interim data were downloaded from the ECMWF website following registration. This research was undertaken with the assistance of resources provided at the NCI National Facility systems at the Australian National University through the National Computational Merit Allocation Scheme supported by the Australian Government. This work was supported through funding by the Australian Government's Australian Antarctic science grant program (FoRCES 4012), the Australian Research Council's Centre of Excellence for Climate System Science (CE110001028), the Commonwealth Department of the Environment ozone summer scholar program. This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

# Appendix A

# Supplementary Notes

## A.1 Measurement Techniques

### A.1.1 MAX-DOAS

Multiple axis DOAS (MAX-DOAS) is a remote sensing technique which uses several DOAS measurements over different viewing paths. In these retrievals, the measurements of light absorption are performed over several elevations in order to add some vertical resolution to the measurement of trace gas concentrations. An example of this is shown in figure A.1, which was taken from Lee et al. (2015). Recently MAX-DOAS has been used to examine HCHO profiles in the clean free troposphere (Franco et al. (2015) and Schreier et al. (2016)) as well as in polluted city air (Lee et al. (2015)). Depending on orography and atmospheric composition (ie. the influence of interfering chemicals), MAX-DOAS can be used to split the tropospheric column into two partial columns; giving a small amount of vertical resolution to HCHO measurements (Franco et al. 2015; Lee et al. 2015, eg.). In Franco et al. (2015), an FTIR spectrometer at Jungfraujoch is compared against both MAX-DOAS and satellite data, with two CTMs; GEOS-Chem and IMAGES v2 used to compare total columns and vertical resolution of each instrument.

## A.2 Data sets

### A.2.1 SPEI

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°.

The Standardised Precipitation Evapotranspiration Index (SPEI) is a measure of drought using TODO *SPEI Drought Index*. This product covers 1901 - 2011, and uses the average over that period as the background, in order to compare drought stressed regions against those with sufficient or excess water *SPEI Drought Index*.

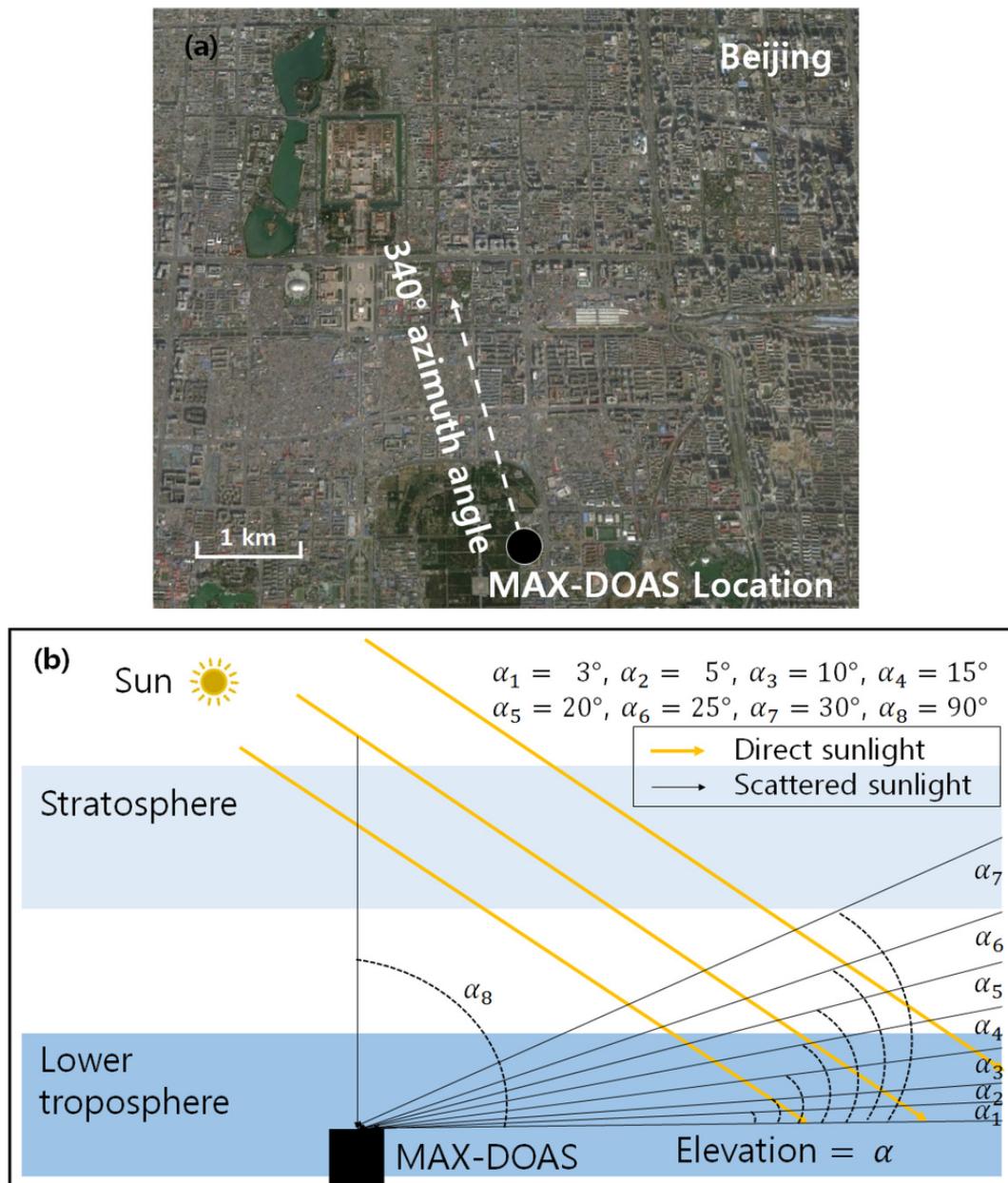


FIGURE A.1: Image from Lee et al. (2015).

### A.2.2 GOME

GOME suffers from similar uncertainties to OMI, as the same general method of DOAS remote measurements are performed. The uncertainty from slant column fitting has been calculated for GOME to be  $4 \times 10^{15}$  molecules cm $^{-2}$  (Chance et al. 2000; Millet et al. 2006). The conversion factor for slant to vertical columns (AMF) calculation also suffers from errors; primarily from surface albedo, HCHO vertical profile a priori, aerosol, and cloud influence (Millet et al. 2006). AMF uncertainties for GOME are calculated to be 1 to  $1.3 \times 10^{15}$  molecules cm $^{-2}$  by Shim et al. (2005).

### A.2.3 NPI

One possible solution to anthropogenic filtering is the national pollution index (TODO:cite:<http://>) which contains industrial HCHO and NO<sub>X</sub> emissions from 2003 to 2014.

## A.3 Chemistry

PATCHEN et al. (2007) examine the branching step where isoprene either forms ROO or RONO<sub>2</sub>, and for specific conditions they determine the reaction rates for each branch. They find the most frequent pathway is the formation of ROO (99.3%). Although the nitrates formation is relatively infrequent, this pathway can lead to NO<sub>X</sub> transport into clean environments (Horowitz et al. 1998). This transport can be exacerbated by fast winds and low OH concentrations, making nitrates an important factor in modelling atmospheric chemistry.

PAN has a relatively long lifetime (against OH, order of 1 day) and is able to transport and release the NO<sub>X</sub> in environments which are quite far from any emissions.

### A.3.0.1 SOA

SOA formation from VOCs in atmospheric CTMs is generally imperfect due to the complicated chemistry and diverse nature of atmospheric conditions. Yields of SOA from VOCs are often lumped together and based on empirical laboratory chamber data. VOC oxidation was not feasible  $\sim$  13 years ago (2005), as chamber studies did not extend over a large enough parameter range and the importance of heterogeneous aerosol chemistry on SOA formation was unquantified (Kanakidou et al. 2005). Monoterpene oxidation by O<sub>3</sub>, OH and NO<sub>3</sub> radicals may also form aerosols, with the reaction with ozone forming the most particles (Kanakidou et al. 2005).

Gas phase emissions with higher vapour pressures can be oxidised into lower vapour pressure products which will partition between gas and particle phase, often called semi or non-volatile. The aerosol products from these gas phase emissions (or the children thereof) are called SOA (Kanakidou et al. 2005). In the Kanakidou et al. 2005 review of global SOA science, uncertainty in radiative forcing of aerosols is highlighted, and 20-90 % of PM mass in the lower troposphere is OA. Less volatile OA also plays a role, although PM production from this source is complicated and makes up only a small fraction ( $\sim$  1%) of the resulting PM (Kroll and Seinfeld 2008; Bei, Li, and Molina 2012). Modelling OA has many uncertainties due to the complexity of SOA formation and various pathways such as aqueous phase oxidation which

can significantly contribute to concentrations. This is further hindered by poor understanding of precursor emissions, and lumping together various compounds, of which only some form SOA (for example ORVOCs in GEIA (back in 2005)). Satellite data requires SOA models to estimate a full vertical profile of aerosols for remote sensing techniques (Kanakidou et al. 2005).

SOA formation from VOCs in atmospheric CTMs is generally imperfect due to the complicated chemistry and diverse nature of atmospheric conditions. Yields of SOA from VOCs are often lumped together and based on empirical laboratory chamber data. VOC oxidation was not feasible  $\sim$  13 years ago (2005), as chamber studies did not extend over a large enough parameter range and the importance of heterogeneous aerosol chemistry on SOA formation was unquantified (Kanakidou et al. 2005).

One of the large uncertainties with OA is the total effect on radiative forcing, 12 years ago it was well understood that most OA cool the atmosphere, with smaller particles having a larger affect due to the size matching the wavelengths of visible light (Kanakidou et al. 2005). Transport and indirect effects complicate matters further, with cloud creation and modification of cloud properties being quite difficult to accurately predict. In the third IPCC report (*Intergovernmental Panel on Climate Change (IPCC): Climate Change: The Scientific Basis* 2001), the uncertainty involved if OA forcing was a factor of 3 times the estimated effect. This has since been improved however OA and cloud formation still remains a large uncertainty in more recent IPCC reports (Forster et al. 2007). Figure A.2 shows the radiative forcing (RF) of various atmospheric constituents, it's clear that OA uncertainty dominates. Figure A.3 shows the same summary updated in chapter 8 of the fifth report, where the SOA uncertainty remains quite large. It's currently understood that SOA plays an indirect and complex role in cloud properties, with a net cooling effect (*IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Chapter 7,8)

(TODO: read more of Kanakidou 2005) The emissions of precursors to SOA was and is quite uncertain, in Kanakidou et al. 2005 they state that these uncertainties range from a factor of 2 to 5. They highlight emissions and flux measurements as well as implementing satellite data in models as a means of improving the emissions inventories. In 2005, (as of Kanakidou et al. 2005,) the knowledge gaps in isoprene and terpene oxidation processes included precursor gases to SOA, impact of NO<sub>x</sub> on SOA formation, heterogeneous reactions between particles and gaseous compounds, aqueous phase chemistry, and complete aerosol compositions. At this time SOA driven nucleation was under debate, as chamber studies showed that SOA led to new particles but only in the particle free laboratory setting. Nucleation of new particles was suppressed by condensation if any seed aerosol was already present. Observed nucleation outside of laboratories was suggested to have arisen from biogenic SOAs, driven by ozonolysis. Kanakidou et al. 2005 concluded that it is very likely that organics contribute to particle growth and formation rates.

Rollins et al. 2009 examine SOA production in a large chemical reaction chamber, over 16 hr in the dark and find first generation mass yield ( $\Delta$ SOA mass/ $\Delta$ isoprene mass) to be less than 0.7%, with further oxidation of initial products (isoprene reacting twice with NO<sub>3</sub>) yielding 14%. This led to an overall mass yield of 2% over the 16 hr experiment. Night NO<sub>3</sub> levels also affect O<sub>3</sub>, TODO: millet et al 2016.

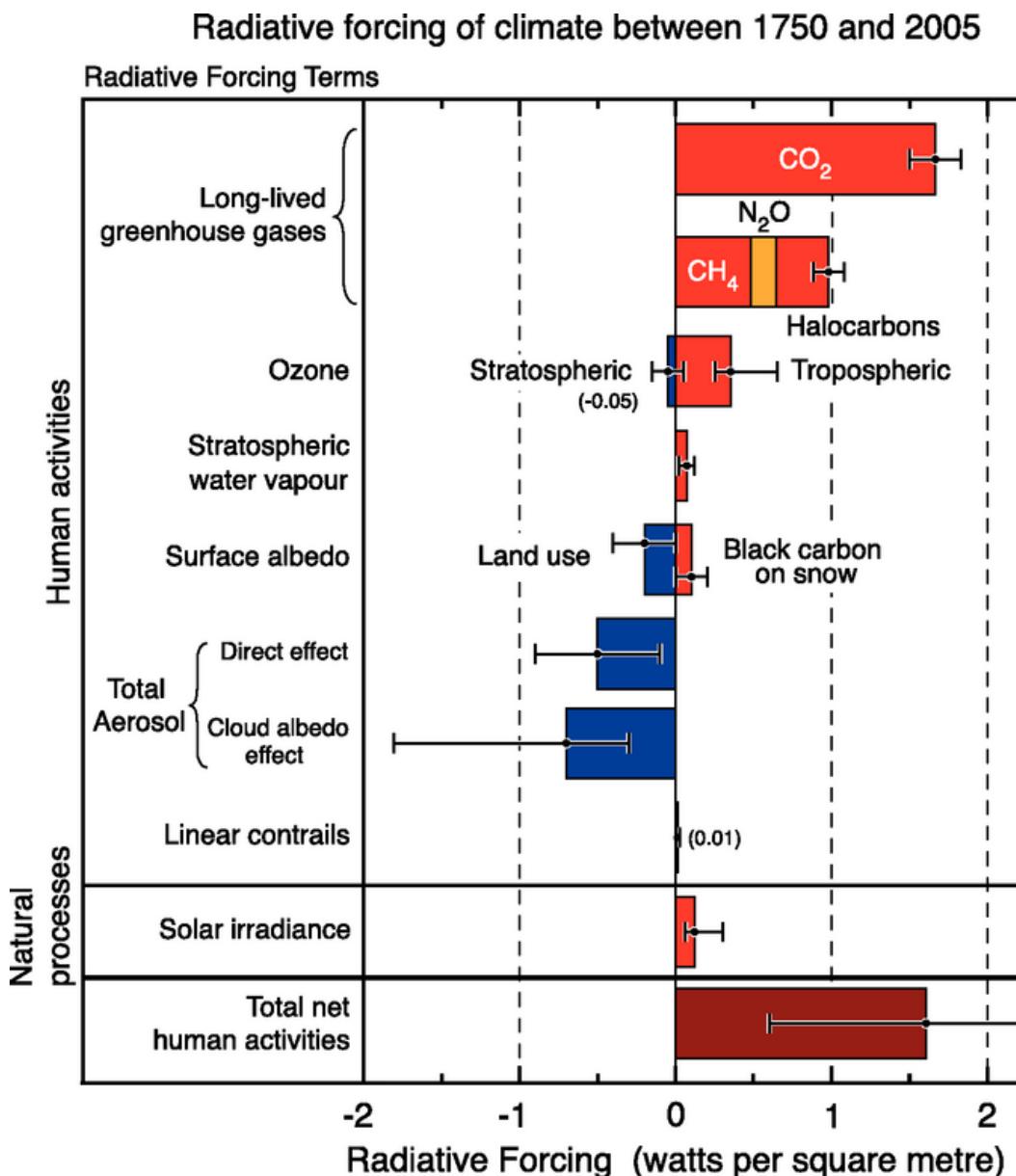


FIGURE A.2: The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from Forster et al. 2007, found at [https://www.ipcc.ch/publications\\_and\\_data/ar4/wg1/en/faq-2-1.html](https://www.ipcc.ch/publications_and_data/ar4/wg1/en/faq-2-1.html).

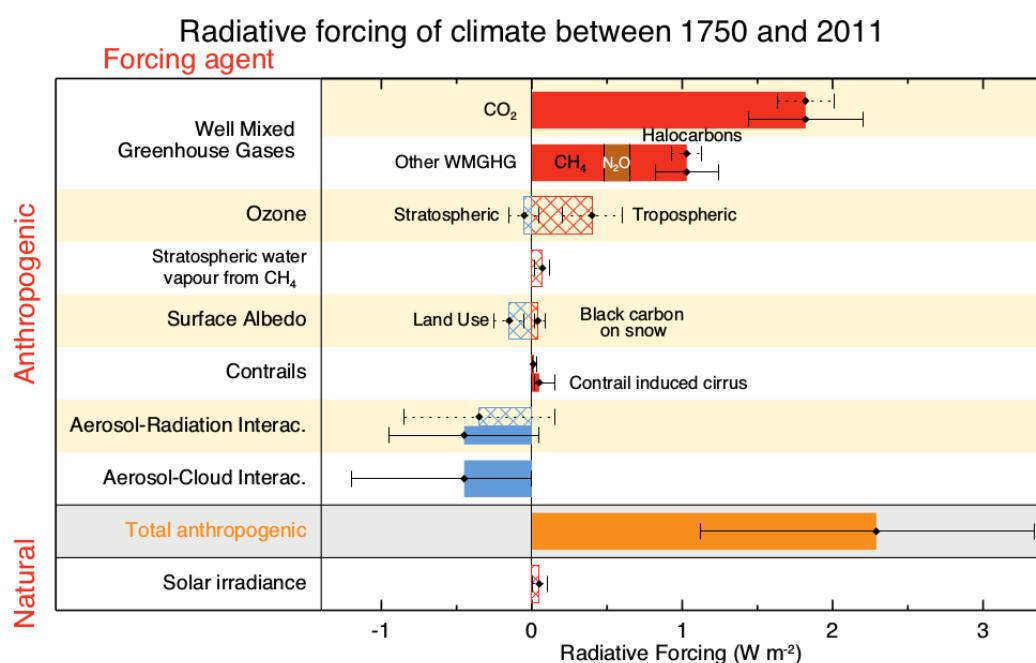


FIGURE A.3: The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 8.

### A.3.1 Relationship to Glyoxyl TODO: remove if never used

Another chemical retrievable from satellite observation is Glyoxyl, which can be used to further determine what sort of precursors to HCHO are being emitted (Stavrakou et al. 2009; Miller et al. 2014; Miller et al. 2016). TODO: go through 2014 paper and see if it's easy to retrieve, then email Dr. Chris Miller. For example Cao et al. 2018 recently used Glyoxyl measurements to improve understanding of biogenic and anthropogenic NMVOC emissions over China. This involved using a method pioneered by Stavrakou et al. 2009 TODO: get this cite and check method out.

Glyoxyl ( $\text{CHOCHO}$ ) is important to us as it shares many properties with HCHO, and may provide additional information in determining isoprene emissions. Glyoxyl is another product of VOC oxidation in the atmosphere, with isoprene being the main source globally. Under high  $\text{NO}_x$  conditions, glyoxyl forms rapidly, similarly to HCHO. However, glyoxyl also forms in low  $\text{NO}_x$  environments both slowly (through isoprene epoxydiols), and rapidly (through di-hydroperoxide dicarbonyl compound photolysis (Crounse et al. 2013)). This process is similar to the proposed mechanisms for hydroperoxyaldehydes by Jozef et al. 2014 and carbonyl nitrates (Müller, Peeters, and Stavrakou 2014). Aromatics which are largely anthropogenic form glyoxyl quickly, while HCHO is produced slower, allowing determination of anthropogenic sources (Cao et al. 2018).

HCHO has been used to estimate isoprene emissions (some examples in Section ??) but many uncertainties exist. One of these uncertainties is the yield of HCHO from isoprene, especially in low  $\text{NO}_x$  environments. Glyoxyl could prove complementary to HCHO in constraining isoprene emissions (TODO: Read and cite Vrekoussis 2009, 2010, Chan Miller 2014, Alvarado 2014) (Miller et al. 2016). Recently Miller et al. 2016 updated GEOS-Chem to include the prompt formation of glyoxyl and compared this with satellite and airplane measurements over the USA. With coming geostationary satellites, which provide greater time resolved measurements of HCHO and  $\text{CHO-CHO}$ , this mechanism could be used to clearly show when low  $\text{NO}_x$  isoprene chemistry is being undertaken (Miller et al. 2016).

## A.4 CAABA/MECCA

CAABA (Chemistry As A Boxmodel Application) estimates the chemical concentrations accounting for J-values (JVAL), simplified and parameterised photolysis (SAP-PHO) and simplified emission and depositions (SEMIDEP). CAABA runs in a single scenario (or box) with given emissions, depositions, and initial concentrations, allowing the examination of chemistry in a very specific environment to be modelled with high temporal resolution. CAABA/MECCA has been implemented for various calculations including ozone chemistry throughout the atmosphere in Zanis et al. 2014. The user manual is available online at [http://www.rolf-sander.net/messy/mecca/caaba\\_mecca\\_manual.pdf](http://www.rolf-sander.net/messy/mecca/caaba_mecca_manual.pdf).

This has been used with an atmospheric chemistry model MECCA (Module Efficiently Calculating the Chemistry of the Atmosphere) which implements tropospheric and stratospheric chemistry for both the gas and the aqueous phases (Sander et al.

2005). MECCA chemical mechanisms include basic O<sub>3</sub>, CH<sub>4</sub>, NO<sub>X</sub>, and HO<sub>X</sub> chemistry, as well as non methane hydrocarbon (NMHC) chemistry, considering gas phase, aqueous phase, and heterogenous reactions. (Sander et al. 2005) For the numerical integration, MECCA uses the KPP software (Sandu and Sander 2006), which takes chemical reactions and their rate coefficients and codes integral solutions to the system. The combination of the CAABA box model with MECCA module is called CAABA/MECCA and is currently at version 3. CAABA/MECCA been implemented for various calculations including ozone chemistry throughout the atmosphere in Zanis et al. 2014.

MECCA could also be used as the chemistry mechanism for a more complex, 3-dimensional model (e.g. Jöckel et al. 2006). The connection is established via the MESSy interface (<http://www.messy-interface.org>) developed by Jöckel, Sander, and Lelieveld 2004 as part of an effort to simplify the framework for modelling the atmospheres at various scales. The user manual is available online at [http://www.rolf-sander.net/messy/mecca/caaba\\_mecca\\_manual.pdf](http://www.rolf-sander.net/messy/mecca/caaba_mecca_manual.pdf).

#### A.4.0.1 CAABA/MECCA outputs

The box model can output in netcdf or text format, TODO: which way am I better off ? Text output from CAABA/MECCA was read using tailored python scripts modified from code written by dr. Luke Surl. Dr. Luke Surl also wrote the code which implements calculations of yield from runs using isoprene injections as described in Section A.4 TODO: update to more specific reference.

#### A.4.1 CAABA/MECCA Box model: isoprene source classifications

Using CAABA/MECCA to examine isoprene to HCHO yield in specific scenarios allows us to determine what environment may be driving the yield calculated by GEOS-Chem. Initially we have three scenarios, grassland, desert, and forest Australia - with each scenario having initial conditions, emission and deposition set as in table TODO:?. Running each scenario with and without a small isoprene injection allows calculation of isoprene lifetimes and HCHO yield for those scenarios.

CAABA runs in a single scenario (or box) with given emissions, depositions, and initial concentrations, allowing the examination of chemistry in a very specific environment to be modelled with high temporal resolution. This has been used with an atmospheric chemistry model MECCA (Module Efficiently Calculating the Chemistry of the Atmosphere) which implements tropospheric and stratospheric chemistry for both the gas and the aqueous phases (Sander et al. 2005). For our purposes it's worth noting that MECCAs chemical mechanism includes basic O<sub>3</sub>, CH<sub>4</sub>, NO<sub>X</sub>, and HO<sub>X</sub> chemistry, as well as non methane hydrocarbon (NMHC) chemistry, considering gas phase, aqueous phase, and heterogenous reactions. (Sander et al. 2005) For the numerical integration, MECCA uses the KPP software (Sandu and Sander 2006), which takes chemical reactions and their rate coefficients and forms efficient code for integral solutions to the system. The combination of the CAABA box model with MECCA module is called CAABA/MECCA and is currently at version 3.

We prescribe parameters values approximating three separate (and relatively broad) scenarios: forest, urban, and scrubland. These parameters are shown in Table A.1! The

TABLE A.1: Parameters for each scenario used in CAABA/MECCA model runs. TODO: fill these values in

Parameter(units)	Forest	Scrublands	Urban
Initial concentrations			
$\text{NO}_X(\text{molec cm}^{-3})$	low	low	high
$\text{O}_3(\text{molec cm}^{-3})$	mid	mid	low
$\text{HO}_X(\text{molec cm}^{-3})$			
Influx			
$\text{NO}_X(\text{molec cm}^{-3} \text{s}^{-1})$	low	low	high
$\text{O}_3(\text{molec cm}^{-3} \text{s}^{-1})$	mid	mid	low
$\text{HO}_X(\text{molec cm}^{-3} \text{s}^{-1})$			
Deposition rates?			

initial concentrations, the influx, (TODO: check this) and the deposition rates of various chemical species or families is perscribed. Chemical creation, destruction and deposition rates are determined every X minutes. This is done by altering a scenarios file which does not change over the lifetime (TODO: 40 days?) of the model runs. For each scenario, two nearly identical model runs are performed, one with an injection of isoprene occurring on (TODO: when/howmuch is this injection). The concentrations of isoprene and HCHO for our three scenarios, with and without the isoprene injection, is plotted over time in figure TODO: CAABA/MECCA scenario figures. The yield of HCHO from isoprene is then calculated by looking at the difference between each parallel run to determine how much of the injected isoprene transformed into HCHO. Isoprene life time can also be calculated using this process, as the time it takes for the extra isoprene to reach  $1/e$  of it's initial value.

TODO: Calculations for this (from Luke, double check these and enter them here). Calculation of the yield follows a calculation of the theoretical maximum carbon production by the amount of injected isoprene:

$$Y_{100} = 10^9 \times \frac{C_{PM} E_{inj} D_{inj}}{(N_A H_{PBL})} \quad (\text{A.1})$$

Where  $Y_{100}$  is the maximum possible carbon yield of isoprene (ppb),  $C_{PM}$  is Carbon per molecule (isoprene=5),  $E_{inj}$  is the emission rate of injected isoprene ( $\text{molec cm}^{-2} \text{s}^{-1}$ ),  $D_{inj}$  is the duration of injection (s),  $H_{PBL}$  is the boundary layer height (cm), and  $N_A$  is the Air number density ( $\text{molec cm}^{-3} \approx 2.5e19$ ). Finding the accumulated increase in HCHO (ppb) from the difference between the perturbed and non perturbed model runs allows calculation of the accumulated extra HCHO (Example: Figure TODO:), which divided by the  $Y_{100}$  gives us the isoprene to HCHO atom C yield:

$$Y_{HCHO} = \frac{\Delta HCHO_{\text{Accumulated}}}{Y_{100}} \quad (\text{A.2})$$

with  $HCHO_{\text{Accumulated}}$  being the accumulated enhanced ppb mixing ratio of HCHO.

Figure TODO: shows the accumulated yield for all three scenarios, which each increase towards a limiting value.

## A.5 Satellite Stuff

### A.5.1 OMI Algorithm BOAS

The following information comes from the OMHCHO dataset documentation at Kurosu and Chance (2014) and Instrument (2002). The method of HCHO total column retrieval depends heavily on measured solar radiation. Radiance is directional radiant flux, expressed in Watts per square metre per steradian (a unit of angle used in three dimensional geometry). Irradiance is radiant flux received by a surface, expressed in watts per square metre. An OMI granule is the sunlit portion of an orbit (one day).

The BOAS algorithm used by OMI is as follows. Slant column abundance can be determined by fitting measured radiance ( $I$ ) at particular wavelengths ( $\lambda$ ), using modelled absorption cross sections ( $\sigma$ ), effective albedo (A) including Rayleigh scattering, a correction for the Ring effect ( $c_R\sigma_R$ ), and a closing polynomial (coefficients  $c_0$ - $c_3$ ).

$$I(\lambda) = AI_0 \exp(-\sum_i S_i \sigma_i) + c_R \sigma_R + c_0 + c_1(\sigma - \bar{\sigma}) + c_2(\sigma - \bar{\sigma})^2 + c_3(\sigma - \bar{\sigma})^3 \quad (\text{A.3})$$

For HCHO, absorption cross sections and number densities for interfering gases are determined beforehand. This is due to HCHO being so optically thin and interferences must be accounted for precisely (Instrument 2002).

In version 3.0 of the OMI satellite data retrievals, HCHO is determined using the spectral window 328.5 nm–356.5 nm. The algorithm used is based on direct fitting of radiances and irradiances. An OMI radiance measurement over the remote Pacific ocean is used instead of an irradiance measurement. This means that the slant columns ( $\Omega_S$ ) are actually the difference with respect to the radiance reference column ( $\Omega_{S_0}$ ).

The model that is fitted to the measurements is made up of the radiance reference attenuated by HCHO contributions, inelastic (rotational Raman) scattering, and interferences from ozone, NO<sub>2</sub>, BrO, and the O<sub>2</sub>-O<sub>2</sub> collision complex. It includes additive and multiplicative closure polynomials and parameters for spectral shift and squeeze, and an undersampling correction and “common mode” spectrum. The spectral fitting results in HCHO slant columns, which are converted to vertical columns through a look-up table of AMFs (see section ??) Undersampling is a problem caused by the wavelength resolution of the instrument. Nyquist theorem requires that the sampling rate be at least twice the highest frequency of the signal in order to uniquely reconstruct it, otherwise the signal is undersampled (contains errors).

There are three main stages in the algorithm:

1. Radiance wavelength calibration, finding the optimum wavelength registration for a representative swath of radiance measurements, and determination of a common wavelength grid for auxiliary data (molecular reference cross sections, etc.).
2. On-line common mode spectrum calculation from residual fits of the central portion of the orbit. This accounts for systematic features not considered in the semi-empirical model.
3. Nonlinear least-squares fitting of all swath lines in the OMI granule. Fitting is performed individually for the 60 cross-track pixels in each swath line.

Cross-track striping is systematically higher or lower column values along a whole track. Several methods are used to reduce cross-track striping of the HCHO columns. These include soft calibration, which is the use of a daily radiance reference, and outlier screening in the fitting residuals.

### A.5.2 AMF recaulcation using 72 level output

The vertical column scattering weights and apriori shape factors provided in the OMH-CHO dataset are defined on 47 levels. In order to reformulate the vertical column using updated GEOS-Chem hcho apriori shape factors I have run GEOS-Chem version 10.01 on the full 72 level vertical grid at 2 by 2.5 (lat by lon) degree monthly resolution. The simulated vertical profiles of HCHO are averaged from 1300-1400 local time in order to match the satellite overpass time of roughly 1330. These vertical profiles then provide the apriori shape factor for the higher horizontally resolved satellite columns, which pick the nearest apriori from the model. TODO: determine which of these is correct! These new model apriori are stored daily, and are briefly compared against the 47 level model output.

A new AMF is determined using equation ??) with the apriori shape factor set by our GEOS-Chem model run.

Pressure dimension from OMI are the surface pressures from each gridbox (offline conversation with Dr. Christopher Miller). Determining the geometric pressure midpoints (here onwards pressure levels) and interpolating to our increased vertical resolution involves a few steps. The lowest level (with highest pressure) in whichever pressure dimension (ours or OMI's) extends to the lowest altitude (or highest pressure) is interpolated upwards to match the lowest level in the other dimension. Secondly, if the OMI dimension has been changed, the scattering weights are interpolated onto this updated dimension. Figure A.4 shows how these first two steps are applied using three fake array comparisons and updating the array with the lower surface level. Finally, once our dimensions match at the surface (we are not so worried about the very top of the atmosphere) we interpolate the scattering weights onto our updated GEOS-Chem pressure dimension.

$S_\sigma(\sigma)$  Is determined after running GEOS-Chem, which outputs vertical profiles of air density and HCHO mixing ratio, at 72 vertical levels with associated metadata such as vertical layer height and pressure, grid box location, height, and surface pressure. Using these outputs the vertical columns ( $\Omega_a, \Omega_v$ ) are calculated for each horizontal grid point (i, j) as follows:

$$\begin{aligned}\Omega_a(i, j) &= \Sigma_z (N_a(i, j, z) \times H(i, j, z)) \\ \Omega_z(i, j) &= \Sigma_z (N_{HCHO}(i, j, z) \times H(i, j, z))\end{aligned}$$

where  $N_a$ , and  $N_{HCHO}$  are the densities of air and HCHO, H is the layer height (for each grid box). Note that HCHO density is determined from the outputted mixing ratio:  $N_{HCHO} = C_{HCHO} \times N_a$ .

$S_\sigma(\sigma)$  is then stored in HDF-EOS5 format, to be used in conjunction with the satellite measurements to calculate an AMF as shown in equation ???. As the GEOS-Chem V10.01 output is in bitpunch format, the code to read the data and create the shape factor is written in IDL, which has many procedures and functions which are already

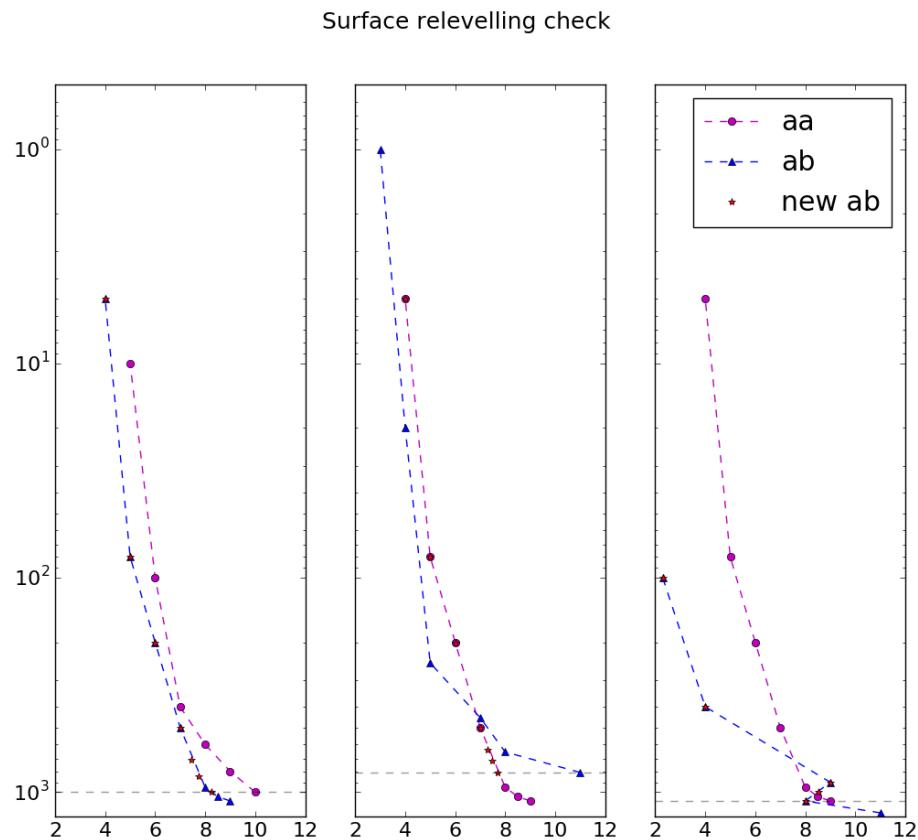


FIGURE A.4: Constructed example of the initial interpolation of OMI's  $\omega$  onto a pressure dimension with mismatched surface pressure.

written to handle reading this format (provided by GAMAP). The code is provided in supplementary TODO: put code into supplement section.

For each OMI slant column, a new AMF is calculated using  $S_\sigma(\sigma)$  and the provided scattering weights  $\omega(\sigma)$  using equation ???. This integral is applied in python by taking the sum of  $S_\sigma(\sigma) \times \omega(\sigma) \times d\sigma$  for each  $\sigma$  determined at 72 levels in GEOS-Chem, with the provided  $\omega$  interpolated linearly to these same levels. An example of these interpolations is shown in figure TODO: interpolation figure with symbols at original points and interpolated line overplotted for both functions over hPa. Globally this reprocessing changed the AMF by TODO: global total percent difference in AMF. In total this caused TODO: total column HCHO change globally/yearly In Summer over Australia the global AMF difference was TODO: Difference summers only. This changed Australia's HCHO amounts from TODO: X to Y Tg per year plus minus one std.

### A.5.3 Old Fire Product MYD14C8H

On board NASA's AQUA satellite, the MODIS instrument is used to detect fire activity. The product used here is called MYD14C8H ((Giglio, Csiszar, and Justice 2006)), which looks at fire activity over eight days on a  $0.5^\circ$  square grid globally. Regridding the product to the native meteorological grid of GEOS5 at  $0.25^\circ$  latitude by  $0.3125^\circ$  longitude is done in python with an interpolator which maps the values of the new grid rectangles to the value of the nearest grid square.

Figure A.5 shows an example of the total column HCHO calculated using GEOS-Chem aprioris ( $\Omega_{GEOS}$ ) before and after using the MYD14C8H product to exclude fire influenced pixels. (TODO: show time series of how many pixels are removed and discuss if this causes any issues down the line)

This filtering ends up removing too much information, and the recalculation of HCHO is too negatively influenced. To deal with this a separate product from the same instrument has been downloaded: MYD14A1, which keeps daily fire counts. Less disruptive filtering can be achieved by removing pixels which coincide with fires on the same day, as shown in figure TODO: which compares the 8 and 1 day filtering. TODO: The script to read and regrid these one day fire counts was adapted from X. Figure (TODO: effect on uncertainty and time series of fire pixels removed) shows the daily filtering effect on uncertainty and time series of fire pixels removed.

An example of the change in resolution is provided in figure A.6, where the grids are shown over a basic map of Tasmania. The direct affect of this interpolation is shown as an example in figure A.7, which is showing the regridded MODIS fire count over Australia from January 2005 (avg of first 8 days) in two subplots.

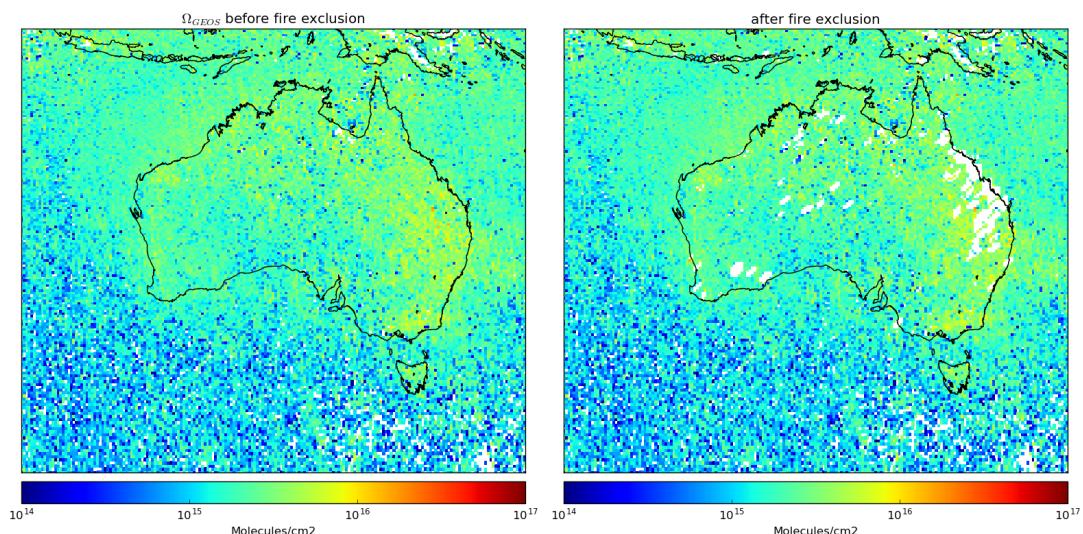
**Eight day average  $\Omega_{GEOS}$  for Jan 1, 2005**

FIGURE A.5: Vertical column HCHO calculated using OMI satellite swaths with GEOS-Chem aprioris, averaged over 1-8 January 2005 with and without fire affected squares removed.

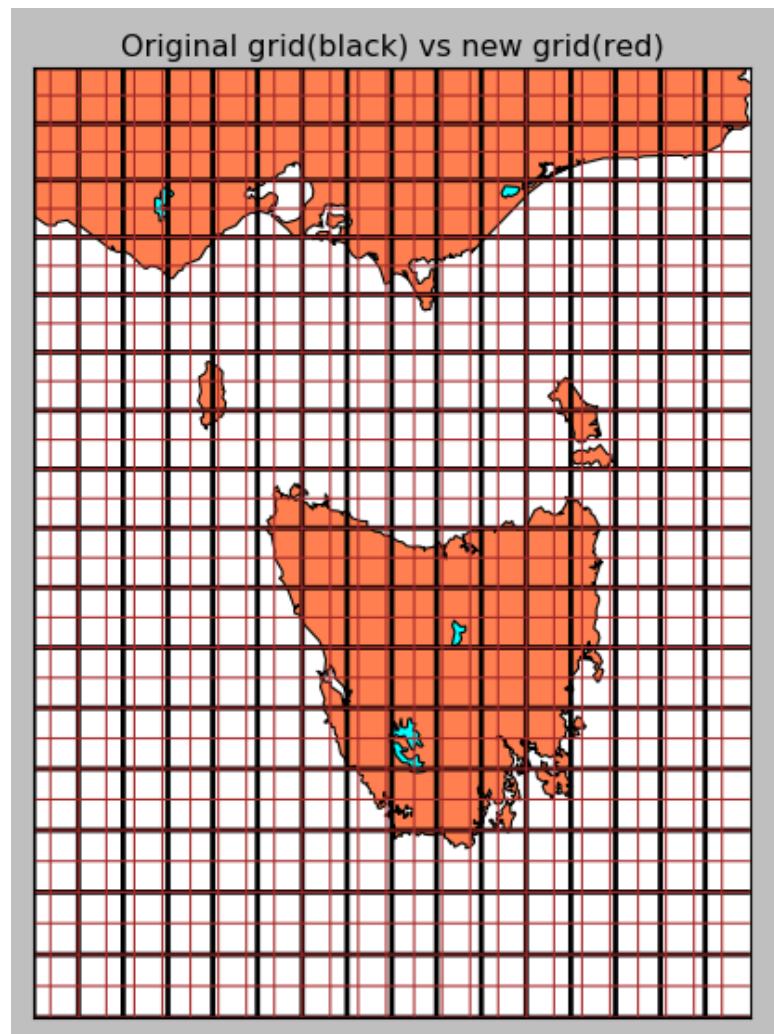


FIGURE A.6: Example of grid space change using  $0.5 \times 0.5$  and  $0.25 \times 0.3125$  latitude by longitude resolution.

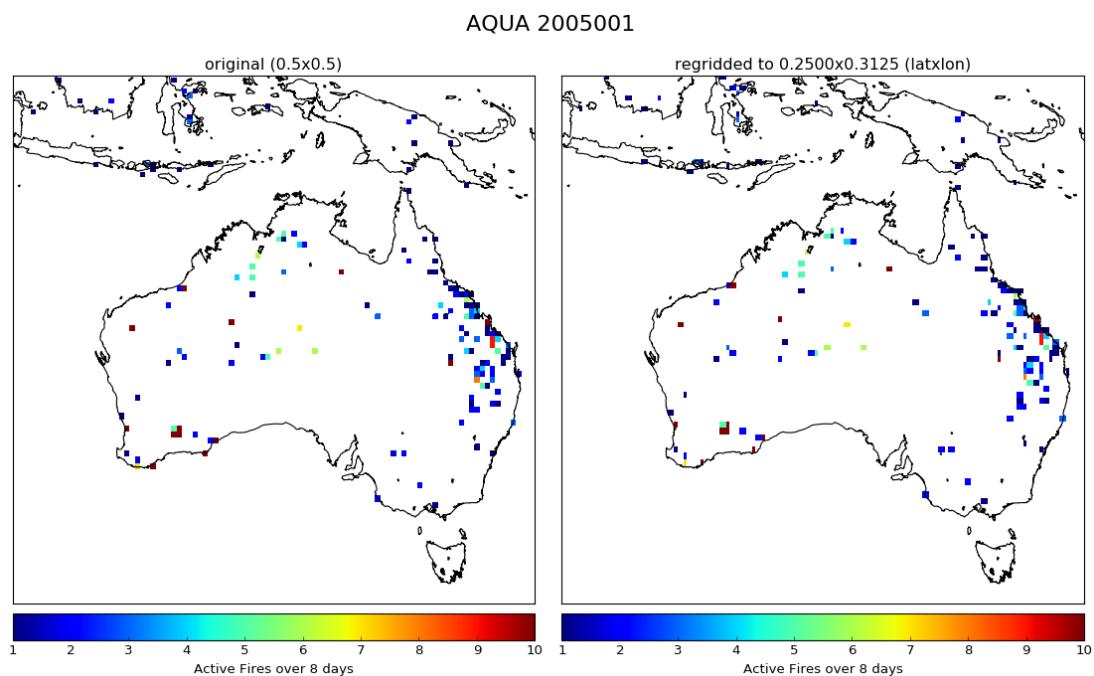


FIGURE A.7: Example of MODIS 8 day grid interpolation from  $0.5 \times 0.5$  to  $0.25 \times 0.3125$  latitude by longitude resolution. This example uses MODIS fire counts for 1-8 January 2005.

## Appendix B

# Appendix A

### B.1 Grid Resolution

Level	z(km)	eta
1	0.058483	0.99250000
2	0.188515	0.97745600
3	0.320427	0.96237000
4	0.453892	0.94728500
5	0.588963	0.93220000
6	0.725697	0.91711600
7	0.864132	0.90203100
8	1.004320	0.88694800
9	1.146330	0.87186400
10	1.290220	0.85678100
11	1.436040	0.84169800
12	1.583870	0.82661600
13	1.758950	0.80902100
14	1.988400	0.78640000
15	2.249380	0.76126500
16	2.517070	0.73613400
17	2.791920	0.71100600
18	3.074460	0.68587800
19	3.439220	0.65447100
20	3.895550	0.61679000
21	4.374520	0.57911500
22	4.879030	0.54144900
23	5.412570	0.50379500
24	5.979530	0.46615300
25	6.585330	0.42852800
26	7.236820	0.39092700
27	7.943070	0.35334900
28	8.845840	0.30985400
29	9.936460	0.26358700
30	11.020800	0.22377200
31	12.085500	0.19006100
32	13.133500	0.16151300
33	14.169600	0.13728700
34	15.198200	0.11669500
35	16.221700	0.09919110
36	17.242600	0.08431270
37	18.727200	0.06655890
38	20.835700	0.04764090
39	23.020500	0.03381390
40	25.306800	0.02375520
41	28.654300	0.01434240
42	34.024400	0.00658754
43	40.166100	0.00281639
44	47.135400	0.00110910
45	54.833600	0.00039878
46	63.052800	0.00012699
47	72.180100	0.00002763

TABLE B.1: The 47 level vertical grid used by GEOS-Chem

## Appendix C

# Frequently Asked Questions

### C.1 How do I change the colors of links?

The color of links can be changed to your liking using:

```
\hypersetup{urlcolor=red}, or  
\hypersetup{citecolor=green}, or  
\hypersetup{allcolor=blue}.
```

If you want to completely hide the links, you can use:

```
\hypersetup{allcolors=}, or even better:  
\hypersetup{hidelinks}.
```

If you want to have obvious links in the PDF but not the printed text, use:

```
\hypersetup{colorlinks=false}.
```



# Bibliography

- Brasseur, Guy P and Daniel J Jacob (2017). *Modeling of Atmospheric Chemistry*. Cambridge University Press. DOI: [10.1017/9781316544754](https://doi.org/10.1017/9781316544754).
- Hsieh, Nan-Hung and Chung-Min Liao (2013). "Fluctuations in air pollution give risk warning signals of asthma hospitalization". In: *Atmospheric Environment* 75, pp. 206–216. DOI: [10.1016/j.atmosenv.2013.04.043](https://doi.org/10.1016/j.atmosenv.2013.04.043). URL: <http://dx.doi.org/10.1016/j.atmosenv.2013.04.043>.
- Avnery, Shiri et al. (2013). "Global crop yield reductions due to surface ozone exposure: 2. Year 2030 potential crop production losses and economic damage under two scenarios of O<sub>3</sub> pollution". In: *Atmospheric Environment* 71.13, pp. 408–409. ISSN: 13522310. DOI: [10.1016/j.atmosenv.2012.12.045](https://doi.org/10.1016/j.atmosenv.2012.12.045). URL: <http://dx.doi.org/10.1016/j.atmosenv.2011.01.002>.
- Yue, Xu et al. (2017). "Ozone and haze pollution weakens net primary productivity in China". In: *Atmospheric Chemistry and Physics* 17.9, pp. 6073–6089. ISSN: 1680-7324. DOI: [10.5194/acp-17-6073-2017](https://doi.org/10.5194/acp-17-6073-2017). URL: <https://www.atmos-chem-phys.net/17/6073/2017/>.
- Myhre, G and D Shindell (2013). *Chapter 8: Anthropogenic and Natural Radiative Forcing, in Climate Change 2013: The Physical Science Basis, Working Group 1 Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 2013. Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 2013.
- Monks, P. S. et al. (2015). "Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer". In: *Atmospheric Chemistry and Physics* 15.15, pp. 8889–8973. ISSN: 1680-7324. DOI: [10.5194/acp-15-8889-2015](https://doi.org/10.5194/acp-15-8889-2015). URL: <https://www.atmos-chem-phys.net/15/8889/2015/>.
- Lelieveld, J. et al. (2013). "Model calculated global, regional and megacity premature mortality due to air pollution". In: *Atmospheric Chemistry and Physics* 13.14, pp. 7023–7037. ISSN: 16807324. DOI: [10.5194/acp-13-7023-2013](https://doi.org/10.5194/acp-13-7023-2013).
- Ayers, James D and William R Simpson (2006). "Measurements of N<sub>2</sub>O<sub>5</sub> near Fairbanks, Alaska". In: *Journal of Geophysical Research: Atmospheres* 111.D14, n/a–n/a. ISSN: 2156-2202. DOI: [10.1029/2006JD007070](https://doi.org/10.1029/2006JD007070). URL: <http://dx.doi.org/10.1029/2006JD007070>.
- Lelieveld, J. et al. (2009). "Severe ozone air pollution in the Persian Gulf region". In: *Atmospheric Chemistry and Physics* 9, pp. 1393–1406. ISSN: 1680-7324. DOI: [10.5194/acp-9-1393-2009](https://doi.org/10.5194/acp-9-1393-2009).
- Ashmore, M R, Lisa. Emberson, and Murray Frank (2003). *Air pollution impacts on crops and forests : a global assessment*. Ed. by Lisa Emberson, Mike Ashmore, and Frank Murray. Imperial College Press London ; River Edge, NJ, xiii, 372 p. : ISBN: 186094292.

- Stevenson, D. S. et al. (2013). "Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP)". In: *Atmospheric Chemistry and Physics* 13.6, pp. 3063–3085. ISSN: 16807316. DOI: [10.5194/acp-13-3063-2013](https://doi.org/10.5194/acp-13-3063-2013).
- Selin, N E et al. (2009). "Global health and economic impacts of future ozone pollution". In: *Environmental Research Letters* 4.4, p. 044014. ISSN: 1748-9326. DOI: [10.1088/1748-9326/4/4/044014](https://doi.org/10.1088/1748-9326/4/4/044014).
- Atkinson, Roger (2000). "Atmospheric chemistry of VOCs and NO(x)". In: *Atmospheric Environment* 34.12-14, pp. 2063–2101. ISSN: 13522310. DOI: [10.1016/S1352-2310\(99\)00460-4](https://doi.org/10.1016/S1352-2310(99)00460-4).
- Fuentes, J. D. et al. (2000). "Biogenic Hydrocarbons in the Atmospheric Boundary Layer: A Review". In: *Bulletin of the American Meteorological Society* 81.7, pp. 1537–1575. ISSN: 00030007. DOI: [10.1175/1520-0477\(2000\)081<1537:BHITAB>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<1537:BHITAB>2.3.CO;2). arXiv: [arXiv : 1011.1669v3](https://arxiv.org/abs/1011.1669v3). URL: [http://journals.ametsoc.org/doi/abs/10.1175/1520-0477\(2000\)081%3C1537%3ABHITAB%3E2.3.CO%3B2](http://journals.ametsoc.org/doi/abs/10.1175/1520-0477(2000)081%3C1537%3ABHITAB%3E2.3.CO%3B2).
- 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).
- Atkinson, Roger and Janet Arey (2003). "Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review". In: *Atmospheric Environment* 37.SUPPL. 2. ISSN: 13522310. DOI: [10.1016/S1352-2310\(03\)00391-1](https://doi.org/10.1016/S1352-2310(03)00391-1).
- Kanakidou, M et al. (2005). "Physics Organic aerosol and global climate modelling : a review". In: *Atmospheric Chemistry and Physics* 5, pp. 1053–1123.
- Stocker, T.F. et al. *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Tech. rep. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. DOI: [10.1017/CBO9781107415324](https://doi.org/10.1017/CBO9781107415324).
- Intergovernmental Panel on Climate Change (IPCC): Climate Change: The Scientific Basis* (2001). Tech. rep. Cambridge University Press. URL: <http://www.ipcc.ch/ipccreports/tar/>.
- Forster, P. et al. (2007). *Changes in Atmospheric Constituents and in Radiative Forcing*. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Man. URL: [https://www.ipcc.ch/publications\\_and\\_data/ar4/wg1/en/ch2.html](https://www.ipcc.ch/publications_and_data/ar4/wg1/en/ch2.html) (visited on 01/14/2016)].
- Hegglin, Michaela I and Theodore G Shepherd (2009). "Large climate-induced changes in ultraviolet index and stratosphere-to-troposphere ozone flux". In: *Nature Geoscience* 2.10, pp. 687–691. DOI: [10.1038/ngeo604](https://doi.org/10.1038/ngeo604). URL: <http://dx.doi.org/10.1038/ngeo604>.
- Jacob, Daniel J (1999). *Introduction to Atmospheric Chemistry*. Ed. by Daniel J Jacob. Princeton University Press. URL: <http://acmg.seas.harvard.edu/people/faculty/djj/book/index.html>.
- 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/>.

- 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.pdf](http://eprints.lancs.ac.uk/88836/1/TOAR_Model_Performance_07062017.pdf) [http://www.igacproject.org/sites/default/files/2017-05/TOAR-Model\\_Performance\\_draft\\_for\\_open\\_comment.pdf](http://www.igacproject.org/sites/default/files/2017-05/TOAR-Model_Performance_draft_for_open_comment.pdf).
- Cape, J. N. (2008). "Surface ozone concentrations and ecosystem health: Past trends and a guide to future projections". In: *Science of the Total Environment* 400.1-3, pp. 257–269. ISSN: 00489697. DOI: [10.1016/j.scitotenv.2008.06.025](https://doi.org/10.1016/j.scitotenv.2008.06.025). URL: <http://dx.doi.org/10.1016/j.scitotenv.2008.06.025>.
- Delmas, R, D Serca, and C Jambert (1997). "Global inventory of NO<sub>x</sub> sources". In: *Nutrient cycling in agroecosystems* 48.x, pp. 51–60. ISSN: 1385-1314. DOI: [10.1023/A:1009793806086](https://doi.org/10.1023/A:1009793806086). URL: <http://link.springer.com/article/10.1023/A:1009793806086>.
- Sillman, Sanford (1999). "The relation between ozone, NO and hydrocarbons in urban and polluted rural environments". In: *Atmospheric Environment* 33. DOI: [https://doi.org/10.1016/S1352-2310\(98\)00345-8](https://doi.org/10.1016/S1352-2310(98)00345-8). URL: [http://www-personal.umich.edu/\\$\sim\\\$sillman/web-publications/Sillmanreview99.pdf](http://www-personal.umich.edu/$\sim\$sillman/web-publications/Sillmanreview99.pdf) <https://www.sciencedirect.com/science/article/pii/S1352231098003458>.
- Stohl, Andreas et al. (2003). "A new perspective of stratosphere-troposphere exchange". In: *Bulletin of the American Meteorological Society* 84.11, pp. 1565–1573+1473. ISSN: 00030007. DOI: [10.1175/BAMS-84-11-1565](https://doi.org/10.1175/BAMS-84-11-1565).
- 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>.
- Jacobson, M C and H Hansson (2000). "Organic atmospheric aerosols: Review and state of the science". In: *Reviews of Geophysics* 38.38, pp. 267–294. ISSN: 87551209. DOI: [10.1029/1998RG000045](https://doi.org/10.1029/1998RG000045). URL: <http://dx.doi.org/10.1029/1998RG000045>.
- Kuang, Shi et al. (2017). "Summertime tropospheric ozone enhancement associated with a cold front passage due to stratosphere-to-troposphere transport and biomass burning: Simultaneous ground-based lidar and airborne measurements". In: *Journal of Geophysical Research: Atmospheres* 122.2, pp. 1293–1311. ISSN: 21698996. DOI: [10.1002/2016JD026078](https://doi.org/10.1002/2016JD026078). URL: <http://doi.wiley.com/10.1002/2016JD026078>.
- Cooper, O. et al. (2004). "On the life cycle of a stratospheric intrusion and its dispersion into polluted warm conveyor belts". In: *Journal of Geophysical Research* 109.23, pp. 1–18. ISSN: 01480227. DOI: [10.1029/2003JD004006](https://doi.org/10.1029/2003JD004006).
- Young, P. J. et al. (2013). "Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP)". In: *Atmospheric Chemistry and Physics* 13.10, pp. 5277–5298. ISSN: 16807316. DOI: [10.5194/acp-13-5277-2013](https://doi.org/10.5194/acp-13-5277-2013).
- Sprenger, Michael, Mischa Croci Maspoli, and Heini Wernli (2003). "Tropopause folds and cross-tropopause exchange: A global investigation based upon ECMWF analyses for the time period March 2000 to February 2001". In: *Journal of Geophysical*

- Research* 108.D12. ISSN: 2156-2202. DOI: [10.1029/2002JD002587](https://doi.org/10.1029/2002JD002587). URL: <http://dx.doi.org/10.1029/2002JD002587>.
- Ojha, Narendra et al. (2016). "Secondary ozone peaks in the troposphere over the Hi-malayas". In: *Atmospheric Chemistry and Physics Discussions* 17.November, pp. 1–25. ISSN: 1680-7375. DOI: [10.5194/acp-2016-908](https://doi.org/10.5194/acp-2016-908). URL: <http://www.atmos-chem-phys-discuss.net/acp-2016-908/>.
- Liu, Junhua et al. (2017). "Causes of interannual variability over the southern hemispheric tropospheric ozone maximum". In: *Atmos. Chem. Phys* 17.5, pp. 3279–3299. ISSN: 1680-7324. DOI: [10.5194/acp-17-3279-2017](https://doi.org/10.5194/acp-17-3279-2017). URL: [www.atmos-chem-phys.net/17/3279/2017/](http://www.atmos-chem-phys.net/17/3279/2017/).
- Lin, Meiyun et al. (2015). "Climate variability modulates western US ozone air quality in spring via deep stratospheric intrusions." In: *Nature communications* 6.May, p. 7105. ISSN: 2041-1723. DOI: [10.1038/ncomms8105](https://doi.org/10.1038/ncomms8105). URL: <http://www.nature.com/ncomms/2015/150512/ncomms8105/full/ncomms8105.html>.
- Marvin, Margaret R. et al. (2017). "Impact of evolving isoprene mechanisms on simulated formaldehyde: An inter-comparison supported by in situ observations from SENEX". In: *Atmospheric Environment* 164, pp. 325–336. ISSN: 13522310. DOI: [10.1016/j.atmosenv.2017.05.049](https://doi.org/10.1016/j.atmosenv.2017.05.049). URL: [https://ac.els-cdn.com/S1352231017303618/1-s2.0-S1352231017303618-main.pdf?\\_tid=3de7eaaa-06ff-11e8-99a9-0000aacb360&acdnat=1517455576\\_09d7334af609ed43470155c1c42fad5fhttp://www.sciencedirect.com/science/article/pii/S1352231017303618](https://ac.els-cdn.com/S1352231017303618/1-s2.0-S1352231017303618-main.pdf?_tid=3de7eaaa-06ff-11e8-99a9-0000aacb360&acdnat=1517455576_09d7334af609ed43470155c1c42fad5fhttp://www.sciencedirect.com/science/article/pii/S1352231017303618).
- Mazzuca, Gina M. et al. (2016). "Ozone production and its sensitivity to NOx and VOCs: Results from the DISCOVER-AQ field experiment, Houston 2013". In: *Atmospheric Chemistry and Physics* 16.22, pp. 14463–14474. ISSN: 16807324. DOI: [10.5194/acp-16-14463-2016](https://doi.org/10.5194/acp-16-14463-2016).
- Stevenson, D S et al. (2006). "Multimodel ensemble simulations of present-day and near-future tropospheric ozone". In: *Journal of Geophysical Research* 111.D8. DOI: [10.1029/2005jd006338](https://doi.org/10.1029/2005JD006338). URL: <http://dx.doi.org/10.1029/2005JD006338>.
- 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).
- Guenther, Alex et al. (1995). "A global model of natural volatile organic compound emissions". In: *Journal of Geophysical Research* 100.D5, pp. 8873–8892. ISSN: 0148-0227. DOI: [10.1029/94JD02950](https://doi.org/10.1029/94JD02950). URL: <http://onlinelibrary.wiley.com/doi/10.1029/94JD02950/full%5Cnhttp://doi.wiley.com/10.1029/94JD02950http://onlinelibrary.wiley.com/doi/10.1029/94JD02950/full>.
- Glasius, Marianne and Allen H. Goldstein (2016). "Recent Discoveries and Future Challenges in Atmospheric Organic Chemistry". In: *Environmental Science and Technology* 50.6, pp. 2754–2764. ISSN: 15205851. DOI: [10.1021/acs.est.5b05105](https://doi.org/10.1021/acs.est.5b05105).
- Yue, X., N. Unger, and Y. Zheng (2015). "Distinguishing the drivers of trends in land carbon fluxes and plant volatile emissions over the past 3 decades". In: *Atmospheric Chemistry and Physics* 15.20, pp. 11931–11948. ISSN: 16807324. DOI: [10.5194/acp-15-11931-2015](https://doi.org/10.5194/acp-15-11931-2015).

- Stavrakou, T. et al. (2014). "Isoprene emissions over Asia 1979-2012: Impact of climate and land-use changes". In: *Atmospheric Chemistry and Physics* 14.9. ISSN: 16807324. DOI: [10.5194/acp-14-4587-2014](https://doi.org/10.5194/acp-14-4587-2014).
- 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/>.
- Aksoyoglu, Sebnem et al. (2017). "Secondary inorganic aerosols in Europe: sources and the significant influence of biogenic VOC emissions, especially on ammonium nitrate". In: *Atmospheric Chemistry and Physics* 17.12, pp. 7757–7773. ISSN: 1680-7324. DOI: [10.5194/acp-17-7757-2017](https://doi.org/10.5194/acp-17-7757-2017). URL: <https://doi.org/10.5194/acp-17-7757-2017>.
- Lelieveld, J et al. (2015). "The contribution of outdoor air pollution sources to premature mortality on a global scale". In: *Nature* 525.7569, pp. 367–371. DOI: [10.1038/nature15371](https://doi.org/10.1038/nature15371). URL: <http://dx.doi.org/10.1038/nature15371>.
- Cooper, O R, S Gilge, and D T Shindell (2014). "Global distribution and trends of tropospheric ozone : An observation-based review". In: pp. 1–28. DOI: [10.12952/journal.elementa.000029](https://doi.org/10.12952/journal.elementa.000029). URL: [https://doi.org/10.12952/journal.elementa.000029/](https://doi.org/10.12952/journal.elementa.000029).
- Hoek, Gerard et al. (2013). "Long-term air pollution exposure and cardio- respiratory mortality: a review". In: *Environmental Health* 12.1, p. 43. DOI: [10.1186/1476-069x-12-43](https://doi.org/10.1186/1476-069x-12-43). URL: <http://dx.doi.org/10.1186/1476-069x-12-43>.
- Krewski, D et al. (2009). "Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality". In: *Res Rep Health Eff Inst* 140, pp. 5–36. ISSN: 1041-5505 (Print) 1041-5505 (Linking). URL: <http://www.ncbi.nlm.nih.gov/pubmed/19627030>.
- Silva, Raquel A et al. (2013). "Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change". In: *Environ. Res. Lett.* 8.3, p. 34005. DOI: [10.1088/1748-9326/8/3/034005](https://doi.org/10.1088/1748-9326/8/3/034005). URL: <http://dx.doi.org/10.1088/1748-9326/8/3/034005>.
- Kroll, Jesse H. and John H. Seinfeld (2008). "Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere". In: *Atmospheric Environment* 42.16, pp. 3593–3624. ISSN: 13522310. DOI: [10.1016/j.atmosenv.2008.01.003](https://doi.org/10.1016/j.atmosenv.2008.01.003). URL: <http://www.sciencedirect.com.ezproxy.uow.edu.au/science/article/pii/S1352231008000253>.
- 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).
- Brown, S. S. et al. (2009). "Nocturnal isoprene oxidation over the Northeast United States in summer and its impact on reactive nitrogen partitioning and secondary organic aerosol". In: *Atmospheric Chemistry and Physics* 9.9, pp. 3027–3042. ISSN: 16807316. DOI: [10.5194/acp-9-3027-2009](https://doi.org/10.5194/acp-9-3027-2009).

- 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](#).
- Arneth, A et al. (2008). "Why are estimates of global terrestrial isoprene emissions so similar ( and why is this not so for monoterpenes )?" In: *Atmos. Chem. Phys* 8.x, pp. 4605–4620. ISSN: 1680-7375. DOI: [10.5194/acpd-8-7017-2008](https://doi.org/10.5194/acpd-8-7017-2008).
- Niinemets, U. et al. (2010). "The emission factor of volatile isoprenoids: Stress, acclimation, and developmental responses". In: *Biogeosciences* 7.7, pp. 2203–2223. ISSN: 17264170. DOI: [10.5194/bg-7-2203-2010](https://doi.org/10.5194/bg-7-2203-2010).
- Lathière, J et al. (2006). "Impact of climate variability and land use changes on global biogenic volatile organic compound emissions". In: *Atmospheric Chemistry and Physics* 6.2003, pp. 2129–2146. ISSN: 16807324. DOI: [10.5194/acp-6-2129-2006](https://doi.org/10.5194/acp-6-2129-2006). URL: [www.atmos-chem-phys.net/6/2129/2006/](http://www.atmos-chem-phys.net/6/2129/2006/).
- 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).
- 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).
- 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#.VqqEubNM61M>.
- Wagner, V (2002). "Are CH<sub>2</sub>O measurements in the marine boundary layer suitable for testing the current understanding of CH<sub>4</sub> photooxidation?: A model study". In: *Journal of Geophysical Research* 107.D3, p. 4029. ISSN: 0148-0227. DOI: [10.1029/2001JD000722](https://doi.org/10.1029/2001JD000722). URL: <http://doi.wiley.com/10.1029/2001JD000722>.
- Messina, Palmira et al. (2016). "Global biogenic volatile organic compound emissions in the ORCHIDEE and MEGAN models and sensitivity to key parameters". In: *Atmospheric Chemistry and Physics* 16.22, pp. 14169–14202. ISSN: 16807324. DOI: [10.5194/acp-16-14169-2016](https://doi.org/10.5194/acp-16-14169-2016). URL: <http://www.atmos-chem-phys.net/16/14169/2016/acp-16-14169-2016.pdf>.
- Nguyen, T. B. et al. (2014). "Overview of the Focused Isoprene eXperiment at the California Institute of Technology (FIXCIT): Mechanistic chamber studies on the oxidation of biogenic compounds". In: *Atmospheric Chemistry and Physics* 14.24. ISSN: 16807324. DOI: [10.5194/acp-14-13531-2014](https://doi.org/10.5194/acp-14-13531-2014).
- Hewitt, C N et al. (2011). "Ground-level ozone influenced by circadian control of isoprene emissions". In: *Nature Geoscience* 4.10, pp. 671–674. DOI: [10.1038/ngeo1271](https://doi.org/10.1038/ngeo1271). URL: <http://dx.doi.org/10.1038/ngeo1271>.
- Fan, Jiwen and Renyi Zhang (2004). "Atmospheric oxidation mechanism of isoprene". In: *Environmental Chemistry* 1.3, pp. 140–149. ISSN: 14482517. DOI: [10.1071/EN04045](https://doi.org/10.1071/EN04045). URL: <http://dx.doi.org/10.1071/en04045>.

- PATCHEN, AMIE K. et al. (2007). "Direct Kinetics Study of the Product-Forming Channels of the Reaction of Isoprene-Derived Hydroperoxy Radicals with NO". In: *International journal of Chemical Kinetics* 31.5, pp. 493–499. ISSN: 13000527. DOI: [10.1002/kin.20248](https://doi.org/10.1002/kin.20248). URL: <http://onlinelibrary.wiley.com/wol1/doi/10.1002/kin.20248/full>.
- 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).
- Wolfe, G. M. et al. (2016). "Formaldehyde production from isoprene oxidation across NO<sub>x</sub> 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/](http://www.atmos-chem-phys.net/16/2597/2016/).
- Crounse, John D et al. (2012). "Atmospheric Fate of Methacrolein. 1. Peroxy Radical Isomerization Following Addition of OH and O<sub>2</sub>". In: *Physical Chemistry m.*
- Nguyen, Tran B. et al. (2016). "Atmospheric fates of Criegee intermediates in the ozonolysis of isoprene". In: *Phys. Chem. Chem. Phys.* 18.15, pp. 10241–10254. ISSN: 1463-9076. DOI: [10.1039/C6CP00053C](https://doi.org/10.1039/C6CP00053C). URL: <http://xlink.rsc.org/?DOI=C6CP00053C>.
- Crounse, John D. et al. (2013). "Autoxidation of organic compounds in the atmosphere". In: *Journal of Physical Chemistry Letters* 4.20, pp. 3513–3520. ISSN: 19487185. DOI: [10.1021/jz4019207](https://doi.org/10.1021/jz4019207). URL: <http://pubs.acs.org/doi/abs/10.1021/jz4019207>.
- Crounse, John D. et al. (2006). "Measurement of gas-phase hydroperoxides by chemical ionization mass spectrometry". In: *Analytical Chemistry* 78.19, pp. 6726–6732. ISSN: 00032700. DOI: [10.1021/ac0604235](https://doi.org/10.1021/ac0604235). URL: <https://pubs.acs.org/doi/abs/10.1021/ac0604235>.
- Yu, Karen et al. (2016). "Sensitivity to grid resolution in the ability of a chemical transport model to simulate observed oxidant chemistry under high-isoprene conditions". In: *Atmospheric Chemistry and Physics* 16.7, pp. 4369–4378. ISSN: 16807324. DOI: [10.5194/acp-16-4369-2016](https://doi.org/10.5194/acp-16-4369-2016). URL: [http://acmg.seas.harvard.edu/publications/2016/Yu\\_ACP\\_2016.pdf](http://acmg.seas.harvard.edu/publications/2016/Yu_ACP_2016.pdf).
- 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>.
- Liu, Yingjun et al. (2016b). "Isoprene photochemistry over the Amazon rainforest". In: *Proceedings of the National Academy of Sciences* 113.22, pp. 6125–6130. ISSN: 0027-8424. DOI: [10.1073/pnas.1524136113](https://doi.org/10.1073/pnas.1524136113). URL: <http://www.pnas.org/content/113/22/6125.abstract>.
- 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).
- 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>.

- Liu, Y. J. et al. (2013). "Production of methyl vinyl ketone and methacrolein via the hydroperoxy pathway of isoprene oxidation". In: *Atmospheric Chemistry and Physics* 13.11, pp. 5715–5730. ISSN: 1680-7324. DOI: [10.5194/acp-13-5715-2013](https://doi.org/10.5194/acp-13-5715-2013). URL: <http://www.atmos-chem-phys.net/13/5715/2013/>.
- Palmer, Paul I et al. (2006). "Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column". In: *J.Geophys.Res.* 111, p. D12315. ISSN: 0148-0227. DOI: [10.1029/2005JD006689](https://doi.org/10.1029/2005JD006689). URL: <http://dx.doi.org/10.1029/2005JD006689>.
- Peeters, Jozef and Jean-Francis Muller (2010). "HOx 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>.
- 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).
- Chan, A. W. H. et al. (2010). "Role of aldehyde chemistry and NO<sub>x</sub> concentrations in secondary organic aerosol formation". In: *Atmospheric Chemistry and Physics* 10.15, pp. 7169–7188. ISSN: 1680-7324. DOI: [10.5194/acp-10-7169-2010](https://doi.org/10.5194/acp-10-7169-2010). URL: <http://www.atmos-chem-phys.net/10/7169/2010/>.
- Surratt, Jason D et al. (2010). "Reactive intermediates revealed in secondary organic aerosol formation from isoprene". In: *PNAS* 107.15, pp. 6640–6645. DOI: [10.1073/pnas.0911114107](https://doi.org/10.1073/pnas.0911114107). URL: <http://www.pnas.org/content/107/15/6640.full.pdf>.
- Lin, Ying-Hsuan et al. (2013). "Epoxide as a precursor to secondary organic aerosol formation from isoprene photooxidation in the presence of nitrogen oxides". In: DOI: [10.1073/pnas.1221150110](https://doi.org/10.1073/pnas.1221150110). URL: <http://www.pnas.org/content/110/17/6718.full.pdf>.
- Peeters, J., T. L. Nguyen, and L. Vereecken (2009). "HOx radical regeneration in the oxidation of isoprene". In: *Physical Chemistry Chemical Physics* 11.28, p. 5935. ISSN: 1463-9076. DOI: [10.1039/b908511d](https://doi.org/10.1039/b908511d). URL: <http://xlink.rsc.org/?DOI=b908511d>.
- Rollins, A W et al. (2009). "Isoprene oxidation by nitrate radical: alkyl nitrate and secondary organic aerosol yields". In: *Atmos. Chem. Phys. Atmospheric Chemistry and Physics* 9, pp. 6685–6703. URL: [www.atmos-chem-phys.net/9/6685/2009/](http://www.atmos-chem-phys.net/9/6685/2009/).
- Fortems-Cheiney, A. et al. (2012). "The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system". In: *Atmospheric Chemistry and Physics* 12.15, pp. 6699–6721. ISSN: 16807316. DOI: [10.5194/acp-12-6699-2012](https://doi.org/10.5194/acp-12-6699-2012). URL: <http://www.atmos-chem-phys.net/12/6699/2012/acp-12-6699-2012.pdf>.
- 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>.
- Andreae, M O (2001). "Emission of trace gases and aerosols from biomass burning". In: *Biogeochemistry* 15.4, pp. 955–966. URL: <http://onlinelibrary.wiley.com/doi/10.1029/2000GB001382/epdf>.

- 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>.
- Fu, Tzung-may et al. (2007). "Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone". In: 112, pp. 1–15. DOI: [10.1029/2006JD007853](https://doi.org/10.1029/2006JD007853).
- Emmerson, Kathryn M. et al. (2016). "Current estimates of biogenic emissions from eucalypts uncertain for southeast Australia". In: *Atmospheric Chemistry and Physics* 16.11, pp. 6997–7011. ISSN: 1680-7324. DOI: [10.5194/acp-16-6997-2016](https://doi.org/10.5194/acp-16-6997-2016). URL: <http://www.atmos-chem-phys.net/16/6997/2016/>.
- Levy, Hiram (1972). "Photochemistry of the lower troposphere". In: *Planetary and Space Science* 20.6, pp. 919–935. ISSN: 00320633. DOI: [10.1016/0032-0633\(72\)90177-8](https://doi.org/10.1016/0032-0633(72)90177-8).
- CRUTZEN, PAUL J., MARK G LAWRENCE, and ULRICH PÖSCHL (1999). "On the background photochemistry of tropospheric ozone". In: *Tellus A* 51.1, pp. 123–146. ISSN: 1600-0870. DOI: [10.1034/j.1600-0870.1999.t01-1-00010.x](https://doi.org/10.1034/j.1600-0870.1999.t01-1-00010.x). URL: <http://dx.doi.org/10.1034/j.1600-0870.1999.t01-1-00010.x>.
- Franco, B. et al. (2015). "Retrievals of formaldehyde from ground-based FTIR and MAX-DOAS observations at the Jungfraujoch station and comparisons with GEOS-Chem and IMAGES model simulations". In: *Atmospheric Measurement Techniques* 8.4, pp. 1733–1756. ISSN: 18678548. DOI: [10.5194/amt-8-1733-2015](https://doi.org/10.5194/amt-8-1733-2015).
- 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).
- Davenport, J. J. et al. (2015). "A measurement strategy for non-dispersive ultra-violet detection of formaldehyde in indoor air : spectral analysis and interferent gases". In: *Measurement Science and Technology* 015802. December 2015, p. 15802. ISSN: 0957-0233. DOI: [10.1088/0957-0233/27/1/015802](https://doi.org/10.1088/0957-0233/27/1/015802). URL: <http://dx.doi.org/10.1088/0957-0233/27/1/015802>.
- Lee, Anita et al. (2006a). "Gas-phase products and secondary aerosol yields from the ozonolysis of ten different terpenes". In: 111.Ci, pp. 1–18. DOI: [10.1029/2005JD006437](https://doi.org/10.1029/2005JD006437).
- Lerner, Brian M. et al. (2017). "An improved, automated whole air sampler and gas chromatography mass spectrometry analysis system for volatile organic compounds in the atmosphere". In: *Atmospheric Measurement Techniques* 10.1, pp. 291–313. ISSN: 18678548. DOI: [10.5194/amt-10-291-2017](https://doi.org/10.5194/amt-10-291-2017). URL: <https://www.atmos-meas-tech.net/10/291/2017/amt-10-291-2017.pdf>.
- Hak, C. et al. (2005). "Intercomparison of four different in-situ techniques for ambient formaldehyde measurements in urban air". In: *Atmospheric Chemistry and Physics Discussions* 5.3, pp. 2897–2945. ISSN: 1680-7316. DOI: [10.5194/acpd-5-2897-2005](https://doi.org/10.5194/acpd-5-2897-2005).

- EUMETSAT (2015). GOME2. URL: <http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/GOME2/index.html>.
- 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).
- Bauwens, M et al. (2013). "Satellite-based isoprene emission estimates (2007–2012) from the GlobEmission project". In: *Proceedings of the ACCENT-Plus Symposium, Atmospheric Composition Change-Policy Support and Science, Urbino*, pp. 17–20.
- 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/>.
- 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>.
- Zhang, Yang et al. (2012). "Impact of gas-phase mechanisms on Weather Research Forecasting Model with Chemistry (WRF/Chem) predictions: Mechanism implementation and comparative evaluation". In: *Journal of Geophysical Research: Atmospheres* 117.D1, n/a-n/a. DOI: [10.1029/2011JD015775](https://doi.org/10.1029/2011JD015775). URL: <http://doi.wiley.com/10.1029/2011JD015775>.
- 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>.
- Miller, C. et al. (2014). "Glyoxal retrieval from the Ozone Monitoring Instrument". In: *Atmospheric Measurement Techniques* 7.11, pp. 3891–3907. ISSN: 1867-8548. DOI: [10.5194/amt-7-3891-2014](https://doi.org/10.5194/amt-7-3891-2014). URL: <http://www.atmos-meas-tech.net/7/3891/2014/>.
- Zeng, G. et al. (2015). "Multi-model simulation of CO and HCHO in the Southern Hemisphere: comparison with observations and impact of biogenic emissions". In: *Atmospheric Chemistry and Physics* 15.13, pp. 7217–7245. ISSN: 1680-7324. DOI: [10.5194/acp-15-7217-2015](https://doi.org/10.5194/acp-15-7217-2015). URL: <http://www.atmos-chem-phys.net/15/7217/2015/>.
- Christian, Kenneth E, William H Brune, and Jingqiu Mao (2017). "Global sensitivity analysis of the GEOS-Chem chemical transport model: ozone and hydrogen oxides during ARCTAS (2008)". In: *Atmos. Chem. Phys* 17, pp. 3769–3784. DOI: [10.5194/acp-17-3769-2017](https://doi.org/10.5194/acp-17-3769-2017). URL: [www.atmos-chem-phys.net/17/3769/2017/](http://www.atmos-chem-phys.net/17/3769/2017/).
- 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).
- Arneth, Almut et al. (2007). "CO<sub>2</sub> inhibition of global terrestrial isoprene emissions: Potential implications for atmospheric chemistry". In: *Geophysical Research Letters* 34.18, p. L18813. DOI: [10.1029/2007GL030615](https://doi.org/10.1029/2007GL030615). URL: <http://doi.wiley.com/10.1029/2007GL030615>.

- Wild, Oliver and Michael J. Prather (2006). "Global tropospheric ozone modeling: Quantifying errors due to grid resolution". In: *Journal of Geophysical Research Atmospheres* 111.11, pp. 1–14. ISSN: 01480227. DOI: [10.1029/2005JD006605](https://doi.org/10.1029/2005JD006605).
- Travis, Katherine R et al. (2016). "Why do models overestimate surface ozone in the Southeast United States?" In: *Atmos. Chem. Phys* 16, pp. 13561–13577. DOI: [10.5194/acp-16-13561-2016](https://doi.org/10.5194/acp-16-13561-2016). URL: [www.atmos-chem-phys.net/16/13561/2016/](http://www.atmos-chem-phys.net/16/13561/2016/).
- Rowntree, P. R. and J. A. Bolton (1983). "Simulation of the atmospheric response to soil moisture anomalies over Europe". In: *Quarterly Journal of the Royal Meteorological Society* 109.461, pp. 501–526. ISSN: 00359009. DOI: [10.1002/qj.49710946105](https://doi.org/10.1002/qj.49710946105). URL: <http://doi.wiley.com/10.1002/qj.49710946105>.
- Chen, Fei and Jimy Dudhia (2001). "Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity". In: *Monthly Weather Review* 129. URL: <https://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%282001%29129%3C0569%3ACAALSH%3E2.0.CO%3B2>.
- 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](http://www.atmos-chem-phys.net/17/12827/2017/acp-17-12827-2017.pdf).
- VanDerA, R J et al. (2008). "Trends seasonal variability and dominant NO x source derived from a ten year record of NO 2 measured from space". In: *J. Geophys. Res.* 113.D4. DOI: [10.1029/2007jd009021](https://doi.org/10.1029/2007jd009021). URL: <http://dx.doi.org/10.1029/2007jd009021>.
- Oltmans, J et al. (2001). "Ozone in the Pacific tropical troposphere from ozonesonde observations". In: *Journal of Geophysical Research* 106.D23, pp. 32503–32525.
- Gloudemans, Annemieke et al. (2007). "Evidence for long-range transport of carbon monoxide in the Southern Hemisphere from SCIAMACHY observations". In: *European Space Agency, (Special Publication)* 33.SP-636, pp. 1–5. ISSN: 03796566. DOI: [10.1029/2006GL026804](https://doi.org/10.1029/2006GL026804).
- Edwards, D. P. et al. (2006). "Satellite-observed pollution from Southern Hemisphere biomass burning". In: *Journal of Geophysical Research* 111.14, pp. 1–17. ISSN: 01480227. DOI: [10.1029/2005JD006655](https://doi.org/10.1029/2005JD006655).
- Pak, B.C.a et al. (2003). "Measurements of biomass burning influences in the troposphere over southeast Australia during the SAFARI 2000 dry season campaign". In: *Journal of Geophysical Research* 108.13, pp. 1–10. ISSN: 0148-0227. DOI: [10.1029/2002JD002343](https://doi.org/10.1029/2002JD002343). URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0742322536&partnerID=40&md5=cafaf03b948fb456696583ed3ab9a5>.
- Liu, Junhua et al. (2016a). "Causes of interannual variability of tropospheric ozone over the Southern Ocean". In: *Atmospheric Chemistry and Physics Discussions* October, pp. 1–46. ISSN: 1680-7316. DOI: [10.5194/ACP-2016-692](https://doi.org/10.5194/ACP-2016-692).
- Winters, Anthony J et al. (2009). "Emissions of isoprene, monoterpene and short-chained carbonyl compounds from Eucalyptus spp. in southern Australia". In: *Atmospheric Environment* 43.19, pp. 3035–3043. ISSN: 13522310. DOI: [10.1016/j.atmosenv.2009.03.026](https://doi.org/10.1016/j.atmosenv.2009.03.026).
- Alexander, S. P. et al. (2013). "High resolution VHF radar measurements of tropopause structure and variability at Davis, Antarctica (69 S, 78 E)". In: *Atmospheric Chemistry*

- and Physics* 13.6, pp. 3121–3132. ISSN: 16807324. DOI: [10.5194/acp-13-3121-2013](https://doi.org/10.5194/acp-13-3121-2013). URL: <http://www.atmos-chem-phys.net/13/3121/2013/>.
- Müller, J.-F. et al. (2008). “Global isoprene emissions estimated using MEGAN ECMWF analyses and a detailed canopy environment model”. In: *Atmospheric Chemistry and Physics Discussions* 7.6, pp. 15373–15407. DOI: [10.5194/acpd-7-15373-2007](https://doi.org/10.5194/acpd-7-15373-2007). URL: <http://dx.doi.org/10.5194/acpd-7-15373-2007>.
- Stavrakou, T et al. (2009). “Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns”. In: *Atmospheric Chemistry and Physics* 9.3, pp. 1037–1060. DOI: [10.5194/acp-9-1037-2009](https://doi.org/10.5194/acp-9-1037-2009). URL: <http://dx.doi.org/10.5194/acp-9-1037-2009>.
- 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.
- Brinksma, E. J. et al. (2002). “Five years of observations of ozone profiles over Lauder, New Zealand”. In: *Journal of Geophysical Research* 107.D14, pp. 1–11. ISSN: 0148-0227. DOI: [10.1029/2001JD000737](https://doi.org/10.1029/2001JD000737). URL: <http://doi.wiley.com/10.1029/2001JD000737>.
- Baray, Jean-Luc et al. (2012). “One year ozonesonde measurements at Kerguelen Island (49.2S, 70.1E): Influence of stratosphere-to-troposphere exchange and long-range transport of biomass burning plumes”. In: *Journal of Geophysical Research* 117.D6. ISSN: 2156-2202. DOI: [10.1029/2011JD016717](https://doi.org/10.1029/2011JD016717). URL: <http://dx.doi.org/10.1029/2011JD016717>.
- 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/acpd-3-895-2003](https://doi.org/10.5194/acpd-3-895-2003). URL: <http://dx.doi.org/10.5194/acpd-3-895-2003>.
- Lamsal, L N et al. (2014). “Evaluation of OMI operational standard NO<sub>2</sub> column retrievals using in situ and surface-based NO<sub>2</sub> 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/](http://www.atmos-chem-phys.net/14/11587/2014/).
- 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/>.
- 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>.
- 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<sub>x</sub> 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>.

- 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/> 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.net/10/1957/2017/](http://www.atmos-meas-tech-discuss.net/amt-2016-420/).
- Lorente, Alba et al. (2017). "Structural uncertainty in air mass factor calculation for NO<sub>2</sub> 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/amt-2016-306/](https://www.atmos-meas-tech.net/10/759/2017/amt-10-759-2017.html).
- Instrument, O M I (2002). "OMI Algorithm Theoretical Basis Document Volume I". In: I.August, pp. 1–50.
- 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).
- 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>.
- Kurosu, T and K Chance (2014). OMIReadme. URL: [https://www.cfa.harvard.edu/atmosphere/instruments/omi/pgereleases/readmes/omhcho\\_readme\\_v3.0.pdf](https://www.cfa.harvard.edu/atmosphere/instruments/omi/pgereleases/readmes/omhcho_readme_v3.0.pdf).
- Texeira, Joao (2013). AIRS/Aqua L3 Daily Standard Physical Retrieval (AIRS-only) 1 degree x 1 degree V006: Accessed 2/Dec/2015. DOI: [doi:10.5067/AQUA/AIRS/DATA303](https://doi.org/10.5067/AQUA/AIRS/DATA303).
- Dee, D P et al. (2011). "The ERA-Interim reanalysis: configuration and performance of the data assimilation system". In: *Quarterly Journal of the Royal Meteorological Society* 137.656, pp. 553–597. ISSN: 1477-870X. DOI: [10.1002/qj.828](https://doi.org/10.1002/qj.828). URL: <http://dx.doi.org/10.1002/qj.828>.
- 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/acpd-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>.
- 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>.
- 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> <https://www.nature.com/articles/ngeo1405.pdf>.
- 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*.
- Guenther, Alex (2016). MEGAN. URL: <http://lar.wsu.edu/megan/>.
- Jin, Xiaomeng et al. (2017). "Evaluating a Space-Based Indicator of Surface Ozone-NO<sub>x</sub>-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).
- 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 NO<sub>2</sub> 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/>.
- 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>.
- Thomas, W et al. (1998). "Detection of biomass burning combustion products in Southeast Asia from backscatter data taken by the GOME spectrometer". In: *Geophysical Research Letters* 25.9, pp. 1317–1320. DOI: [10.1029/98GL01087](https://doi.org/10.1029/98GL01087). URL: <http://onlinelibrary.wiley.com/doi/10.1029/98GL01087/epdf>.
- Abbot, Dorian S. (2003). "Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space". In: *Geophysical Research Letters* 30.17, pp. 1999–2002. ISSN: 0094-8276. DOI: [10.1029/2003GL017336](https://doi.org/10.1029/2003GL017336). URL: <http://doi.wiley.com/10.1029/2003GL017336>.
- Stavrakou, T. et al. (2015). "How consistent are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI?" English. In: *Atmospheric Chemistry and Physics* 15.20, pp. 11861–11884. ISSN: 1680-7324. DOI: [10.5194/acp-15-11861-2015](https://doi.org/10.5194/acp-15-11861-2015). URL: <http://www.atmos-chem-phys.net/15/11861/2015/acp-15-11861-2015.html>.
- Curci, G. et al. (2010). "Estimating European volatile organic compound emissions using satellite observations of formaldehyde from the Ozone Monitoring Instrument". In: *Atmospheric Chemistry and Physics* 10.23, pp. 11501–11517. ISSN: 16807316. DOI: [10.5194/acp-10-11501-2010](https://doi.org/10.5194/acp-10-11501-2010).
- Lee, Anita et al. (2006b). "Gas-phase products and secondary aerosol yields from the photooxidation of 16 different terpenes". In: *Journal of Geophysical Research Atmospheres* 111.17, pp. 1–18. ISSN: 01480227. DOI: [10.1029/2006JD007050](https://doi.org/10.1029/2006JD007050).
- Pegoraro, E. et al. (2004). "Effect of drought on isoprene emission rates from leaves of *Quercus virginiana* Mill." In: *Atmospheric Environment* 38.36, pp. 6149–6156. ISSN: 13522310. DOI: [10.1016/j.atmosenv.2004.07.028](https://doi.org/10.1016/j.atmosenv.2004.07.028). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1352231004007198>.
- Škerlak, B, M Sprenger, and H Wernli (2014). "A global climatology of stratosphere-troposphere exchange using the ERA-Interim data set from 1979 to 2011". In: *Atmospheric Chemistry and Physics* 14.2, pp. 913–937. DOI: [10.5194/acp-14-913-2014](https://doi.org/10.5194/acp-14-913-2014). URL: <http://www.atmos-chem-phys.net/14/913/2014/>.

- Liu, Junhua et al. (2015). "Origins of tropospheric ozone interannual variation over Réunion: A model investigation". In: *Journal of Geophysical Research*, pp. 1–19. DOI: [10.1002/2015JD023981](https://doi.org/10.1002/2015JD023981). URL: <http://onlinelibrary.wiley.com/doi/10.1002/2015JD023981/abstract>.
- Thompson, A. M. et al. (2014). "Tropospheric ozone increases over the southern Africa region: Bellwether for rapid growth in Southern Hemisphere pollution?" In: *Atmospheric Chemistry and Physics* 14.18, pp. 9855–9869. ISSN: 16807324. DOI: [10.5194/acp-14-9855-2014](https://doi.org/10.5194/acp-14-9855-2014).
- Danielsen, Edwin F. (1968). *Stratospheric-Tropospheric Exchange Based on Radioactivity, Ozone and Potential Vorticity*. DOI: [10.1175/1520-0469\(1968\)025<0502:STEBOR>2.0.CO;2](https://doi.org/10.1175/1520-0469(1968)025<0502:STEBOR>2.0.CO;2).
- Lefohn, Allen S. et al. (2011). "The importance of stratospheric-tropospheric transport in affecting surface ozone concentrations in the western and northern tier of the United States". In: *Atmospheric Environment* 45.28, pp. 4845–4857. ISSN: 13522310. DOI: [10.1016/j.atmosenv.2011.06.014](https://doi.org/10.1016/j.atmosenv.2011.06.014). URL: <http://dx.doi.org/10.1016/j.atmosenv.2011.06.014>.
- Langford, A. O. et al. (2012). "Stratospheric influence on surface ozone in the Los Angeles area during late spring and early summer of 2010". In: *Journal of Geophysical Research* 117.3, pp. 1–17. ISSN: 01480227. DOI: [10.1029/2011JD016766](https://doi.org/10.1029/2011JD016766).
- Zhang, L et al. (2014). "Sources contributing to background surface ozone in the US Intermountain West". In: *Atmospheric Chemistry and Physics* 14.11, pp. 5295–5309. DOI: [10.5194/acp-14-5295-2014](https://doi.org/10.5194/acp-14-5295-2014). URL: <http://dx.doi.org/10.5194/acp-14-5295-2014>.
- Lin, Meiyun et al. (2012). "Springtime high surface ozone events over the western United States: Quantifying the role of stratospheric intrusions". In: *Journal of Geophysical Research* 117.19, pp. 1–20. ISSN: 01480227. DOI: [10.1029/2012JD018151](https://doi.org/10.1029/2012JD018151).
- Galani, E. (2003). "Observations of stratosphere-to-troposphere transport events over the eastern Mediterranean using a ground-based lidar system". In: *Journal of Geophysical Research* 108.D12, pp. 1–10. ISSN: 0148-0227. DOI: [10.1029/2002JD002596](https://doi.org/10.1029/2002JD002596). URL: <http://www.agu.org/pubs/crossref/2003/2002JD002596.shtml>.
- Terao, Yukio et al. (2008). "Contribution of stratospheric ozone to the interannual variability of tropospheric ozone in the northern extratropics". In: *Journal of Geophysical Research* 113.D18. DOI: [10.1029/2008jd009854](https://doi.org/10.1029/2008jd009854). URL: <http://dx.doi.org/10.1029/2008jd009854>.
- Tang, Q. and M. J. Prather (2012). "Five blind men and the elephant: What can the NASA Aura ozone measurements tell us about stratosphere-troposphere exchange?" In: *Atmospheric Chemistry and Physics* 12.5, pp. 2357–2380. ISSN: 16807316. DOI: [10.5194/acp-12-2357-2012](https://doi.org/10.5194/acp-12-2357-2012). URL: <http://dx.doi.org/10.5194/acpd-11-26897-2011>.
- Frey, W. et al. (2015). "The impact of overshooting deep convection on local transport and mixing in the tropical upper troposphere/lower stratosphere (UTLS)". In: *Atmospheric Chemistry and Physics* 15.11, pp. 6467–6486. ISSN: 1680-7324. DOI: [10.5194/acp-15-6467-2015](https://doi.org/10.5194/acp-15-6467-2015). URL: <http://www.atmos-chem-phys.net/15/6467/2015/>.
- Das, Siddarth Shankar et al. (2016). "Influence of tropical cyclones on tropospheric ozone: possible implications". In: *Atmospheric Chemistry and Physics* 16, pp. 4837–

4847. DOI: [10.5194/acp-16-4837-2016](https://doi.org/10.5194/acp-16-4837-2016). URL: [www.atmos-chem-phys.net/16/4837/2016/](http://www.atmos-chem-phys.net/16/4837/2016/).
- Mihalikova, M et al. (2012). "Observation of a tropopause fold by MARA VHF wind-profiler radar and ozonesonde at Wasa, Antarctica: comparison with ECMWF analysis and a WRF model simulation". In: *Annales Geophysicae* 30.9, pp. 1411–1421. DOI: [10.5194/angeo-30-1411-2012](https://doi.org/10.5194/angeo-30-1411-2012). URL: <http://www.ann-geophys.net/30/1411/2012/>.
- Vaughan, G., J. D. Price, and A. Howells (1993). "Transport into the troposphere in a tropopause fold". In: *Quarterly Journal of the Royal Meteorological Society* 120.518, pp. 1085–1103. ISSN: 00359009. DOI: [10.1002/qj.49712051814](https://doi.org/10.1002/qj.49712051814).
- Beekmann, M. et al. (1997). "Regional and global tropopause fold occurrence and related ozone flux across the tropopause". In: *Journal of Atmospheric Chemistry* 28.1-3, pp. 29–44. ISSN: 01677764. DOI: [10.1023/A:1005897314623](https://doi.org/10.1023/A:1005897314623).
- Baray, J. L. et al. (2000). "Planetary-scale tropopause folds in the southern subtropics". In: *Geophysical Research Letters* 27.3, pp. 353–356. ISSN: 00948276. DOI: [10.1029/1999GL010788](https://doi.org/10.1029/1999GL010788).
- Zanis, P. et al. (2014). "Summertime free-tropospheric ozone pool over the eastern Mediterranean/middle east". In: *Atmospheric Chemistry and Physics* 14.1, pp. 115–132. ISSN: 16807316. DOI: [10.5194/acp-14-115-2014](https://doi.org/10.5194/acp-14-115-2014).
- Akritidis, Dimitris et al. (2016). "On the role of tropopause folds in summertime tropospheric ozone over the eastern Mediterranean and the Middle East". In: *Atmospheric Chemistry and Physics* 16.21, pp. 14025–14039. DOI: [10.5194/acp-16-14025-2016](https://doi.org/10.5194/acp-16-14025-2016). URL: <http://www.atmos-chem-phys.net/16/14025/2016/>.
- Tyrlis, Evangelos et al. (2014). "On the linkage between the Asian summer monsoon and tropopause fold activity over the eastern Mediterranean and the Middle East". In: *Journal of Geophysical Research* 119.6, pp. 3202–3221. ISSN: 2169897X. DOI: [10.1002/2013JD021113](https://doi.org/10.1002/2013JD021113). URL: <http://doi.wiley.com/10.1002/2013JD021113>.
- Price, J. D. and G. Vaughan (1993). "The potential for stratosphere-troposphere exchange in cut-off-low systems". In: *Quarterly Journal of the Royal Meteorological Society* 119.510, pp. 343–365. DOI: [10.1002/qj.49711951007](https://doi.org/10.1002/qj.49711951007). URL: <http://onlinelibrary.wiley.com/doi/10.1002/qj.49711951007/abstract>.
- Wirth, Volkmar (1995). "Diabatic heating in an axisymmetric cut-off cyclone and related stratosphere-troposphere exchange". In: *Quarterly Journal of the Royal Meteorological Society* 121.521, pp. 127–147. ISSN: 00359009. DOI: [10.1002/qj.49712152107](https://doi.org/10.1002/qj.49712152107). URL: <http://doi.wiley.com/10.1002/qj.49712152107>.
- Trickl, T. et al. (2014). "How stratospheric are deep stratospheric intrusions?" In: *Atmospheric Chemistry and Physics* 14.18, pp. 9941–9961. ISSN: 16807324. DOI: [10.5194/acp-14-9941-2014](https://doi.org/10.5194/acp-14-9941-2014).
- Mze, N. et al. (2010). "Climatology and comparison of ozone from ENVISAT/GOMOS and SHADOZ/balloon-sonde observations in the southern tropics". In: *Atmospheric Chemistry and Physics* 10.16, pp. 8025–8035. ISSN: 16807316. DOI: [10.5194/acp-10-8025-2010](https://doi.org/10.5194/acp-10-8025-2010).
- Smit, Herman G J et al. (2007). "Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE)". In: *Journal of Geophysical Research* 112.19, pp. 1–18. ISSN: 01480227. DOI: [10.1029/2006JD007308](https://doi.org/10.1029/2006JD007308).

- WMO, World Meteorological Organization (1957). "Meteorology A Three-Dimensional Science". In: *Geneva, Second Session of the Commission for Aerology* 4, pp. 134–138.
- Bethan, S., G. Vaughan, and S. J. Reid (1996). "A comparison of ozone and thermal tropopause heights and the impact of tropopause definition on quantifying the ozone content of the troposphere". In: *Quarterly Journal of the Royal Meteorological Society* 122.532, pp. 929–944. ISSN: 00359009. DOI: [10.1002/qj.49712253207](https://doi.wiley.com/10.1002/qj.49712253207). URL: <http://doi.wiley.com/10.1002/qj.49712253207>.
- Tomikawa, Yoshihiro, Yashiro Nishimura, and Takashi Yamanouchi (2009). "Characteristics of Tropopause and Tropopause Inversion Layer in the Polar Region". In: *SOLA* 5, pp. 141–144. DOI: [10.2151/sola.2009-036](https://doi.org/10.2151/sola.2009-036). URL: <http://dx.doi.org/10.2151/sola.2009-036>.
- Lelieveld, Jos and Frank J. Dentener (2000). "What controls tropospheric ozone?" In: *Journal of Geophysical Research* 105.D3, pp. 3531–3551. ISSN: 01480227. DOI: [10.1029/1999JD901011](https://doi.wiley.com/10.1029/1999JD901011). URL: <http://doi.wiley.com/10.1029/1999JD901011>.
- Struthers, H. et al. (2004). "Past and future simulations of NO<sub>2</sub> from a coupled chemistry-climate model in comparison with observations". In: *Atmospheric Chemistry and Physics* 4.8, pp. 2227–2239. ISSN: 1680-7324. DOI: [10.5194/acp-4-2227-2004](https://doi.org/10.5194/acp-4-2227-2004). URL: <http://www.atmos-chem-phys.net/4/2227/2004/>.
- Bey, Isabelle et al. (2001). "Global Modeling of Tropospheric Chemistry with Assimilated Meteorology: Model Description and Evaluation". In: *Journal of Geophysical Research* 106, pp. 73–95. ISSN: 0148-0227. DOI: [10.1029/2001JD000807](https://doi.org/10.1029/2001JD000807).
- Eastham, Sebastian D., Debra K. Weisenstein, and Steven R H Barrett (2014). "Development and evaluation of the unified tropospheric-stratospheric chemistry extension (UCX) for the global chemistry-transport model GEOS-Chem". In: *Atmospheric Environment* 89, pp. 52–63. ISSN: 13522310. DOI: [10.1016/j.atmosenv.2014.02.001](https://doi.org/10.1016/j.atmosenv.2014.02.001). URL: <http://dx.doi.org/10.1016/j.atmosenv.2014.02.001>.
- Rienecker, Michele (2007). "File Specification for GEOS-5 DAS Gridded Output". In: pp. 1–54. URL: [https://gmao.gsfc.nasa.gov/products/documents/GEOS-5.1.0\\_File\\_Specification.pdf](https://gmao.gsfc.nasa.gov/products/documents/GEOS-5.1.0_File_Specification.pdf).
- Press, William H et al. (1992). *Numerical Recipes in C (2Nd Ed.): The Art of Scientific Computing*. New York, NY, USA: Cambridge University Press. ISBN: 0-521-43108-5.
- Tang, Q. and M. J. Prather (2010). "Correlating tropospheric column ozone with tropopause folds: The Aura-OMI satellite data". In: *Atmospheric Chemistry and Physics* 10.19, pp. 9681–9688. ISSN: 16807316. DOI: [10.5194/acp-10-9681-2010](https://doi.org/10.5194/acp-10-9681-2010).
- Jaffe, Daniel a. and Nicole L. Wigder (2012). "Ozone production from wildfires: A critical review". In: *Atmospheric Environment* 51, pp. 1–10. ISSN: 13522310. DOI: [10.1016/j.atmosenv.2011.11.063](https://doi.org/10.1016/j.atmosenv.2011.11.063). URL: <http://dx.doi.org/10.1016/j.atmosenv.2011.11.063>.
- Edwards, D. P. (2003). "Tropospheric ozone over the tropical Atlantic: A satellite perspective". In: *Journal of Geophysical Research* 108.D8, p. 4237. ISSN: 0148-0227. DOI: [10.1029/2002JD002927](https://doi.wiley.com/10.1029/2002JD002927). URL: <http://doi.wiley.com/10.1029/2002JD002927>.
- Sinha, Parikhit et al. (2004). "Transport of biomass burning emissions from southern Africa". In: *Journal of Geophysical Research* 109, p. D20204. ISSN: 01480227. DOI: [10.1029/2004JD005044](https://doi.org/10.1029/2004JD005044).
- Mari, C H et al. (2008). "Tracing biomass burning plumes from the Southern Hemisphere during the AMMA 2006 wet season experiment, Atmos". In: *Atmospheric*

- Chemistry and Physics* 8, pp. 3951–3961. ISSN: 1680-7324. DOI: [10.5194/acpd-7-17339-2007](https://doi.org/10.5194/acpd-7-17339-2007).
- Reutter, P. et al. (2015). "Stratosphere-troposphere exchange (STE) in the vicinity of North Atlantic cyclones". In: *Atmospheric Chemistry and Physics* 15.19, pp. 10939–10953. ISSN: 16807324. DOI: [10.5194/acp-15-10939-2015](https://doi.org/10.5194/acp-15-10939-2015).
- Škerlak, Bojan et al. (2015). "Tropopause folds in ERA-Interim: Global climatology and relation to extreme weather events". In: *Journal of Geophysical Research* 120.10, pp. 4860–4877. ISSN: 21698996. DOI: [10.1002/2014JD022787](https://doi.org/10.1002/2014JD022787).
- Wauben, Wiel M F, J Paul F Fortuin, and Peter F J Van Velthoven (1998). "Comparison of modeled ozone distributions observations". In: *Journal of Geophysical Research* 103, pp. 3511–3530.
- Hu, Lu et al. (2017). "Global budget of tropospheric ozone: evaluating recent model advances with satellite (OMI), aircraft (IAGOS), and ozonesonde observations". In: *Atmospheric Environment*, pp. 1–36. DOI: [10.1016/j.atmosenv.2017.08.036](https://doi.org/10.1016/j.atmosenv.2017.08.036).
- Roelofs, Geert Jan and Jos Lelieveld (1997). *Model study of the influence of cross-tropopause O<sub>3</sub> transports on tropospheric O<sub>3</sub> levels*. DOI: [10.1034/j.1600-0889.49.issue1.3.x](https://doi.org/10.1034/j.1600-0889.49.issue1.3.x).
- Elbern, H, J Hendricks, and A Ebel (1998). "A Climatology of Tropopause Folds by Global Analyses". In: *Theoretical and Applied Climatology* 59.3, pp. 181–200. ISSN: 1434-4483. DOI: [10.1007/s007040050023](https://doi.org/10.1007/s007040050023). URL: <http://dx.doi.org/10.1007/s007040050023>.
- Olsen, Mark a. (2003). "A comparison of Northern and Southern Hemisphere cross-tropopause ozone flux". In: *Geophysical Research Letters* 30.7, p. 1412. ISSN: 0094-8276. DOI: [10.1029/2002GL016538](https://doi.org/10.1029/2002GL016538). URL: <http://doi.wiley.com/10.1029/2002GL016538>.
- Lee, Hanlim et al. (2015). "Investigations of the Diurnal Variation of Vertical HCHO Profiles Based on MAX-DOAS Measurements in Beijing: Comparisons with OMI Vertical Column Data". In: *Atmosphere*. URL: [10.3390/atmos6111816](https://doi.org/10.3390/atmos6111816).
- Schreier, Stefan F. et al. (2016). "Estimates of free-Tropospheric NO<sub>2</sub> and HCHO mixing ratios derived from high-Altitude mountain MAX-DOAS observations at mid-latitudes and in the tropics". In: *Atmospheric Chemistry and Physics* 16.5. ISSN: 16807324. DOI: [10.5194/acp-16-2803-2016](https://doi.org/10.5194/acp-16-2803-2016).
- SPEI Drought Index. URL: <http://spei.csic.es/home.html> (visited on 12/19/2017).
- Chance, K. et al. (2000). "Satellite observations of formaldehyde over North America from GOME". In: *Geophysical Research Letters* 27.21, pp. 3461–3464. ISSN: 00948276. DOI: [10.1029/2000GL011857](https://doi.org/10.1029/2000GL011857). URL: <http://dx.doi.org/10.1029/2000gl011857>.
- Bei, N, G Li, and L T Molina (2012). "Uncertainties in SOA simulations due to meteorological uncertainties in Mexico City during MILAGRO-2006 field campaign". In: *Atmospheric Chemistry and Physics* 12.23, pp. 11295–11308. DOI: [10.5194/acp-12-11295-2012](https://doi.org/10.5194/acp-12-11295-2012). URL: <http://www.atmos-chem-phys.net/12/11295/2012/>.
- Cao, Hansen et al. (2018). "Adjoint inversion of Chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal". In: *Atmospheric Chemistry and Physics (Discussions)*. DOI: [10.5194/acp-2017-1136](https://doi.org/10.5194/acp-2017-1136). URL: <https://www.atmos-chem-phys-discuss.net/acp-2017-1136/acp-2017-1136.pdf>.

- Müller, J. F., J. Peeters, and T. Stavrakou (2014). "Fast photolysis of carbonyl nitrates from isoprene". In: *Atmospheric Chemistry and Physics* 14.5, pp. 2497–2508. ISSN: 16807316. DOI: [10.5194/acp-14-2497-2014](https://doi.org/10.5194/acp-14-2497-2014).
- Sander, R et al. (2005). "Technical note: The new comprehensive atmospheric chemistry module MECCA". In: *Atmospheric Chemistry and Physics* 5.2, pp. 445–450. ISSN: 1680-7324. DOI: [10.5194/acp-5-445-2005](https://doi.org/10.5194/acp-5-445-2005). URL: <http://www.atmos-chem-phys.net/5/445/2005/>.
- Sandu, A and R Sander (2006). "Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1". In: *Atmospheric Chemistry and Physics* 6.1, pp. 187–195. DOI: [10.5194/acp-6-187-2006](https://doi.org/10.5194/acp-6-187-2006). URL: <http://www.atmos-chem-phys.net/6/187/2006/>.
- Jöckel, P et al. (2006). "The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere". In: *Atmospheric Chemistry and Physics* 6.12, pp. 5067–5104. DOI: [10.5194/acp-6-5067-2006](https://doi.org/10.5194/acp-6-5067-2006). URL: <http://www.atmos-chem-phys.net/6/5067/2006/>.
- Jöckel, Patrick, Rolf Sander, and Jos Lelieveld (2004). "Technical Note: The Modular Earth Submodel System (MESSy) – a new approach towards Earth System Modeling". In: *Atmospheric Chemistry and Physics Discussions* 4.6, pp. 7139–7166. ISSN: 1680-7324. DOI: [10.5194/acpd-4-7139-2004](https://doi.org/10.5194/acpd-4-7139-2004).
- Giglio, Louis, Ivan Csiszar, and Christopher O. Justice (2006). "Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors". In: *Journal of Geophysical Research: Biogeosciences* 111.2, pp. 1–12. ISSN: 01480227. DOI: [10.1029/2005JG000142](https://doi.org/10.1029/2005JG000142).