

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**

January 29, 2019



## 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

Ozone in the troposphere is a toxic pollutant which causes respiratory and agricultural damage. The two main precursors are biogenic emissions of volatile organic compounds (BVOCs) and transport from the stratosphere. Most tropospheric ozone is formed through chemical reactions involving nitrogen oxides, the hydroxy radical, and volatile organic compounds (such as isoprene). Atmospheric chemistry transport models (CTMs) have uncertain BVOC emissions in Australia, affecting our ability to model and predict ozone production along with other important atmospheric processes. Isoprene, formaldehyde, and ozone in the troposphere are linked by oxidative chemistry and are all important to air quality, climate, and radiation budgets. My thesis has three aims: recalculating formaldehyde amounts over Australia seen by satellite using a global CTM (GEOS-Chem), determination of isoprene emissions using modelled formaldehyde yields along with satellite formaldehyde amounts, and attribution of ozone in the troposphere (from either chemical production following isoprene emissions or transported from the stratosphere). Model and observations are combined for each aim in this thesis.

To quantify isoprene emissions (the dominant BVOC), formaldehyde (one of the main products of isoprene oxidation) is used as a proxy. Formaldehyde observed by satellite over Australia is calculated based on modelled a priori vertical distributions using both CTMs and radiative transfer models. In order to compare satellite formaldehyde products against models or other measurements, corrections are required to remove the influence of the a priori profile. Impacts on formaldehyde levels from anthropogenic and pyrogenic sources are determined and filtered out to determine the biogenic footprint and minimise bias in isoprene emissions quantification.

Isoprene is predominantly emitted by trees and shrubs, and Eucalypts are potentially very high emitters along with being close to population centres in Australia. Subsequent oxidation reactions form formaldehyde which has a sufficiently long lifetime in the atmosphere to establish chemical equilibrium. Using a simple linear model and assuming minimal transport and a chemical steady state allows an estimate of the yield of formaldehyde from isoprene emissions. This yield is modelled over Australia and then applied to the recalculated satellite formaldehyde to create a new estimate of Australian isoprene emissions. This technique is used to improve isoprene emissions estimates without the need for extensive measurement campaigns. Results are compared against existing measurement campaigns for validation and analysis. Isoprene emissions from Australian forests appear substantially lower than previous estimates have predicted.

The second most abundant source of tropospheric ozone is the stratosphere, which occasionally mixes into the troposphere bringing ozone-rich air masses down towards the earth's surface. Analysing the local weather patterns and ozone seasonality, most transport is seen to occur during low pressure frontal weather systems. This work provides a novel technique using a Fourier filter on ozone profiles for estimation and quantification of tropospheric ozone transported from the stratosphere. An estimate encompassing three measurement stations over the Southern Ocean near Australia of about  $7.2 \times 10^{17}$  molec cm<sup>-2</sup> per year is derived.

Overall, this work should improve the knowledge of tropospheric ozone and its precursors for Australia.

## *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>viii</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 . . . . .	17
1.3.3.2 High NO <sub>x</sub> pathway . . . . .	17
1.3.3.3 Low NO <sub>x</sub> pathway . . . . .	18
1.3.3.4 Night time processes . . . . .	19
1.4 Formaldehyde . . . . .	19
1.4.1 Sources and sinks . . . . .	19
1.4.2 Measurement techniques . . . . .	21
1.4.2.1 Satellite measurements . . . . .	22
1.5 Atmospheric Chemistry Modelling . . . . .	23
1.5.1 Box models . . . . .	24
1.5.2 Chemical transport models . . . . .	24
1.5.3 Emissions . . . . .	25
1.5.4 Uncertainties . . . . .	26
1.5.4.1 Emissions Inventories . . . . .	26
1.5.4.2 Resolution . . . . .	27
1.5.4.3 Chemistry mechanisms . . . . .	27
1.5.4.4 Clouds . . . . .	28
1.5.4.5 Soil Moisture . . . . .	28
1.6 Australia and the southern hemisphere . . . . .	28

1.6.1	Ozone	30
1.6.2	VOCs	31
1.6.3	Measurements	32
1.7	Aims	32
<b>2</b>	<b>Data and Modelling</b>	<b>35</b>
2.1	Introduction	35
2.2	Datasets	36
2.2.1	Satellite	36
2.2.1.1	Formaldehyde	37
2.2.1.2	Nitrogen dioxide	37
2.2.1.3	Aerosol optical depth	38
2.2.1.4	Active fires	40
2.2.1.5	Carbon monoxide	40
2.2.1.6	Uncertainties	40
2.2.2	Model datasets	41
2.2.2.1	GEOS-Chem output	41
2.2.2.2	Meteorological reanalysis	41
2.2.2.3	Surface temperatures	41
2.2.3	Campaign datasets	42
2.2.3.1	Measurements of Urban, Marine and Biogenic Air (MUMBA)	42
2.2.3.2	Sydney Particle Studies (SPS1, SPS2)	45
2.2.3.3	Ozonesondes	45
2.2.3.4	Wollongong	47
2.2.3.5	Uncertainties	47
2.3	Satellite formaldehyde	47
2.3.1	Pixel filtering	50
2.3.2	DOAS	50
2.3.3	Air mass factor (AMF)	52
2.3.4	LIDORT	53
2.3.5	Uncertainty	54
2.4	GEOS-Chem	56
2.4.1	Overview	56
2.4.2	Installing and running GEOS-Chem	57
2.4.3	Chemical Mechanism	57
2.4.4	GEOS-Chem isoprene	58
2.4.4.1	Oxidation	59
2.4.4.2	Nitrogen oxide impacts	60
2.4.4.3	OH	60
2.4.5	Emissions from MEGAN	61
2.4.6	Nitrogen oxides	62
2.4.7	GEOS-Chem simulations	63
2.4.7.1	GEOS-Chem outputs	63
2.4.7.2	GEOS-Chem runs	67
2.4.7.3	UCX vs tropchem	68
2.5	Calculating an AMF	73
2.6	Recalculation of OMI HCHO	76

2.6.1	Outline . . . . .	76
2.6.2	Creating new shape factors . . . . .	78
2.6.3	Reading satellite data . . . . .	80
2.6.4	Creating the new AMF . . . . .	80
2.6.5	Recalculating the AMF using PP code . . . . .	81
2.6.6	Vertical columns from AMF . . . . .	82
2.6.7	Reference sector correction . . . . .	82
2.6.8	Binning the results daily . . . . .	85
2.6.9	Difference between new and old OMI HCHO columns . . . . .	86
2.7	Filtering Data . . . . .	86
2.7.1	Fire and smoke . . . . .	89
2.7.1.1	Checking that fire masks are influencing pyrogenic HCHO	94
2.7.2	NO <sub>x</sub> . . . . .	101
2.8	Data Access . . . . .	101
<b>3</b>	<b>Biogenic Isoprene emissions in Australia</b>	<b>109</b>
3.1	Introduction . . . . .	109
3.1.1	Aims . . . . .	110
3.1.2	Existing emissions estimates . . . . .	110
3.1.3	Top-down isoprene emissions estimates . . . . .	111
3.1.3.1	Bayesian . . . . .	111
3.1.3.2	Linear . . . . .	112
3.2	Methods . . . . .	113
3.2.1	Outline . . . . .	113
3.2.2	Masks and reprocessed satellite HCHO . . . . .	115
3.2.3	GEOS-Chem emissions . . . . .	117
3.2.4	Relationship between isoprene emissions and formaldehyde . . . . .	117
3.2.5	Calculation of modelled slope . . . . .	118
3.2.6	Calculation of Emissions . . . . .	120
3.2.7	Accounting for smearing . . . . .	124
3.2.7.1	Calculation of smearing . . . . .	125
3.2.7.2	Sensitivity to smearing . . . . .	126
3.2.7.3	Smearing length scale . . . . .	128
3.2.7.4	NO <sub>x</sub> dependence . . . . .	130
3.2.8	Running GEOS-Chem using a postiori emissions . . . . .	130
3.3	Results . . . . .	133
3.3.1	Updated emissions . . . . .	135
3.3.2	A postiori GEOS-Chem results . . . . .	135
3.3.3	Trends . . . . .	143
3.3.4	Literature comparisons . . . . .	143
3.3.5	Comparison with available measurements . . . . .	143
3.3.6	HCHO Products and yield . . . . .	147
3.4	Uncertainty . . . . .	147
3.4.1	Summary . . . . .	147
3.4.2	Top down emissions . . . . .	149
3.4.3	Model Uncertainty . . . . .	150
3.4.4	Satellite Uncertainty . . . . .	151

3.4.4.1	OMI Retrieval . . . . .	151
3.4.4.2	Satellite vertical column recalculation . . . . .	152
3.4.5	Fire Filtering . . . . .	152
3.5	Conclusions and implications . . . . .	153
3.5.1	Effects from scaling emissions . . . . .	153
3.5.1.1	HCHO levels . . . . .	153
3.5.1.2	Ozone levels . . . . .	153
<b>4</b>	<b>Stratospheric ozone intrusions</b>	<b>157</b>
4.1	Foreword . . . . .	157
4.2	Introduction . . . . .	157
4.3	Data and Methods . . . . .	159
4.3.1	Ozonesonde record in the Southern Ocean . . . . .	159
4.3.2	Model description . . . . .	162
4.3.3	Characterisation of STT events and associated fluxes . . . . .	163
4.3.4	Biomass burning influence . . . . .	165
4.3.5	Classifying synoptic conditions during STT events . . . . .	166
4.4	STT event climatologies . . . . .	167
4.5	Simulated ozone columns . . . . .	170
4.6	Stratosphere-to-troposphere ozone flux from STT events . . . . .	175
4.6.1	Method . . . . .	175
4.6.2	Results . . . . .	176
4.6.3	Comparison to literature . . . . .	176
4.7	Sensitivities and limitations . . . . .	181
4.7.1	Event detection . . . . .	181
4.7.2	Flux calculations . . . . .	182
4.8	Conclusions . . . . .	183
4.9	Contributions and Acknowledgements . . . . .	184
<b>5</b>	<b>Conclusions</b>	<b>187</b>
5.1	Aims . . . . .	187
5.2	Isoprene emissions . . . . .	188
5.3	Ozone over Australia . . . . .	189
5.4	Outputs and future work . . . . .	190
<b>A</b>	<b>Supplementary Notes</b>	<b>195</b>
A.1	Measurement Techniques . . . . .	195
A.1.1	MAX-DOAS . . . . .	195
A.2	Data sets . . . . .	195
A.2.1	SPEI . . . . .	195
A.2.2	GOME . . . . .	197
A.2.3	NPI . . . . .	197
A.3	Chemistry . . . . .	197
A.3.0.1	SOA . . . . .	197
A.3.1	Relationship to Glyoxyl TODO: remove if never used . . . . .	201
A.4	CAABA/MECCA . . . . .	201
A.4.0.1	CAABA/MECCA outputs . . . . .	202

A.4.1	CAABA/MECCA Box model: isoprene source classifications	202
A.5	Satellite Stuff	204
A.5.1	OMI Algorithm BOAS	204
A.5.2	AMF recaulcation using 72 level output	205
A.5.3	Old Fire Product MYD14C8H	207
A.6	Modelling	207
A.6.1	Emissions drivers	207
A.7	Campaigns	208
A.7.1	Daintree	208
<b>B</b>	<b>Appendix A</b>	<b>211</b>
B.1	Grid Resolution	211
<b>C</b>	<b>Frequently Asked Questions</b>	<b>213</b>
C.1	How do I change the colors of links?	213
	<b>Bibliography</b>	<b>215</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 emitted 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 ( $N_2$ : ~ 78%), oxygen ( $O_2$ : ~ 21%), and argon ( $Ar$ : ~ 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 ( $CO_2$ : ~ 0.4%), ozone ( $O_3$ : .000001 to 0.001%), and methane ( $CH_4$ : ~ 0.4%) (Brasseur and Jacob 2017, Ch. 2). Trace gases in the atmosphere can have a large impact on conditions for life on earth. 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 ~50-60 ppbv over eight hours or ~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 (equivalent to US dollars in the year 2000) per year by 2030 (Avnery et al. 2013), along with detrimental health outcomes equivalent to ~\$USD<sub>2000</sub> 11.8 billion per year by 2050 (Selin et al. 2009). Recently Yue et al. (2017) showed that the net effect of near-surface ozone on is a ~ 14% decrease in



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>.

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.

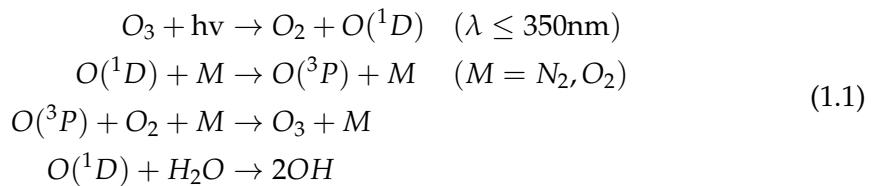
### 1.1.2 Composition and chemistry

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 '90s 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 large.

It has been known for quite a while that our understanding of VOC emissions needs to be improved in order to better capture radiative forcing (Kanakidou et al. 2005). VOC emissions affect ozone along with several atmospheric parameters which directly and indirectly alter radiative forcing rates (eg. Arneth et al. 2008) 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, and that satellite datasets should be used to improve emissions inventories. This is even more important in Australia where VOCs are so poorly represented by contemporary modelling (Emmerson et al. 2016).

## 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 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:





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.

The high energy photons ( $\lambda < 242$  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 ( $\lambda < 1180$  nm) 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, with current southern hemisphere trends lacking full explanation (Zeng et al. 2017).

### 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  $NO_2$  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.

$NO_X$  ( $\equiv NO_2 + NO$ ) is another important chemical family in the atmosphere which interacts with ozone and regulates the atmospheric oxidative capacity.  $NO_X$  or VOC emissions affect the tropospheric ozone equilibrium and can lead to enhanced ozone formation, shown in figure 1.4.  $NO_X$  compounds are short lived, with emissions (Power generation and combustion transport) being the main driver of concentrations (Delmas, Serca, and Jambert 1997).  $NO_X$  is removed primarily by conversion to nitric acid ( $HNO_3$ ) followed by wet or dry deposition (Ayers and Simpson 2006).

$NO_X$  and  $O_3$  relative concentrations during the day are regulated by the following reactions (Sillman 1999; Atkinson 2000):

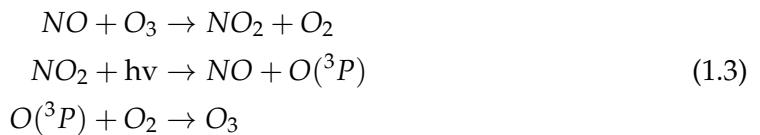




FIGURE 1.4: Tropospheric ozone processes, Figure 1 in Young et al. (2017). DOI: <https://doi.org/10.1525/elementa.265.f1>



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).

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

### 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 occur as tongues of stratospheric air descend and get disconnected from the stratosphere, potentially due to low pressure systems and jet streams (Sprenger, Croci Maspoli, 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 NO<sub>x</sub> sources (such as lightning and car exhaust emissions) and isoprene emissions (Christian et al. 2018).

#### 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 NO, NO<sub>2</sub>, CO, and VOCs such as 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 CO<sub>2</sub> 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).



FIGURE 1.6: Ozone production rate dependent on  $\text{NO}_X$  and VOC concentrations (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 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, evaporating 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 (removal) 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 ( $PM_{2.5}$ ) 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  $NO_x$ , OH,  $O_3$ , 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 ( $R$ ). VOCs are removed mainly by photolysis and oxidation, but also by wet and dry deposition, reaction with  $NO_3$ , 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 classified as either anthropogenic, biogenic (BVOC), or pyrogenic. 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, non-methane VOCs (NMVOC) are often grouped together. 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). 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).

The main non-methane BVOC emissions are isoprene (44%) and monoterpenes (11%) (Guenther et al. 2000; Kefauver, Filella, and Peñuelas 2014). 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). Major emitters are broadleafs (notably Eucalyptus), and shrubs (Guenther et al. 2006; Arneth et al. 2008; Niinemets et al. 2010; Monks et al. 2015). NMVOC emissions are a byproduct of photosynthesis, and also a response to environmental conditions such as drought or sunlight. Emissions are affected by many factors including temperature, atmospheric

$\text{CO}_2$ , soil moisture, drought stress, etc. 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). In this thesis I focus on emissions of isoprene in Australia.

Globally around 710 - 1150 Tg C  $\text{yr}^{-1}$  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 “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 has recently been analysed using 30 years of meteorological reanalysis information by Sindelarova et al. (2014). They estimate emissions of BVOCs to be 760 Tg C  $\text{yr}^{-1}$ , 70% (532 Tg C  $\text{yr}^{-1}$ ) of which is isoprene. This is similar to isoprene emission estimates from MEGAN itself, of 400-600 Tg C  $\text{yr}^{-1}$  (Guenther et al. 2006). Another model (ORCHIDEE, with inputs similar to MEGAN) estimates  $752 \pm 16$  Tg C  $\text{yr}^{-1}$ , 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 ~566 Tg C  $\text{yr}^{-1}$ . This estimate decreases simulated OH concentrations by ~10%, to  $9.5\text{e}5$  molec  $\text{cm}^{-3}$ .

### 1.3.2 Isoprene

Isoprene, or 2-methylbuta-1,3-diene, is a VOC with the chemical formula  $\text{C}_5\text{H}_8$ . 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  $\text{NO}_x$  and  $\text{HO}_x$  cycling, see for example formulae 1.1, 1.3. In the presence of  $\text{NO}_x$ , 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)).

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 isoprene (and other organic hydrocarbons), which could reduce some of the high uncertainties involved with VOC chemistry (Nguyen et al. 2014). Uncertainties in measurements exist both structurally (between different techniques) per measurement due to the difficulty of detecting isoprene and its high reactivity.

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

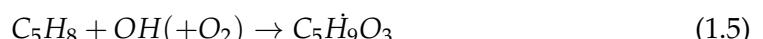
community as a global estimate of isoprene emissions, at roughly 500-600 Tg yr<sup>-1</sup>, emitted mostly during the day. Recently an estimate of global isoprene emissions, of around 465 Tg C yr<sup>-1</sup>, has been made using a completely different model (Messina et al. 2016). The global emission 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. This leads to increased uncertainty for Australian emissions estimates. Due to the highly reactive nature of isoprene, modelling is sensitive to uncertainties, for example the diurnal pattern of isoprene emissions affects modelled ground level ozone (Hewitt et al. 2011; Fan and Zhang 2004).

Isoprene emissions estimates are still fairly uncertain, as global measurements are difficult and regional emissions and chemistry 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 (Guenther et al. 2012; Jiang et al. 2018). The lack of accuracy in BVOC emissions measurements (in general) prevents accurate determinations of the sources and distribution of pollutants including ozone and organic aerosols.

### 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<sup>·</sup>). 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 (Izci and Hosgün 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<sup>·</sup>) (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).



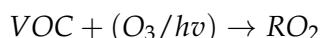
The many children processes and products which begin with isoprene oxidation are often called the isoprene (photochemical) cascade (Paulot2012; eg. Crounse et al. 2012; Wolfe et al. 2016).



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

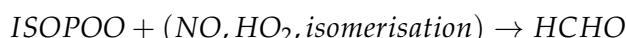
### 1.3.3.1 Oxidation

The primary sink for isoprene is oxidation by OH. First isoprene has its double bond replaced by OH, as summarised by the equation from Izci and Hosgün (2007):  $R-CH=CH-R' + OH \longrightarrow R-CH(OH)CH-R'$  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.

After oxidation by OH, the adducted OH then reacts with  $O_2$  to produce ISOPOO, which can be any of six different isomers (Izci and Hosgün 2007). ISOPOO reacts with  $HO_2$  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  $NO_x$ : the high and low  $NO_x$  pathways are dominated by NO and  $HO_2$  reactions respectively.  $HO_2$  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  $HO_2$ , 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 $NO_x$ pathway

In the presence of  $NO_x$ , ISOPOO reacts with NO and forms ISOPN, which affect levels of both  $HO_x$  (H, OH, peroxy radicals) and  $NO_x$ . ISOPN generally act as a sink of  $HO_x$ , and can be a sink or reservoir for  $NO_x$  (Mao et al. 2013). A portion of the ISOPN are recycled back to  $NO_x$ , serving as a reservoir of nitrogen and allow its transport to the boundary layer of remote regions (Izci and Hosgün 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  $NO_2$  are the main radical sink in high- $NO_x$  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 ISOOPOH (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, ISOOPOH 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 ISOOPOH 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-NO<sub>x</sub> 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 ISOOPOH and reduce local concentrations of OH and HO<sub>2</sub> (Guenther et al. 2000; Paulot et al. 2009b). However this reduction was not seen in measurements and HO<sub>x</sub> levels have been shown to be largely unaffected by isoprene concentrations (Paulot et al. 2009b). HO<sub>x</sub> is recycled through dihydroperoxides (IEPOX), formed from ISOOPOH oxidation, and some HO<sub>x</sub> is produced in the formation of MACR and MVK (Paulot et al. 2009b). Paulot et al. (2009b) estimated that  $95 \pm 45$  Tg yr<sup>-1</sup> 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 HO<sub>x</sub> concentrations against several campaigns. Peeters and Muller (2010) showed that HO<sub>2</sub> 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 NO<sub>x</sub> conditions as many observations of OH were still quite under-predicted in models (Mao et al. 2012). Due to OVOC interference, measurements in low NO<sub>x</sub> 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 Liu et al. (2013), but only when cold-trapping methods are employed. Mao et al. (2012) showed that many instruments were generating OH internally, creating anomalous VOC readings due to within-instrument oxidation.

Improved understanding of both the chemistry and instrument sensitivities has helped closed the gap between model predictions and detected concentrations of VOCs

and OH (Mao et al. 2012). But even with the recent boom in analysis, uncertainties remain in 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 O<sub>2</sub> 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 NO<sub>X</sub> exist, isoprene is consumed by nitrate radicals (NO<sub>3</sub>), which joins to one of the double bonds and produces organic nitrates in high yield (65% to 85%) (Mao et al. 2013). NO<sub>3</sub> are largely formed through ozone reactions, as in equation 1.3. A build up of NO<sub>3</sub> radicals can be seen at night, when photolysis is not removing them (Atkinson 2000; Brown et al. 2009).

In areas with high NO<sub>X</sub> levels, greater than 20% of the isoprene emitted late in the day ends up being oxidised by the NO<sub>3</sub> radical overnight (Brown et al. 2009). At night isoprene affects on both NO<sub>X</sub> 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 NO<sub>3</sub> 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 NO<sub>X</sub> removal in high latitude winters, when photolysis and NO reactions are reduced (Ayers and Simpson 2006).

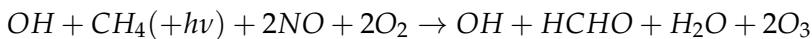
## 1.4 Formaldehyde

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 NO<sub>X</sub> concentrations, as it affects the yield from isoprene oxidation. HCHO yield is higher in the high-NO<sub>X</sub> pathway (compared to the low-NO<sub>X</sub> pathway) from isoprene reactions (Marais et al. 2012). HCHO measurements are often used as a check on how well isoprene reactions are simulated, as HCHO levels depend on initial VOCs and oxidants (which can be prescribed) (Marvin et al. 2017).

#### 1.4.1 Sources and sinks

Background levels of HCHO in the atmosphere are driven by the oxidation of methane (CH<sub>4</sub>) by the hydroxyl radical (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. CH<sub>4</sub> concentrations are relatively well constrained in models, with the ACCMIP comparison showing only ~ 3% inter-quartile range (Young et al. 2013). There is a complex relationship between VOCs, HO<sub>X</sub>, and NO<sub>X</sub>: with higher levels of NO<sub>X</sub> 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 NO<sub>X</sub> (Wagner 2002; Millet et al. 2006; Kefauver, Filella, and Peñuelas 2014). The total contribution from NMVOC oxidation is ~ 358 Tg yr<sup>-1</sup> (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).

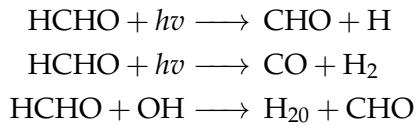
Other terpenoids (monoterpenes, sesquiterpenes, etc.) can also produce HCHO, although generally to a lesser extent than isoprene, methane and biomass burning (Guenther et al. 2012). Many of the HCHO yields from terpenoids are estimated through chamber studies which examine molecular mass and charge after mixing the compound of choice into a known volume of air (eg. Nguyen et al. 2014). These conditions generally do not 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 ??).

In the past, HCHO levels were underestimated by models, often with large discrepancies, due to the poor understanding of methyl peroxy radical (CH<sub>3</sub>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) with reactions as follows (Ayers et al. 1997):



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.

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 a priori 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

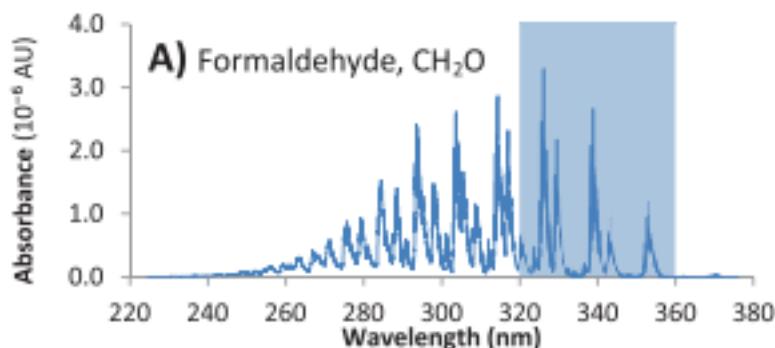


FIGURE 1.8: HCHO spectrum, with a typical band of wavelengths used for DOAS path measurements. This is a portion of an image from Davenport et al. (2015).

four techniques (MAX-DOAS, FTIR, chromatographic, and flourimetric), 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).

#### 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

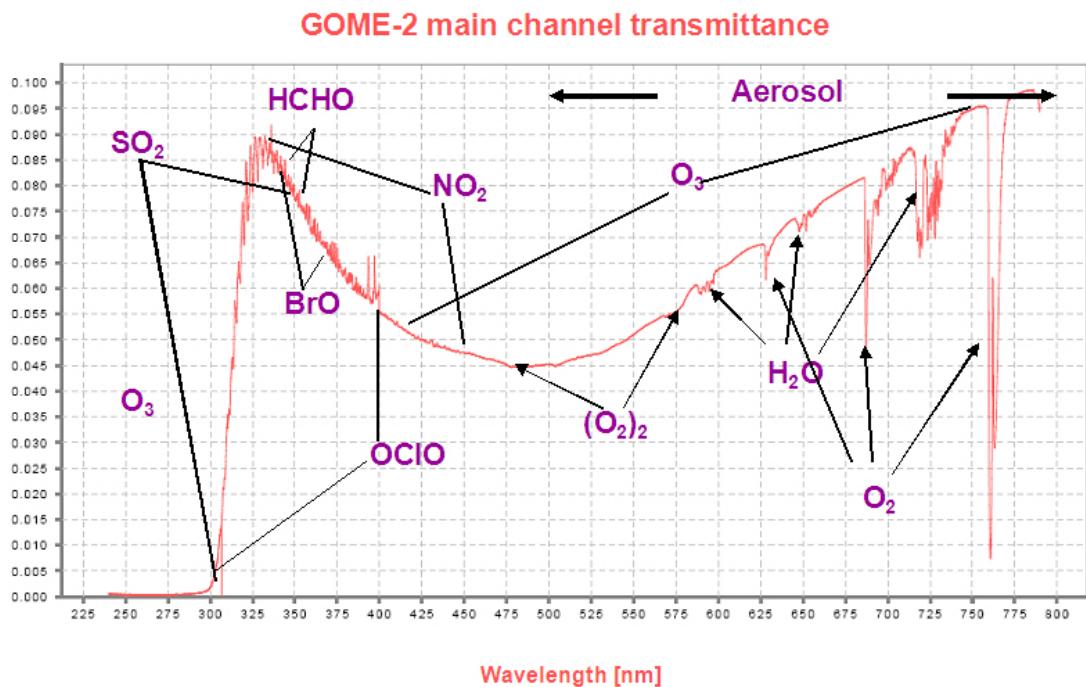


FIGURE 1.9: An example spectrum showing interferences used for species concentration measurements by GOME-2. Image by EUMETSAT and ESA (EUMETSAT 2015).

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 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 ??.

## 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.

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. Eulerian models use examine equations and transport within and between volumes in a spatial coordinate systems, while Lagrangian models look at behaviour within a potentially changing frame of reference (for example within a cloud). 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

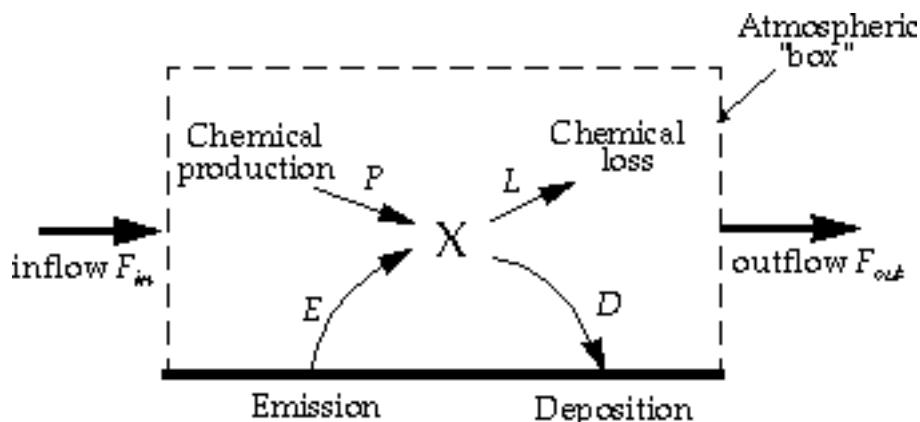


FIGURE 1.10: Standard box model parameters, image taken from Jacob (1999).

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. The general formula governing modelled emissions  $E$  for a species  $i$  (from Brasseur and Jacob (2017)) is as follows:

$$E_i = A \times EF_i \times S_i$$

with  $A$  the activity rate (eg. how many trees in an area),  $EF_i$  being the emission factors (eg. isoprene emitted per tree per year), and  $S_i$  is a scaling factor accounting for meteorology and other effects not included in  $A$  or  $F$  (eg. seasonal temperature).

Isoprene is emitted by trees or shrubs, depending on several parameters such as leaf area index (LAI), 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)). MEGAN has estimated  $\sim 1150 \text{ Tg C yr}^{-1}$  BVOC emissions globally, of which  $\sim 465\text{--}500 \text{ Tg C yr}^{-1}$  is isoprene (Guenther et al. 2006; Messina et al. 2016). Since little data exists with which to verify 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 a priori 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. Atmospheric chemical models by necessity require various simplifications of real world processes, and also utilise information which may be itself uncertain or extrapolated. Uncertainty is introduced through both of these channels as well as through computational limitations and non-linear non-continuous system solution approximations.

#### 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  $\sim 2$  (Christian, Brune, and Mao 2017).

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) have 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). Both are sensitive to light and temperature parameterisations.

The concentration of NO<sub>x</sub> is an important factor in determining the yield of HCHO and ozone from BVOCs. Travis et al. (2016) show how modelled surface ozone is overestimated due to high estimates of NO<sub>x</sub> emissions, which affect oxidative capacity and VOC reactions in the US. NO<sub>x</sub> and isoprene emissions are shown to be the most significant sources of uncertainty for ozone concentrations near the surface in GEOS-Chem over the US, while isoprene derived products and lightning NO<sub>x</sub> drives uncertainty in the upper atmosphere (Christian, Brune, and Mao 2017).

#### 1.5.4.2 Resolution

Atmospheric chemistry 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. Yu et al. (2016) show how only at higher resolution (0.25 by 0.3125°) does isoprene oxidise under the correct NO<sub>x</sub> scheme (through high or low NO<sub>x</sub> pathways, see section 1.3.3.1) in variable NO<sub>x</sub> environments. This leads to an increase of high NO<sub>x</sub> pathway oxidation of isoprene at the lower resolutions, which leads to an overestimation of HCHO but not ozone at coarser resolutions. However, for many global scale analyses, errors from resolution are less important than those from chemistry, meteorology, and emissions (Christian, Brune, and Mao 2017; Christian et al. 2018).

#### 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). Recently Christian, Brune, and Mao (2017) analysed GEOS-Chem (A global CTM; see section 2.4) for ozone and oxidant (OH and HO<sub>2</sub>) sensitivity to the driving processes and inputs. They 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.

Many models lack in-situ measurements with which to verify their chemical mechanisms, leading to large discrepancies (Marvin et al. 2017). 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) HCHO concentrations are underestimated in a way that can not be easily fixed through reaction rate changes. They 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). Five models (GEOS-Chem, CB05, CB6r2, MCMv3.2, and MCMv3.3.1) all are found to underestimate HCHO concentrations (by 15 – 30%).

#### 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.

#### 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 (Jiang et al. 2018). Jiang et al. (2018) 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

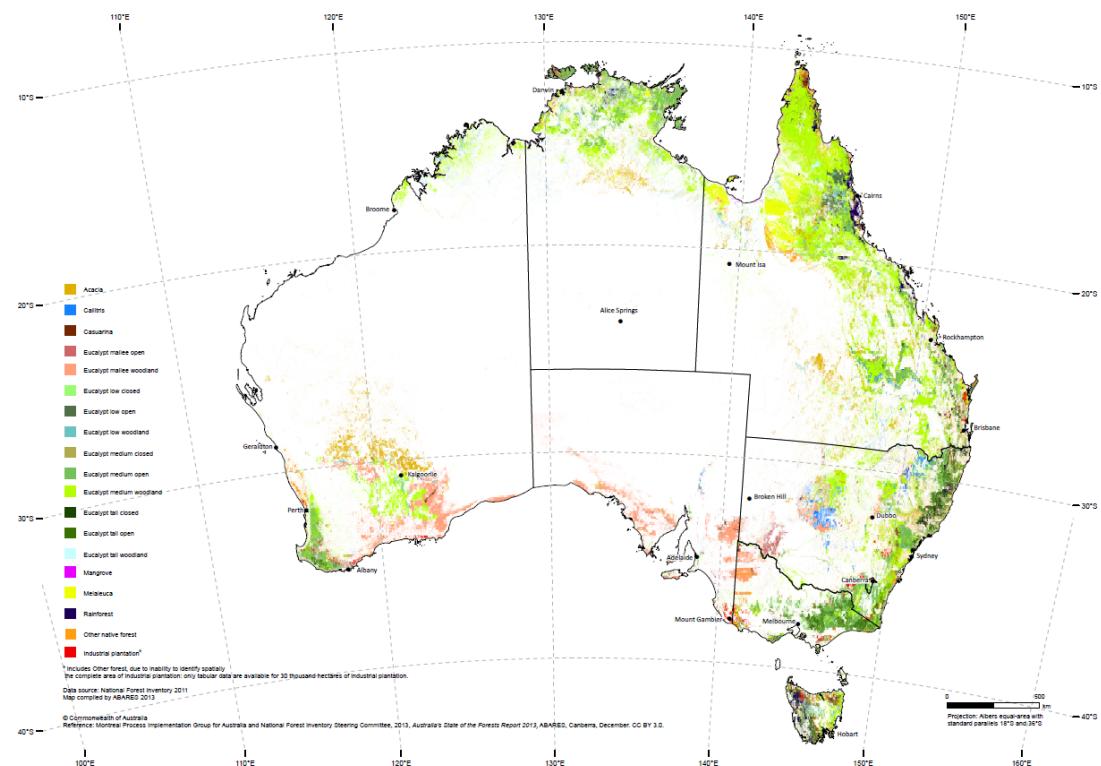
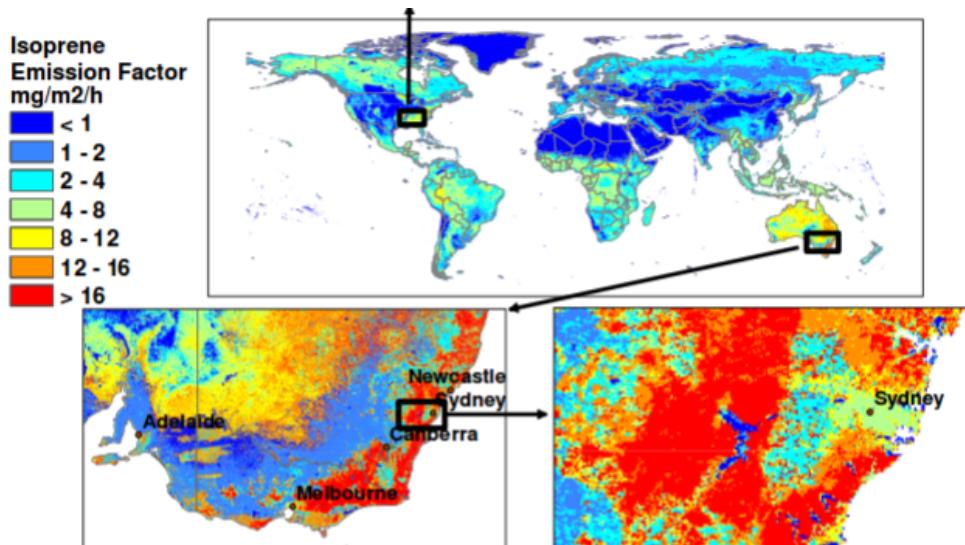


FIGURE 1.11: Forest types in Australia (<http://www.agriculture.gov.au/abares/forestsaustralia/australias-forests>)

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.

Ozone enhancements above the background levels are most sensitive to emissions (of precursor gases), with meteorology, and atmospheric composition also important. Anthropogenic emissions of ozone precursors are important but relatively stable, while pyrogenic sources are greatly variable and dependent on weather, fuel, and fire intensity (Lawson2017). Emissions from burning 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.



**Fig. 2.** Global distribution of landscape-average isoprene emission factors ( $\text{mg isoprene m}^{-2} \text{ h}^{-1}$ ). 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.

2007; Edwards et al. 2006), particularly from July to December (Pak et al. 2003; Liu et al. 2016a). Local fires are even more influential and the burning season for Australia can be all year, with severity depending on regional vegetation, recent and current weather, and El-niño.

It has been estimated by MEGAN that the Australian outback is among the world's 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. These metrics are measured by varying numbers of monitoring stations in each state. Measurement stations are generally located in population centres, and do not 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  $\text{O}_3$  and other pollutants.

Generally STT of ozone 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). The contribution of STT to overall tropospheric ozone budgets remains uncertain, especially in the southern hemisphere (SH) (Škerlak, Sprenger, and Wernli 2014). STT can enhance surface ozone concentrations above legal air quality limits (e.g. Lelieveld et al. 2009; Lin et al. 2015). It is easier to determine tropospheric ozone enhancements over the relatively clean southern ocean atmosphere. However measurements of tropospheric ozone over this region are relatively sparse (Škerlak, Sprenger, and Wernli 2014), and quantification of transported ozone is difficult without large scale extrapolations. Ozone enhancements over the southern ocean signify either transported pollution or stratospheric influx (Jacobson and Hansson 2000). Quantifying ozone processes over the southern ocean would be helpful towards understanding chemistry in the “clean background environment”, which is important when validating models and satellite datasets.

### 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 (Jiang et al. 2018).

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 thresholds, 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

Isoprene and many of its products can be difficult to measure accurately due to their short lifetimes, high reactivity, and optically thin natures. There are relatively few measurements in the southern hemisphere, including MUMBA (PatonWalsh2013), SPS(Dunne et al. 2018), and Tumbarumba (Emmerson et al. 2016) These campaigns focus on air quality or biogenic emissions and use several different instruments (including PTR-MS and GC-FID) to detect metrics such as air particulates, HCHO, isoprene, and meteorological information. An air-flight campaign (HIPPO) measuring over the Pacific ocean, with one flight passing along the NSW coastline, also provides isoprene and ozone concentrations in November 2009 (Wolfsy2011). Wollongong also has 20 years of DOAS calculated HCHO measurements from a solar FTS. Satellite total columns suffer from orography amongst other limitations when compared to this data (Demol2010). For further details on these campaigns see Section ??.

Detecting ozone from the surface up to the top of the stratosphere requires different techniques such as remote sensing and ozonesonde releases. Ozonesondes are weather balloons (with attached ozone detectors) which detect ozone concentrations up to the mid stratosphere ( $\sim$  30 km), providing a vertical profile over a single location. 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 (including Davis, Macquarie Island, and Melbourne) is available from the world ozone and ultraviolet radiation data centre <http://woudc.org/data/explore.php> and is used in Chapter 4 to examine stratospheric impacts on tropospheric ozone (see ?? for more info on these ozonesondes).

## 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.

Estimation of BVOC emissions in Australia can be improved through satellite measurements of one of the primary oxidation products HCHO. 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 are uncertain over Australia. Satellite measurements use modelled a priori vertical profiles of HCHO to estimate total column amounts. **I aim to recalculate satellite vertical columns of HCHO using updated model a priori information.** In this effort I aim to improving the understanding of the importance of relevant parameters (within GEOS-Chem) in calculating vertical columns of HCHO measured by satellite. This includes an examination of how well GEOS-Chem simulates several species such as NO<sub>X</sub>, isoprene, and HCHO compared to both in-situ and remote measurement data that exists for Australia. Additionally I detail the construction and effects of satellite data filters. The work towards this aim is in Chapter 2.

The technique of determining isoprene emissions from satellite detected HCHO is called satellite inversion. **I aim to determine isoprene emissions in Australia using a top-down inversion of satellite HCHO, through an estimated yield from isoprene to HCHO** HCHO amounts and the yield of isoprene to HCHO over Australia is required to create top-down estimates. This process also requires careful examination of when the assumptions required within the inversion process are not valid. Due to the low availability of in-situ data over most of the Australian continent, a combination of modelled and satellite data could reduce the uncertainties of isoprene 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 towards fulfilling 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 ~ 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.

**I aim to describe relative importance of sources of tropospheric ozone in Australia, as well as seasonality.** I will describe how modelled ozone is affected by updated isoprene emissions, comparing changes in GEOS-Chem outputs. Trends of isoprene emissions and their relationship to tropospheric ozone trends could provide new insight into the future of tropospheric ozone in Australia.



## 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 interpret measurements, estimate chemical concentrations at any scale, and predict atmospheric composition in the future. In remote sensing measurements, modelling is required in order to produce useful outputs. Models of ozone in the atmosphere are used broadly for international assessments of ozone precursor emissions, and estimating effects from related processes (such as radiation) (Young et al. 2017). Models provide an estimate of many trace gas concentrations, however verification is required, and generally performed using results from measurement campaigns. In situ measurements from campaigns or measurement stations can be used to examine what is happening at a particular location. These data are used to determine how accurate models or estimates are - however the utility is limited to where and when the measurements took place. In this thesis data from several campaigns are compared against model outputs and satellite datasets. Satellite datasets provide large amounts of data over most of the planet. However, they can have high uncertainty due to instrument limitations. Many datapoints can be averaged in order to reduce uncertainty. In this chapter several satellite datasets are combined to estimate biogenic HCHO amounts over Australia.

The first goal is to analyse Australia-specific HCHO concentrations measured by satellite, and determine isoprene sensitivity and any model bias. This leads into Chapter 3 where biogenic HCHO columns are used to estimate isoprene emissions. The second goal is to quantify ozone transported from the stratosphere down into the troposphere (Chapter 4). The focus in this chapter is to describe and analyse model outputs and measurements along with how they are recalculated and compared. Section 2.2 details satellite and campaign datasets, and additionally describes model outputs. Measurement techniques used to retrieve the most utilised satellite dataset are outlined in Section 2.3. Section 2.4 describes the GEOS-Chem model, how it is run and what setup and outputs are used in this thesis. In Section 2.5 the process of using model outputs to recalculate satellite vertical columns is defined and analysed. In order to compare satellite data with other datasets, some work must be undertaken to avoid introducing bias (e.g., Palmer et al. 2001; Eskes and Boersma 2003; Marais et al. 2012; Lamsal et al. 2014). One key step is to recalculate the satellite information using modelled data, detailed in Section 2.6. The effects of these recalculations on

satellite HCHO is also examined. The creation and effects of filters used to remove non-biogenic influences are described in Section 2.7.

## 2.2 Datasets

This section describes the datasets used in this thesis, along with an overview of the measurement techniques used for each. This includes modelled output, satellite measurements, and measurement campaigns. These datasets serve four purposes:

1. Model output validation in this chapter
2. Calculation of biogenic HCHO distribution over Australia in this chapter
3. Recalculated OMI formaldehyde columns are used as a basis for estimating isoprene emissions in Chapter 3
4. Extrapolation of ozone transport in Chapter 4

I will also give details on filtering and interpolations which are undertaken when reading data, as each dataset has its own resolution. 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.

Uncertainty (or error) is present in each dataset and where possible the causes are explained. There are two types of error: systematic and random. Arguably the worst of these is systematic error (or bias). Bias 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 often reported as some function of a datasets variance, or uncertainty. It can be reduced through averaging either spatially or temporally. Temporal and/or spatial averaging decreases uncertainty by a factor of  $1/\sqrt{N}$  where N is the number of observations being averaged.

### 2.2.1 Satellite

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

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 (e.g., 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/>).

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 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 measurable properties seen by remote sensing devices such as a satellite. Although instruments will be more or less sensitive to various properties depending on altitude, radiation, and other parameters (e.g., Martin et al. 2002b).

Difficulties can arise when aerosols interfere with recorded spectra (e.g., 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 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. Cloud filtering is performed on several satellite products used in this thesis, due to the uncertainty introduced by cloud interference. This has been seen to introduce a clear-sky bias in monthly averages since measurements do not include cloudy days (Surl, Palmer, and Abad 2018).

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 much more satellite data will be available in the form of geostationary satellite measurements (Kwon et al. 2017). 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.

### 2.2.1.1 Formaldehyde

OMI spectra are used in several products used in this thesis, including OMHCCHO, OMNO2d, and OMAERUVd. Satellite based formaldehyde measurements from the OMI instrument on board AURA are stored in the OMHCCHO product. OMHCCHO data is used and modified extensively throughout this thesis, and so is discussed in more detail in Section 2.3. Calculation of column density and AMF are discussed respectively in sections 2.3.2 and 2.3.3.

HCHO products can be found in 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 leads to different estimates between instruments as described in Lorente et al. (2017), and both validation and comparison become more important when using these remotely sensed data.

### 2.2.1.2 Nitrogen dioxide

OMNO2d is a gridded daily level three NO<sub>2</sub> product with good satellite pixels averaged into 0.25x0.25° horizontally resolved bins. An example figure from Jan 29, 2005 is



FIGURE 2.1: Example of NO<sub>2</sub> tropospheric columns taken from the OMNO2d product.

shown in Figure 2.1, while an average for 2005 (global) is shown in Figure 2.2. OMNO2 pixel resolution is 40 km by 130 km. NO<sub>2</sub> measured by OMI is used to check whether NO<sub>2</sub> is well represented by GEOS-Chem (see Section 2.4.6 for the comparison between this product and GEOS-Chem calculations). It is also used to form the anthropogenic influence filter for OMHCHO (See Section 2.7.2).

Like other satellite products, OMNO2d is influenced by a priori modelling which is required to convert slant path radiance to vertical columns. These models are generally low resolution ( $\sim 110$  km by 110 km), which leads to column smearing and difficulty detecting point sources of high NO emissions (Goldberg et al. 2017). Uncertainty in this product arises mostly from the calculation of the AMF (up to 50% of total error) (Lorente et al. 2017).

### 2.2.1.3 Aerosol optical depth

Aerosols in the atmosphere can be seen through their affects on light. Smoke and dust can be seen as an increase in aerosol optical depth (AOD) (see Section 2.3.2). This is due these particles scattering and absorbing UV radiation (Ahn2008). A data product provided by Earthdata ([https://disc.gsfc.nasa.gov/datasets/OMAERUVd\\_V003/summary](https://disc.gsfc.nasa.gov/datasets/OMAERUVd_V003/summary)) called OMAERUVd (DOI: 10.5067/Aura/OMI/DATA3003) is used in this thesis.

OMAERUVd provides a useful dataset which allows detection of areas which may be smoke affected. The product contains AOD and aerosol absorption optical depths (AAOD) at three wavelengths (354, 388, and 500 nm), along with UV aerosol index



FIGURE 2.2: Average 2005 tropospheric NO<sub>2</sub> from OMNO2d with pixels screened for < 30% cloud cover.

(UVAI). The OMAERUVd product is level three gridded daily data, based on quality filtered level two swath pixels which are then gridded by averaging. The product is most sensitive to error in the form of sub-pixel scale cloud interference, so I select AAOD as the basis for my smoke filter as it is least affected by clouds ([Ahn2008](#)).

In this work AAOD is mapped from  $1 \times 1^\circ$  horizontal resolution to  $0.25^\circ \times 0.3125^\circ$  using nearest value mapping. The AAOD at 500 nm wavelength is used to determine smoke influence, although any of the provided wavelengths would be affected by smoke plumes and could also be used. This daily AAOD is compared to a threshold to create a daily smoke filter, any areas with  $\text{AAOD} > 0.03$  are considered to be potentially smoke plume affected (see Section [2.7.1](#)).

#### 2.2.1.4 Active fires

MOD14A1 is a gridded daily satellite based dataset of fire counts at  $1 \times 1 \text{ km}^2$  horizontal resolution. Fire observations are performed four times daily from Terra (10:30 LT, 22:30 LT) and Aqua (01:30 LT, 13:30 LT). The fire pixels are detected based on parameters including apparent pixel temperature and the nearby background temperature. The dataset is obtained from NASA Earth Observations (NEO) that is part of the EOS Project Science Office at the NASA Goddard Space Flight Center [https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MOD14A1\\_M\\_FIRE](https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MOD14A1_M_FIRE). This product is downloaded and binned into a lower resolution (using the sum of fire pixels) to create an active fire influence mask (see Section [2.7.1](#)).

#### 2.2.1.5 Carbon monoxide

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](#)). CO is used as a proxy for biomass burning plumes, and used to qualitatively attribute ozone intrusion events as explained in Section [4.3.5](#). This is a separate method of detecting fire influence near specific sites through visual analysis.

#### 2.2.1.6 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 into 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 calculation of the air mass factor (AMF) which converts slanted light path concentrations into a vertical profile ([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 (Lorente et al. 2017). 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 Lorente et al. (2017).

A common way of reducing satellite uncertainty is through oversampling or temporal averaging. This is done frequently for 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. Another example of this can be seen in Dufour et al. (2008), where monthly averaging is used to decrease the measurements uncertainty at the cost of temporal resolution.

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.

## 2.2.2 Model datasets

### 2.2.2.1 GEOS-Chem output

GEOS-Chem model output is used extensively in this thesis and is discussed in more detail in Section 2.4. Section 2.4.7.1 specifically describes the model outputs used in this thesis. These are generally resolved to 47 vertical levels from the ground up to 0.01 hPa, at  $2^\circ \times 2.5^\circ$  horizontal resolution.

### 2.2.2.2 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. The version used was ERA-Interim, which was the most up to date at the time (2016) but has since been superseded by ERA5.

### 2.2.2.3 Surface temperatures

The Climate Prediction Center (CPC) provides a product with maximum daily land-surface temperature at  $0.5 \times 0.5^\circ$  horizontal resolution. This data is used to check the correlation between HCHO and temperature at a higher resolution than is provided by GEOS-Chem output. A full description of the data can be found at <https://www.esrl.noaa.gov/psd/data/gridded/data.cpc.globaltemp.html>. CPC Global Temperature data is provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their web site at <https://www.esrl.noaa.gov/psd/>. An example of one day of land temperature output is shown in Figure 2.3.

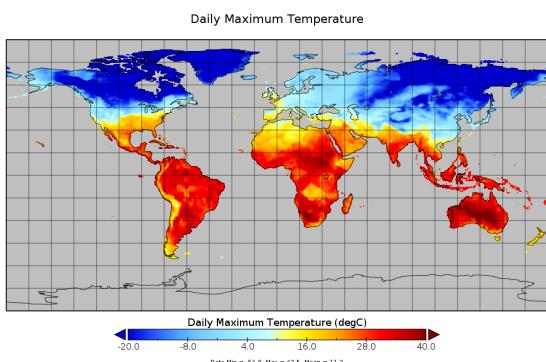


FIGURE 2.3: CPC daily maximum temperature dataset output for 1, Jan, 2005.

TABLE 2.1: Detection limits for MUMBA

Dates	HCHO (ppb)	Isoprene (ppb)	Ozone (ppb)
21/Dec/2012 - 29/Dec/2012	0.205	0.003	0.5
29/Dec/2012 - 18/Jan/2013	0.105	0.005	0.5
19/Jan/2013 - 15/Feb/2013	0.186	0.003	0.5

### 2.2.3 Campaign datasets

In this thesis data from several measurement campaigns are used to examine accuracy of modelled data at specific sites. Figure 2.4 shows the locations of BVOC measurement sites in the top panel, and release sites for ozonesondes in the bottom panel. These took place over disparate times, and are in situ measurements which are hard to directly compare against GEOS-Chem output which is averaged over a large horizontal space.

The campaign datasets provide three separate time series for brief periods of both isoprene and formaldehyde. Figure 2.5 shows these along with the detection limits and also shows isoprene measurements superimposed over a single year. It is apparent that more measurements are required to see more than the daily cycles.

#### 2.2.3.1 Measurements of Urban, Marine and Biogenic Air (MUMBA)

The MUMBA campaign (Paton-Walsh et al. 2017) measured various compound abundances including isoprene, formaldehyde, and ozone from 21 December 2012 to 15 February 2013. These measurements took place in Wollongong, 10 m above ground level (40 m above sea level). Ozone was measured by Thermo UV absorption with 1-minute time resolution averaged into hourly outputs. Isoprene and HCHO were measured by Ionicon Proton-Transfer-Reaction Mass spectrometer (PTR-MS), with a time resolution of 3-minutes, averaged each hour. Detection limits varied due to instrument conditions, and are listed in Table 2.1. The full dataset has been published on PANGAEA (DOI:10.1594/PANGAEA.871982) (Guérette et al. 2018).

In this thesis we assume uncertainty in this product is as estimated by Dunne et al. (2018) at (50%). Although the uncertainty determined through calibration measurements was only 15% (Guérette et al. 2018), this does not account for competing trace

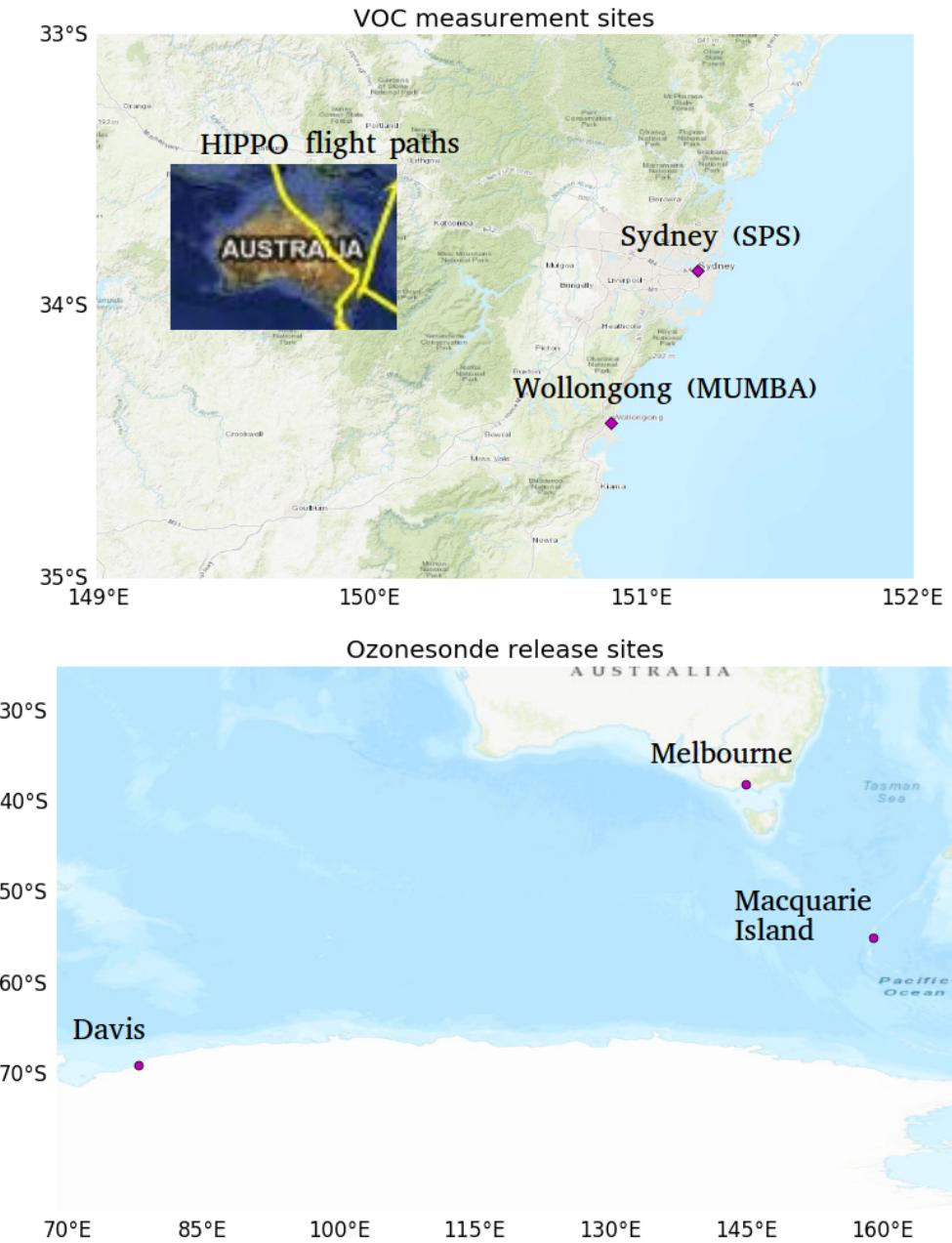


FIGURE 2.4: Locations of VOC measurements (top panel) and ozonesonde release sites (bottom panel). Inlaid in top panel is the flight paths over Australia of the HIPPO campaign.

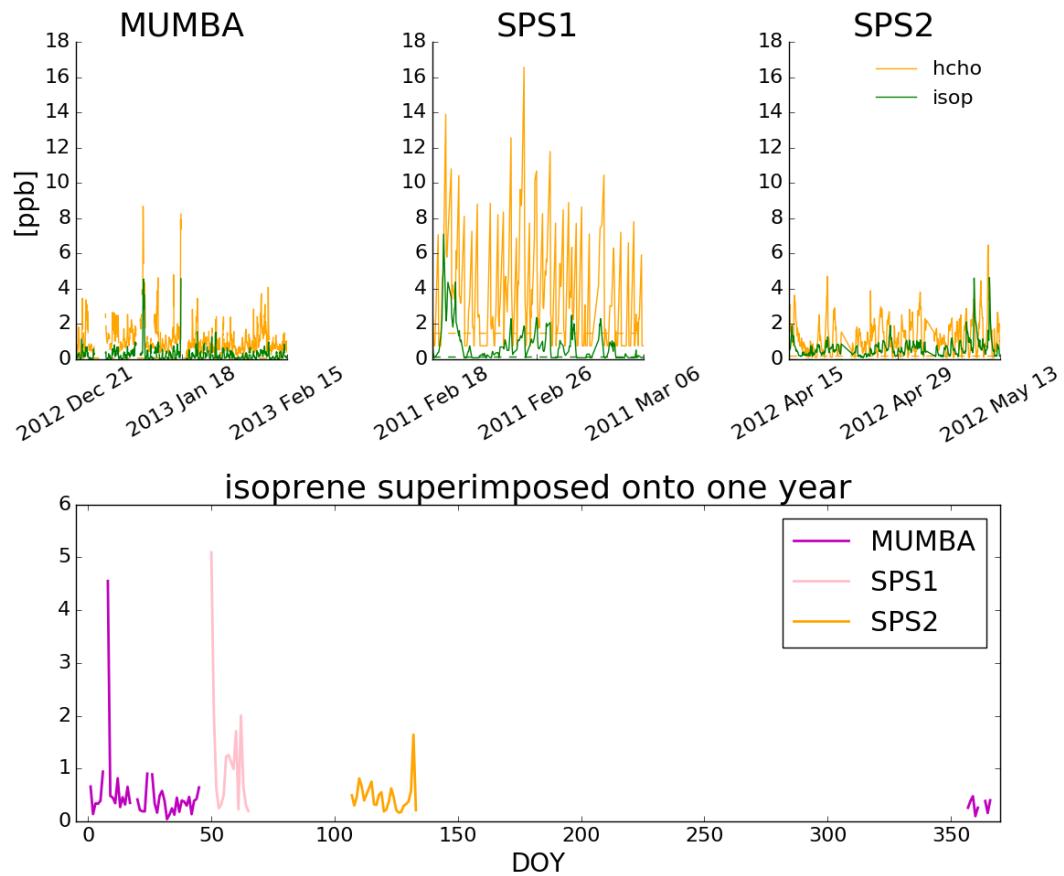


FIGURE 2.5: Top: MUMBA, SPS1, and SPS2 time-series for HCHO (orange) and isoprene (magenta), along with detection limits (dashed).  
Bottom: isoprene measurements superimposed onto a single year.

gas interference (such as furan). The readings are re-sampled to hourly averages. Measurements below the detection limit are set to half of the detection limit when reading this dataset.

### 2.2.3.2 Sydney Particle Studies (SPS1, SPS2)

Two trace gas measurement campaigns took place at the Westmead air quality station. Stage 1 (SPS1) from 5 February to 7 March, 2011 and stage 2 (SPS2) from 16 April to 14 May, 2012. Two instruments measured VOC concentrations: a PTR-MS, and a gas chromatographer (GC) with a 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 other VOCs, further measurement specifics can be found in Dunne et al. (2018).

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 is 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 start to become frequent (Guérette et al. 2018). The GC-FID analysed samples collected in multi-absorbent tubes, with lower temporal resolution but no interference. Further details for this method can be found in Cheng et al. (“Factors controlling volatile organic compounds in dwellings in Melbourne, Australia”). GC-FID data is averaged from 0500-1000 LT, and 1100-1900 LT, while PTR-MS data is averaged hourly. This includes significant differences between measurement devices when detecting isoprene, potentially due to interfering compounds in the PTR-MS (Dunne et al. 2018).

Figure 2.6 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 (see Section 2.4) a daily average and an midday time (13:00-14:00 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 Lawson et al. (2015). This should minimise any introduced bias.

### 2.2.3.3 Ozonesondes

Ozonesonde data come from the World Ozone and Ultraviolet Data Centre (WOUDC). Ozonesondes are weather balloons which measure from the surface to around 35km. Ozonesondes provide a high vertical resolution profile of ozone, temperature, pressure, and humidity. Generally the instrument will perform 150-300 measurements in the troposphere with ozone mixing ratios quantified by an electrochemical concentration cell (<http://www.ndsc.ncep.noaa.gov/organize/protocols/appendix5/>).

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). Melbourne, a major city with more than 4 million residents (ABS 2013), may be affected by anthropogenic pollution in the lower troposphere. Actual releases are north of the central business district in the Broadmeadows suburb. Macquarie Island is in the remote Southern Ocean and unlikely to be affected by any local pollution events. Davis (on the coast of Antarctica)

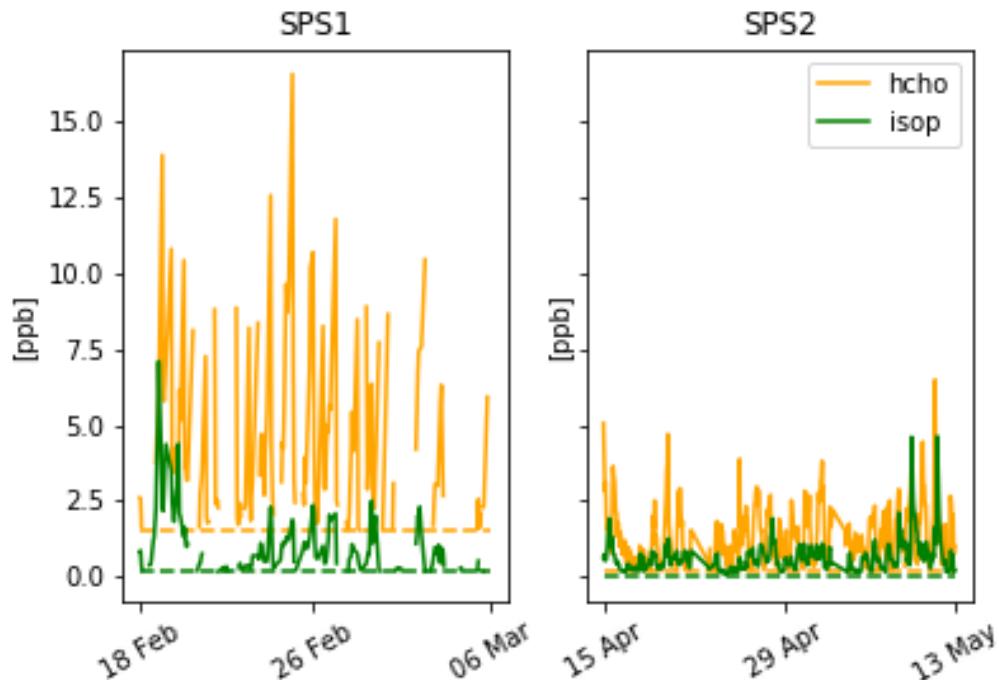


FIGURE 2.6: SPS HCHO (yellow) and isoprene (green) time series, along with detection limits (dashed). SPS 1 (left) took place in late summer 2011, while SPS 2 (right) occurred during autumn 2012.

is also unlikely to experience the effects of anthropogenic pollution. More information on this dataset is given in Section 4.3.

#### 2.2.3.4 Wollongong

Upon the roof of the Chemistry building in Wollongong lies a solar Fourier transform infra-red spectrometer (FTIR) which measures HCHO (amongst other gases) on the path of light between the instrument and the sun. TODO: details about these measurements.

Figure TODO: shows Wollongong FTIR measurement profiles of HCHO with an assumed averaging kernel and then one calculated from GEOS-Chem, along with the two averaging kernel mean and standard deviation. TODO: Discuss this conversion and results seen here... The bottom panel shows the time series of total column HCHO of both GEOS-Chem and the FTIR instrument in molecules  $\text{cm}^{-2}$ .

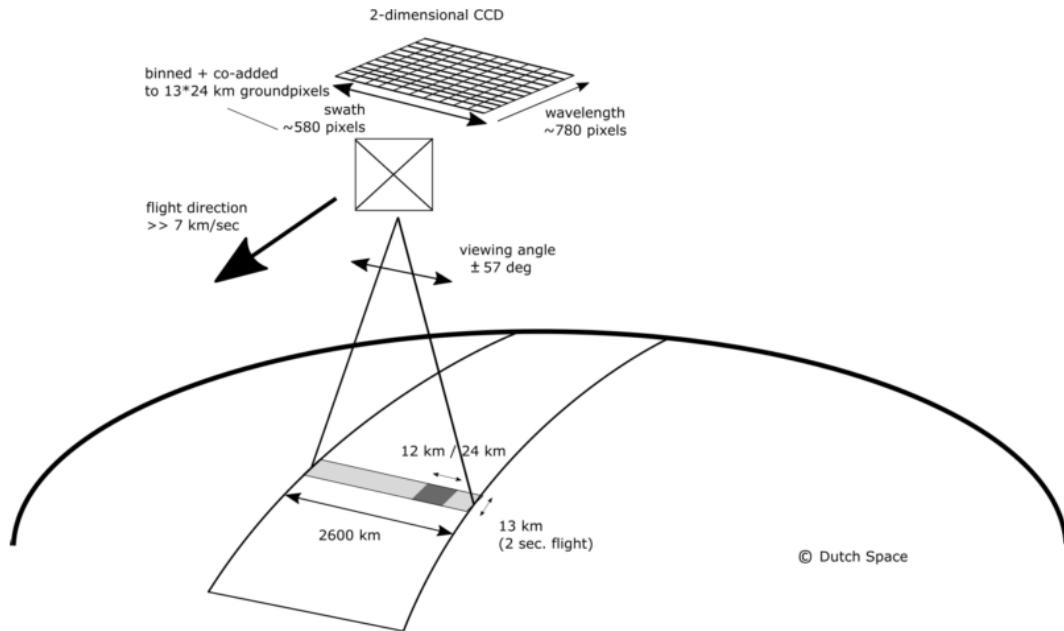
#### 2.2.3.5 Uncertainties

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 (e.g., Dunne et al. 2018; Guérette et al. 2018). Overall isoprene uncertainty in measurements analysed by Dunne et al. (2018) was a factor of 1.5 to 2. This can feed into uncertainties in modelling and satellite retrievals, as verification and correlations are affected.

### 2.3 Satellite formaldehyde

One satellite product used extensively in this thesis is named OMHCHO: from NASA's Earth Observing System's "Aura", which provides several other 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. 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 (among other quantities) and the formaldehyde product is detailed here.

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. Figure 2.7 shows the details of OMI's detector array and measurement resolutions. OMI measures atmospheric trace gases including  $\text{NO}_2$ ,  $\text{SO}_2$ ,  $\text{BrO}$ , HCHO,  $\text{O}_3$ , and aerosols. OMI measurements occur from right to left on a band covering  $115^\circ$ , resulting in swaths of around 2600 km, with pixel sizes from  $13 \times 24 \text{ km}^2$  at nadir to  $26 \times 135 \text{ km}^2$  at the swath edges (Gonzalez Abad et al. 2015). The swaths cover Earth daily, both on the light and dark side of the planet, only daytime measurements provide useful near-UV/Visible information. 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).



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 2.7: 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  $7 \text{ kms}^{-1}$  and an exposure time of 2 s lead to a ground pixel size of 13 km in the flight direction. The viewing angle of  $114^\circ$  leads to a swath width on the ground of 2600 km.” The table shows the optical properties for OMIs three channels.

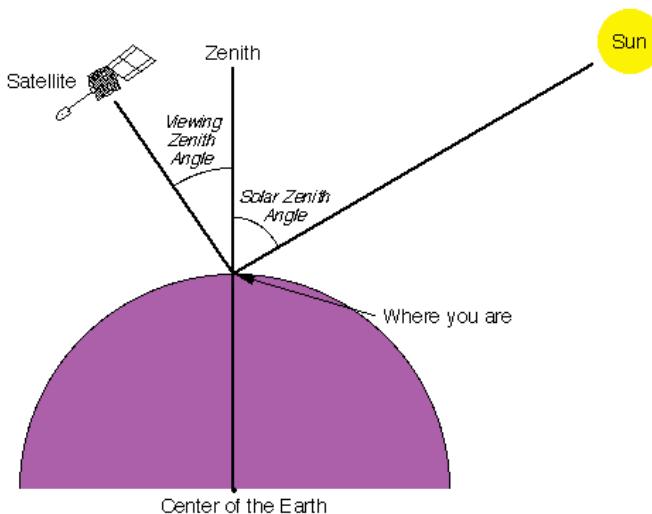


FIGURE 2.8: Solar and viewing zenith angles, image copied from Wikipedia (2016), originally from a NASA website.

The latest OMHCHO algorithm uses a shape factor determined from GEOS-Chem using 47 vertical levels at monthly temporal resolution and  $2^\circ \times 2.5^\circ$  latitude by longitude horizontal resolution (Gonzalez Abad et al. 2015). The GEOS-Chem model has been substantially updated since then. In this thesis the more recent version V10.01 is used to recalculate the vertical column HCHO (details are shown in Section 2.6).

OMI uses a Differential Optical Absorption Spectroscopy (DOAS) based technique to read HCHO along the path of light which reaches the satellite instrument. The first step is to determine how much HCHO is in the path of light between the sun and detector. Measurements done using DOAS often apply a forward radiative transfer model (RTM) such as LIDORT (see Section 2.3.4) in order to determine a trace gas' radiative properties at various altitudes. The forward RTM used for satellite data products also involve functions representing extinction from Mie and Rayleigh scattering, and the effect of these on spectra. These RTM are also required to account for (often estimated) atmospheric parameters such as albedo. The next step is to transform the calculated amounts along the non-vertical light path into vertical column amounts. This is done by applying an AMF. 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 2.8. However, in the UV-VIS region of the spectrum, Rayleigh and Mie scattering (see Section 2.3.2) must be accounted for.

Atmospheric HCHO detected by satellite requires that other trace gases with similar features near the HCHO affected wavelengths 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 formed from the spectra differential with respect to the radiance reference column over the Pacific. The full method details for slant column retrieval by OMI are outlined in supplemental Section A.5.1, or in the technical document (DOI: 10.5067/Aura/OMI/DATA2015). Slant columns range from  $\sim 4 \times 10^{15}$  to  $\sim 6 \times 10^{16}$  molec cm $^{-2}$ , with uncertainties from 30% (larger columns) to over 100% (smaller columns) (Gonzalez Abad et al. 2015).

### 2.3.1 Pixel filtering

This thesis uses the level 2 product swath output from the NASA earth data web portal. OMHCHO level two data includes 14-15 daily swaths of measurements. Each swath contains roughly  $9 \times 10^4$  pixels, each of which includes latitude, longitude vertical column HCHO, along with all the ancillary data required to make the vertical column and several data quality metrics. The OMHCHO dataset has a quality flag which can be used to remove unlikely or poor satellite measurements. The states represented by this quality flag are shown in Table 2.2 which is taken from Kurosu and Chance (2014). Filtering bad or missing measurement pixels is performed prior to any other filtering, this includes the datapoints affected by the row anomaly. This anomaly (`rowanomaly_url`) 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 removed before any further processing.

Each  $\sim 90$  minutes the AURA satellite sweeps over the sunny side of the planet of which around 50 k – 80 k of the roughly 90 k pixels 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 using a modelled a priori shape factor rather than NASA's included a priori shape factor. Each pixel includes an estimate of the cloud fraction created using 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.

One more filter is applied before any calculations, to remove unreasonable column amounts. Due to numerous highly negative vertical columns (beyond what is expected) a screen is applied to remove any pixels with vertical columns outside the range of  $-5 \times 10^{15}$  to  $1 \times 10^{17}$ . This has been performed previously in TODO

### 2.3.2 DOAS

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

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

TABLE 2.2: 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	<p>Caution advised because one or more of the following conditions are present:</p> <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 (1e19 molec cm<math>^{-2}</math>)</li> </ul>
2	Bad	<p>Avoid using as one of the following conditions are present:</p> <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

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 ( $\eta, \text{molec cm}^{-3}$ ) 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 (s) substituting Equation 2.3 into Equation 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.

### 2.3.3 Air mass factor (AMF)

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

changes with trace gas optical depths at specific altitudes (Lorente et al. 2017). Calculation of the AMF is important as it is multiplied against the estimated slant columns in order to give vertical column amounts.

DOAS column retrievals 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 prescribed altitudes using an RTM. These properties are combined into a single array called the AMF. Two examples of this are GOME-2 (on the MetOp-A satellite) products ([http://atmos.caf.dlr.de/gome/product\\_hcho.html](http://atmos.caf.dlr.de/gome/product_hcho.html)), and OMI products which respectively use LIDORT combined with IMAGESv2 and GEOS-Chem for processing (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 largest error sources in remote sensing of BVOCs (Palmer et al. 2001; Millet et al. 2006)). Lam-sal 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.

Related to the AMF is the averaging kernal (AK), which is used to handle instrument measurements which are sensitive to gas concentrations at different altitudes through 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 satellite AK (or AMF), which takes into account the vertical profile of the modelled trace gas and instrument sensitivity to the trace gas (Eskes and Boersma 2003; Palmer et al. 2001). This process is called deconvolution ( $\Omega = AK \times VC_{satellite} + \times(I - AK)VC_{apriori}$ ) of the 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:

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

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).

#### 2.3.4 LIDORT

LIDORT is a model of LInearized Discrete Ordinate Radiative Transfer, used to determine backscatter intensities and weighting functions at arbitrary elevation angles (Spurr, Kurosu, and Chance 2001). 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 (e.g., Palmer et al. 2001; Martin et al. 2002a; De Smedt et al. 2015; Gonzalez Abad et al. 2015).

### 2.3.5 Uncertainty

Uncertainty in a single pixel for OMHCHO is roughly the same magnitude as HCHO background levels. Each pixel has  $\sim 2 \times 10^{14}$  molec  $\text{cm}^{-2}$  uncertainty, which is  $5\times$  higher than GOME. However, there are  $\sim 100 - 200\times$  as many measurements allowing a greater reduction of uncertainty with averaging. This is due to the smaller footprint and better temporal resolution of OMI (Instrument 2002; Millet et al. 2008). The finer nadir resolution of OMI (13 by 24  $\text{km}^2$ ) compared to other satellites also reduces cloud influence (Millet et al. 2006; Millet et al. 2008). The top row in Figure 2.9 shows OMI HCHO columns binned to at  $0.25^\circ \times 0.3125^\circ$  longitude by latitude averaged over one day and one month (with and without filtering). Row two shows uncertainty of the satellite data after averaging. It is clear that one day of satellite data is too uncertain when binned at  $0.25^\circ \times 0.3125^\circ$  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. High resolution low detection limit estimates can be built up using “oversampling”, which averages satellite measurements over time (e.g., Zhu et al. 2014). Uncertainty in satellite recalculations, along with other factors is analysed in Section ??.

In this thesis, HCHO columns (pixels) with cloud fractions over 40% are filtered as done in Palmer et al. (2001), which introduces a clear-sky bias to any monthly averages. This is due to HCHO being lower on unrecorded cloudy days. This bias has been measured as a 13% positive monthly mean bias (Palmer et al. 2001; Surl, Palmer, and Abad 2018).

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. 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. Their analysis of OMI concludes that data is still of high quality and will deliver useful information for 5-10 more years, with radiance only changing by 1 – 2% outside of RA impacted areas. An analysis of the row anomaly by Huang et al. (2017) state that measurements remain suitable for scientific use, with recommendation for further evaluation. 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).

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 artefact 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

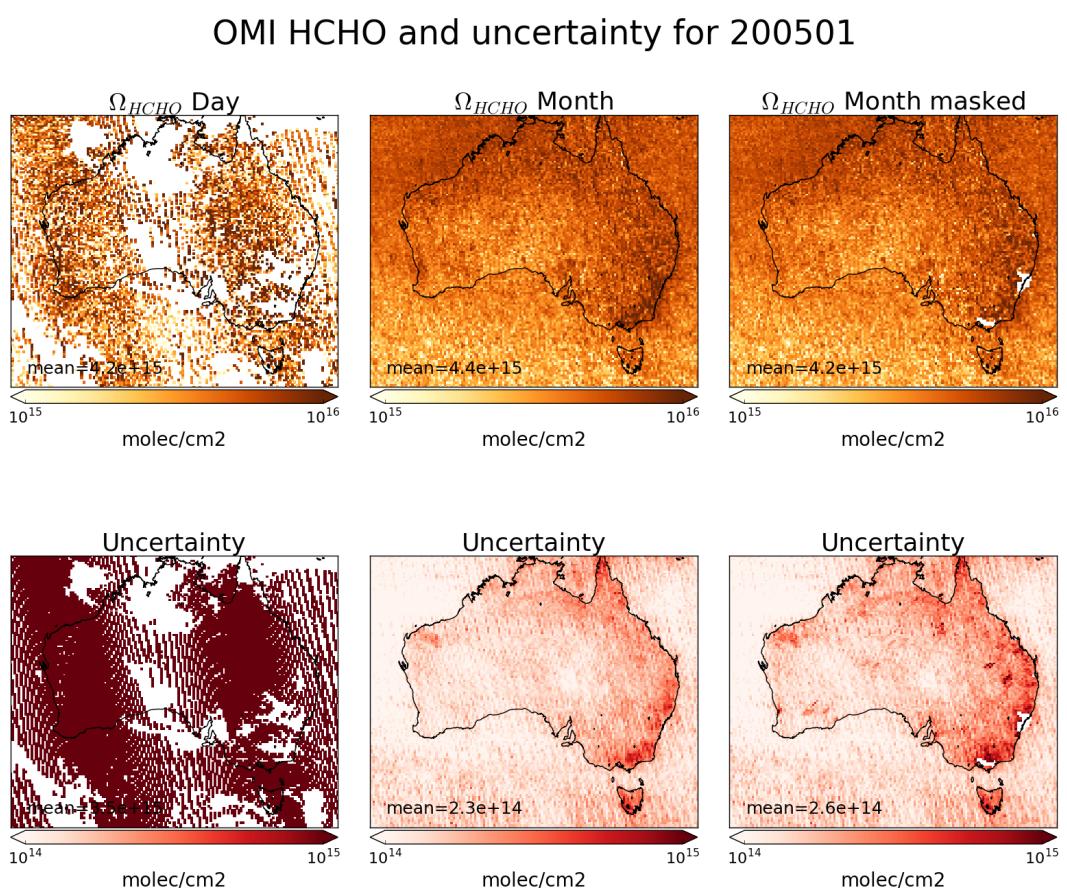


FIGURE 2.9: Top row shows  $0.25^\circ \times 0.3125^\circ$ binned OMHCHO columns with one day, one month, and one month with non-biogenic masking applied from left to right respectively. Bottom row shows the uncertainty for each gridsquare after averaging.

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 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 comes from 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).

## 2.4 GEOS-Chem

### 2.4.1 Overview

GEOS-Chem is a well supported global, Eulerian CTM (see Section 1.5.2) 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 time steps within a global set of 3-D boxes. Emissions are either prescribed by inventories or modelled (e.g., fire emissions are pulled from the global fire emissions database GFED4).

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  $2^\circ \times 2.5^\circ$  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. 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 global fire emissions database (GFED4) product (Giglio, Rander-son, and Van Der Werf 2013). Anthropogenic VOC emissions come from the Emission Database for Global Atmospheric Research (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.

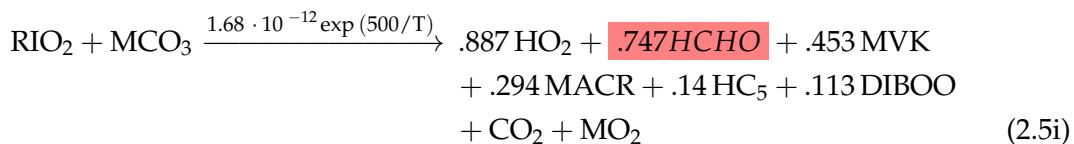
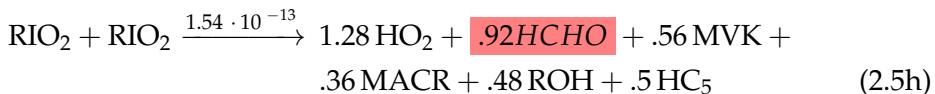
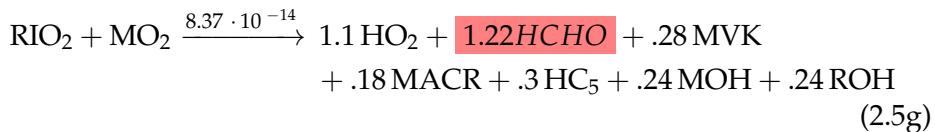
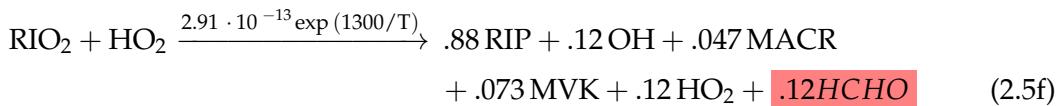
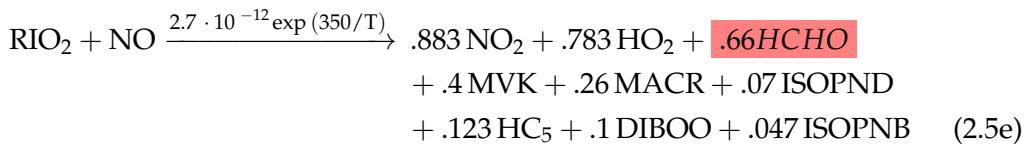
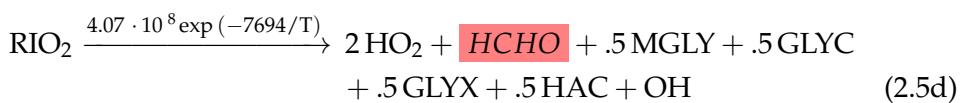
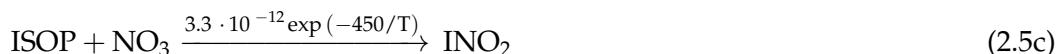
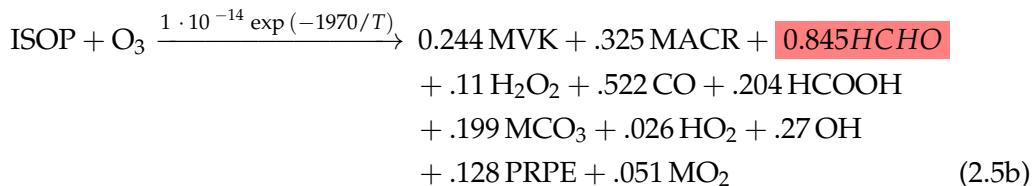
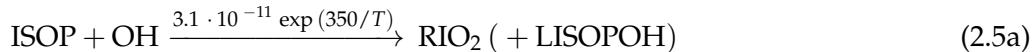
### 2.4.2 Installing and running GEOS-Chem

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 (for us  $\sim 70$  computation hours per month). I installed GEOS-Chem onto a suitably configured workspace on the Raijin supercomputer at 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 Chemical Mechanism

A chemical mechanism is a closed system of chemical reactions and their associated rate constants. Chemical reactions are represented by systems of differential equations to be solved for each gridbox in GEOS-Chem. Simplifications are required due to the large number of reactions which occur in the atmosphere, and the coupled and stiff nature of these reactions which slow down computation of the solutions (Brasseur and Jacob 2017). Stiffness in chemical systems of differential equations is due to the massively differing reaction time scales - for instance hydroxyl radicals react within seconds while methane has an atmospheric lifetime of 8-10 years (Wuebbles and Hayhoe 2002).

Some of the important reactions involving isoprene are reproduced here, including reaction rates ( $k$ ) in the form  $k = A \exp -E/RT$ , where  $T$  is temperature,  $E$  is activation energy, and  $R$  is the gas constant,  $A$ ,  $E$ , and  $R$  are predefined for each reaction. Equation 2.5 lists the main isoprene and sundry reactions, with terms defined in 2.3. LISPOH is added in order to allow the model to keep track of how much isoprene is oxidised by OH. The mechanisms used in version 10.01 (and its history) are described online at [http://wiki.seas.harvard.edu/geos-chem/index.php/NOx-Ox-HC-Aer-Br\\_chemistry\\_mechanism](http://wiki.seas.harvard.edu/geos-chem/index.php/NOx-Ox-HC-Aer-Br_chemistry_mechanism).



#### 2.4.4 GEOS-Chem isoprene

The isoprene reactions simulated by GEOS-Chem were originally based on Horowitz et al. (1998). The mechanism was subsequently updated by Mao et al. (2013), who changed the isoprene nitrate yields and added products based on Paulot et al. (2009a) and Paulot et al. (2009b). They used the yields and reactions of various positional isomers of isoprene nitrates, and their oxidation products. Further mechanistic properties, like isomerisation rates, are based on results from four publications: Peeters, Nguyen, and Vereecken (2009), Peeters and Muller (2010), Crounse et al. (2011), and Crounse et al. (2012).

TABLE 2.3: Species or classes from the GEOS-Chem mechanism.

Name	Definition
ACTA	Acetic acid: CH <sub>3</sub> C(O)OH
DIBOO	Dibble peroxy radical
GLYC	Glycoaldehyde: HOCH <sub>2</sub> CHO
GLYX	Glyoxal: CHOCHO
HAC	Hydroxyacetone: HOCH <sub>2</sub> C(O)CH <sub>3</sub>
INO2	RO <sub>2</sub> from ISOP+NO <sub>3</sub>
ISNP	an isoprene nitrate
ISOPNB	$\beta$ isoprene nitrates
ISOPND	$\delta$ isoprene nitrates
MACR	Methacrolein: CH <sub>2</sub> =C(CH <sub>3</sub> )CHO
MCO3	Peroxyacetyl radical: CH <sub>3</sub> C(O)OO
MEK	Methyl ethyl ketone: RC(O)R
MGLY	Methylglyoxal: CH <sub>3</sub> COCHO
MO2	Methylperoxy radical: CH <sub>3</sub> O <sub>2</sub>
MOH	Methanol: CH <sub>3</sub> OH
MVK	Methylvinylketone: CH <sub>2</sub> =CHC(=O)CH <sub>3</sub>
PRPE	>C <sub>3</sub> alkenes: C <sub>3</sub> H <sub>6</sub> , ...
RIO2	isoprene peroxy radical: ROO
ROH	>C <sub>2</sub> alcohols

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. Formation of isoprene nitrates (ISOPN) affect ozone levels through NO<sub>x</sub> sequestration, the yields and fate of these nitrates was analysed in Paulot et al. (2009a). 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; Wolfe 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).

#### 2.4.4.1 Oxidation

Crounse et al. (2011) examined the isomerisations associated with the oxidation of isoprene to six different isomers of ROO formed in the presence of oxygen. The primary oxidation pathway of isoprene is reaction with the OH radical (*ISOP* + OH → *RIO<sub>2</sub>*) shown in Equation 2.5a. Isoprene undergoes OH addition at the 1 and 4 positions, becoming  $\beta$  (71%) or  $\delta$  (29%) ROO, although these are not distinguished in the GEOS-Chem mechanism. Secondary oxidative pathways are from ozonolysis (Equation 2.5b) and reaction with NO<sub>3</sub> (Equation 2.5c). These pathways are much slower but can be important in particular scenarios such as inside a pollution plume or at night when NO<sub>3</sub> radicals can build up.

Following isoprene oxidation, seven potential ROO reactions (*RIO<sub>2</sub>* in Equation 2.5), can occur depending on reactant concentrations and local temperature. Crounse

et al. (2011) determined rates and uncertainties involved in these reactions, and studied the rate of formation of C<sub>5</sub>-hydroperoxyenals (C5-HPALD) by isomerisation. Reactions 2.5d - 2.5j compete to determine the fate of ROO', which is often grouped into high or low NO<sub>x</sub> concentration pathways.

#### 2.4.4.2 Nitrogen oxide impacts

GEOS-Chem reactions account for high and low NO<sub>x</sub> scenarios as determined in Mao et al. (2013), based on Paulot et al. (2009a). High NO<sub>x</sub> has been defined as conditions where NO reactions are the main cause of losses for ROO' (Reaction 2.5e) (Palmer 2003). High NO<sub>x</sub> concentrations are roughly 1 ppb, while low NO<sub>x</sub> concentrations are around 0.1 ppb.

In low NO<sub>x</sub> ROO' losses occur from isomerisation or reaction with several potential compounds. NO<sub>x</sub> + HO<sub>2</sub> (Reaction 2.5f) produces mostly hydroxy hydroperoxides (ISOPOOH), and some HCHO. ISOPOOH can be oxidised (by OH) to produce epoxydiols, recycling OH (Paulot et al. 2009b). Isomerisation of ROO' (Reaction 2.5d) largely produces HCHO, while recycling OH and producing HO<sub>2</sub>. This isomerisation accounts for 1,5-H shifts producing MACR, MVK, HCHO, or 1,6-H shifts producing HPALDs. HPALDs can photolysis to regenerate OH and small VOCs (Crounse et al. 2011; Wolfe et al. 2012; Jozef et al. 2014). Under low NO<sub>x</sub> conditions the expected production of HCHO, MVK, and MACR is 4.7%, 7.3%, and 12% respectively. Refer to Section 1.3.3 for more information. Less frequently, ROO' reacts with itself, MO<sub>2</sub>, or MCO<sub>3</sub>.

Under high NO<sub>x</sub> conditions, the fate of ROO' differs depending on how it was formed. 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%). These two pathways are combined in Reaction 2.5e, which shows the production from ROO' and NO.

#### 2.4.4.3 OH

Until recently, models struggled to represent OH in the atmosphere, due to missing chemistry and issues in measurement techniques. 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. OH regeneration through photolysis of HPALDs in areas with high isoprene emissions are included from Peeters and Muller (2010). Photolysis of photolabile peroxy-acid-aldehydes generates OH, improving 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 sink of OH (Peeters and Muller 2010; Taraborrelli et al. 2012). Isoprene chemistry in GEOS-Chem includes OH regeneration from oxidation of epoxydiols (not shown in Equation 2.5) and slow isomerisation of ROO' (Equation 2.5d) (Mao et al. 2013). This regeneration dealt with the problem seen in older models where ISOPOOH production titrated OH, which was not backed up by measurements (Paulot et al. 2009b; Mao et al. 2013). Mao et al. (2013) showed that drastically lowering (by a factor of 50) the rate constant for ROO' isomerisation lead

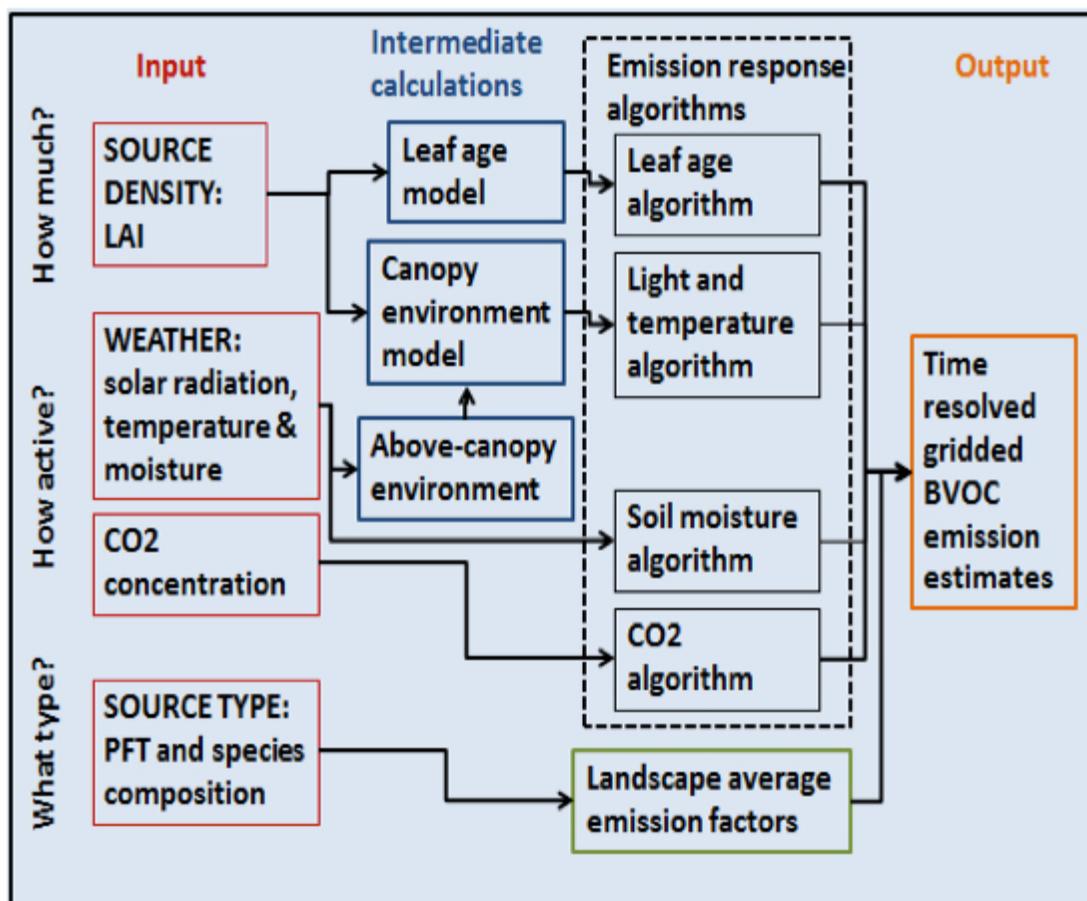


FIGURE 2.10: MEGAN schematic, from <http://lar.wsu.edu/megan/>

to better organic nitrate agreements with measurements. These chemical updates led to more accurate modelling of OH concentrations, especially in low NO<sub>x</sub> conditions common in remote forests. 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 GEOS-Chem version 11.

#### 2.4.5 Emissions from MEGAN

GEOS-Chem runs the Model of Emissions of Gases and Aerosols from Nature (MEGAN) to determine biogenic emissions globally. GEOS-Chem V10.01 uses MEGAN V2.1 with biogenic emissions from Guenther et al. (2012). MEGAN is itself a global model with resolution of around 1 km, which uses globally estimated or measured (including remote sensed satellite data) inventories of parameters to estimate emissions. MEGAN uses leaf area index (LAI), global meteorological data, plant functional types (PFT), and photosynthetic photon flux density to simulate terrestrial isoprene emissions (Kefauver, Filella, and Peñuelas 2014). The schematic for MEGAN, from <http://lar.wsu.edu/megan/>, is shown in Figure 2.10.

MEGAN was developed as a replacement for two earlier canopy-environment emission models (the biogenic emission inventory system, and the global emissions

initiative), initially including a simple canopy radiative transfer model which parameterised sun-lit and shaded conditions through a canopy. Early models did not account for abiotic stresses, such as drought, prior rainfall and land use changes. These stresses influence species-specific emissions by more than an order of magnitude (Ninemets 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). 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). 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. GEOS-Chem uses an embedded version of MEGAN 2.1.

GEOS-Chem computes some emissions using predefined emission factor ( $\epsilon$ ) maps from MEGAN source code, and others using PFT maps and associated emission factors. For isoprene (the focus in this thesis) MEGAN calculates emissions online, using local meteorological conditions. Emissions  $E$  of species  $i$  are calculated using these EF for classes of plant types  $j$  and associated grid-box coverage  $\xi$  and an activity factor  $\gamma$  which accounts for response to environmental conditions. The following equation is reproduced from Guenther et al. (2012) showing how  $E_i$  are determined in MEGAN:

$$E_i = \gamma_i \sum_j \epsilon_{i,j} \chi_j \quad (2.6)$$

For example: isoprene emissions are tied to sunshine, temperature, plant emission strength ( $\epsilon$ ), emitting species coverage, etc., for each grid box.

#### 2.4.6 Nitrogen oxides

$\text{NO}_x$  concentrations affect atmospheric oxidative capacity, which changes many factors important in estimating isoprene emissions including isoprene to HCHO yield, isoprene lifetime, and isoprene oxidation pathways. In GEOS-Chem,  $\text{NO}_x$  concentrations are regulated by  $\text{O}_3$ , VOC,  $\text{HO}_x$ , Bromine and aerosols reactions.  $\text{NO}_x$  emissions (from power generation and combustion transport) drive enhancements, while conversion to nitric acid ( $\text{HNO}_3$ ) followed by deposition is the primary removal mechanism (Delmas, Serca, and Jambert 1997; Ayers and Simpson 2006). Anthropogenic emissions (TODO total  $\text{Tg N a}^{-1}$ ) are from the Emissions Database for Global Atmospheric Research (EDGAR), for more details visit [http://wiki.seas.harvard.edu/geos-chem/index.php/EDGAR\\_v4.2\\_anthropogenic\\_emissions](http://wiki.seas.harvard.edu/geos-chem/index.php/EDGAR_v4.2_anthropogenic_emissions). Other  $\text{NO}_x$  emissions arise from sources including soil emissions ( $\sim 10 \text{ Tg N a}^{-1}$ ), and lightning. Soil emissions are layed out in **Hudson2012** parameterised using biome specific emission factors, and an explicit fertiliser dataset. Lightning based  $\text{NO}_x$  production is created based on inventories of lightning flash rates scaled within GEOS-Chem.

If GEOS-Chem is misrepresenting  $\text{NO}_x$ , yields and the effects of transport may be incorrectly accounted for. In order to determine the accuracy of GEOS-Chem simulated  $\text{NO}_x$  over Australia, modelled  $\text{NO}_2$  amounts are compared to satellite data

(where available) for 2005. Figure 2.11 shows tropospheric NO<sub>2</sub> columns from GEOS-Chem, OMNO2d, and their differences over four seasons in 2005. Simulated GEOS-Chem tropospheric NO<sub>2</sub> columns averaged from 13:00-14:00 local time (LT) are compared against OMNO2d data which are averaged into seasonal 2° × 2.5°bins. Both data sets show Sydney and Melbourne as NO<sub>2</sub> hotspots, however Sydney itself is underestimated by GEOS-Chem throughout the year, likely due to low or poorly aligned emission maps (shown in Figure 2.12). Over much of the country GEOS-Chem overestimates NO<sub>2</sub> by 10-60%, except in the northern areas where up to 50% underestimation occurs in Summer. An underestimation in southern NSW can also be seen in the later half of the year. Notably GEOS-Chem anthropogenic NO does not have any seasonality (Figure 2.12), while soil shows clear differences in emission hotspots between summer and winter.

Figure 2.13 shows scatter plots with one data point for each land square over Australia, for each season, coloured by total NO emissions. The correlation between model and satellite NO<sub>2</sub> columns is reasonable throughout the year over Australia, with a general positive bias throughout the year in modelled amounts. A comparison between the bias (GEOS-Chem - OMNO2d) with anthropogenic or soil emissions (columns 2 and 3 respectively) shows no real correlation, suggesting the bias is not driven by either anthropogenic or soil NO emissions in any season. Without a clear link between emissions and biases, alterations would be too subjective, and a change to NO<sub>x</sub> chemistry is beyond the scope of this thesis. Additionally, high satellite NO<sub>2</sub> columns are filtered in later calculations which assume biogenic air masses as they suggest anthropogenic influence. 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 measurements over Australia. For this reason modelled NO emissions are not modified in model runs in this thesis.

#### 2.4.7 GEOS-Chem simulations

GEOS-Chem is run five times independently in this thesis, with different outputs from each simulation used to determine specific information. The different output types are first described in Section 2.4.7.1. Following this is the list of model runs, including a summary of the run, outputs created, and a summary of how they are used (Section 2.4.7.2). Finally a brief comparison between a subset of the runs is performed (Section 2.4.7.3).

##### 2.4.7.1 GEOS-Chem outputs

There are various outputs available when running GEOS-Chem, some of which require manipulation in order to compare against observations. GEOS-Chem in this thesis is run with a 15 minute time step for transport and 30 minutes for chemistry, at 2° × 2.5°horizontal resolution over 47 vertical levels. Output is the average of these time steps. Optionally one can save high temporally resolved data for a single (or list of) column(s). This has been used here to compare modelled ozone with ozonesonde profiles at three sonde release sites discussed in Chapter 4.

**Midday output** is output from averaging over a window of local time for each grid-box. Output averaged between 13:00-14:00 local time is saved to match with

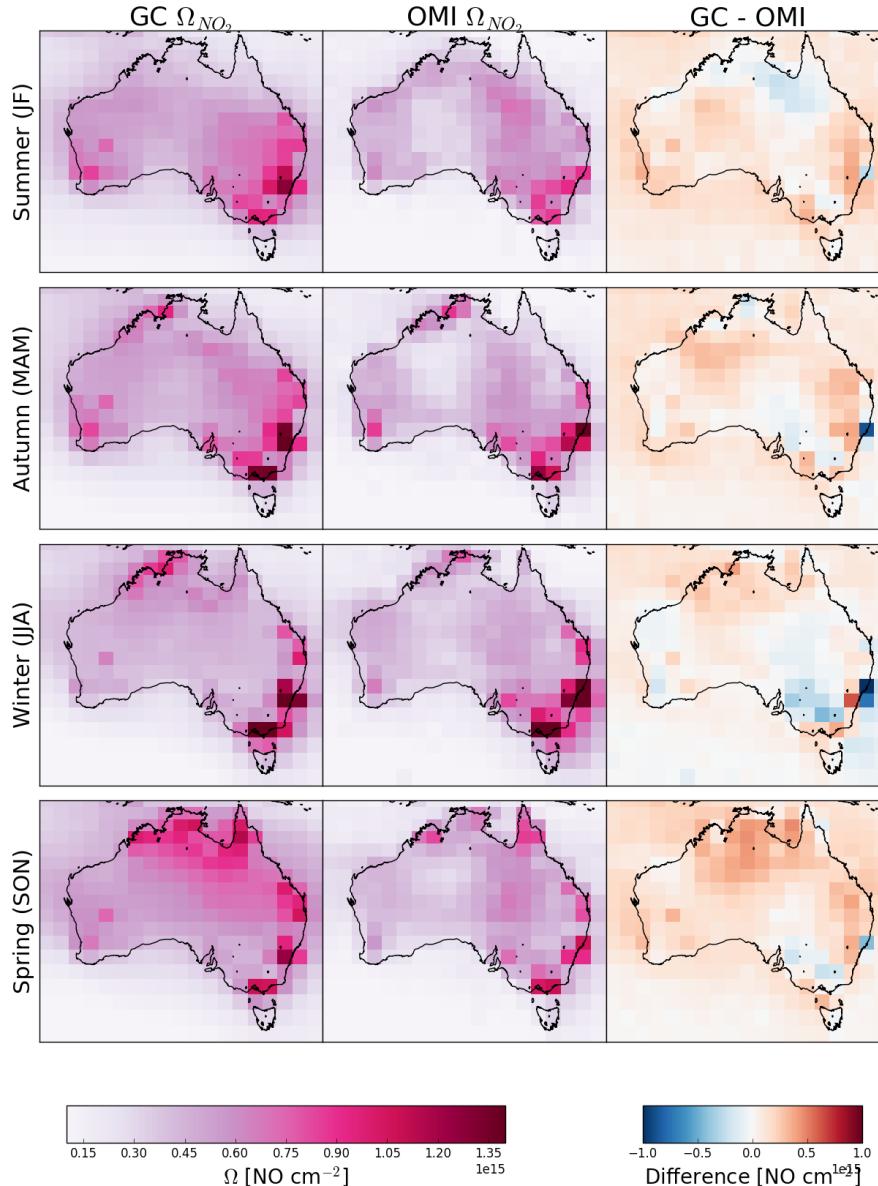


FIGURE 2.11: Left to right columns show tropospheric  $NO_2$  columns ( $\Omega_{NO_2}$ ; molec  $cm^{-2}$ ) from GEOS-Chem (GC), OMNO2d (OMI), and the differences. Each row shows one season from 2005, the left two columns use the left colour scale, while the third column uses the right colour scale.

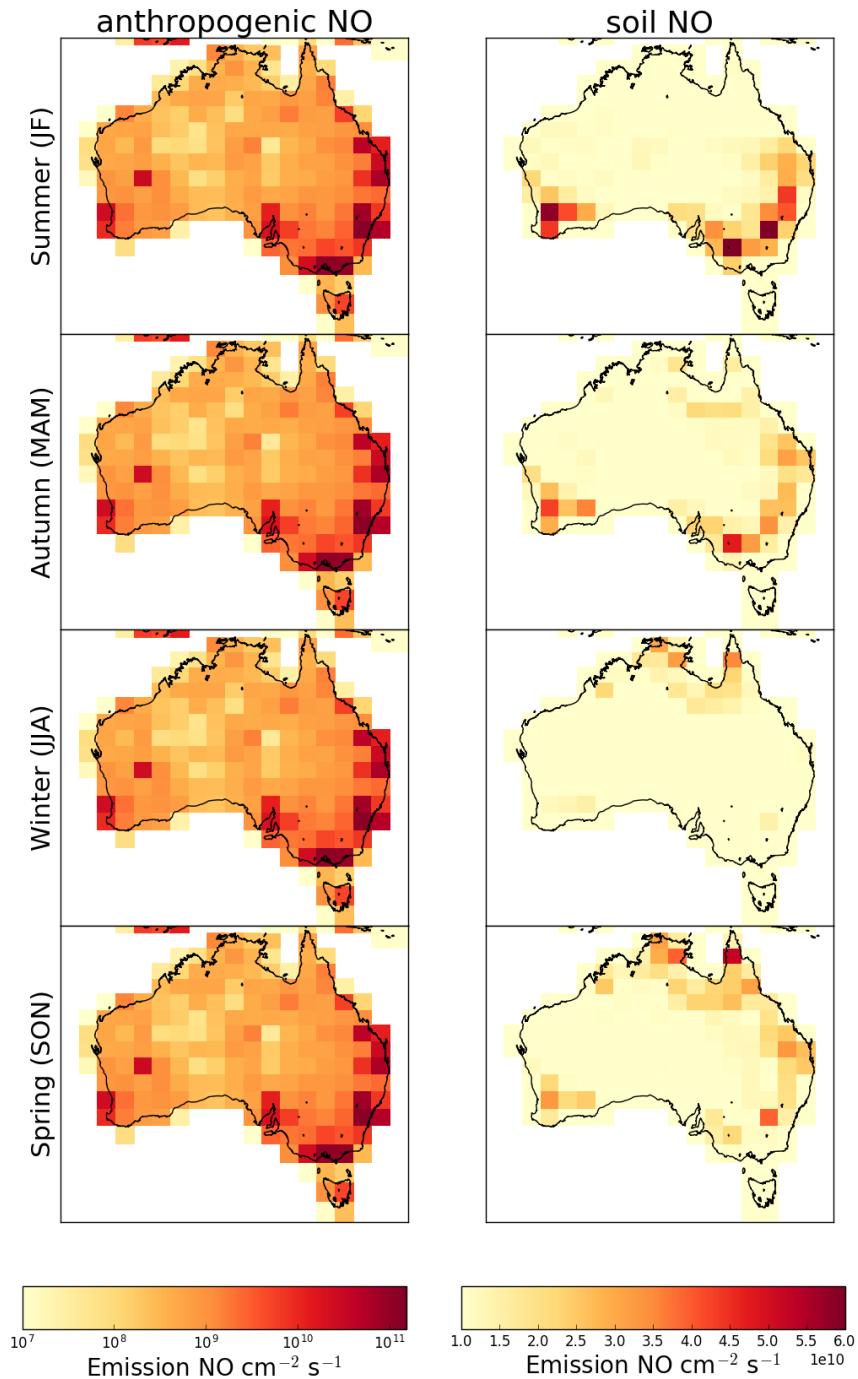


FIGURE 2.12: Left and right columns show anthropogenic and soil emissions of NO respectively from GEOS-Chem in molec  $\text{cm}^{-2} \text{s}^{-1}$ . Each row shows one season from 2005. Anthropogenic and soil emissions use a logarithmic and linear colour scale respectively.

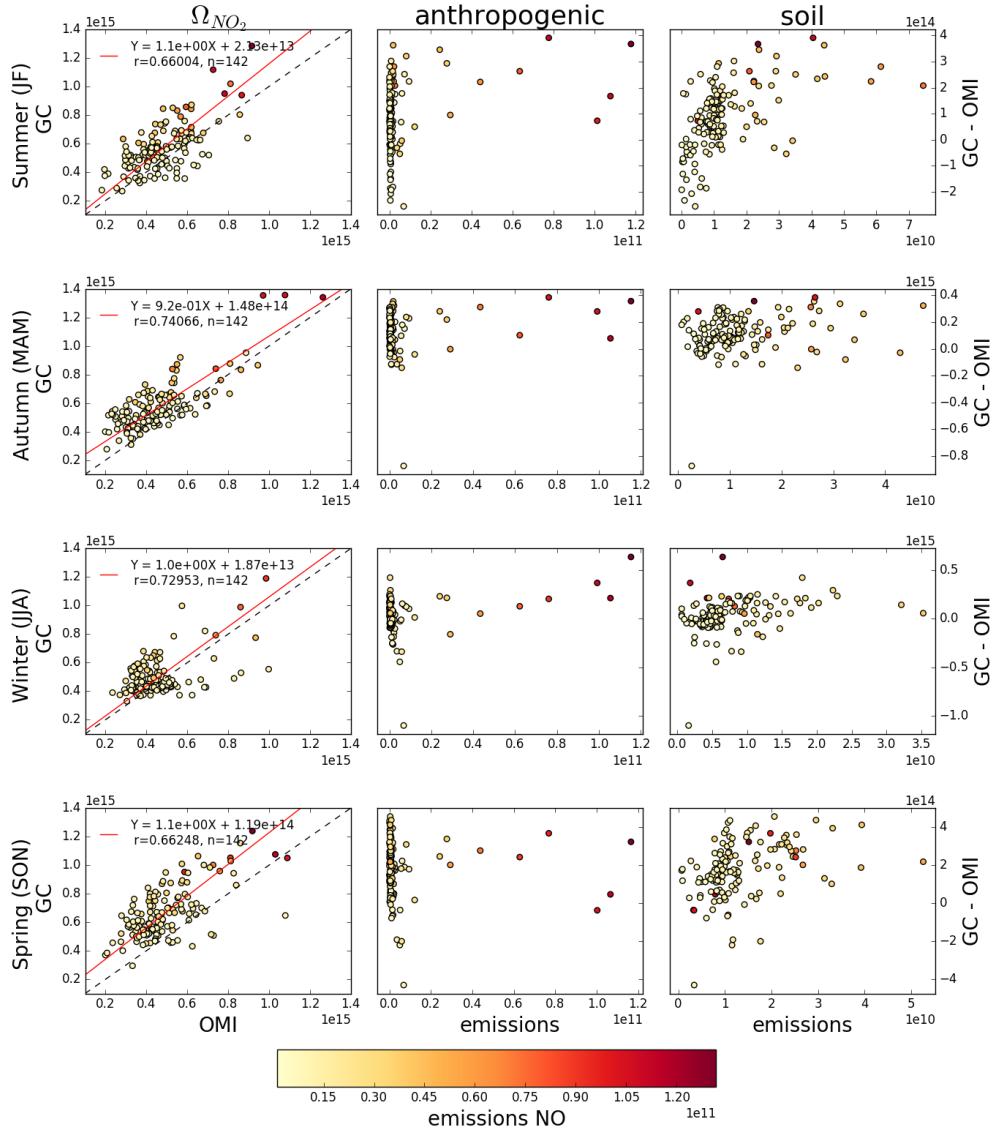


FIGURE 2.13: The first column shows the scatter plots of tropospheric  $NO_2$  columns between GEOS-Chem (y-axis) and OMNO2d (x-axis) at  $2^\circ \times 2.5^\circ$ . The reduced major axis linear regression is drawn in red, and the equation, regression coefficient ( $r$ ), and number of grid squares ( $n$ ) used is inlaid as a legend. The second and third columns show the scatter between emissions and the bias between GEOS-Chem and OMNO2d (GC-OMI), for anthropogenic and soil emissions respectively. These two columns share the far right axis, however emissions are from anthropogenic and soil sources respectively. All scatter points are coloured by the sum of anthropogenic and soil NO emissions (from GEOS-chem), as per the colour bar shown at the bottom.

Aura satellite measurements, as Aura overpasses at  $\sim$ 1330 local time each day. Midday output is saved both for comparison with, and recalculation of, satellite measurements.

**HEMCO diagnostics** :the Harvard-NASA Emissions Component (HEMCO) deals with emissions inventories used in GEOS-Chem. When working with globally gridded data, handling local time offsets becomes more important. The hourly averaged emissions of isoprene saved using GMT which needs to be offset based on longitude in order to retrieve local time. This is done 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  $15^\circ$ (as  $360^\circ$  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+local time (modulo 24) with the desired hour on this grid over the 24 GMT hours.

**Tracer averages** are daily or monthly averaged gridbox concentrations.

**Time series** are vertical columns with diagnostics saved at a temporal resolution of up to 15 minutes.

#### 2.4.7.2 GEOS-Chem runs

The following list summarises each model run as well as enumerating the outputs (described above), and how the run is used in the thesis.

**UCX** The model was run using the Universal tropospheric-stratospheric Chemistry eXtension (UCX) mechanism with 72 vertical levels from the surface to the top of the atmosphere (TOA $\sim$ 0.1 hPa). UCX runs a chemistry mechanism with combined tropospheric and stratospheric reactions, with an increased number of stratospheric calculations performed online (Eastham, Weisenstein, and Barrett 2014).

1. Outputs: Midday output, daily tracer averages, and time series over three stations.
2. The Midday output is used to check how shape factors for AMF recalculation are affected by vertical resolution and stratospheric chemistry. TODO: Make a plot of AMF with and without 72 vertical levels
3. This run shows what influence the stratospheric chemistry additions have over tropospheric isoprene or HCHO concentrations.
4. The daily tracer averages are used to determine ozone intrusion quantification (Section 4.6.1), and ozone concentration seasonality (Section 4.5).
5. Time series outputs are compared against ozonesonde releases (Section 4.5) both over time and vertically.

**Tropchem (standard)** default settings for GEOS-Chem 10.01, using 47 vertical levels at  $2^\circ \times 2.5^\circ$ horizontal resolution. Additional midday output is created to allow AMF recalculation code to run on OMI satellite measurements (Section 2.6.5).

1. Outputs: Midday output, daily tracer averages, and HEMCO diagnostics
2. These are used in recalculation of the satellite AMF (Section 2.5), and the modelled background HCHO over the remote Pacific which is used in the reference sector correction for OMI column retrievals (Section 2.6.7). Additional diagnostic outputs are used hard coded to allow AMF recalculation.
3. Midday output is combined from two different runs in order to determine smearing (Section 3.2.7)
4. Total yearly ozone concentrations are compared before and after scaling isoprene emissions using the top-down estimate.

**Tropchem (isoprene emissions halved)** identical to standard tropchem except isoprene emissions are halved.

1. Outputs: Midday output, and monthly tracer averages
2. Check modelled ozone sensitivity to isoprene emissions TODO: plot ozone between the two runs and reference here
3. Combined with standard run to determine model sensitivity transport (Section 3.2.7)

**Tropchem (biogenic emissions only)** identical to standard tropchem except all non-biogenic emissions inventories are disabled.

1. Outputs: Midday output, and hourly biogenic emissions from MEGAN
2. Midday and HEMCO outputs are used to determine isoprene to HCHO yield, after removing days with high smearing (Section 3.2.4)
3. HEMCO diagnostics are compared against top-down estimations of isoprene emissions (Section 3.3.1)

**Tropchem (altered MEGAN scaling factor)** Identical to standard tropchem except isoprene emissions are scaled to match multi-year monthly averaged top-down estimates.

1. Midday output, time series, daily averaged tracers
2. Compare to campaign datasets after altering isoprene emissions (see Chapter 3)
3. Compare against satellite column HCHO as a sanity check on improving isoprene emissions

#### 2.4.7.3 UCX vs tropchem

Here we compare the model output with and without enabling the Universal tropospheric-stratospheric Chemistry eXtension (UCX). Both runs use  $2^\circ \times 2.5^\circ$  latitude by longitude, however the UCX mechanism is run with 72 vertical levels from the surface to the top of the atmosphere (TOA  $\sim 0.1$  hPa), while the standard (tropchem) run uses 47 levels. The reduced vertical levels are created by lumping together sets of the higher resolved levels from around 70 hPa to the top of the atmosphere. For both runs the

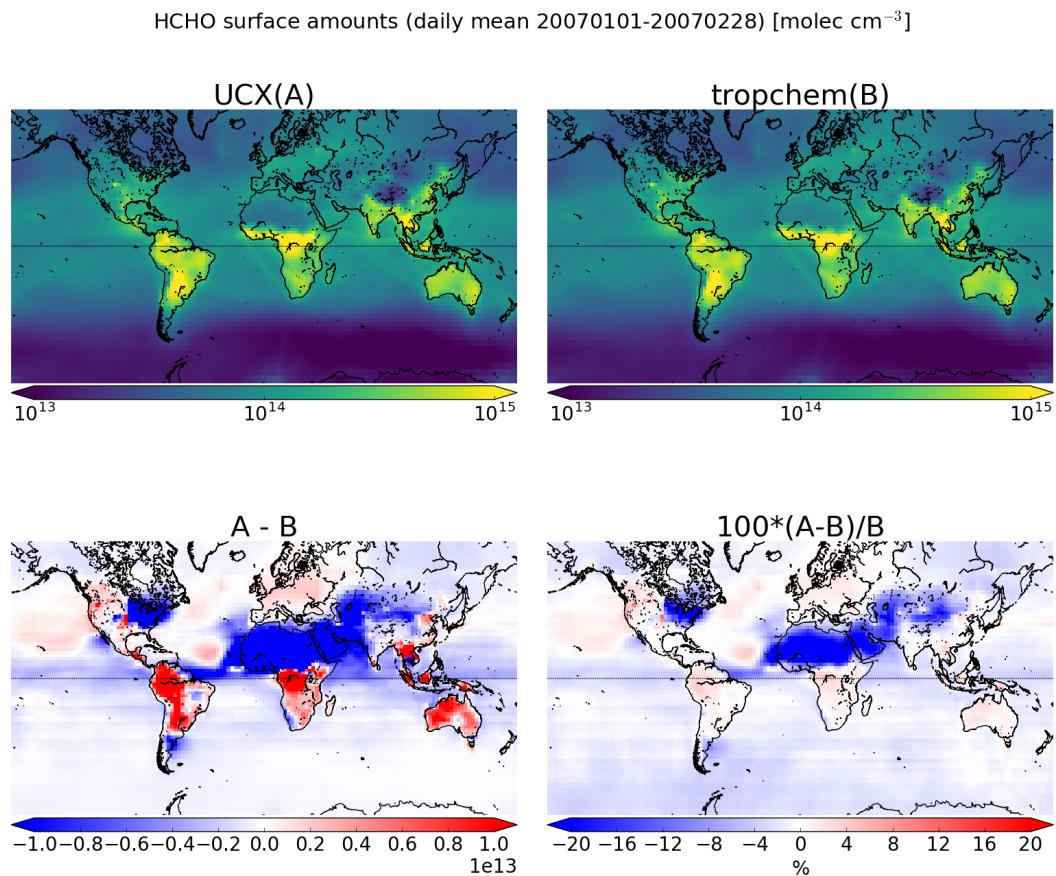


FIGURE 2.14: Surface HCHO simulated by GEOS-Chem with UCX (top left), and without UCX (top right), along with their absolute and relative differences (bottom left, right respectively). Amounts are the average of all times between 1 Jan and 28 Feb 2007.

input parameters such as MEGAN emissions and GEOS-5 meteorological fields are identical.

Figure 2.14 shows an example of surface HCHO amounts, averaged over Jan and Feb, 2007, with and without (A, B respectively) the UCX mechanism enabled. Surface HCHO (first model level; up to  $\sim 100$  m) does not differ much between runs. The differences do not exceed 3% over Australia, and absolute differences are minor (note the scale in A-B). The major notable difference occurs over northern Africa, where HCHO is around 20% lower in the UCX run. Additionally a slight ( $< 8\%$ ) decrease in HCHO over the oceans can be seen. The comparison is repeated using total columns (instead of surface values) in Figure 2.15, showing that differences affecting HCHO between the model run are spread over the entire vertical column.

Figure 2.16 shows the differences in surface isoprene concentrations over Australia, averaged over 1, Jan to 28, Feb, 2007. Here we start to see a higher relative

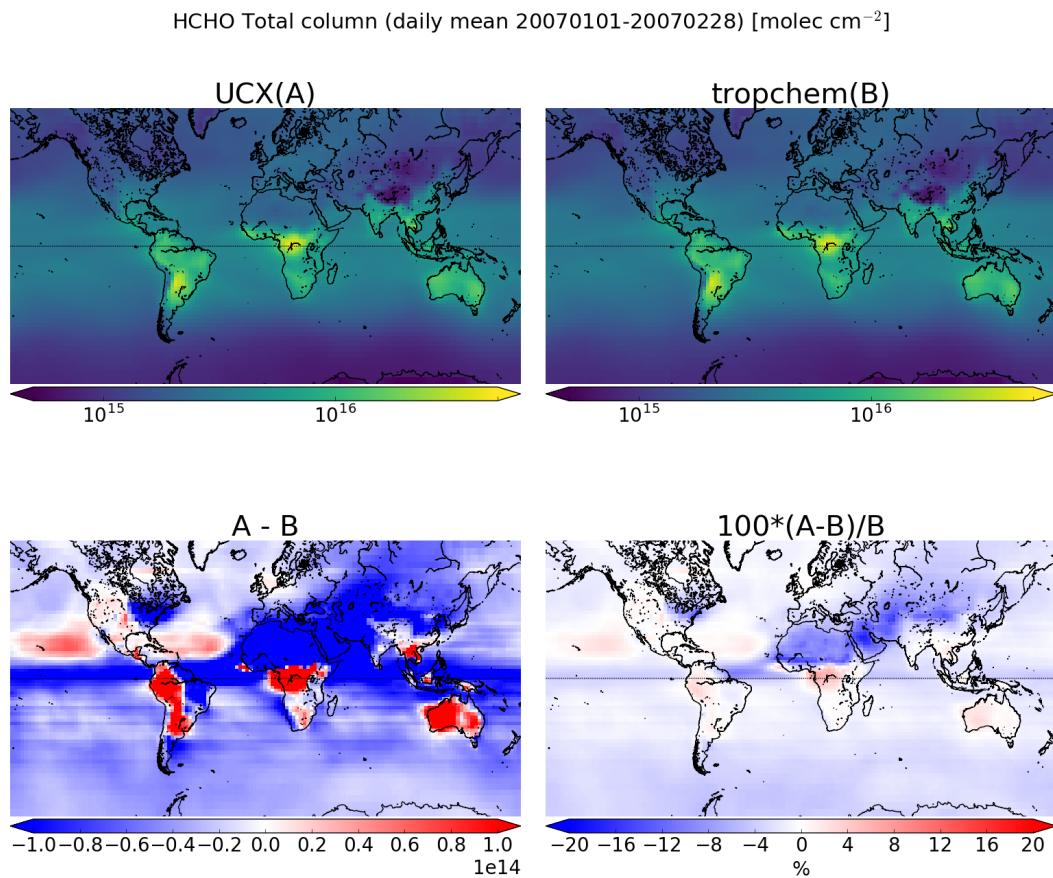


FIGURE 2.15: As Figure 2.14 using total column amounts instead of surface concentrations.

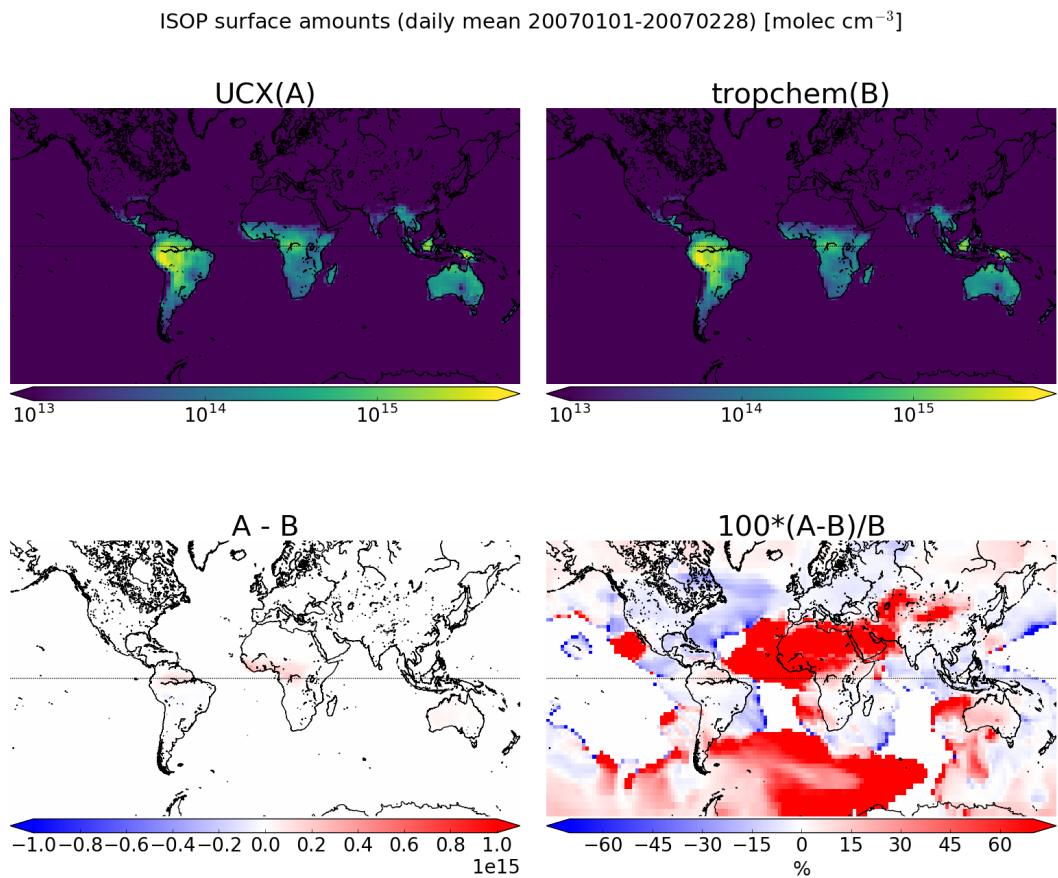


FIGURE 2.16: As Figure 2.14, except showing isoprene surface concentrations.

difference in concentrations, although this is generally over the areas with less absolute concentrations. Very little isoprene is seen away from the continents (4-5 orders of magnitude less), due to its short lifetime and lack of oceanic sources. Generally isoprene is 0-30% higher over mid to western Australia when the UCX mechanism is turned on, however this increase is lower in the regions with high isoprene emissions (north-east to south-east coastline). This enhancement can be seen throughout the entire tropospheric column as shown by Figure 2.16. There is a greater effect in Africa and South America in the tropics, with high relative differences in many regions with low absolute amounts.

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. Figure 2.18 shows the total column ozone between UCX and non-UCX run. This shows that UCX ozone is lower everywhere except for a thin band just north and south of the equator. With higher stratospheric ozone levels, less radiation would reach the troposphere. This would slow photolysis limited reactions, such as the splitting of Ozone which leads to OH production in the troposphere,

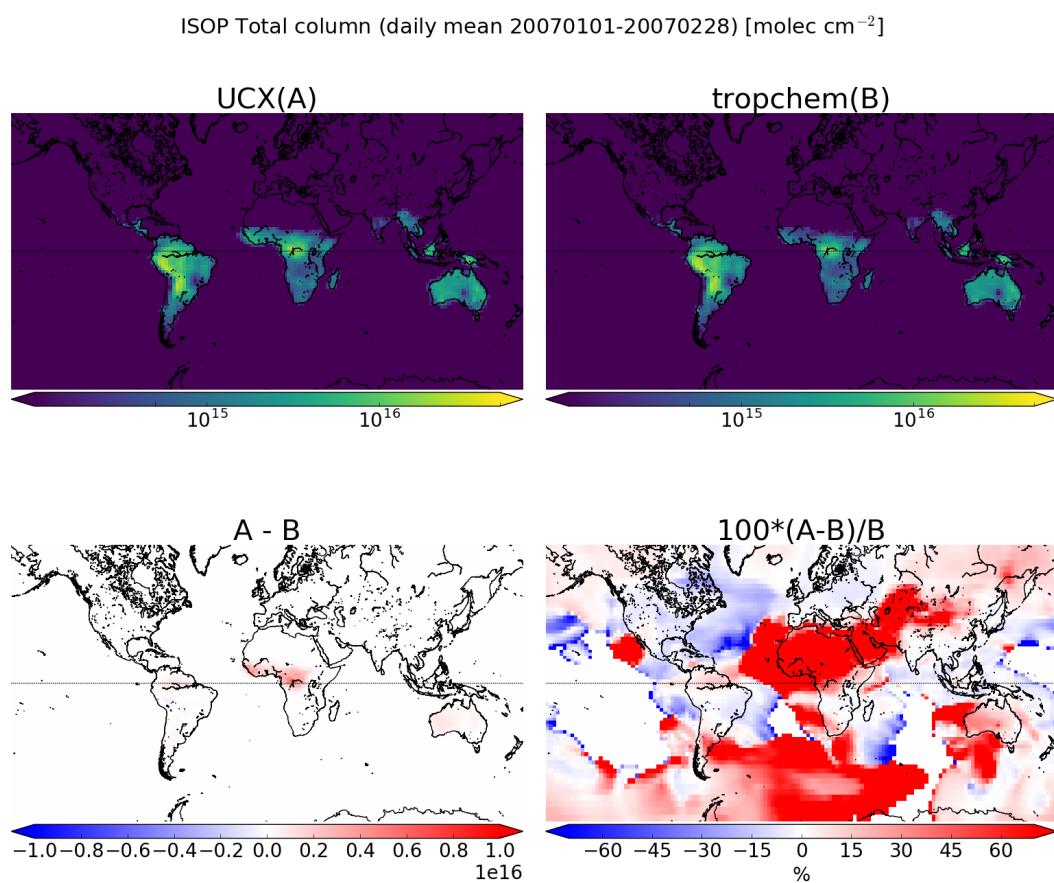


FIGURE 2.17: As Figure 2.16, except showing isoprene total column amounts.

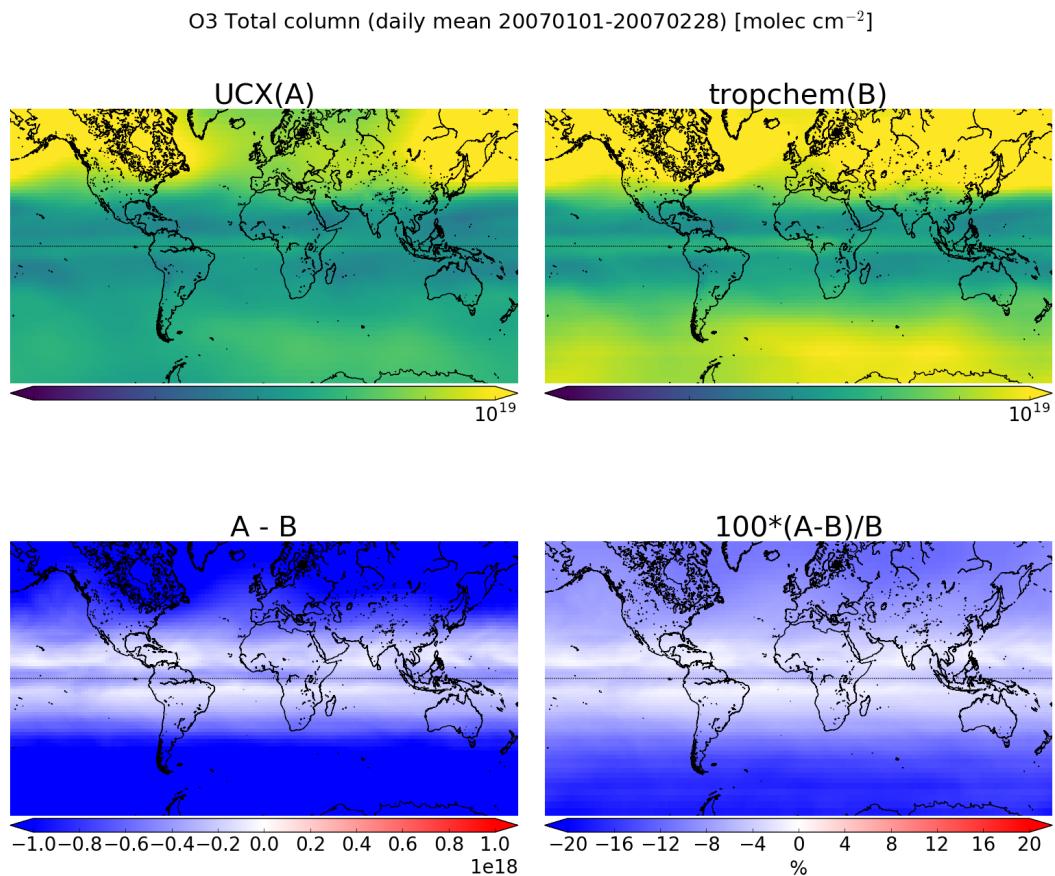


FIGURE 2.18: Total column ozone simulated by GEOS-Chem with UCX (top left), and without UCX (top right), along with their absolute and relative differences (bottom left, right respectively). Amounts are the average of all times between 1 Jan and 28 Feb 2007.

in turn slowing the isoprene loss to OH reactions.

Overall the UCX mechanism and increased vertical resolution lead to slightly higher isoprene and HCHO, and slightly lower ozone over Australia. These differences are on the order of 0-10% and as shown later (see Section 3.4) are minimal compared to other uncertainties in both AMF calculation and emissions estimation. Running UCX requires roughly twice the computation (and real) hours, and for the majority of this thesis the normal tropchem runs are used.

## 2.5 Calculating an AMF

The air mass factor (AMF) is needed to transform slant columns (SC) viewed by a satellite into vertical columns ( $\Omega$  in molecules  $\text{cm}^{-2}$ ): The AMF is the ratio of the slant

column ( $\Omega_s$ ) to the vertical column ( $\Omega_v$ )

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

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 the Aura satellite. The spectra provide backscattered intensity ( $I_B$ ) at various wavelengths (see Section 2.3), 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 2.7 to give an expression for the AMF which includes scattering:

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

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

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

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 2.9 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

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

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 2.3.2).

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 2.10 we obtain:

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

Often the integrand of this AMF formula (Equation 2.11) is broken apart into two defining terms: the scattering weights  $\omega(z)$  and the shape factor  $S(z)$ , described here:

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

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

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

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

$$S(z) = \frac{\eta(z)}{\Omega_v} \quad (2.13)$$

Plugging Equations 2.12 and 2.13 into Equation 2.11 gives us:

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

Since we are using the  $\omega$  provided by OMI, the  $AMF_G$  term is removed from this calculation as it is not part Equation 2.12 leading to

$$AMF = \int_0^\infty \omega(z) S(z) dz \quad (2.15)$$

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):

$$S_\sigma(\sigma) = \frac{\Omega_a}{\Omega_v} C_{HCHO}(\sigma) \quad (2.16)$$

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. We can integrate over the sigma coordinates using the hydrostatic relation  $P = -\rho_a g z$ , with  $\rho_a$  being air density, and  $g$  the gravitational acceleration:

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

Substitution into 2.15 gives AMF using the sigma coordinates:

$$AMF = \int_0^1 w(\sigma) S_\sigma(\sigma) d\sigma \quad (2.17)$$

## 2.6 Recalculation of OMI HCHO

OMI HCHO vertical columns are calculated using modelled a priori HCHO profiles (see Section 2.3). When comparing satellite measurements against models it is important to recognise the impact of this a priori on the total column values. This is complicated by how OMI is differently sensitive to HCHO (and other trace gases) 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 ( $AMF_{GC}$ ) uses the original satellite scattering weights while recalculating the shape factor, whereas the second ( $AMF_{PP}$ ) also recalculates scattering weights.  $AMF_{PP}$  is created using code initially written by Professor P. Palmer (see sections ?? 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: the original ( $\Omega_{OC}$ ), one using GEOS-Chem shape factors ( $\Omega_{GC}$ ), and one from Palmer's code ( $\Omega_{PC}$ ).

### 2.6.1 Outline

An outline in computational order of what takes place when recalculating the  $\Omega$  from OMI follows:

1. GEOS-Chem satellite overpass output (see Section 2.4.7.1) 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), air density (molec  $\text{cm}^{-3}$ ), and box heights ( $H(z)$ )
  - (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 ( $\text{AMF}_{GC}$ ) 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:

$$\text{AMF}_z = \Sigma_z (\omega(z) \times S_z(z) \times H(z)) \quad (2.18)$$

$$\text{AMF}_s = \Sigma_\sigma (\omega(\sigma) \times S_\sigma(\sigma) \times d\sigma) \quad (2.19)$$

4. These two AMFs represent the same thing using different coordinates, and either one can be used as the  $\text{AMF}_{GC}$
5. The  $\text{AMF}_{PP}$  (created separately) and  $\text{AMF}_{GC}$  are used to determine the new vertical columns ( $\Omega_P$ ,  $\Omega_G$  respectively):  $\Omega = SC / \text{AMF}$
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  $140^\circ\text{W}$  to  $160^\circ\text{W}$ ) are averaged monthly and longitudinally to provide a modelled reference sector  $\Omega_0[\text{lat}]$
  - (b) The modelled reference slant columns (MRSC) are calculated using  $MRSC = \Omega_0 \times \text{AMF}$  for each AMF
  - (c) For each satellite pixel between  $140^\circ\text{W}$  and  $160^\circ\text{W}$ , the correction is calculated as the measured SC minus the MRSC at the nearest latitude:

$$\text{corr}[\text{lat}, \text{track}] = SC[\text{lat}, \text{track}] - MRSC[\text{lat}]$$

- (d) These corrections are binned by satellite detector (track: 1-60), and latitude ( $0.36^\circ$ ; 500 latitudes from  $90^\circ\text{S}$  to  $90^\circ\text{N}$ )

- (e) The median entry of each bin is determined and this forms the RSC[lat,track] (e.g., Figure 2.22)
- 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  $0.25^\circ \times 0.3125^\circ$  grid along with how many pixels have been binned, and the average pixel uncertainty in product OMHCHORP

Figure 2.19 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 priors. 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 2.15). The shape factor is calculated using GEOS-Chem midday output (see Section 2.4.7.1) which provide simulated HCHO concentration profiles ( $\eta(z)$ ) and total columns ( $\Omega$ ) at  $2^\circ \times 2.5^\circ$  horizontal resolution. Using Equation 2.13 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  $S_\sigma$  is calculated (through Equation 2.16) to confirm that these shape factors are equivalent. Comparing the resulting AMFs created by Equations 2.17 and 2.15 for each pixel provides confidence in the unit conversions (and other

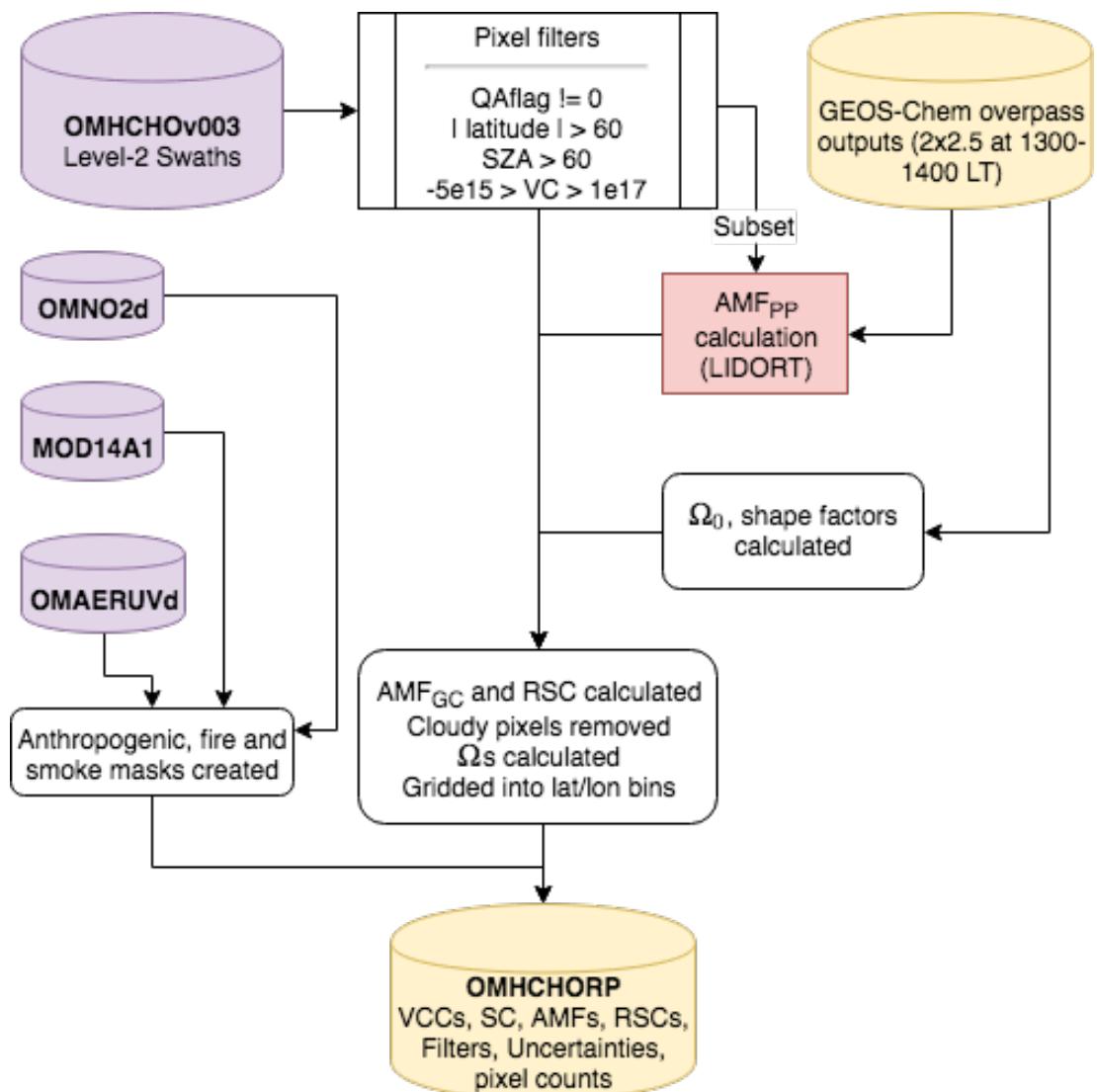


FIGURE 2.19: Flow diagram showing how OMHCHO level two swath data is read, processed, and gridded in this thesis

factors) applied. For example see figure TODO: plot showing correlation between AMFz and AMFs - or state that these are exactly one to one if that is the case..

### 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 ??). 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 2.19 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, e.g.). This filtering removes highly uncertain pixels, along with those for which instrument problems such as the OMI row anomaly (see Section ??) 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 2.20 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.

Additional information is added to each pixel, including the new AMFs (see the following subsections). Each pixel and its relevant data are saved in a long list for subsequent gridding. 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 2.15 we have:

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

then using the  $\omega(z)$  from satellite data, 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 performed using a simple rectangular method, which multiplies the integrand midpoints by the change in height, and then takes the sum for each vertical box. This assumes that the provided scattering weights and shape factors are linear between the 47 resolved values.

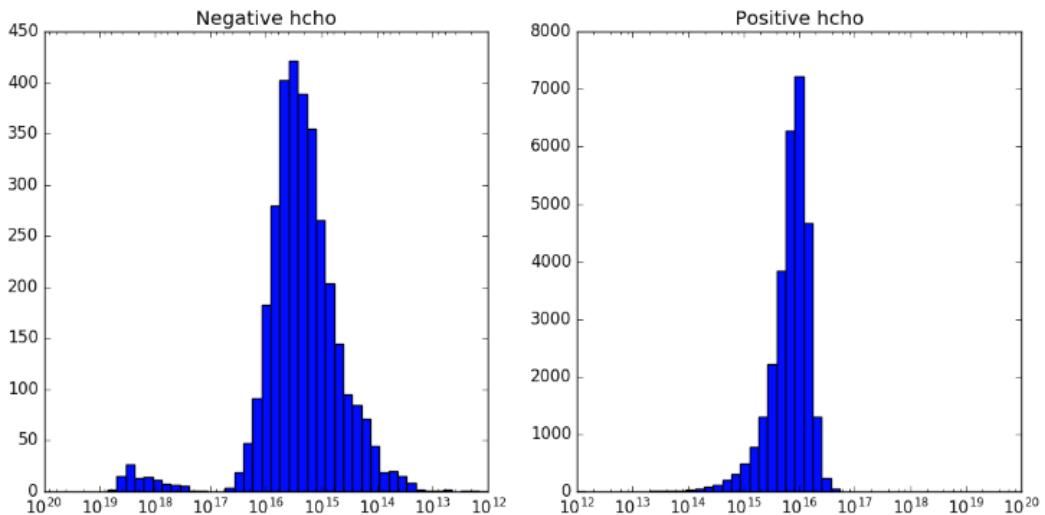


FIGURE 2.20: Column density histograms for a subset of OMI swaths over Australia on the 18th of March 2013. Negative entries are shown in the left panel, positive in the right, note the different scale between negative and positive panels.

### 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, and Randal Martin, subsequently 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. Running this code is computationally expensive, and is subset to within the region (50-10° S, 160° W-160° E). The code uses a combination of GEOS-Chem a priori profile information and satellite measurement data to calculate the AMF by using LIDORT radiative transfer calculations to determine scattering. The instrument sensitivity (or scattering weights;  $\omega$ ) and shape factors for each pixel are calculated within the PP code, which outputs  $\text{AMF}_{PP}$ .

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 layed 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 are required, and provided by modifying the GEOS-Chem diagnostic output.

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 (<https://github.com/LukeSurl/amf581g>). In order to run, output is required from GEOS-Chem: the HCHO concentration profile averaged between 1300 and 1400 LT, including optical depths

at 550 nm, 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), and read by the PP code in conjunction with the GEOS-Chem outputs for each day. The PP code then produces a list of recalculated AMFs which is read by my python code and associated with the corresponding pixel (outlined in 2.19).

### 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 2.7:  $\Omega = \text{SC} / \text{AMF}$ .

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 its divergence from a modelled reference sector. HCHO products from OMI and SCIAMACHY both use a median daily remote Pacific ocean radiance reference spectrum, over 15°S-15°N, 140°-160°W where it is assumed that the only significant source of HCHO is methane oxidation (De Smedt et al. 2008; Barkley et al. 2013; Kurosu and Chance 2014).

The 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).

Here a new reference sector correction (RSC) is created using modelled and measured HCHO columns over the remote Pacific, to produce corrected vertical columns (VCC). This follows Abad et al. (2016), and defines the remote Pacific as the band between 140°W to 160°W. It uses the difference between slant columns ( $\Omega_S$ ) and reference slant columns ( $\Omega_{S_0}$ ) divided by the AMF, plus the modelled reference sector column ( $\Omega_{V_B}$ ):

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

This method is used in several works, including De Smedt et al. (e.g., 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).

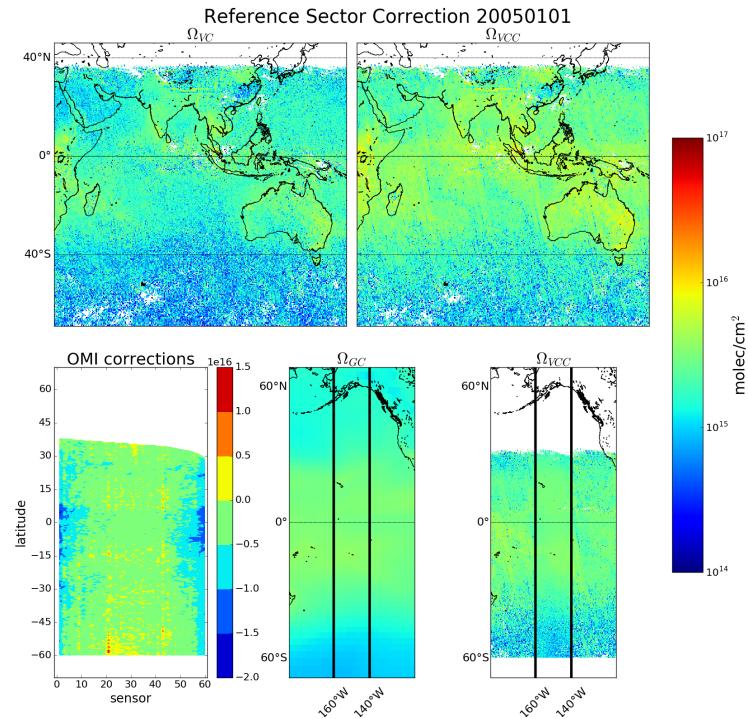


FIGURE 2.21: Example of remote Pacific RSC using 8-day average measurements and one month modelled data.  $\Omega_{VC}$  shows the uncorrected vertical columns, while  $\Omega_{VCC}$  shows the corrected vertical columns. OMI corrections shows the correction applied globally based on latitude and OMI 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.

To get the modelled slant columns, each of the AMFs (calculated in prior sections) is applied to the modelled vertical columns using Equation ???. The longitudinal average is taken within the remote Pacific, as corrections are assumed to be longitudinally invariant. The modelled reference sector is averaged over the month and interpolated latitudinally to 500 equidistant bins. Figure 2.21 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, satellite measurements (pixels) over the remote Pacific 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 columns AMF.

For OMI swaths, each row of measured data contains 60 “Across track”(track) measurements. The track index ( $i$ ) relates the measurement to one of the 60 columns of data. Corrections (molecules cm<sup>-2</sup>) for each measurement are calculated by taking the difference between the measured slant column and the a priori slant column as follows:

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

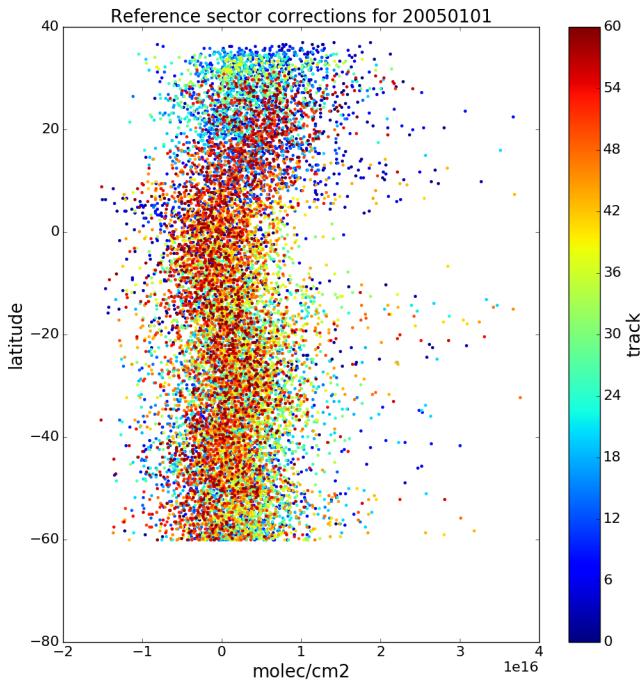


FIGURE 2.22: Example of track correction interpolations for January 1st 2005, points represent the difference between satellite slant column measurements and modelled slant columns over the remote Pacific.

where  $j$  represents a latitude index and  $VC_{GEOS-Chem}(lat)$  represents the a priori reference sector vertical column HCHO at the latitude corresponding to  $j$ . 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 for each of the three AMFs.

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 2.22 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 2.21, 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:

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

For each day, good satellite measurements and associated SC, VC, VCC, and AMFs are listed.

### 2.6.8 Binning the results daily

Finally the pixels are binned into a gridded dataset named OMHCHORP, as shown in Figure 2.19. The resolution is chosen to match the native resolution of GEOS-Chem ( $0.25^\circ \times 0.3125^\circ$ ) and the GEOS met data fields. A bin entry count is used to allow easy rebinning, and can be used to check for sparse data days due to filtering or poor weather. 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 ( $\text{AMF}_{PP}$ )
9. GEOS-Chem RSC VC based on  $\text{AMF}_{PP}$
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)

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,  $\sim 162$  GB ( $142 + 16 + 4$  OMHCHO, MOD14A1, and OMNO2d respectively) of daily data was downloaded and then transformed into  $\sim 8$  GB (per year) 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 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  $0.25^\circ \times 0.3125^\circ$  horizontal resolution. These calculations are compared over Australia in figure(s) 2.23, TODO: regression, and time series. Figure 2.23 shows how the recalculated columns compare to the original (OMI, left). Recalculation increases the January column amounts while also slightly flattening the distributions (row 3). The effect of not recalculating the  $\omega_z$  can also be seen in Figure 2.24 which shows the altered satellite vertical columns using each method.

Figure 2.24 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  $0.25^\circ \times 0.3125^\circ$  resolution) has been created by binning the recalculated satellite pixels within the month. The average number of pixels per land square is inset as text, changing due to how the fire filter is applied. Each row has a stricter fire filter applied from top to bottom, with no fire filter on the first row up to filtering pixels from squares with fires up to 8 days prior. This figure looks at March 2005 with biomass burning filtered differently in each row. Active fires over the last 0, 1, 2, 4, and 8 days are filtered as the row number increases.

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

TODO: compare how many nans from palmers code to my code The AMF calculated using professor Palmer's produces TODO: MORE OR FEWER reasonable (within screening range of  $-5 \times 10^{15}$  to  $1 \times 10^{17}$ ) vertical column amounts. 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. Biomass burning can be a large local or transported (via smoke plumes) source of HCHO, CHOCHO, glyoxal, and other compounds which influence levels of both HCHO and isoprene. Anthropogenic emissions from power generation, transport, and agriculture can influence these levels as well. Where possible these influences need to be removed so that calculations of purely biogenic emissions are not biased. In GEOS-Chem we can simply turn off pyrogenic and anthropogenic emissions, however in satellite datasets we need to mask potentially affected pixels.

Influence from biomass burning can be removed through measurements of acetonitrile and CO (e.g., 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). Marais et al. (2012) remove pixels colocated with non zero

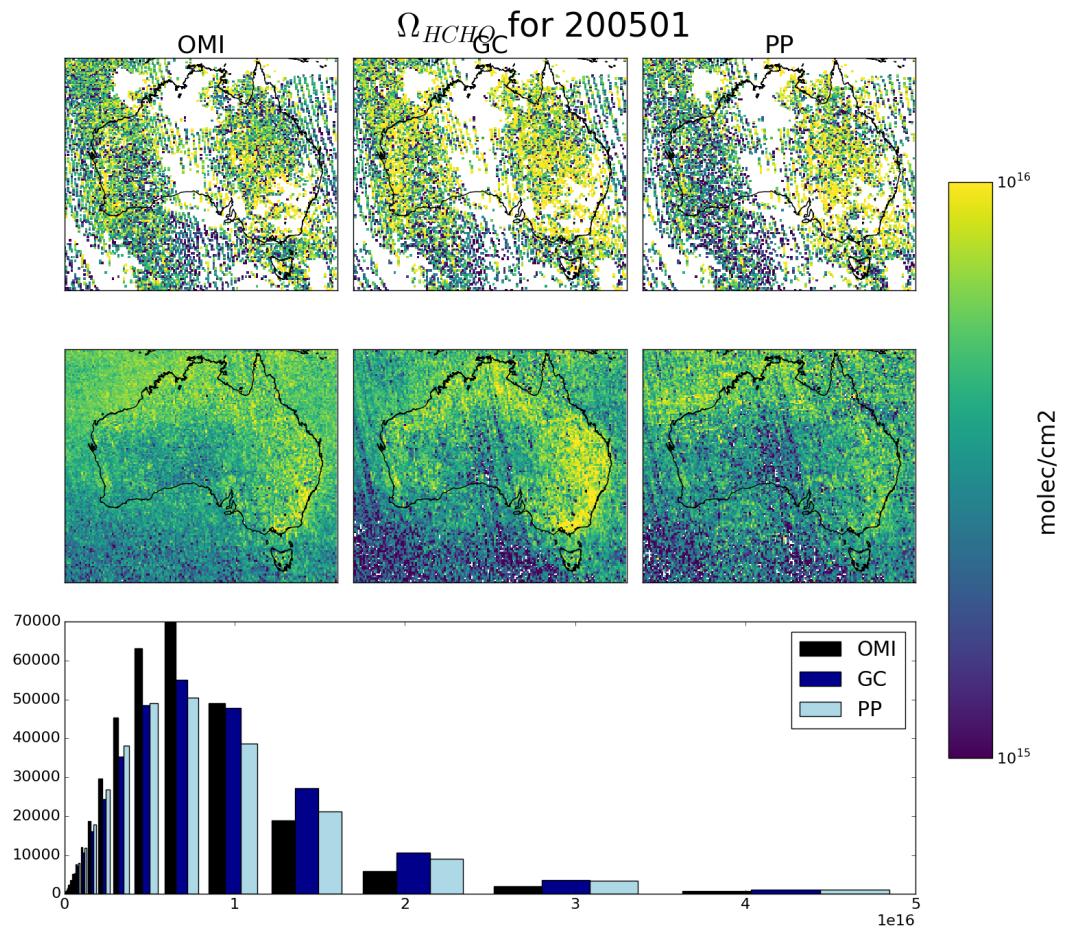


FIGURE 2.23: Row 1: regridded corrected  $\Omega_{HCHO}$  from OMHCHO on January 1, 2005: original (left), recalculated using new shape factors (middle), and additionally using updated scattering weights (right). Row 2: shows the monthly average for January 2005. Row 3: shows the distribution over the month for each of the three column amounts.

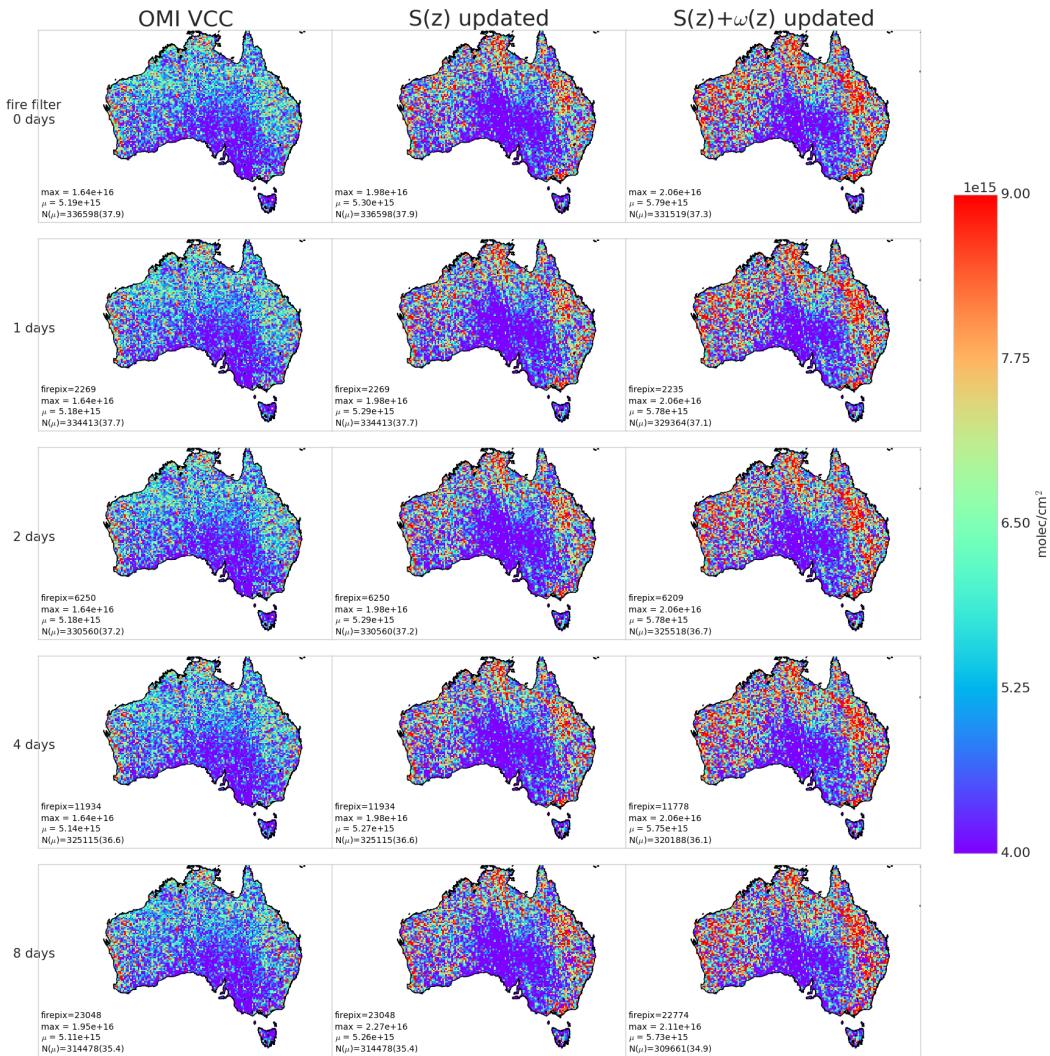


FIGURE 2.24: Column 1: Reference sector corrected HCHO vertical columns  $\Omega$  from OMHCHOv003. Column 2:  $\Omega$  with recalculated a priori shape factors using GEOS-Chem v10.01. Column 3:  $\Omega$  with recalculated a priori shape factors and scattering weights using GEOS-Chem v10.01 and LIDORT. Row 1-5: increasing number of prior days which have active fires included when masking fire influence.

TABLE 2.4: How many satellite pixels are filtered by pyrogenic and anthropogenic masking. Left to right the columns display year, how many land pixels are read over Australia, how many of these pixels are removed by the pyrogenic filter, how many are removed by the anthropogenic filter, and how many are removed in total. In parenthesis are the portion of pixels filtered.

Year	Pixels	Pyro	Anthro	Total
2005	3.9e+06	4.1e+05(10.7%)	5.0e+04( 1.3%)	4.4e+05(11.5%)
2006	3.8e+06	5.1e+05(13.5%)	9.2e+04( 2.4%)	5.6e+05(14.7%)
2007	3.7e+06	4.4e+05(11.9%)	7.5e+04( 2.0%)	4.9e+05(13.0%)
TODO				
TODO				

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  $0.25^\circ \times 0.3125^\circ$ . Wolfe et al. (2016) disregard HCHO measurements when acetonitrile > 210 pptv and CO > 300 ppbv, while acetonitrile > 200 pptv is used to determine fire influence in Miller et al. (2016). TODO: look at yearly correlation, compare to exponential curve and look for fire outliers As seen in TODO: citation, HCHO concentrations scale exponentially with temperature. This allows another method for detecting the influence of non-biogenic HCHO emission/creation by looking for outliers above the curve at low temperature. Zhu et al. (2013) has a similar analysis over south-eastern USA showing an exponential correlation of  $HCHO = \exp(0.15 \times T - 9.07)$ .

I use satellite data to account for anthropogenic and pyrogenic influences on the OMHCHO satellite HCHO columns. MODIS fire counts are used in conjunction with smoke AAOD enhancements (from OMI) to remove data points which may be affected by fires or fire smoke plumes. OMI NO<sub>2</sub> measurements are used to mask potential anthropogenic influence. These masks negatively affect uncertainty, as fewer measurements are available to be averaged. This section describes the creation and effects of filters used on satellite data.

A quick summary of how much data is filtered over Australian land squares is provided in Table 2.4, and an quick check of how many pixels are filtered in January 2006 can be seen in Figure 2.25. The anthropogenic filter completely removes grid squares over Sydney and Melbourne, and high removal rates over Brisbane. Other major cities in Australia either do not emit enough NO<sub>2</sub> or are too spread out and do not breach the threshold to be filtered as anthropogenic.

### 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 AAOD. 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 are filtered using smoke AAOD from product OMAERUVd, with the threshold determined through analysis of Australian AAOD distributions.

Anthro and Fire filters applied on 20060101-20060131

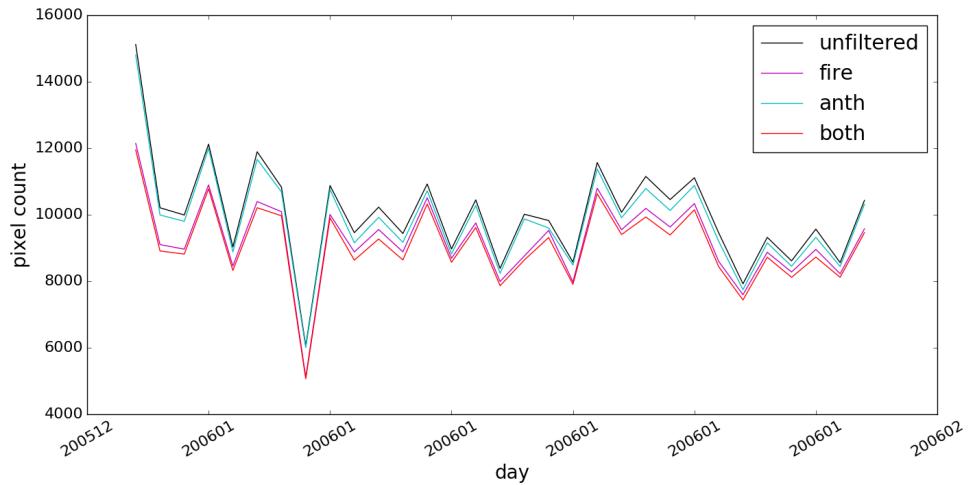
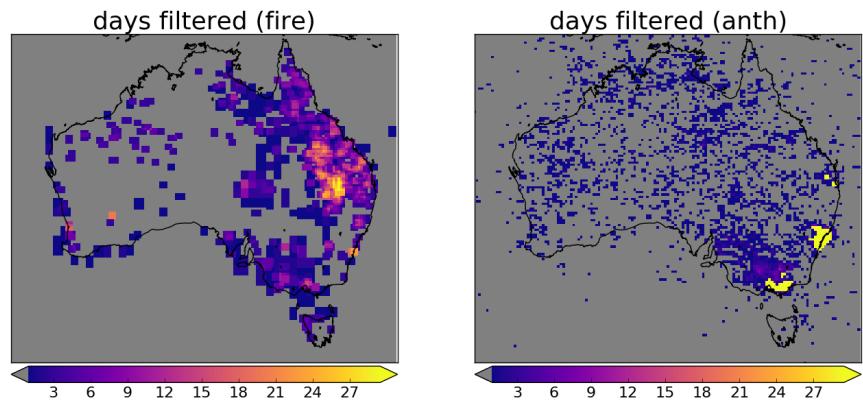


FIGURE 2.25: Top row shows grid squares filtered out by pyrogenic(left) and anthropogenic(right) influence masks during January 2006. Along the bottom is the time series of total pixels over Australian land squares with and without filtering the data.

OMHCHO total column HCHO  $\Omega$  is processed into a  $0.25^\circ \times 0.3125^\circ$  horizontal daily grid. Pyrogenic filters are interpolated to the same horizontal resolution as  $\Omega$  to simplify application. The following steps are performed in order to create the pyrogenic influence mask:

1. MOD14A1 daily gridded Aqua/Terra combined fire counts ( $1 \times 1 \text{ km}^2$ ) are read, and binned into  $0.25^\circ \times 0.3125^\circ$  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  $0.25^\circ \times 0.3125^\circ$  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 (see Figure 2.26).
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 less sensitive to cloud contamination than AOD, and I use AAOD from the daily gridded level 3 satellite product OMAERUVd (Ahn2008) described in Section 2.2.1.3 to provide a filter for smoke plumes. Although removing gridsquares with dust reduces how much data is available to analyse, 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  $0.25^\circ \times 0.3125^\circ$  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 2.26 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. Figure 2.27 shows how the smoke filter compares to the fire filter over 2005. The left panel shows how many days are filtered using the AAOD threshold, while the right panel shows how many are filtered using the combined smoke and active fire masks. The time series shows pixels available for analysis before being filtered (black), after smoke filtering (cyan) and after the full fire filter is applied (magenta). The Ayre basin is the dominant area filtered by smoke detection, likely due to dust influence in the arid region.

Figure ?? shows what portion of pixels are filtered out by the pyrogenic filter. The top panel shows the spatial distribution of fire masks, with most pixels removed along

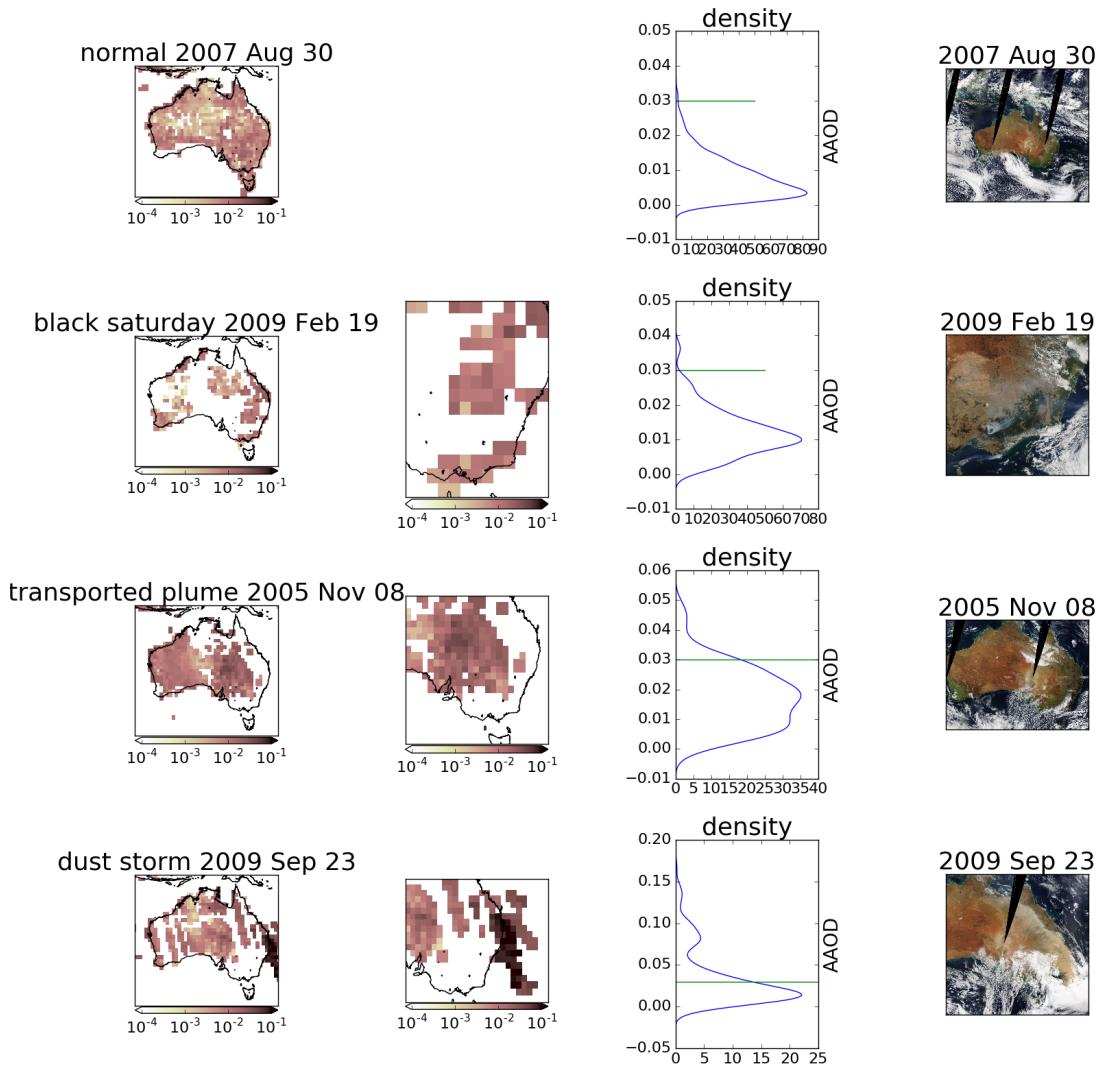


FIGURE 2.26: AAOD from OMAERUVd (columns 1, 2, 3) over Australia for four different scenarios (rows 1-4). Scenes from the same day are taken from the EOS Worldview website <https://worldview.earthdata.nasa.gov/>.

Fire and smoke filters applied on 20050101-20051231

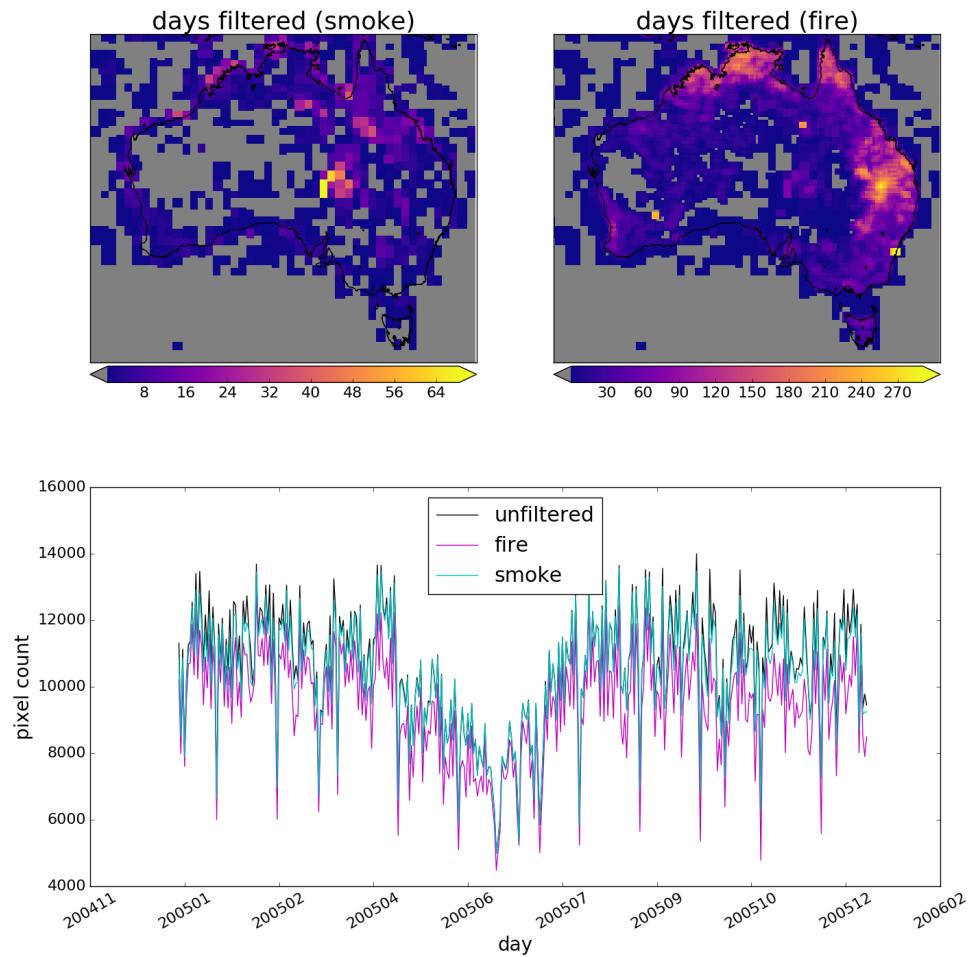


FIGURE 2.27: Smoke (top left) filtered days using AAOD > 0.03 compared to the combined (smoke and active fires) fire filter (top left). Time series shows how many pixels are available for analysis before and after filtering.

the northern and eastern coastlines. A large portion of the filtered areas appear to correspond with forested areas (see Figure ??), which suggests that forest fires are being masked properly. Central Australia is largely unmasked, which could be due to a lack of sufficient vegetation to create a large enough fire to be seen by satellite. The other potential cause of pyrogenic filtering is the proliferation of petrol or gas wells (see Figure 2.29 and Figure 2.30). The filtering shown here is for 2005, when 388 gas wells existed in Queensland, however more than 2000 wells (cumulative) were approved by 2013, so this may cause more filtering over the course of this thesis' timeline (Carlisle 2012). To check this the filtering portion for 2012 is also plotted in Figure TODO: plot map of filtered squares for 2012 when it has been run by NCI. One clear hotspot is located over port Kembla (south of Sydney), most likely due to the flame which burns over the blast furnace stack throughout the year. Another hotspot can be seen in Western Australia over Kalgoorlie, where a large open cut gold mine "super pit" is always open and blasting daily. In Western Queensland over Mount Isa there is again a mining related hotspot. A large area in southern Queensland/northern NSW is also heavily filtered, potentially due to gas flaring in the Surat Basin, which has thousands of petrol and gas wells.

### 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 2.34 - 2.36 show the relationship between modelled temperature, and satellite 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  $m$  terms is shown in the embedded plot, with one datapoint available for each grid square where the regression is performed.

Figures 2.32 and 2.33 show the regressions between OMI HCHO total columns and temperature from GEOS-Chem output and CPC daily maximum temperatures respectively. Comparing against GEOS-Chem modelled surface temperatures first requires deresolution from  $0.25^\circ \times 0.3125^\circ$  to  $2^\circ \times 2.5^\circ$  latitude by longitude resolution. The left column in Figure 2.32 shows scatter and RMA correlation within a single gridbox over 2 months from Jan 1 to Feb 28, 2005, without having applied either the fire nor anthropogenic masks to OMI HCHO columns. The right column shows the same correlation after applying the fire filter, affected datapoints are marked in teal (matching the red marked points in the left column). The analysis is repeated for Sydney, Canberra, and three gridsquares to the north-west, west, and southwest of Sydney (w1, w2, w3 respectively). Figure 2.33 shows the same analysis at higher resolution using CPC daily maximum temperatures (see 2.2).

One cause of high HCHO at lower temperatures is direct or transported emissions

## Pyrogenic filter: 20050101-20060101

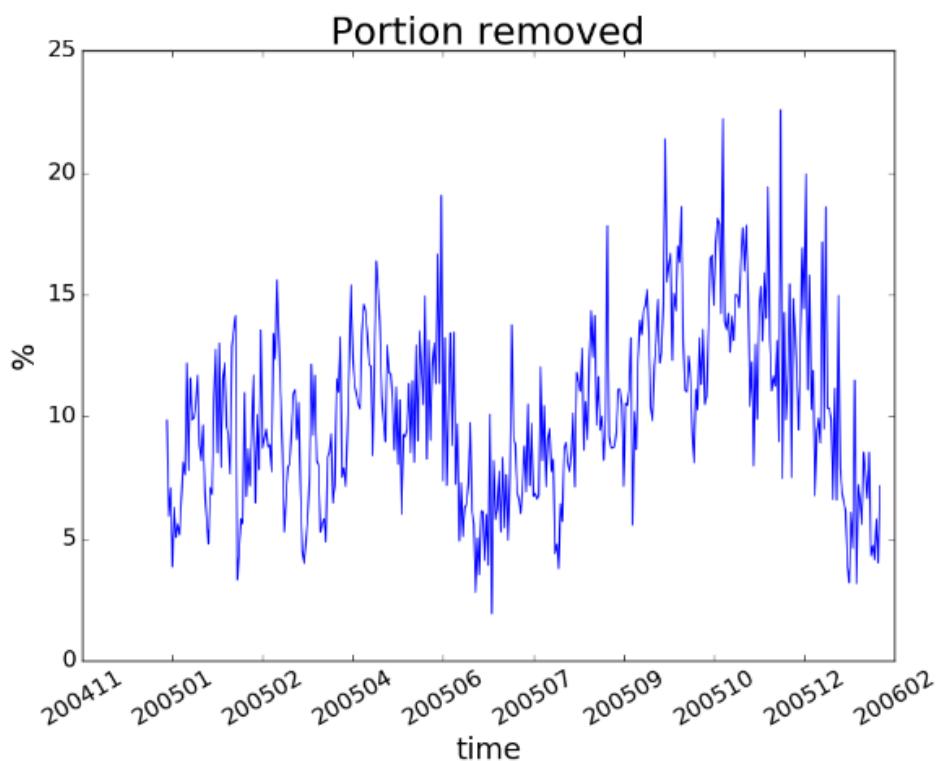
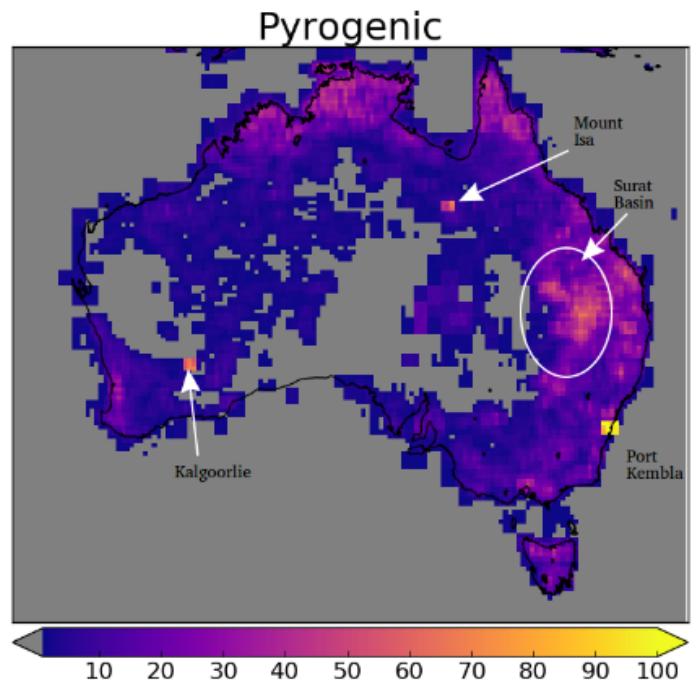


FIGURE 2.28: Top: Portion of 2005 filtered out by fire and smoke masks.  
Bottom: portion filtered out each day from land squares in Australia.

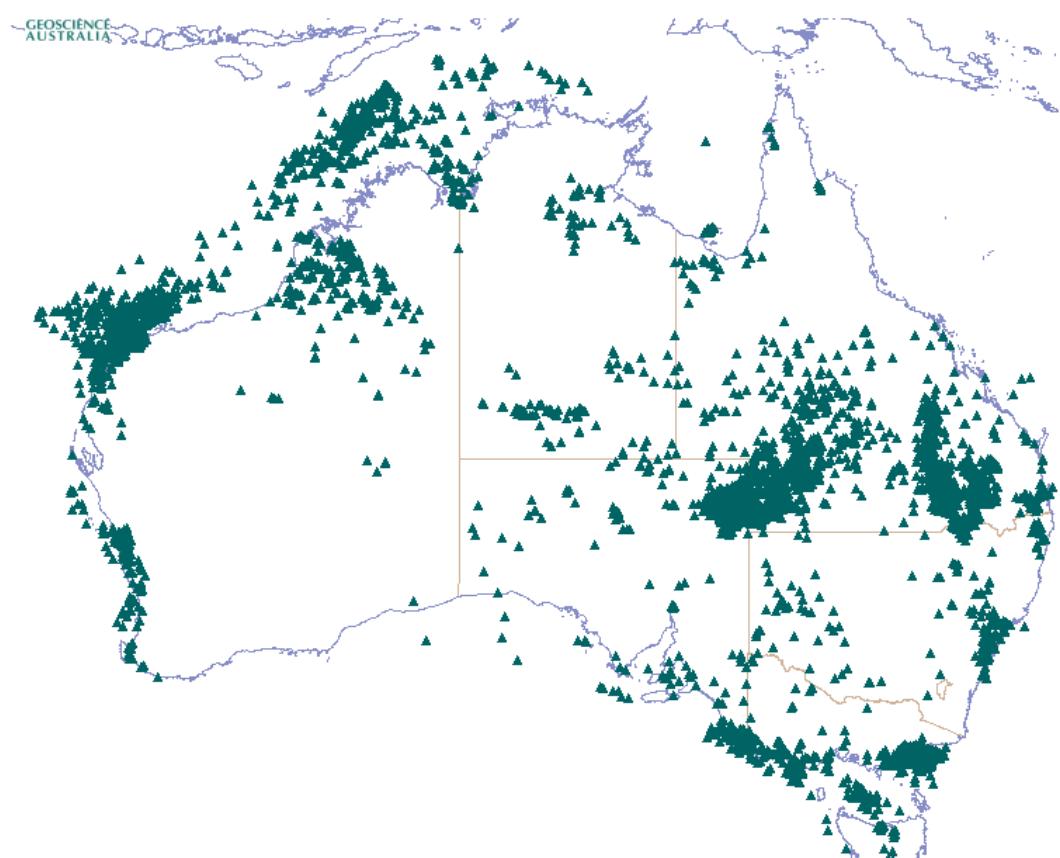


FIGURE 2.29: Petrol Well locations over Australia (current-2018) (<http://dbforms.ga.gov.au/www/npm.well.search>)

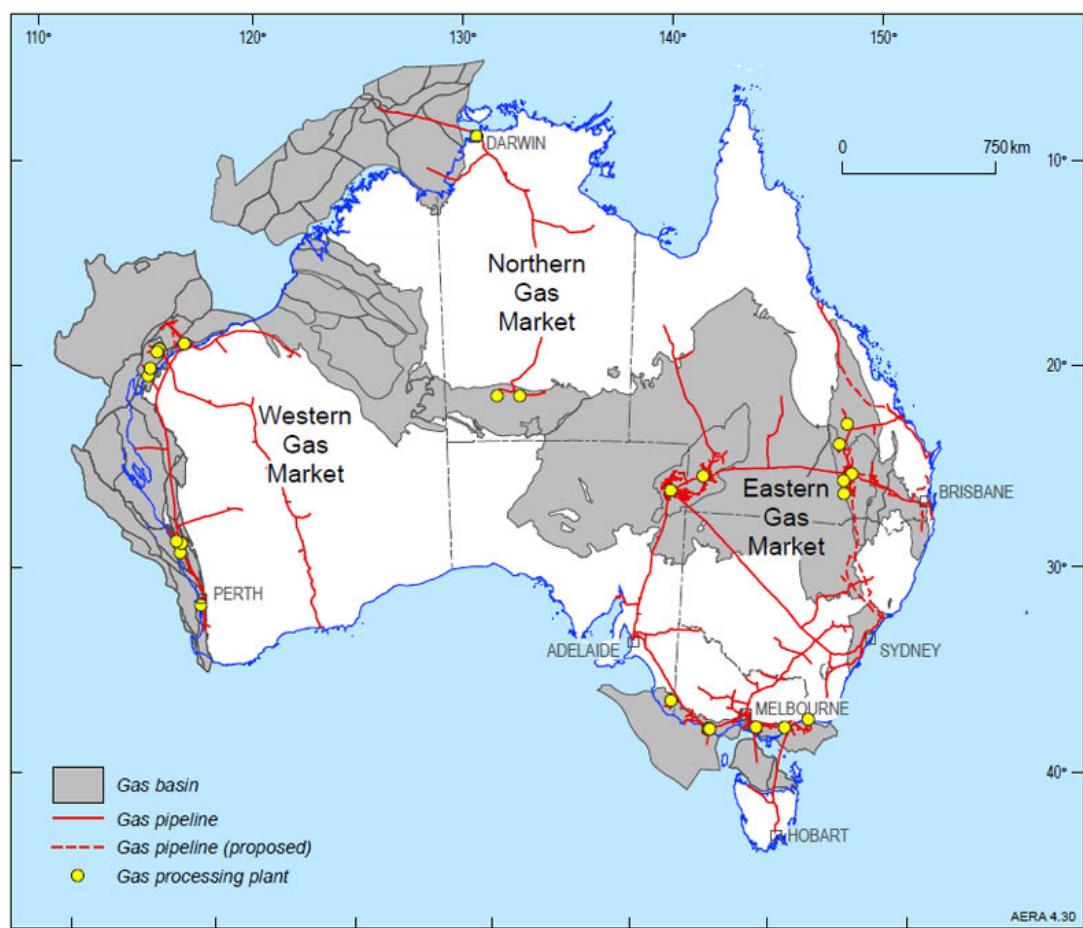


FIGURE 2.30: Gas fields and pipelines (2018) for Australia  
(<http://www.ga.gov.au/scientific-topics/energy/resources/petroleum-resources/gas>)

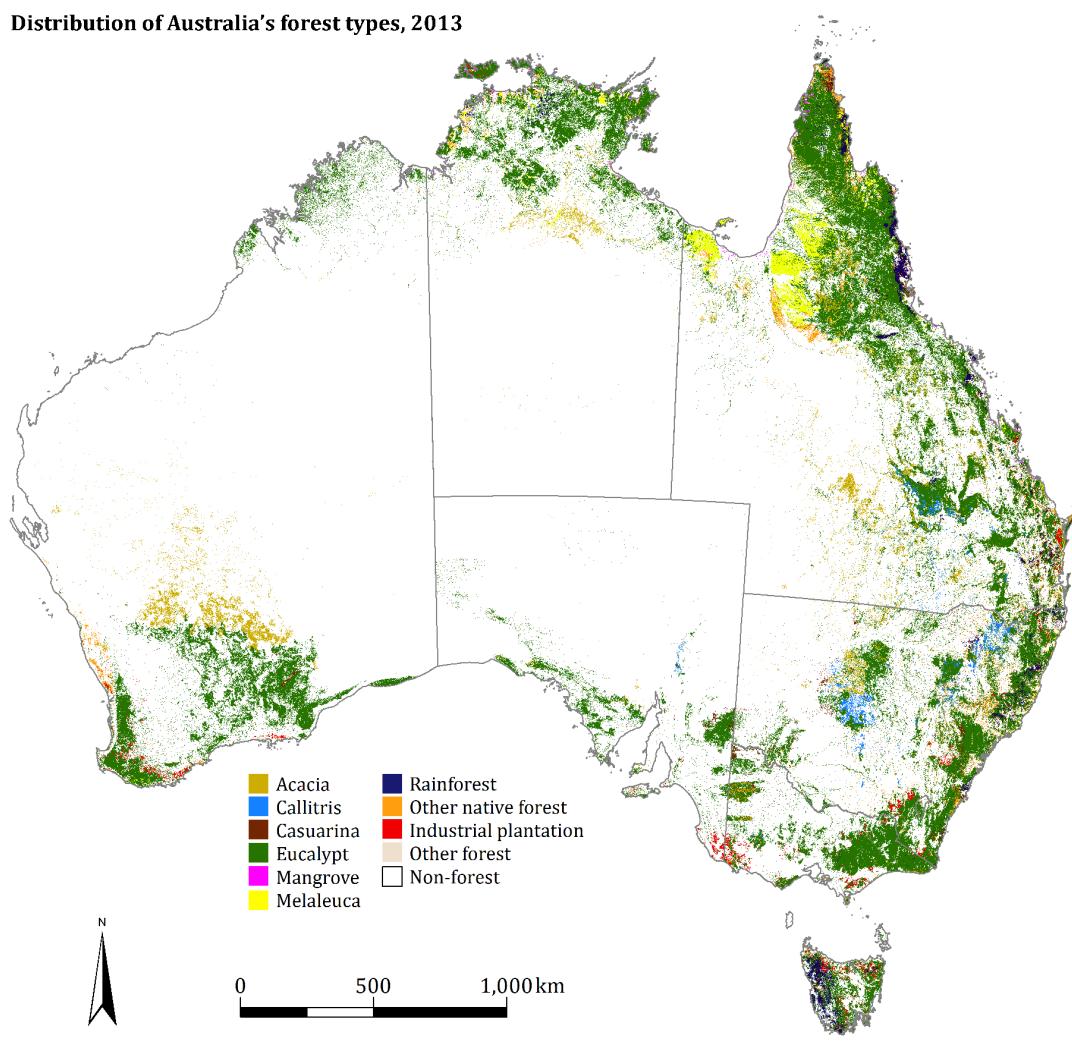


FIGURE 2.31: Forest coverage, coloured by predominant tree species.

??

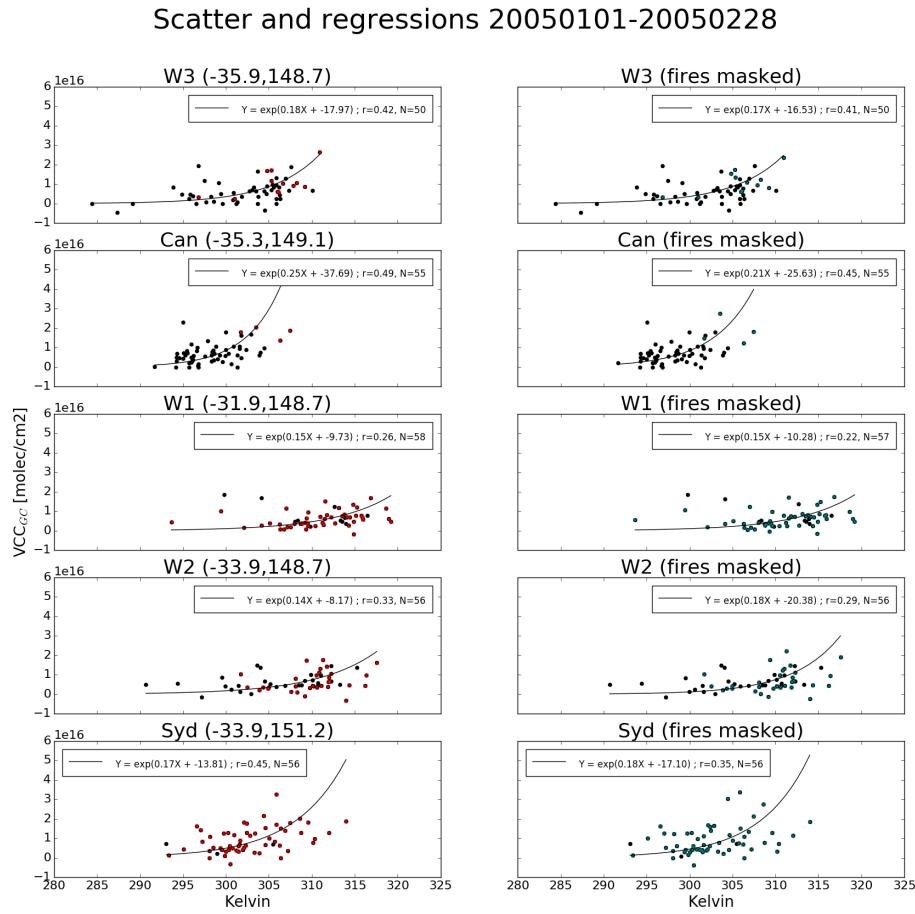


FIGURE 2.32: HCHO vs GEOS-Chem daily midday temperatures  
TODO add longer caption

## Scatter and regressions (CPC temperature) 20050101-20050228

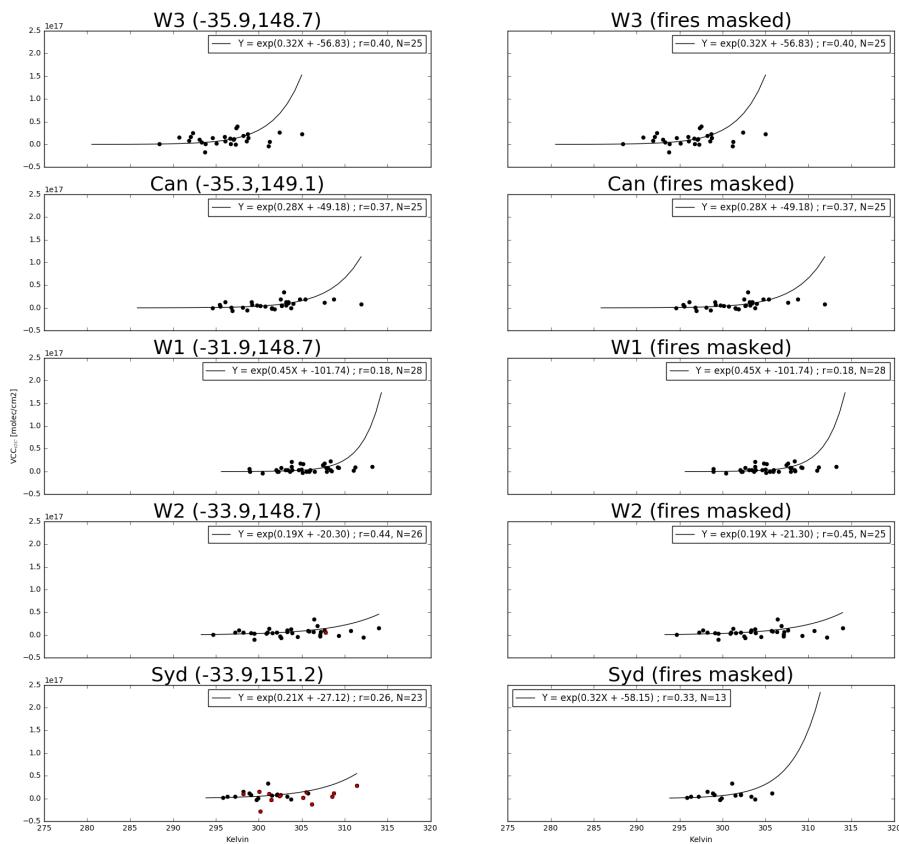


FIGURE 2.33: HCHO vs CPC daily maximum temperatures TODO add longer caption

and subsequent products from biomass burning. One potential problem with showing this is that days with fire enhanced HCHO are also likely to be hot. Another problem with correlating heat and HCHO is that increased temperature accelerates HCHO destruction (Zheng et al. 2015). 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 ( $2^\circ \times 2.5^\circ$ ) 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 tropospheric  $\text{NO}_2$  seen by Aura is shown in Figure 2.37. The average tropospheric  $\text{NO}_2$  column averaged within all of Australia and then each region shown in this figure is listed in table TODO 2.5.

Anthropogenic influences on the  $\text{NO}_2$  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  $\text{NO}_2$  greater than  $10^{15}$  molec  $\text{cm}^{-2}$  are flagged as anthropogenic.
2. After taking the yearly average over Australia, any gridsquares greater than  $1.5 \times 10^{15}$  molec  $\text{cm}^{-2}$  are flagged for the whole year.

This removes both the gridsquares close enough to cities to be affected by their emissions year round, as well as effects from transported pollution plumes. The affects of applying this filter to the OMNO2d product itself can be seen in Figure 2.38

The same regions as in Figure 2.37 are shown again in Figure 2.39, with  $\text{NO}_2$  pixels densities for each region shown, along with the threshold of  $1 \times 10^{15}$  molec  $\text{cm}^{-2}$ . This led to a reduction of TODO gridsquares from the total available measurement space over Australia. The removal of gridsquares which went above the yearly averaged limit of  $1.5 \times 10^{15}$  molec  $\text{cm}^{-2}$  further reduced the available data by TODO gridsquares.

## 2.8 Data Access

TODO: ADD MORE HERE

**OMHCHO** Satellite swaths of HCHO slant columns downloaded from <https://search.earthdata.nasa.gov>, with DOI 10.5067/Aura/OMI/DATA2015

Temperature 20050101-20050228

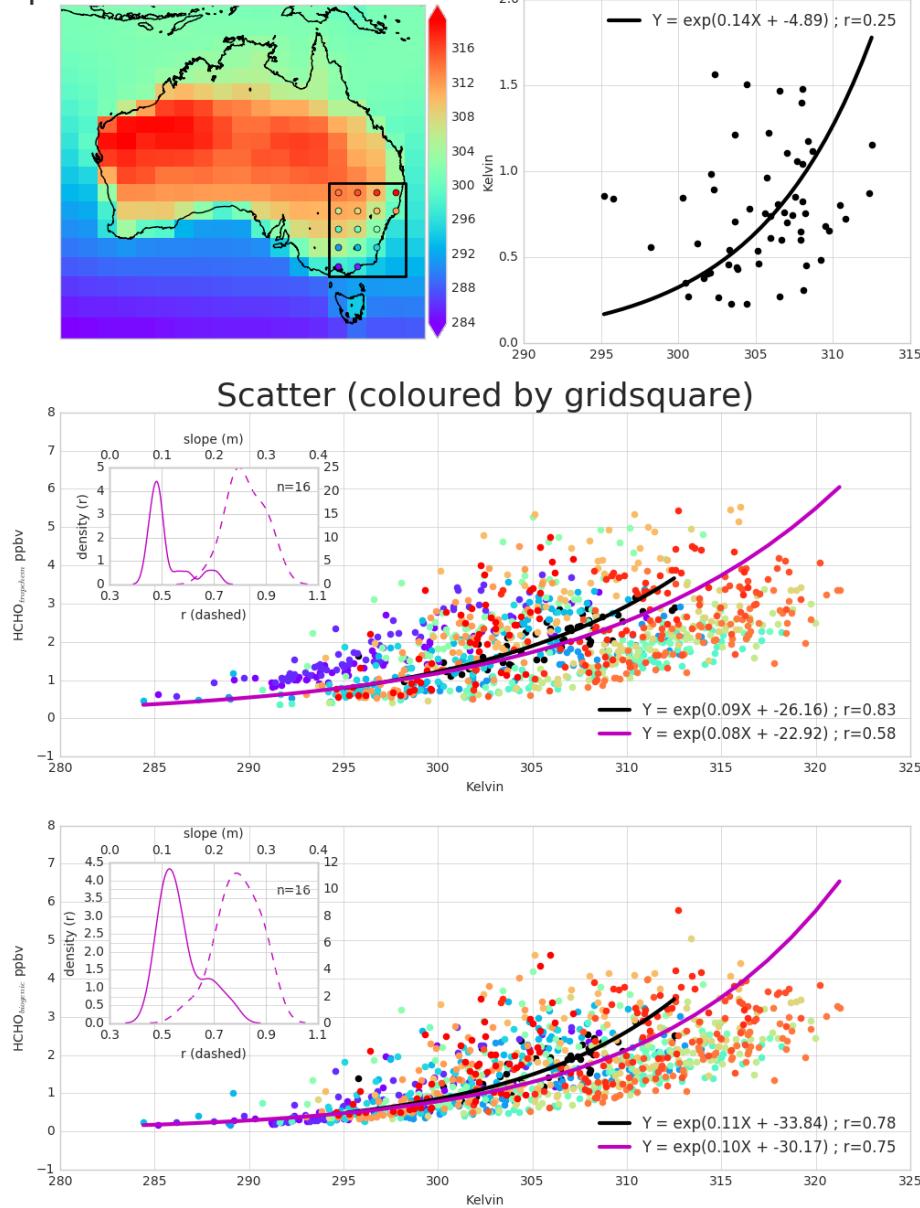


FIGURE 2.34: Top panel: surface temperature averaged over January and February 2005. Bottom panel: surface temperature correlated against temperature over, with different colours for each gridbox, and the combined correlation. A reduced major axis regression is used within each gridbox (shown in top 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.

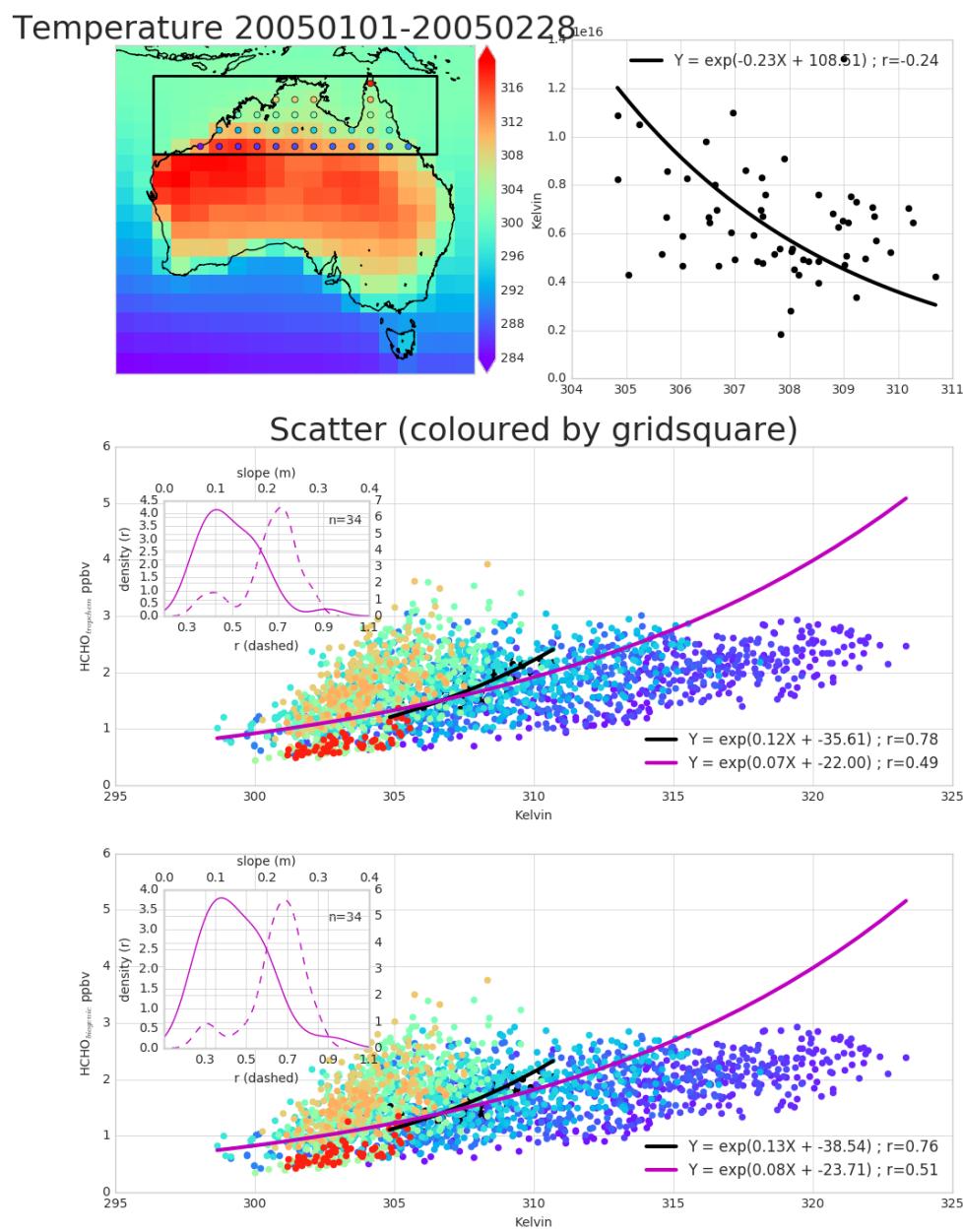


FIGURE 2.35: As Figure 2.34 but for northern Australia.

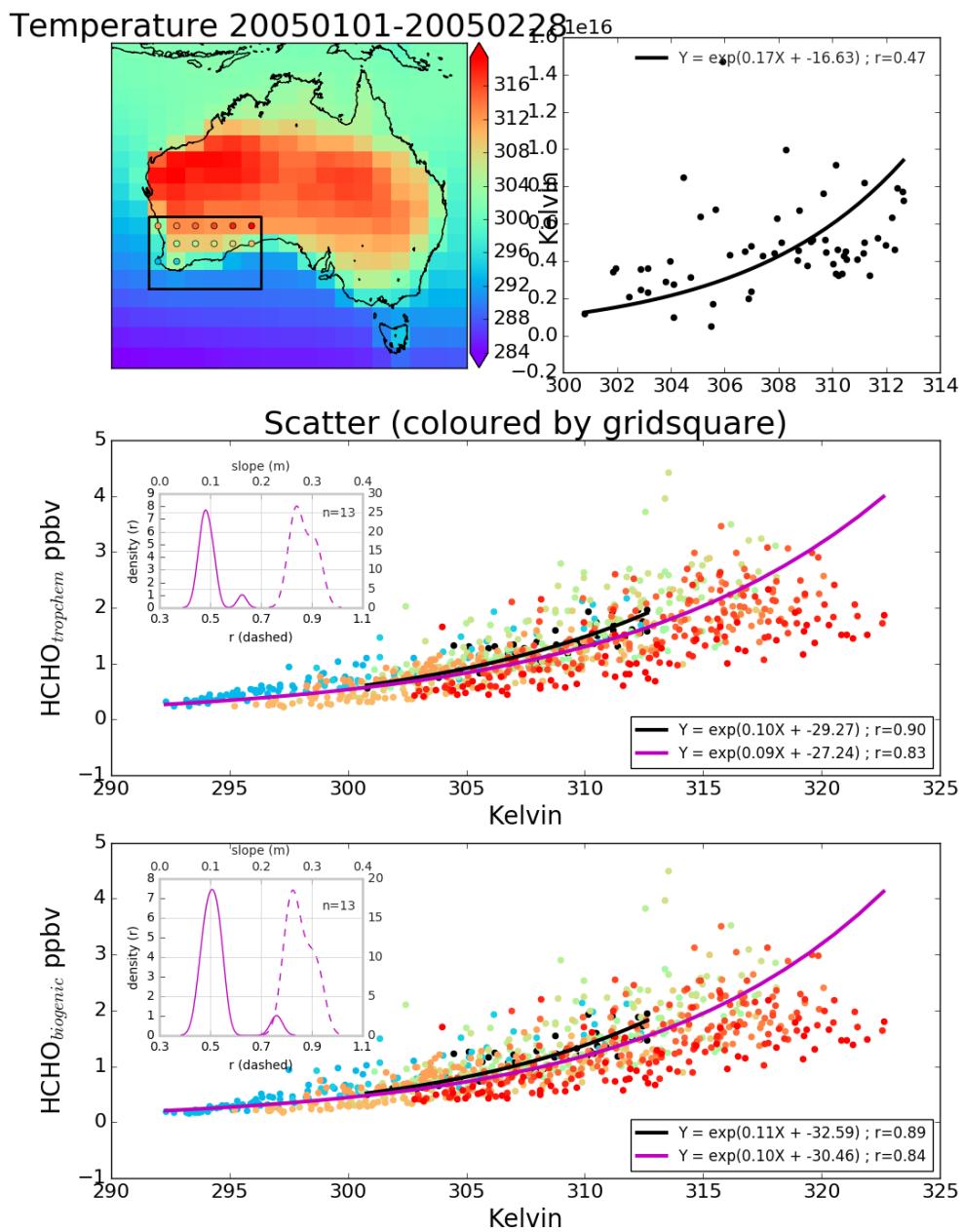


FIGURE 2.36: As Figure 2.34 but for south-western Australia.

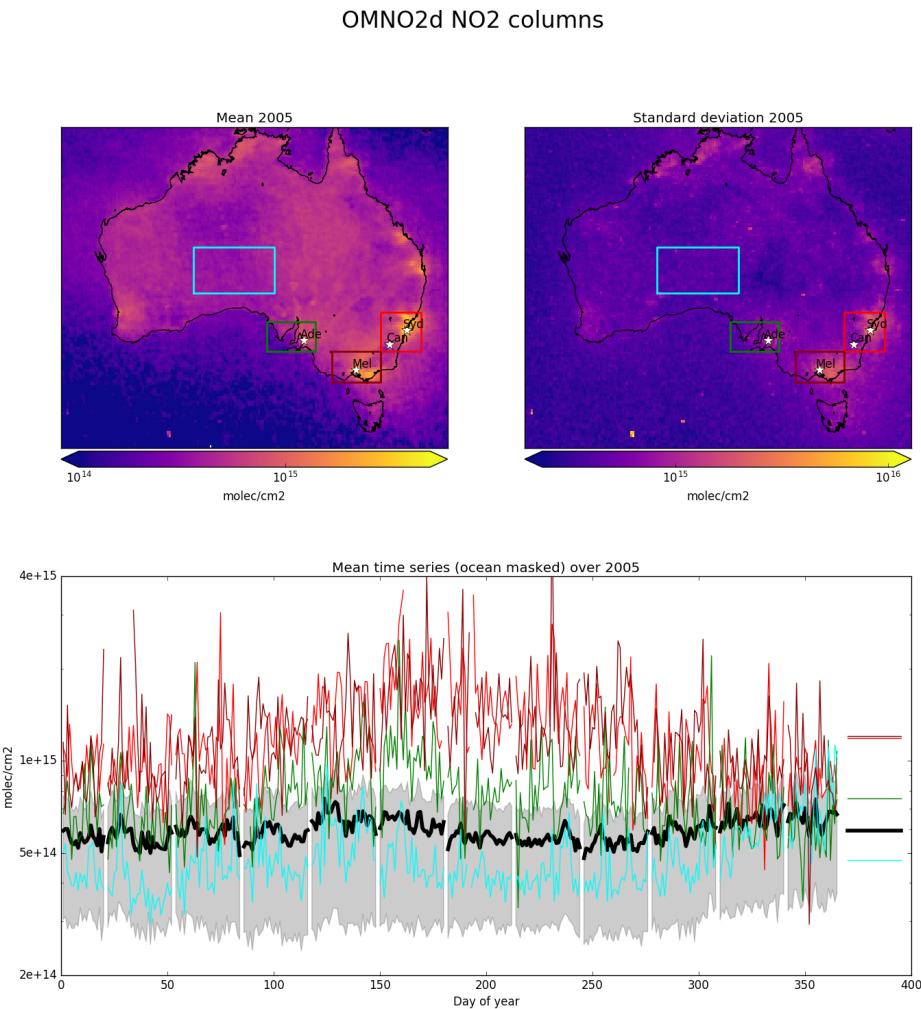


FIGURE 2.37: Mean (top left) and standard deviation (top right) of OMNO2d daily  $0.25 \times 0.25^\circ$  tropospheric cloud filtered NO<sub>2</sub> columns. Time series for Australia, and each region (by colour) shown in the bottom panel, with mean for that region shown on the right. A grey shaded 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.

TABLE 2.5: 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

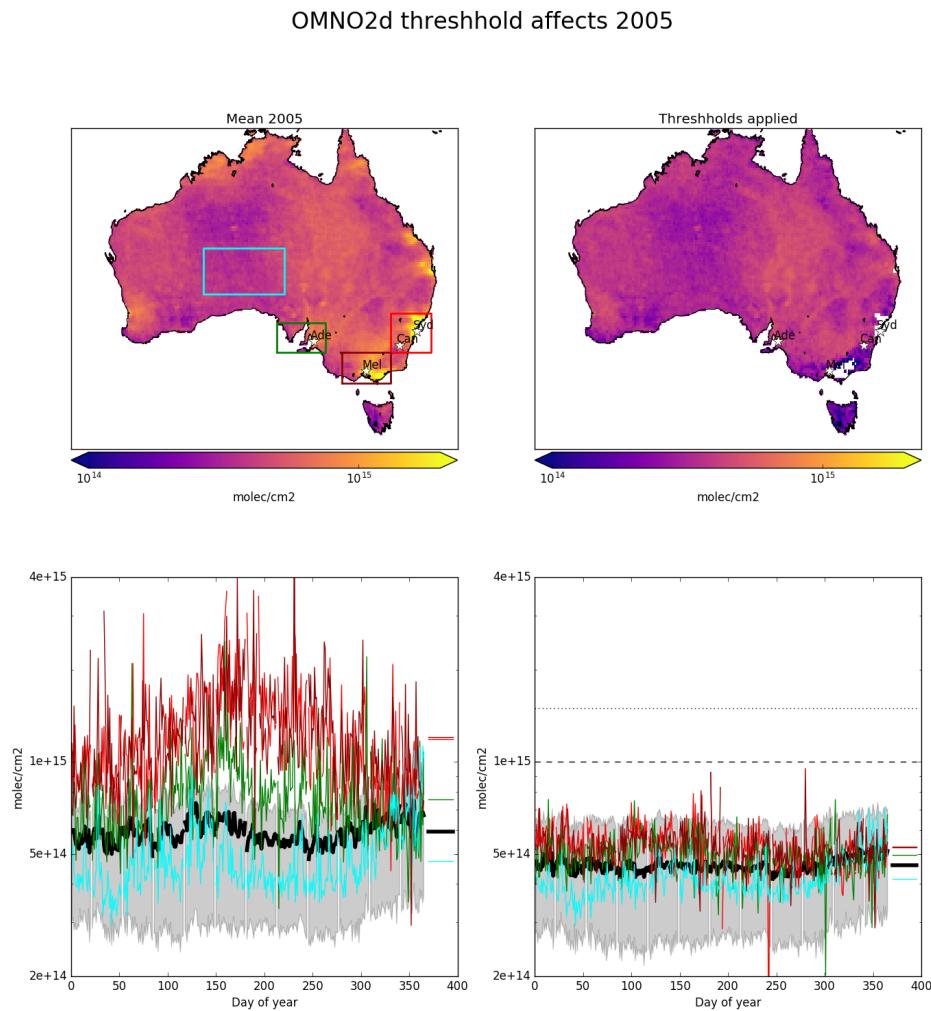


FIGURE 2.38: 2005 OMNO2d NO $_2$  column mean before (left) and after (right) applying the threshold filters as described in the text. Time series for Australia, and each region (by colour) shown in the bottom panel, with mean for that region shown on the right.

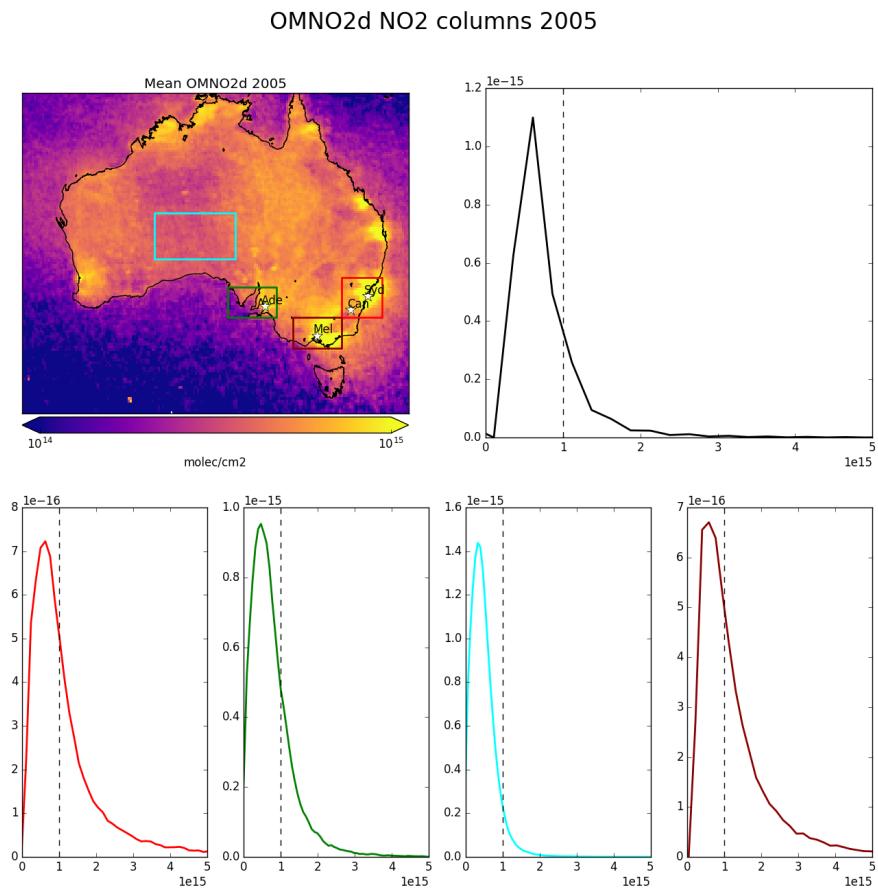


FIGURE 2.39: 2005 OMNO2d NO<sub>2</sub> column means (top left), along with column amount distributions for Australia (top right) and each region shown in the area map (by colour)

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

**OMAERUVd** Gridded satellite based AAOD measurements downloaded from <https://search.earthdata.nasa.gov>, DOI 10.5067/Aura/OMI/DATA3003.

**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

## Chapter 3

# Biogenic Isoprene emissions in Australia

### 3.1 Introduction

Biogenic volatile organic compounds (BVOC) affect the oxidative capacity of the atmosphere and are largely driven by and sensitive to what type of vegetation is in the area (Kefauver, Filella, and Peñuelas 2014). In the troposphere, BVOC emissions affect the hydroxyl radical (OH) cycling, ozone ( $O_3$ ) and secondary organic aerosol (SOA) production, and methane longevity. Australian forests are strong emitters of isoprene, the primary BVOC emitted globally (Guenther et al. 2006; Messina et al. 2016). However; emissions are poorly understood due to poor measurement coverage. The lack of measurements makes it difficult to estimate the subsequent atmospheric processes. Isoprene has a large impact on the oxidative properties of the atmosphere, as it reacts quickly with the OH radical.

Emission models used to derive these an estimate of isoprene fluxes are based understanding of different plant species (phenotypes) in varying conditions, however these are seldom well understood within Australian forests. The chemistry involved is complex and still suffers from relatively large uncertainties in both measurement and chemistry mechanisms. Isoprene emissions may be overestimated in Australia since they are based on measurements taken from a few heavily emitting young eucalyptus trees which are not representative (Winters et al. 2009; Fortems-Cheiney et al. 2012). Emissions estimates can improve boundary conditions for atmospheric chemistry models without requiring expensive measurement campaigns over the large data sparse continent of Australia. Guenther et al. (2012) Estimated global biogenic isoprene emissions at roughly  $535 \text{ Tg yr}^{-1}$ , while Sindelarova et al. (2014) estimated around  $411 \text{ Tg yr}^{-1}$ .

In this chapter the we describe and use a technique using satellite measurements of HCHO to calculate surface isoprene emissions. HCHO is the dominant product of most BVOCs and is measured by satellites via remote sensing techniques. In situ isoprene concentration measurements are costly and sparse within Australia, while satellite HCHO data are plentiful and freely available, making this technique very attractive. Such techniques have informed isoprene emission inventories in North America (Abbot 2003; Palmer 2003; Palmer et al. 2006; Millet et al. 2006; Millet et al. 2008), South America (Barkley et al. 2013), Europe (Dufour et al. 2008; Curci et al. 2010), Africa (Marais et al. 2012), Asia (Fu et al. 2007; Stavrakou et al. 2014), India

(Surl, Palmer, and Abad 2018), and even globally (Shim et al. 2005; Fortems-Cheiney et al. 2012; Bauwens et al. 2016). In this thesis the OMHCHO dataset from the OMI instrument (see Section 2.3) is used as the basis for HCHO amounts.

### 3.1.1 Aims

In Chapter 2, the OMI HCHO total columns are recalculated using an updated estimate of HCHO profiles from GEOS-Chem v10.01 . These estimates are compared to available datasets of isoprene or HCHO (SPS1, SPS2, MUMBA, Wollongong FTIR) in Section 3.3.5 Sensitivity satellite AMF calculation is examined and quantified for some scenarios. In this chapter we outline why current isoprene emissions estimates are inadequate and how they can be improved. We discuss literature which shows how the estimates may be too high, and describe how emissions may be calculated using satellite datasets. Section 3.2 lays out how new isoprene emissions are estimated, with results examined in Section 3.3. Uncertainties for each step along the way are quantified in Section 3.4. In Section 3.5 we examine how these changes in emissions would affect ozone concentrations in Australia, along with some other chemical processes.

Recent work suggests that modelled emissions may be overestimated in Australia (Emmerson et al. 2016). This work tries to improve the understanding of isoprene emissions over the whole of Australia, clarifying the spatial distribution of bias and how these biases impact modelled chemistry. We estimate isoprene emissions in Australia using top-down estimates based on OMI HCHO measurements and modelled yields. This a posteriori top-down estimate is used to determine if model fit against sparse available ground-based measurements can be improved. GEOS-Chem model output is examined before and after updating isoprene emissions which are used as inputs. 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 Existing emissions estimates

The Model of Emissions of Gases and Aerosols from Nature (MEGAN) is one of the most popular emissions inventories for biogenic isoprene. Along with other models which rely on measured plant emission rates, it is poorly calibrated for Australian conditions. Emissions of isoprene ( $C_5H_8$ ) appear to be overestimated in some regions within Australia (Sindelarova et al. 2014; Stavrakou et al. 2014; Emmerson et al. 2016). A lack of measurements of isoprene emission rates in Australia, along with missing soil moisture parameterisation are the most likely culprits TODO: cite. Sindelarova et al. (2014) showed how the isoprene emissions could be as much as halved by accounting for lower soil moisture. Stavrakou et al. (2015) saw isoprene emissions overestimated by a factor of 2-3 in January. 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. Additionally, emissions of monoterpenes ( $C_{10}H_{16}$ , two units of isoprene) appear to be underestimated (Emmerson et al. 2016).

Recently Bauwens et al. (2016) estimated isoprene emissions with a top-down technique using the IMAGESv2 global CTM. They calculate emissions which create the closest match between model and satellite vertical columns, and compare these a posteriori data with their a priori (satellite data) and independent data sets. They examine global emissions seen by three models and a top-down inversion. For Australia they found emissions ranging from 26–94 Tg C yr<sup>-1</sup>. In this thesis we prioritise the analysis of a top-down emissions estimate compared against MEGAN, along with flow on effect of changed emissions on modelled ozone levels.

### 3.1.3 Top-down isoprene emissions estimates

In the remote troposphere HCHO production is dominated by methane oxidation, while in the continental boundary layer production is largely due to non-methane VOCs (NMVOCs) (Abbot 2003; Kefauver, Filella, and Peñuelas 2014). This leads to a causal relationship between enhanced HCHO concentrations and NMVOC emissions. In the continental boundary layer, 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 yield in many reactions beginning with isoprene oxidation, discussed in more detail in Section 1.3.3.

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 measurements were first used to measure HCHO over Asia (Thomas et al. 1998), satellites have been used to provided a total column measurement of HCHO, allowing isoprene emissions estimates by proxy (Palmer et al. 2001; Bauwens et al. 2016). Using satellite information to improve biogenic emissions is pinpointed as a valuable use of satellite derived datasets (Streets et al. 2013). In this work we use NASA's OMHCHO product, using measurements from OMI on board the AURA satellite (see Section 2.3).

There are two top-down isoprene emission estimation techniques, Bayesian and linear, which are discussed briefly here. Both the linear and Bayesian techniques assume that modelled chemistry is accurate and only try to correct precursor emissions, which may be a problem if the chemistry is uncertain. Uncertainties in the techniques are discussed in Section 3.4.2.

#### 3.1.3.1 Bayesian

Bayesian inversion corrects biased biogenic isoprene emissions by optimising emission parameters in order to reduce the differences between measured and modelled HCHO. Inversions can be set up to account for effects from transport and allow source attribution (Fortems-Cheiney et al. 2012).

In general a model (the forward model) is used to determine the relationship between HCHO ( $y$ ) and the state variable  $x$ , which represents isoprene and other variable

parameters of interest:

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

where  $\epsilon$  are the (assumed) independent errors in measurements.  $K$  is the Jacobian matrix determined from the forward model representing the sensitivity of  $y$  to the state variable  $x$ . Essentially the  $K$  matrix is the modelled estimation of how  $y$  responds to each of the driving parameters represented by elements of  $x$ . This  $K$  matrix is used in conjunction with error covariance in  $x$  to determine the most likely (maximum a posteriori) solution to  $x$ , given what we know about  $y$ .

For example this method is used by Shim et al. (2005) who optimise GEOS-Chem isoprene emissions in areas with high HCHO concentrations to improve comparison against GOME HCHO observations. They show that the original model underestimates isoprene emissions and HCHO concentrations by 14-46%, with the corrected VOC emissions reducing model biases to 3-25%. The Bayesian inversion is also used in Curci et al. (2010), where a regional CTM (CHIMERE) simulates HCHO (this is the forward model), which is compared against OMI observed HCHO and shown to be regionally biased. The model is run initially with emissions of BVOCs and reactive anthropogenic VOCs turned off in order to work out the background ( $b$ ) values of these compounds. Curci et al. (2010) uses CHIMERE as the forward model to determine the relationship between HCHO and  $x$ , which in this case is isoprene and reactive anthropogenic VOCs, then they calculate the most likely values for  $x$  given known measured HCHO and modelled  $K$ .

An advantage of the Bayesian method is that it can account for pyrogenic and anthropogenic emissions rather than filtering them out. Biases may still arise due to errors in modelled emission estimation (Curci et al. 2010). The Bayesian method is computationally expensive due to the requirement that model runs take place using many permutations of changed inputs. In this work we do not use the Bayesian method due to the heavy computational costs involved.

### 3.1.3.2 Linear

This technique is the simplest, and is performed in this thesis. Using vertical columns of biogenic HCHO one 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 one assumes fast HCHO yield, so that the effect of transport is minimal, and that HCHO and isoprene are at steady states, then one may calculate local yield from a CTM.

By assuming a simple linear steady-state relationship between HCHO and its precursors (Palmer 2003; Palmer et al. 2006; Millet et al. 2006), one may calculate isoprene emissions using measured HCHO. The calculation requires reaction rates and yields from isoprene to HCHO, which can be determined most readily using chemical modelling. The methodology for calculating isoprene emissions from HCHO is laid out in Palmer (2003), and takes into account the expected lifetime and reaction rates of the precursor VOCs and HCHO. In their work 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 and the Global Emissions Inventory Activity. An outline of how this method is applied in this thesis is given in Section 3.2.1.

This technique assumes fast HCHO yield and no precursor transport, which is unrealistic in several scenarios; e.g. high winds can lead to transported precursors, or low NO<sub>x</sub> concentrations can slow down HCHO yield. Filtering out potentially problematic scenarios can limit the utility of this technique depending on environmental factors. As we are estimating biogenic emissions of isoprene, we must also filter out areas where HCHO may be coming from anthropogenic or pyrogenic sources. On the plus side, the simple nature of the inversion requires very little computational power after acquiring satellite and model datasets, even over large amounts of gridded data.

## 3.2 Methods

We broadly follow the method of Palmer et al. (2001) to create an emissions estimate of isoprene over Australia. In this work yield is calculated from the modelled slope between isoprene emissions and HCHO tropospheric columns within each gridbox over Australia, as performed in Palmer (2003), using modelled values between 1300-1400 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 this yield we use a reduced major axis (RMA) regression between modelled average values of the total columns and isoprene emission rates, an example is shown in Figure 3.1 shows the modelled regression between emissions and tropospheric columns for January, 2005. Also shown is the time series for these two quantities averaged over Australia, and the squared correlation coefficient along with a sample from four grid squares.

### 3.2.1 Outline

This section provides an overview of the steps involved in creating a top-down emissions estimate. This process is summarised in Figure 3.2.

1. Corrected vertical columns ( $\Omega$ ; saved in OMHCHORP dataset) are calculated (see Section 2.6) using level two OMI HCHO satellite data (see Section 2.3), along with GEOS-Chem model runs (see Section 2.4.7) at  $0.25^\circ \times 0.3125^\circ$  horizontal resolution.
2. Level three satellite data are used to make anthropogenic, fire, and smoke influence masks (see Section 2.7). These are applied to remove  $\Omega$  which may be influenced by pyrogenic or anthropogenic sources.
3. GEOS-Chem modelled biogenic emissions of isoprene ( $E_{GC}$ ) along with total biogenic columns of HCHO ( $\Omega_{GC}$ ), both averaged over  $2^\circ \times 2.5^\circ$  horizontally and 13:00-14:00 LT temporally, are used to calculate a reduced major axis linear regression slope ( $\Omega_{GC} = SE_{GC} + \Omega_{GC,0}$ ). Calculation of this modelled slope is explained in Section 3.2.4.

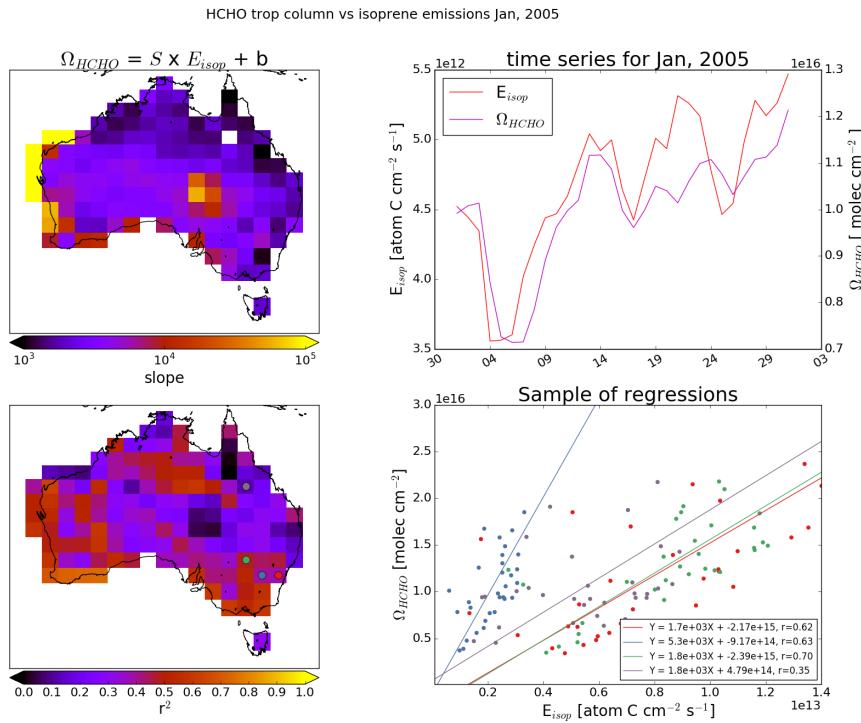


FIGURE 3.1: Top left: RMA slope between modelled tropospheric column HCHO ( $\Omega_{GC}$ ) and isoprene emissions ( $E_{GC}$ ) using midday (13:00-14:00 LT) values over for January 2005, per grid square at  $2^\circ \times 2.5^\circ$  horizontal resolution. Top right: Australia-wide average of midday emissions and tropospheric columns. Bottom left: Squared RMA correlation coefficient for regression in top left. Coloured dots correspond to colour of regressions shown in bottom right panel. Bottom right: Sample of correlations from four grid squares.

4. Satellite measured HCHO columns (See section 2.3) are used to create a top-down estimate of biogenic isoprene emissions ( $E_{OMI}$  atoms C cm $^{-2}$  s $^{-1}$ ). This product is our a posteriori, and calculation details are given in Section 3.2.6.
5. Modelled slope ( $S$ ) calculations depend on several assumptions which are not always valid. A mask is created for where the HCHO production is not dominated by local isoprene emissions. This is determined by calculating smearing over Australia using two model runs with differing isoprene emissions. This smearing value is determined as  $\hat{S} = \Delta\Omega_{GC}/\Delta E_{GC}$ : the ratio of the differences between model runs of HCHO columns and isoprene emissions. The acceptable range for  $\hat{S}$  over Australia is determined as 800 - 4600 s. A full description of the creation of this smearing filter is given in Section 3.2.7
6. A posteriori top-down emissions  $E_{OMI}$  are compared against a priori emissions, as well as analysed in conjunction with concentration measurement campaigns (MUMBA, and SPS). Results are examined in Section 3.3.
7. GEOS-Chem is run using the a posteriori emissions (see Section 3.2.8), and HCHO, O<sub>3</sub>, isoprene, and NO<sub>x</sub> outputs are compared to campaign and satellite measurements where possible.

### 3.2.2 Masks and reprocessed satellite HCHO

Satellite data pixels are read from the OMHCHO level 2 satellite dataset, AMFs are recalculated, and then pixels are gridded daily into 0.25° × 0.3125° horizontal bins. This forms the intermediate product OMHCHORP, which is fully described in Section 2.6.1. This dataset includes gridded satellite HCHO columns ( $\Omega$ ), along with pixel counts (how many satellite datapoints were used for each gridbox) to allow averaging and resolution changes. In this thesis we use OMI (on board the AURA satellite) as it has better temporal coverage and increased pixel counts when compared to GOME (on board the ERS-2 satellite).

In order to determine biogenic HCHO enhancements from  $\Omega$ , we require filters for non-biogenic sources. While one source of HCHO production is methane oxidation, the linear regression used to estimate isoprene emissions effectively ignores this source as part of the background, which means a methane filter is not required. Anthropogenic, fire, and smoke influence masks are created from three satellite products: NO<sub>2</sub> from OMNO2d, fire counts from MOD14A1, and AAOD from OMAERUVd respectively.

1. The fire mask is created daily using non-zero (MODIS) fire counts over the prior 2 days which occur in local or adjacent grid squares at 0.25° × 0.3125° horizontal resolution.
2. Influence from transported smoke plumes is removed where OMI aerosol absorption optical depth (AAOD, from OMAERUVd) is greater than 0.03.
3. A filter for anthropogenic influence is created daily using OMNO2d NO<sub>2</sub> tropospheric column amounts; masking any grid squares with greater than 2 ×

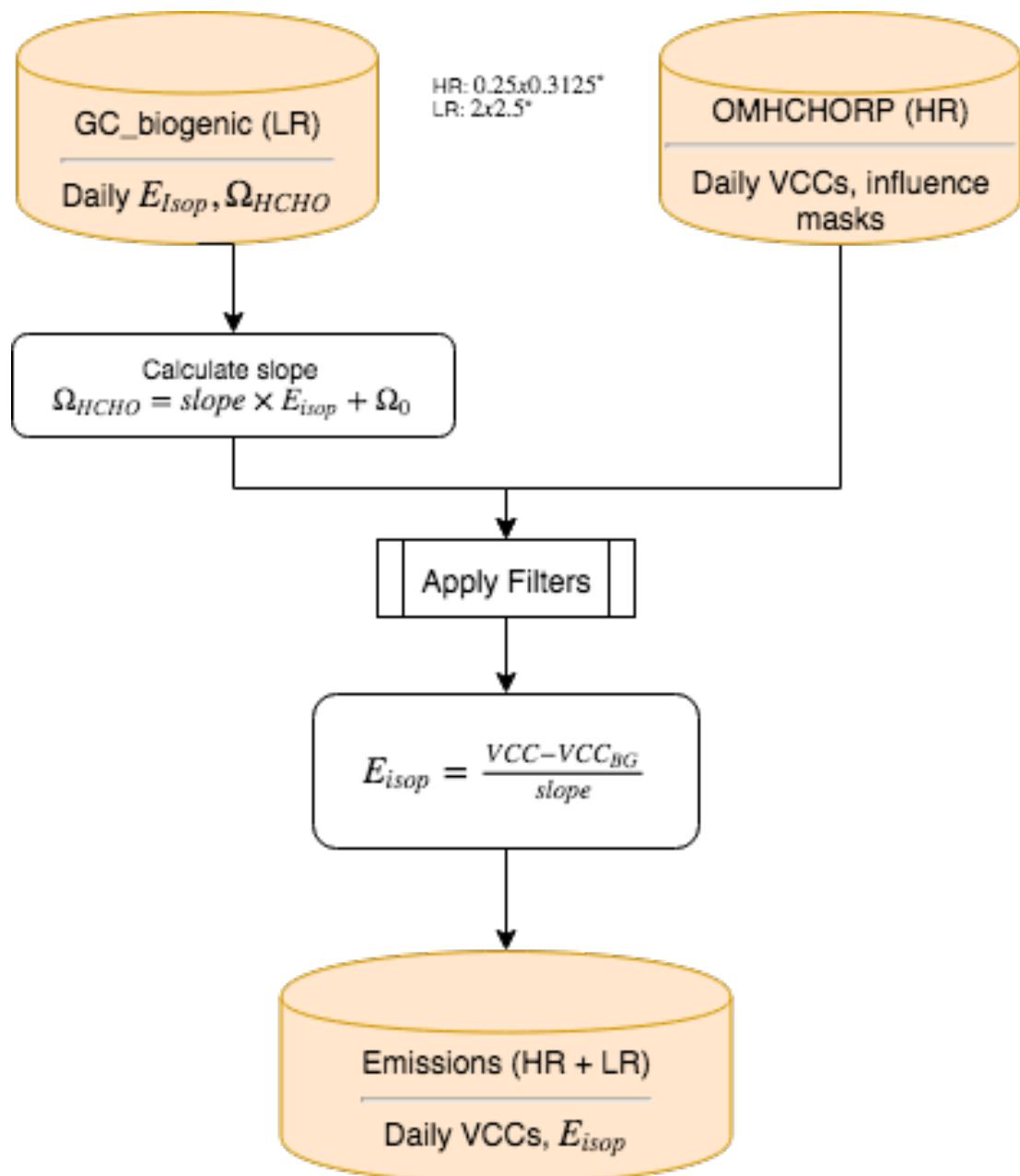


FIGURE 3.2: Creation of isoprene emissions dataset using OMH-CHORP and biogenic GEOS-Chem outputs.

$10^{15}$  molec cm $^{-2}$  on any particular day, along with grid squares where the yearly average is above  $1.5 \times 10^{15}$  molec cm $^{-2}$ .

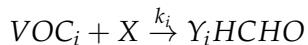
These masks are binned using daily averaged or summed values within our  $0.25^\circ \times 0.3125^\circ$  horizontal resolution, see Section 2.7 for further details. The recalculated corrected vertical columns are saved to OMHCORP dataset both before and after applying the filters, to allow filter analysis.

### 3.2.3 GEOS-Chem emissions

Global atmospheric studies often use models and inventories of various chemical emissions, along with a chemical transport model (CTM) together to examine transport, emission, deposition, and other chemical processes in the atmosphere. Emissions of Biogenic 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. In this work MEGAN is run as a module within GEOS-Chem, a global CTM which uses emissions inventories and meteorological data to simulate atmospheric gas concentrations and transport. Isoprene emissions from the default tropchem simulation are referred to as the a priori emissions.

### 3.2.4 Relationship between isoprene emissions and formaldehyde

Generally HCHO production is due to the oxidation of VOC precursor species ( $VOC_i$ ). Background concentrations are driven by methane; the longer lived VOC. Over continents the variability in HCHO is driven by precursor emissions. The intermediate steps are considered negligible as HCHO is produced quickly from short-lived intermediates:



where  $X$  is an oxidant,  $Y_i$  is HCHO yield (per C atom in  $VOC_i$ ), and  $k_i$  is the reaction rate. In specific conditions, HCHO total columns ( $\Omega$  molec cm $^{-2}$ ) are linearly related to isoprene emissions.

The isoprene to HCHO relationship is derived using several assumptions that are important to understand. The first assumption is that HCHO is at steady state, which implies productions ( $P_{HCHO}$ ) and losses ( $L_{HCHO}$ ) are equivalent.

$$\frac{d\Omega}{dt} = 0 = P_{HCHO} - L_{HCHO} \quad (3.2)$$

This is reasonable during midday when isoprene emissions are steady and  $\Omega$  has had time to stabilise. The second assumption is that loss ( $k_{HCHO}$ ) is only due to photolysis and oxidation which is first order.

$$L_{HCHO} = k_{HCHO}\Omega \quad (3.3)$$

Loss due to transport depends on grid box size and wind speed. This assumption is reasonable for large enough grid boxes as transport becomes negligible in relation to the linear (first order) terms. Production and loss are on the order of minutes, and grid

box sizes in this work are rectangular with  $\sim 200$  km edge lengths. Transport can still be an issue however, and is handled in Section 3.2.7.

These assumptions mean  $\Omega$  production above the background level is due only to precursor emissions ( $E_i$ ; atoms  $\text{C cm}^{-2}\text{s}^{-1}$ ) multiplied by their yields to HCHO ( $Y_i$ ):

$$P_{\text{HCHO}} = \sum_i Y_i E_i \quad (3.4)$$

By combining Equations 3.2, 3.3, and 3.4, we have the link between  $\Omega$  and precursor emissions:

$$\begin{aligned} k_{\text{HCHO}} \Omega &= \sum_i Y_i E_i \\ \Omega &= \frac{1}{k_{\text{HCHO}}} \sum_i Y_i E_i \end{aligned}$$

Finally we assume isoprene emissions are driving changes in  $\Omega$  (as performed elsewhere, e.g. Palmer 2003; Millet et al. 2008; Marais et al. 2014; Stavrakou et al. 2015) and lump other terms into the background  $\Omega_0$ :

$$\begin{aligned} \sum_i Y_i E_i &= Y_{\text{isop}} E_{\text{isop}} + \sum_{i \neq \text{isop}} Y_i \text{VOC}_i \\ &= Y_{\text{isop}} E_{\text{isop}} + \Omega_0 \end{aligned}$$

This leads to the linear relationship between isoprene emissions and  $\Omega$ , by equating  $P_{\text{HCHO}}$  and  $L_{\text{HCHO}}$  from Equations 3.4 and 3.3:

$$\begin{aligned} \Omega &= \frac{Y_{\text{isop}}}{k_{\text{HCHO}}} \times E_{\text{isop}} + \Omega_0 \\ &= S \times E_{\text{isop}} + \Omega_0 \end{aligned} \quad (3.5)$$

Here  $S$  is the slope:  $S \equiv \frac{Y_{\text{isop}}}{k_{\text{HCHO}}}$ . This assumption can be false when pyrogenic or anthropogenic emissions occur, however these scenarios are filtered as much as possible (see Section 2.7).

### 3.2.5 Calculation of modelled slope

To determine  $S$ ; the link between biogenic isoprene and midday column HCHO, we use GEOS-Chem. The term  $E_{\text{GC}}$  is used when discussing isoprene emissions estimated within GEOS-Chem, and  $\Omega_{\text{GC}}$  is used to represent simulated HCHO. The method to calculate  $S$  using GEOS-Chem follows roughly the following three stages:

1. Hourly gridded model output  $E_{\text{GC}}$  atoms  $\text{C cm}^{-2} \text{s}^{-1}$  at 13:00 LT daily are extracted, along with  $\Omega_{\text{GC}}$  molec  $\text{cm}^{-2}$  output.
2. Grid square days which are likely to be affected by smearing are removed (see Section 3.2.7).
3. A reduced major axis regression slope is determined between  $\Omega_{\text{GC}}$  and  $E_{\text{GC}}$  using a month of modelled output (one value per day) for each grid square (e.g. see Figure 3.3)

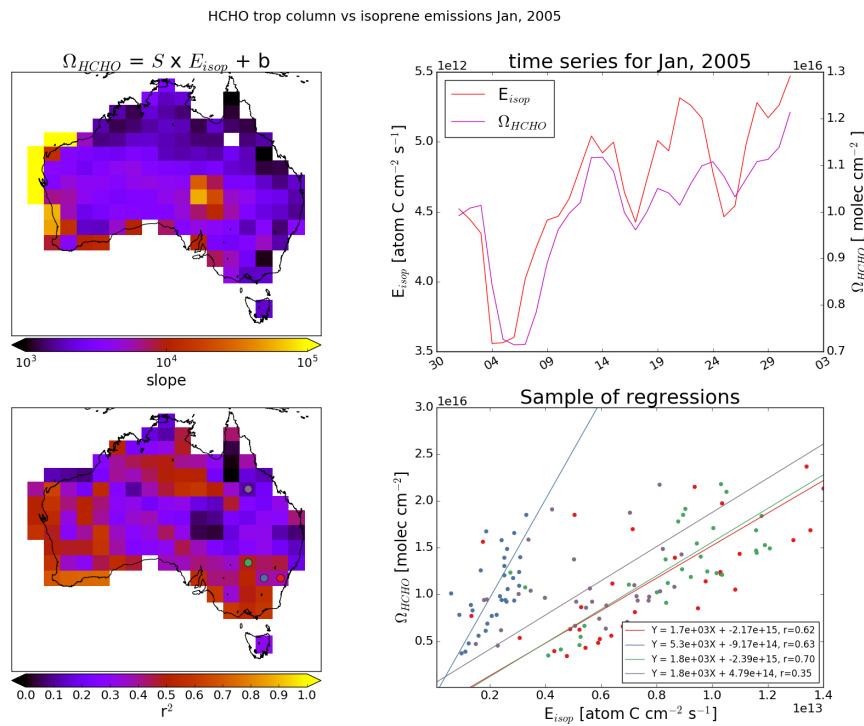


FIGURE 3.3: Top left: the RMA regression slope between modelled HCHO columns and isoprene emissions calculated within each grid box over January 2005. Bottom left: the square of the regression coefficients. Top right: Australia averaged time series Bottom right: a sample of the regressions coloured by latitude and longitude pairs, matching dots in the bottom left panel.

We look at each  $2^\circ \times 2.5^\circ$  grid box from daily GEOS-Chem (biogenic only) output of  $\Omega_{HCHO} \equiv \Omega_{GC}$  and  $E_{isop}$  within Australia, and calculate the (RMA) slope for each month. Modelled background concentrations can be ignored here as they do not affect slope calculation. This effectively gives us monthly gridded isoprene to HCHO slope ( $S$ ) in units of molecules HCHO C<sup>-1</sup> s. Figure 3.3 (top left) shows the RMA regression slope between modelled HCHO columns and isoprene emissions calculated within each grid square over January 2005, averaged between 1300-1400 LT each day. Also shown are the regression coefficients (bottom left), Australian  $E_{isop}$  and  $\Omega_{GC}$  time series (top right), and a sample of the regressions (bottom right). Each regression is for a single grid box, over the course of one month, and is coloured by location (matching dots shown in the bottom left panel). The slopes shown in the bottom right panel range widely due to differences in emission and yield parameters, which plays a role in the smearing filters applied in Section 3.2.7. Due to the  $2^\circ \times 2.5^\circ$  horizontal resolution of GEOS-Chem, 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 equation 3.5.

One issue with slope calculation is the potential transport of isoprene emissions before the resulting HCHO is formed. The effects of this are dealt with by forming a

smearing filter (see section 3.2.7 Accounting for smearing). In order to decide if this filter should be applied before or after calculating the monthly slope, a quick analysis is performed on how the filter affects monthly slope, correlation, and uncertainty. Figure 3.4 shows the calculated monthly slope for 2005-2012, along with its 95% confidence interval over Sydney. The left column does not apply the smearing filter prior to slope calculation, while the right column does. The bottom two rows show how the slope calculation looks when using multiple years of data for each month. This plot has been repeated for several grid squares over Australia (not shown). When calculating top-down emissions the smear filtered slope ( $S$ ) is used for each grid square month. The multiple year monthly averaged slope is used instead when the regression coefficient ( $r$ ) is less than 0.4, or number of data points used in the regression ( $n$ ) is less than 10. When  $r$  for the multiple year slope is also lower than 0.4, no estimation is performed, this is also true for negative values of  $S$ .

There are two simple ways to determine the modelled background HCHO, one of which involves running the model with no isoprene emissions, which shows how much they alter vertical column 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 uses HCHO over the remote Pacific at matching latitudes and times, which emulates how the background is determined for the satellite measured HCHO. The difference between these two methods is negligible, so for consistency with satellite data, we determine backgrounds using the remote Pacific. Figure 3.5 shows GEOS-Chem total column HCHO with and without isoprene emissions along with amounts over the remote Pacific at the same latitudes. Background HCHO for any latitude in this thesis are calculated by averaging longitudinally ( $140^{\circ}\text{W}$  to  $160^{\circ}\text{W}$ ) the matching latitudes over the remote Pacific.

### 3.2.6 Calculation of Emissions

Top-down emissions estimates are calculated using OMHCHO (see Section 2.3) slant columns and an updated AMF calculated using code from Paul Palmer's group (see Section 2.6.5). These emissions are referred to as the a posteriori from here on, or  $E_{OMI}$  in formulae.

As performed elsewhere Palmer (e.g. 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, smoke, and anthropogenic emissions. Isoprene emissions are calculated using the linear relationship described in Section 3.2.4 using the modelled slope  $S$  calculated in the prior section and satellite HCHO columns:

$$\begin{aligned} \Omega &= S \times E_{OMI} + \Omega_0 \\ E_{OMI} &= \frac{\Omega - \Omega_0}{S} \end{aligned} \quad (3.6)$$

This is the same equation 3.5, except now we use the modelled slope along with satellite HCHO ( $\Omega$ , and  $\Omega_0$ ).  $\Omega_0$  is calculated using measurements in the remote Pacific averaged monthly and longitudinally, for each  $\Omega$ . This leaves  $E_{OMI}$  as the only

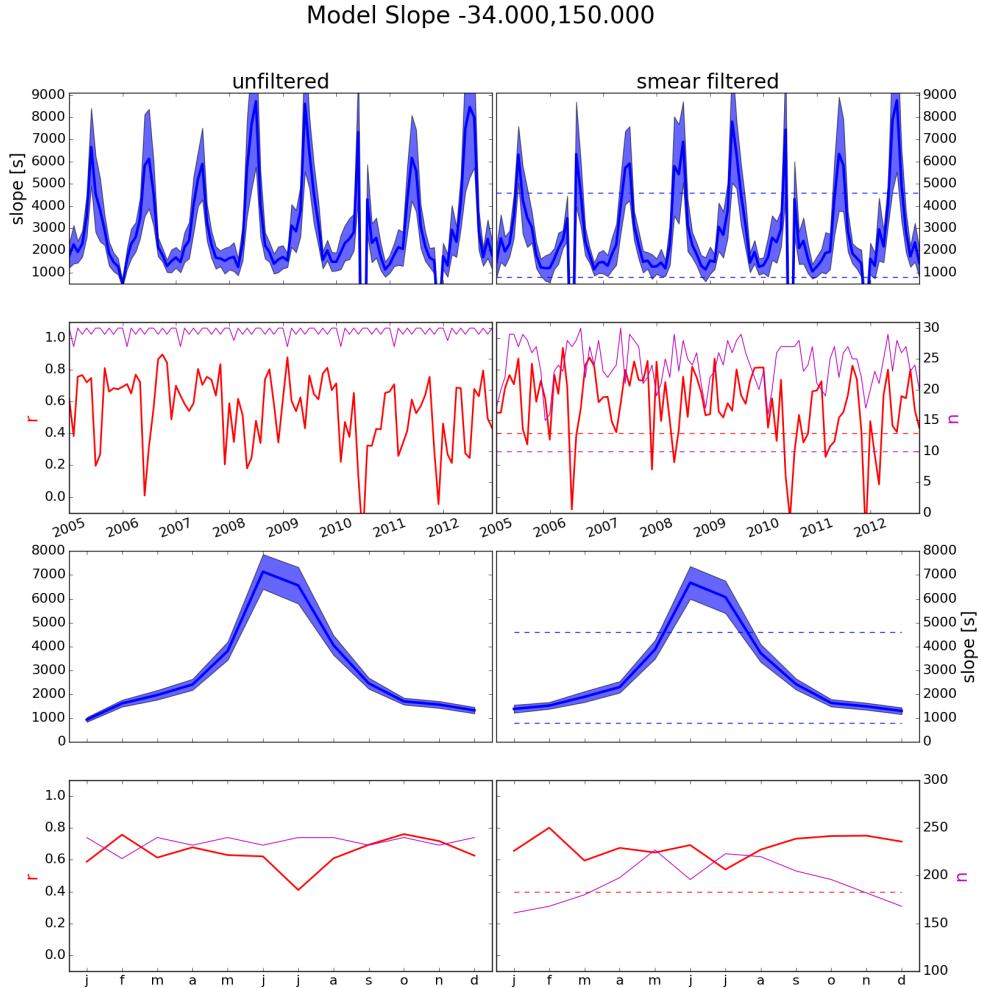


FIGURE 3.4: Row 1: monthly slope along with 95% confidence interval both before (left) and after (right) applying the smear filter for the model grid square containing Sydney over 2005-2012. Limits used in creation of the smear filter are shown with dashed lines. Row 2: regression coefficient and data-point counts for slope shown in row 1. Additionally limits for  $r$  and  $n$  used in slope utilisation (see text) are shown with dashed lines. Row 3: slope and confidence interval using the multi-year dataset for each month. Row 4: regression coefficient and data-point counts for row 3.

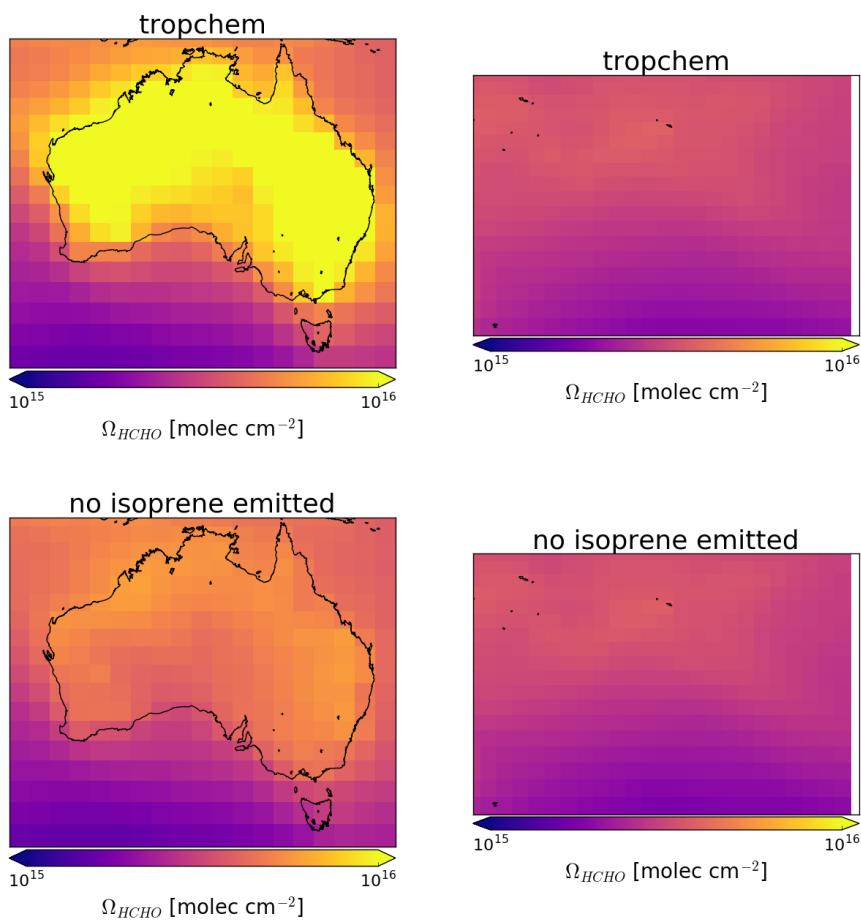


FIGURE 3.5: Total column HCHO over Australia (left) and the remote Pacific region (right) using GEOS-Chem with (top) and without (bottom) isoprene emissions.

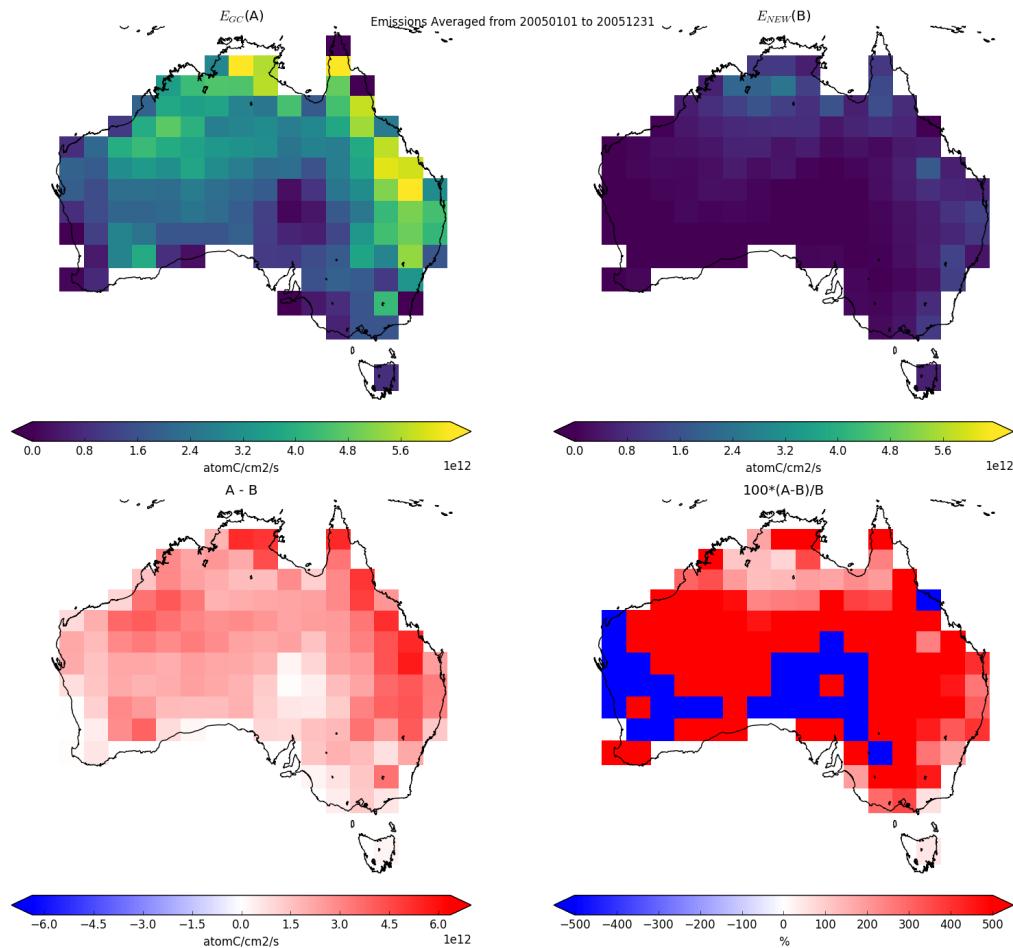


FIGURE 3.6: Top row is isoprene emissions for the month of January, in 2005, from GEOS-Chem and estimated from OMI respectively. Bottom row shows the absolute and relative differences between the two.

unknown once the satellite measurements are processed to match the temporal and horizontal resolution of  $S$ . Figure 3.6 shows the emissions calculated this way along with GEOS-Chem simulated emissions ( $E_{GC}$ ) averaged over January, 2005. This figure gives a single month of output as an example, analysis of the full record is discussed in Section 3.3.

One potential issue is the low number of valid measurements, which occurs due to the higher zenith angles during winter or at higher latitudes. When calculating the  $E_{OMI}$  from our modelled slope, negative emissions result wherever measured columns are lower than background amounts (as  $E_{OMI} = \frac{\Omega - \Omega_0}{S}$ ). These are set to zero, which increases monthly Australian  $E_{OMI}$  averages by TODO:  $\sim X - Y\%$  with the highest increases occurring during (TODO winter?).

Top-down emission rates calculated in this work are in units of molecules cm<sup>-2</sup> s<sup>-1</sup>. In order to calculate the emissions in kg, each grid square is multiplied by the area, and then daily emissions are assumed to follow a sine wave peaking at  $E_{OMI}$ . Figure

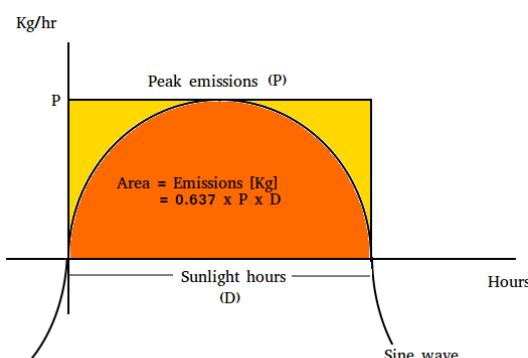


FIGURE 3.7: Total daily isoprene emissions (in kg) is represented by the area under the sine wave.

3.7 shows how the daily approximation of total emitted isoprene per grid square is calculated. Daytime hours are estimated per month, from 14 hrs (Jan) to 10 hr(Jul). This approximation is required since OMI observations occur at midday, when isoprene emissions are at their diurnal peak.

### 3.2.7 Accounting for smearing

In high  $\text{NO}_x$  ( $>\sim 1 \text{ ppb}$ ) 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 (Palmer 2003). In low  $\text{NO}_x$  conditions, isoprene has a longer lifetime (hours-days) and may form HCHO further from the source area. Horizontal transport *smears* the HCHO signal so that source location would need to be calculated using wind speeds and loss rates (Palmer et al. 2001; Palmer 2003). Smearing is a measure of how much HCHO in a given grid box was produced from isoprene emitted in a different (upwind) grid box. Smearing affects emissions estimates as HCHO enhancements may be wrongly attributed as coming from areas downwind of where they occurred. Once smearing sensitive grid squares are filtered out, application of equation 3.5 is used to estimate isoprene emissions across Australia.

Over Australia  $\text{NO}_x$  levels are generally not high enough to ensure quick HCHO formation and we must take care to account for resultant smearing. Smearing limits the horizontal resolution of the top-down inversion process used in this Thesis, as a finer resolution leads to increased issues with transport. The relatively coarse horizontal resolution ( $2^\circ \times 2.5^\circ$ ) used by GEOS-Chem is advantageous in this aspect. Smearing has been analysed in various publications (e.g. Martin et al. 2003; Palmer 2003; Millet et al. 2006; Stavrakou et al. 2009; Marais et al. 2012; Barkley et al. 2013; Zhu et al. 2014; Wolfe et al. 2016; Surl, Palmer, and Abad 2018), and is often calculated using the method designed in Palmer (2003). This involves using two almost identical model runs, one of which has isoprene emissions scaled globally by a factor (generally from 0.5 to 2). Another method (e.g. Stavrakou et al. 2009) involves the analysis of an adjoint CTM, however this is computationally expensive and is not pursued here.

### 3.2.7.1 Calculation of smearing

In order to understand the smearing calculation the underlying equations and assumptions must first be understood. From Section 3.2.4, equation 3.5 we have the formulation of a modelled slope ( $S$ ) being the yield of HCHO per C of emitted isoprene divided by HCHO loss rate per second ( $S = \frac{Y_{isop}}{k_{HCHO}}$ ). Using two runs of GEOS-Chem with isoprene emissions being the only difference we have:

$$\begin{aligned} Run_1 : \Omega_{HCHO} &= SE_{isop} + \Omega_0 \\ Run_2 : \Omega'_{HCHO} &= S'E'_{isop} + \Omega'_0 \end{aligned} \quad (3.7)$$

There are several assumptions which need to be understood, as these are what is tested by the smearing calculation. The initial assumption is that the system is in a steady state, with no transport of isoprene affecting HCHO columns, this is the basis for equations 3.7. It is assumed that background values ( $\Omega_0$ ) are from oxidation of methane and other long lived VOCs, so that  $\Omega_0 = \Omega'_0$ . Between these two runs we are only changing the  $E$  term, we do not change any chemistry and so we can expect that the yield and loss rate is not changing between the two runs  $S = S' = \frac{Y_{isop}}{k_{HCHO}}$ . So that Equations 3.7 may be combined as follows:

$$\begin{aligned} Run_1 - Run_2 : \Omega_{HCHO} - \Omega'_{HCHO} &= SE_{isop} - S'E'_{isop} + \Omega_0 - \Omega'_0 \\ \Omega_{HCHO} - \Omega'_{HCHO} &= S(E_{isop} - E'_{isop}) \\ \Delta\Omega_{HCHO} &= S\Delta E_{isop} \\ \hat{S} &\equiv \frac{\Delta\Omega_{HCHO}}{\Delta E_{isop}} \approx \frac{Y_{isop}}{k_{HCHO}} \end{aligned} \quad (3.8)$$

This allows the combination of outputs from the two runs to be used as a check against  $\hat{S}$  being excessively different from expected values for  $S$ .

In order to filter potential smearing, a daily modelled value for  $\hat{S} \approx Y_{isop}/k_{HCHO}$  is determined. By assuming midday HCHO lifetime typically falls within 1.5 to 4 hrs (as seen in the USA), and isoprene to HCHO yield (HCHO per isoprene carbon emitted) lies within the range 0.2 to 0.4 (scenarios estimated in Palmer (2003)): one can set a simple bound on  $\hat{S}$  of  $[0.2 \times 1.5, 0.4 \times 4]$  hrs or 1080 to 5760 seconds. As  $\text{NO}_x$  levels across Australia are relatively low, the yield is likely lower than seen in Palmer (2003): and here we reduce the bounds by 20% and round to the nearest hundred to get a bounding range of 800 to 4600 for  $\hat{S}$ . This range strikes a balance between unlikely modelled yields and how much data is lost to filtering. A better approximation of lifetimes for HCHO is required to properly account for seasonality and regional  $\text{NO}_x$  concentrations. Table 3.1 shows the smearing filters or typical slopes seen in other works.

Similarly to smearing sensitivity calculations in Marais et al. (2012), we run GEOS-Chem with isoprene emissions halved, then calculate  $\hat{S} = \frac{\Delta\Omega_{HCHO}}{\Delta E_{isop}}$  from Equation 3.8.

Here  $\Delta$  represents the departure (daily over 1300-1400 LT) from default run values. If  $\hat{S}$  sits outside an acceptable range then we remove that grid square day from both  $S$  and subsequent  $E_{isop}$  calculations. A relatively large change in  $\Omega_{HCHO}$  compared to local emissions ( $\hat{S} > 4600$ ) suggests that HCHO is being formed from non-local isoprene

TABLE 3.1: Smearing filters or slopes ( $S$ ) seen in literature.

Publication	min. (s)	max. (s)	Notes
Palmer (2003)	1270	2090	Slope ranges seen in north America summer
Marais et al. (2012)		4000	Smearing limits for Africa
Barkley et al. (2013) <sup>a</sup>	1300	1800	Smearing limits for South America
Surl, Palmer, and Abad (2018)	2200	4900	Slope ranges seen in India
In this Thesis	800	4600	Smearing limits for Australia

a: Assumed HCHO lifetime of 2.5 hours implies yields from 0.14 to 0.2 per C, consistent with box modelling.

Slope ranges are observed or modelled  $S$ , while smearing limits are the applied acceptable limits for  $S$  ranges.

emissions. Alternatively, a relatively low value of  $\hat{S}$  ( $\hat{S} < 800$ ) suggests emissions from the local grid square are being exported before they form HCHO, meaning local HCHO levels are not due to local emissions.

Smearing is sensitive to how  $E_{isop}$  is determined, Figure 3.8 shows smearing over two seasons. Isoprene emissions ( $E_{isop}$ ) is defined in two ways in this figure: left column from daily averaged flux, and right column from and midday (1300-1400 LT) flux. Essentially the midday isoprene emissions are at the peak of their daily cycle (shown later in Figure 3.19) which means the effect of smearing is relatively smaller during these hours.

### 3.2.7.2 Sensitivity to smearing

Figure TODO: shows averaged isoprene emissions with added markers showing when the threshold of 800-4600 affects at least one day within the season (cyan or pink diamonds) and where it removes all data for that gridbox (blue or red x). 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 can vary greatly both spatially and temporally. Large smearing values can be seen near coastlines as only a fraction of these grid squares actually emit isoprene, which makes transported isoprene relatively more important in these gridboxes. During summer data-loss from smearing is only minor (todo: Y%), however data-loss peaks in winter (TODO X %), especially in higher southern latitudes.

When limiting smearing ( $\hat{S}$ ) to within 800-4600 s, GEOS-Chem correlation between isoprene emissions and HCHO columns should improve (TODO analysis). A problem arises due to the loss of data used to create monthly grid square regressions.  $S$  and associated regression coefficients ( $r$ ) are calculated monthly, if  $r < 0.4$  then the regression is calculated using multiple years of data for that month. This multiple year regression is discarded completely if it also fails the  $r < 0.4$  threshold, leaving no slope for the grid square for the month. Figure TODO shows GEOS-Chem midday HCHO columns compared against GEOS-Chem emissions of isoprene, over land

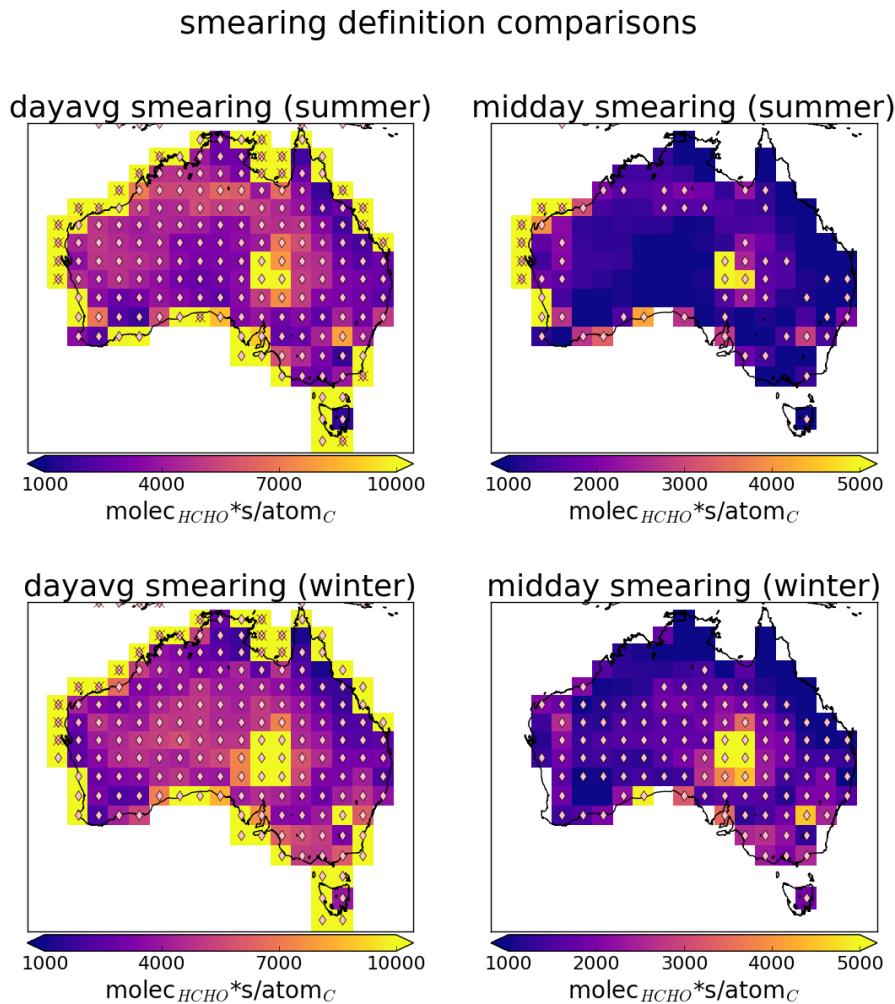


FIGURE 3.8: Smearing ( $\hat{S}$ , see text) in summer (DJF, top row) and winter (JJA, bottom row) using averaged isoprene emissions daily (left column) and 1300-1400 LT (right column). The scale changes between left and right columns.

squares with (red) and without (grey) filtering for smearing. This smearing range captures isoprene to HCHO yields of around 0.16 to 0.32 C per C if HCHO lifetime is assumed to lie within 1.5 to 4 hours. These ranges mean we only use data which is not outside the feasible bounds of yields or lifetimes in Australia.

### 3.2.7.3 Smearing length scale

The expected horizontal transport (prior to reaction) of a precursor can be calculated using the smearing length (Palmer 2003). The distance travelled ( $L$ ) downwind ( $d$ ) by a precursor ( $i$ ) before forming HCHO can be estimated through:

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

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

$$\frac{1}{k_{HCHO} - k_i} \left( k_{HCHO} \exp \left[ \frac{-k_i L_{s,i}}{U} \right] - k_i \exp \left[ \frac{-k_{HCHO} L_{s,i}}{U} \right] \right) = \frac{1}{e} \quad (3.9)$$

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  $k_{HCHO} \ll k_i$ . Figure TODO: shows a rough estimate of isoprene smearing length( $L_{isop}$ ) in Australia using wind speeds from TODO:, and reaction rates  $k_{isop}$ ,  $k_{HCHO}$  from GEOS-Chem.

The half-life of HCHO is around 1 hr depending on environmental conditions (Kaden DA, Mandin C, Nielsen GD 2010). This would make the expected lifetime ( $\tau = \text{half-life} / \ln 2$ ) around 1.4 hours. Over the majority of Australia conditions are relatively clean (low  $\text{NO}_x$ ) which extends the expected lifetime. The estimated loss rate of HCHO in GEOS-Chem is up to three times higher in summer and along the north and eastern regions associated with denser forest regions, when compared against other regions. This is largely due to loss rates being proportional to concentrations.

Isoprene yields can vary greatly, and are dependent on  $\text{NO}_x$  levels, some prior literature values are shown in Table 3.2. Increasing  $\text{NO}_x$  from 0.1 ppbv to 2.0 ppbv can triple the prompt yield of HCHO, due to isoprene, and double background HCHO (Wolfe et al. 2016).

GEOS-Chem daily averaged HCHO lifetime ( $\tau$ ) is shown for 2005 in Figure 3.9. This lifetime  $\tau$  is calculated using daily averaged surface loss rates ( $L_{HCHO}$ ) and concentrations of HCHO:

$$\tau = \frac{[HCHO]}{L_{HCHO}}$$

The expected lifetime of HCHO is determined by assuming loss is linear (first order) and dividing grid box daily averaged concentrations of GEOS-Chem HCHO ( $[HCHO]$  in molecules  $\text{cm}^{-3}$ ) by their modelled losses ( $L_{HCHO}$  in molecules  $\text{cm}^{-3} \text{ s}^{-1}$ ). For each grid square over Australia this daily averaged surface lifetime in summer (Jan., Feb.) and winter (Jul., Aug.) is shown in Figure 3.9. Additionally lifetimes coloured by land grid squares (dots in top right panel) are shown over time in the bottom panel.

TABLE 3.2: Isoprene to HCHO yields and lifetime.

HCHO Yield (molar % )	Lifetime	NO <sub>x</sub> background	Source
315±50		High	a
285±30		High	a
225	35 min	High	b
150		Low	b
150		Low	d
450		High	d
235		1 ppbv	e
150		0.1 ppbv	e

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.

The problem with this approximation is that we are not interested in daily averaged lifetime, but the midday (13:00-14:00 LT) lifetime. This figure highlights the seasonal nature of HCHO loss rates, although midday numbers are expected to have less seasonality. Another highlighted issue is the potential latitudinal dependence of HCHO lifetimes, since there is less total insolation leading to lower HCHO loss rates at higher latitudes.

### 3.2.7.4 NO<sub>x</sub> dependence

Isoprene production of HCHO depends on several factors, importantly NO<sub>x</sub> levels directly affect the fate of VOCs in the atmosphere. Isoprene first reacts with the hydroxy radical, producing an organic peroxy radical (RO<sub>2</sub>). At higher NO mixing ratios (at least a few hundred pptv), RO<sub>2</sub> react mostly with NO. At low NO (less than 50 pptv), RO<sub>2</sub> is more likely to either isomerise, or react with HO<sub>2</sub>, or another RO<sub>2</sub>. In low NO<sub>x</sub> environments, reported HCHO yields from isoprene are around 0.2 - 0.3 C per C, while in high NO<sub>x</sub> environments this value becomes two to three times higher (Palmer 2003; Wolfe et al. 2016).

NO<sub>2</sub> from the OMNO2d product provides daily mid-day measurements which we compare to output from GEOS-Chem NO<sub>2</sub> (see Section 2.4.6). The effect of NO<sub>2</sub> on smearing can be seen in Figure 3.10. This plot shows how smearing over Australia compares to satellite NO<sub>2</sub>, with smearing distributions binned by NO<sub>2</sub> both with and without applying a filter for smearing. At lower NO<sub>2</sub> levels the smearing is often 2-4 orders of magnitude above the upper threshold. This abruptly decreases at around  $5 \times 10^{14}$  molec cm<sup>-2</sup> NO<sub>2</sub>. There is also a higher number of data points below the lower threshold before that same NO<sub>2</sub> level, suggesting that transport is a bigger issue at NO<sub>2</sub> <  $5 \times 10^{14}$  molec cm<sup>-2</sup>.

### 3.2.8 Running GEOS-Chem using a posteriori emissions

After creating  $E_{isop}$ , our a posteriori isoprene emissions estimate, we run GEOS-Chem again with biogenic emissions scaled to match the new estimate. First all the new midday (13:00-14:00 LT) emissions (per grid box) and forming a multi-year monthly mean, which can be compared to the unaltered GEOS-Chem equivalent. A monthly factor ( $\alpha$  at  $2^\circ \times 2.5^\circ$ ) that scales MEGAN towards our top-down estimate is then read by the emissions module in GEOS-Chem, and applied whenever isoprene emissions are calculated. This scaling allows the shorter time-scale variability and dependencies to be retained, while moving the monthly mean value towards the multi-year monthly averaged a posteriori. Short time-scale variability is largely due to meteorological parameters such as temperature and light, etc.. TODO: Figure X shows an example of this process for January: with the multi-year average emissions mapped over Australia (rows one and two) being used to create the  $\alpha$  in the third row.

The scaling factor  $\alpha$  is the ratio between the multi-year averaged monthly emissions from GEOS-Chem  $E_{GC}$  and the a posteriori  $E_{OMI}$ :

$$\alpha = \frac{E_{OMI}}{E_{GC}} \quad (3.10)$$

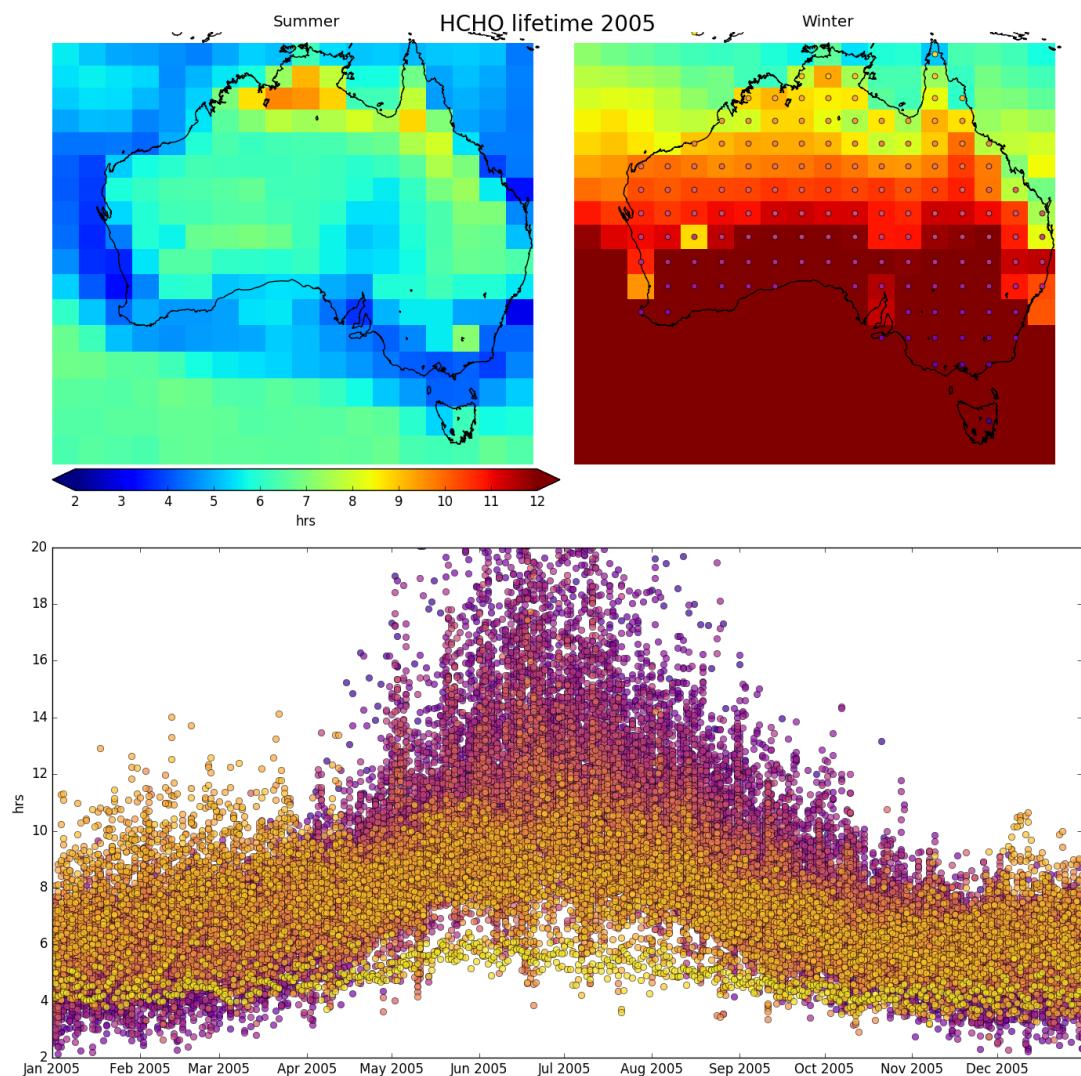


FIGURE 3.9: Top left, right: summer (Jan., Feb.) and winter (JJA) averaged daily surface HCHO lifetime ( $\tau$ ). Bottom panel:  $\tau$  over the year, coloured by grid square (see dots in top right panel).

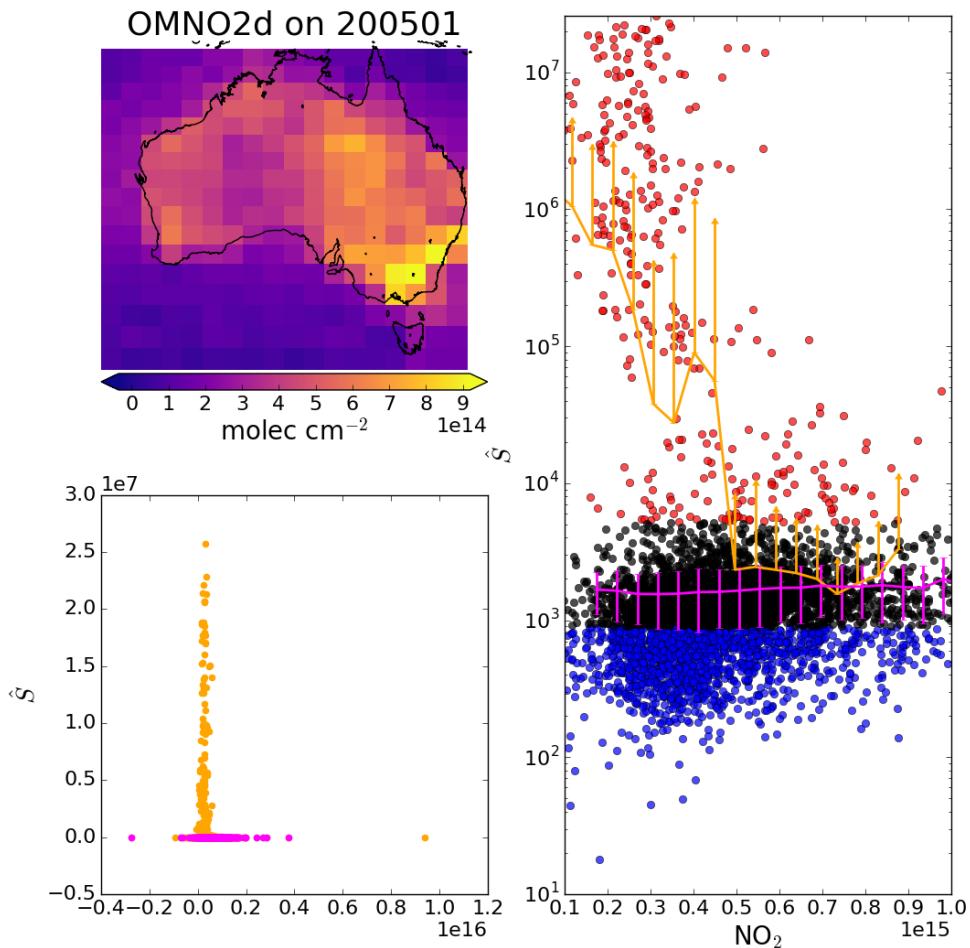


FIGURE 3.10: Top left: OMNO2d tropospheric NO<sub>2</sub> columns (NO<sub>2</sub>: molec cm<sup>-2</sup>) averaged into  $2^\circ \times 2.5^\circ$  horizontal bins for Jan, 2005. Right: Scatter plot of NO<sub>2</sub> against smearing calculations from GEOS-Chem ( $\hat{S}$ ), with points above and below the smearing threshold range of 900–5200 s coloured red and blue respectively. Points are binned by NO<sub>2</sub> with and without the smearing filter applied (orange and magenta respectively). Overplotted is the mean and standard deviation (error bars) within each bin. Due to the logarithmic Y scale we only show the positive direction of standard deviations for unfiltered data. Bottom left: Daily NO<sub>2</sub> scattered against smearing with (magenta) and without (orange) applying the smearing filter. This plot is a zoomed out version of the right panel.

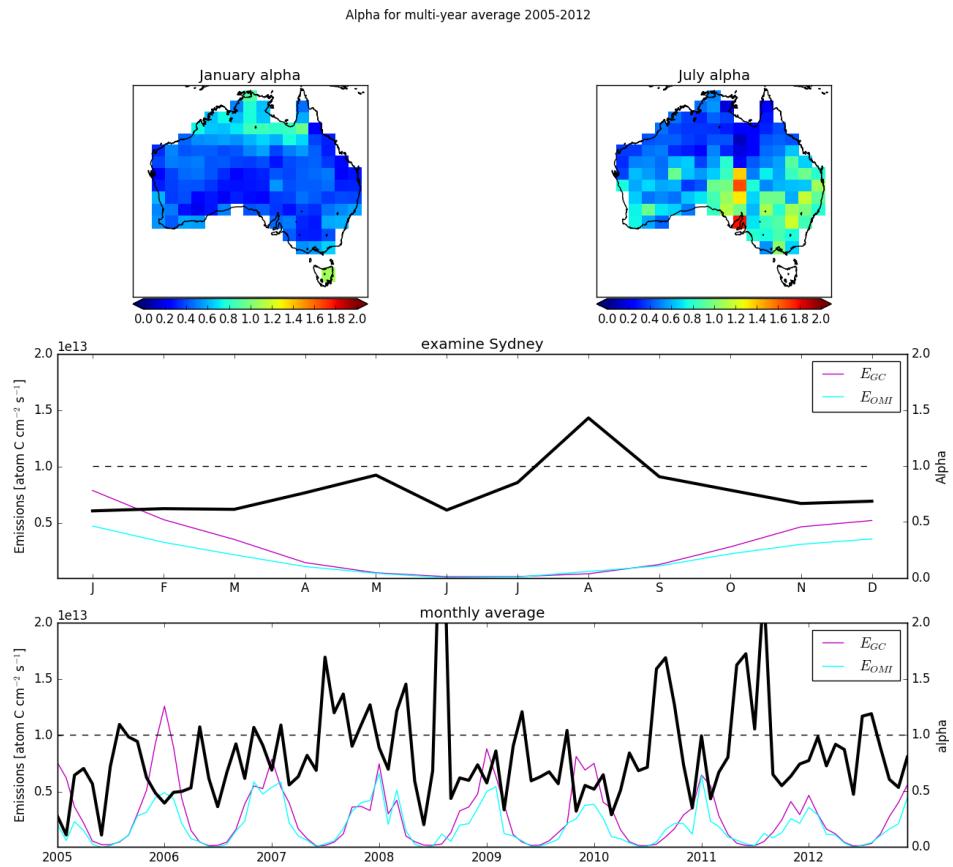


FIGURE 3.11: Row 1:  $\alpha$  for the average January (left) and June (right) over 2005-2012. Row 2:  $E_{GC}$  (xxx, left axis),  $E_{OMI}$  (xxx, left axis), and  $\alpha$ (black, right axis) multi-year monthly averages calculated for Sydney. Row 3: Monthly averages of the same terms in Row 2.

A small modification of GEOS-chem source code multiplies biogenic emissions by  $\alpha$  for each grid-square based on the default emission factors and meteorology. Initially  $\alpha$  is uniformly set to 1 globally. Where we have top down emissions and  $E_{GC}$  is non-zero, we can set  $\alpha$  using Equation 3.10. Missing values for  $\alpha$  when  $E_{GC}$  are zero are an issue, however this is negligible since the dominant discrepancies between estimates occurs during summer when high emission rates are overestimated. Figure 3.11 shows  $\alpha$  for the average January and June over 2005-2012, along with a time series of  $E_{GC}$  and  $E_{OMI}$  and  $\alpha$  calculated for Sydney, and their multi-year seasonal average.

### 3.3 Results

Australia is roughly  $7.7 \times 10^6 \text{ km}^2$ , with heterogeneous environmental conditions. The results presented in this section are frequently split into five regions which are differentiated by colour, as shown in Figure 3.12.

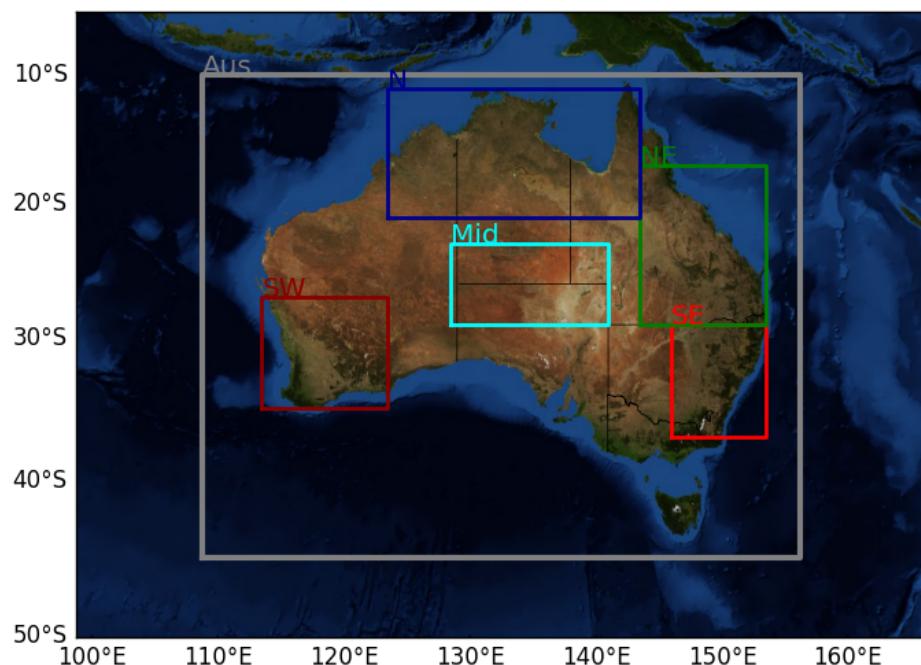


FIGURE 3.12: Sub-regions used in subsequent figures. Australia-wide averages will be black or grey, while averages from within the coloured rectangles will match the colour shown here.

### 3.3.1 Updated emissions

Figure 3.13 shows emissions over Australia calculated using the OMI top down estimate (column 1:  $E_{OMI}$ ) and GEOS-Chem simulated emissions (column 2:  $E_{GC}$ ). The first row shows the time series (daily midday averages) and the final row shows the absolute differences ( $E_{GC} - E_{OMI}$ ). This figure is repeated using monthly means (of the daily midday estimates) in Figure 3.14. Over the entirety of Australia the seasonal cycle of emissions is shown to be overestimated by the a priori. The difference between a priori and a posteriori estimates differs between regions, and some noise can be seen in the monthly time series due to data loss (TODO: check data counts). While most regions generally adhere to this behaviour of overestimation, the northern region of Australia follows a different cycle of bias. Northern Australia appears to be overestimated throughout the year, however occasional low points occur in summer in which the a priori is near or lower than the a posteriori. One potential reason is the onset of the wet season (November-April) which both disrupts the ecosystems and limits satellite measurement (due to cloud coverage).

Figure 3.15 shows the multi-year monthly mean and inter-quartile range (IQR) of daily midday isoprene emissions estimates, averaged regionally (see Figure 3.12). Generally months outside of May to August show the a posteriori lower than the a priori, except in the south eastern portion of Australia.

Figures 3.16 and 3.17 show how the distributions of top-down emissions compare to those of MEGAN in each region during summer months (DJF) with zeros and negatives removed from both distributions. Figure 3.16 shows the daily midday distributions, binned hexagonally to show data-point frequency. There is only weak correlation apparent between daily top down and MEGAN estimations, however daily values suffer from large uncertainty. Figure 3.17 displays the regressions between monthly averages of the same data. In the monthly averages more correlation is apparent, with regression coefficients ranging from 0.48 to 0.78 across regions. The portion of this correlation due to seasonality is examined by re-running the regression after subtracting the multi-year monthly average from each dataset. TODO: run regression after removing seasonal average and see how r and slope are affected.

After running GEOS-Chem with isoprene emission scaled to match the a posteriori, the first preliminary check on the modelled output checks total column HCHO ( $\Omega_{HCHO}$ ) over January in 2005 in Figure 3.18. In every region the new modelled  $\Omega_{HCHO}$  is closer to the satellite measured amounts with biases lowering from 50-120% down to 30-50%.

### 3.3.2 A posteriori GEOS-Chem results

This section describes how outputs from GEOS-Chem change after applying a scaling factor  $\alpha$  (see Section 3.2.8) to biogenic isoprene emissions. Figure 3.19 shows the multi-year monthly mean daily cycle of isoprene emissions from GEOS-Chem, along with the top-down midday estimate, for an example grid box located 2.5° W of Sydney. TODO: Analysis of this and lead into more analysis of differences

TODO: how isoprene is linked to ozone in modelling Figure 5.4 shows surface level (up to ~ 150 m altitude) ozone concentration over 2005 before and after scaling modelled isoprene emissions. Reducing isoprene emissions lowers surface ozone

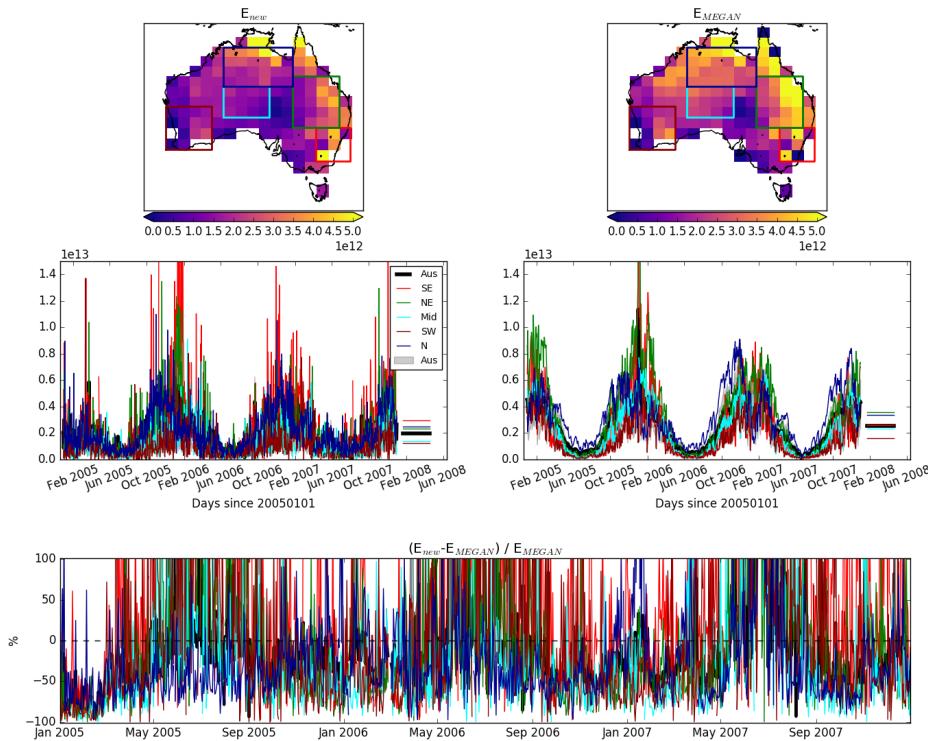


FIGURE 3.13: Emissions of isoprene estimates using OMI top down inversion (column 1:  $E_{OMI}$ ) and MEGAN (column 2:  $E_{GC}$ ). Row 1: overall averaged daily midday emission rates, from 1, Jan, 2005 to 1, May, 2013. Row 2: time series of daily midday averages for all of Australia along with several subregions (shown in row 1). Row 3: relative differences:  $\frac{E_{MEGAN} - E_{OMI}}{E_{MEGAN}}$ . The black lines and grey areas show the Australian mean and inter-quartile range respectively, while the coloured lines show the mean within the rectangles (of matching colours) shown in the first row.

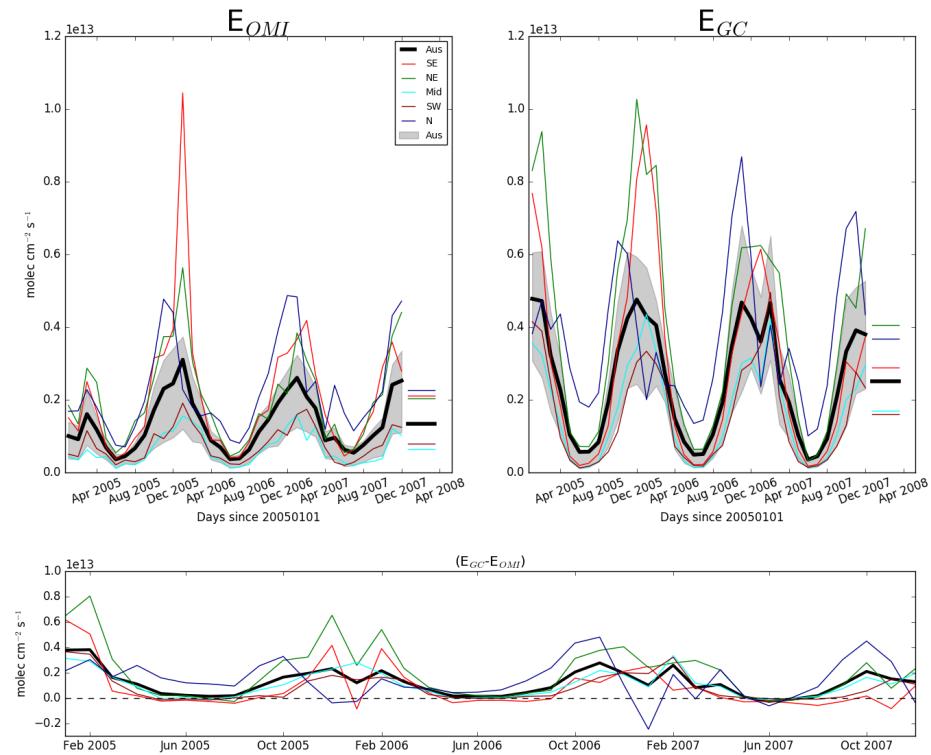


FIGURE 3.14: As Figure 3.13 using monthly medians of the daily mid-day emissions estimates.

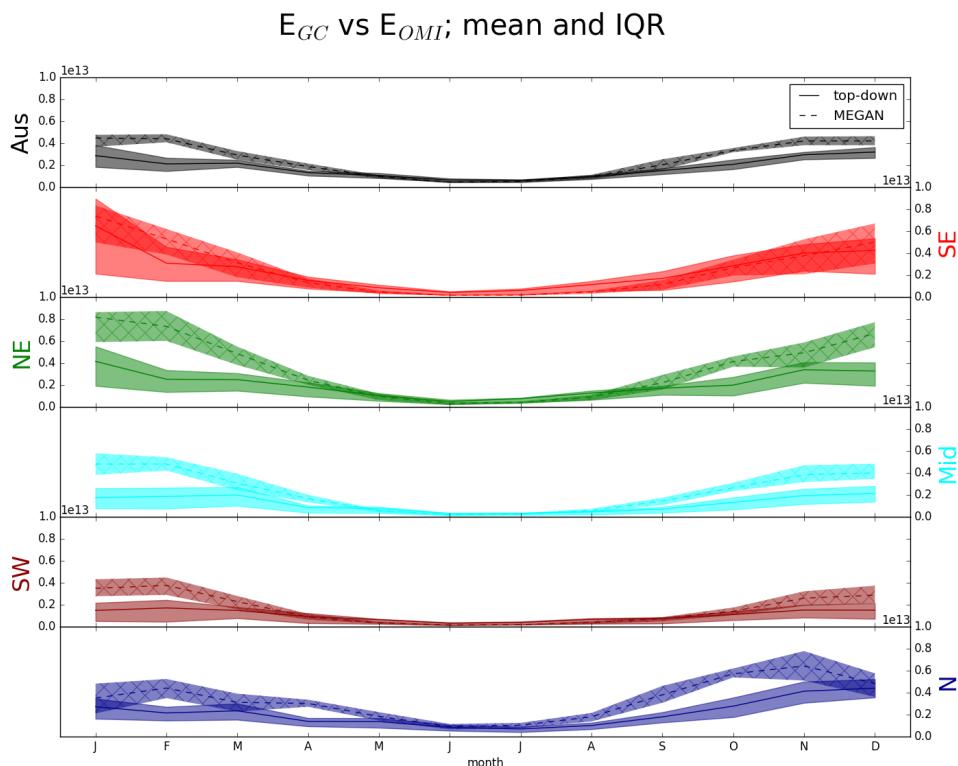


FIGURE 3.15: The multi-year monthly mean (lines) and IQR (shaded) of midday (13:00-14:00 LT) isoprene emissions estimates. Estimates come from MEGAN run by GEOS-Chem ( $E_{GC}$ ), and the OMI top-down technique ( $E_{OMI}$ ). The mean  $E_{GC}$  is shown by the dashed lines and hatched shaded areas show the IQR.  $E_{OMI}$  means are shown using the solid lines, with IQR shown by unhatched shaded areas. Colours denote the region over which the monthly average was taken, as shown in Figure 3.12.

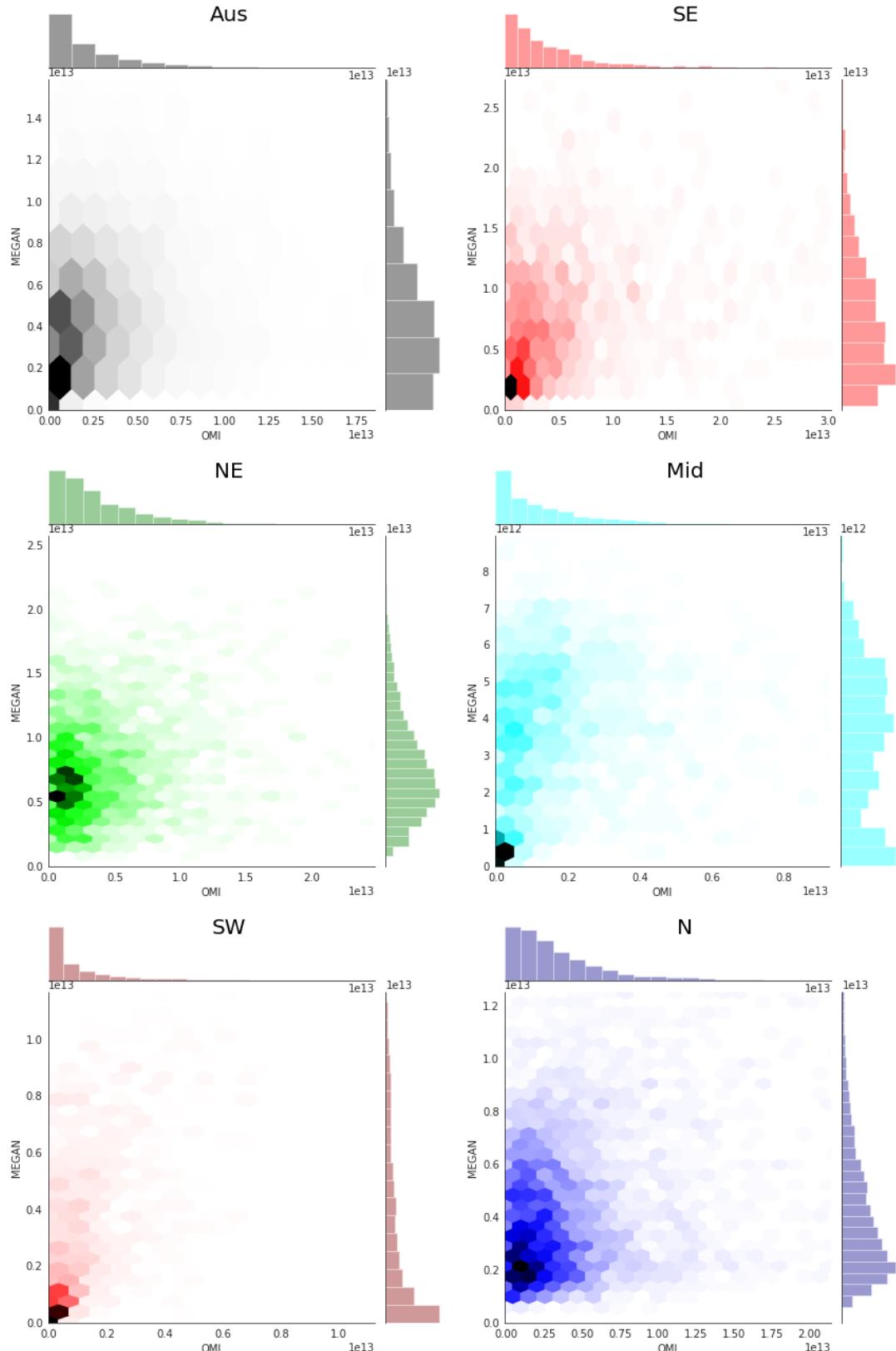


FIGURE 3.16: Scatter plot (binned hexagonally to show data-point frequency) along with the distributions of MEGAN (y axis) and the top down estimate (x axis). This figure is based on summer (DJF) midday values over multiple years. Coloured by regions shown in Figure 3.12.

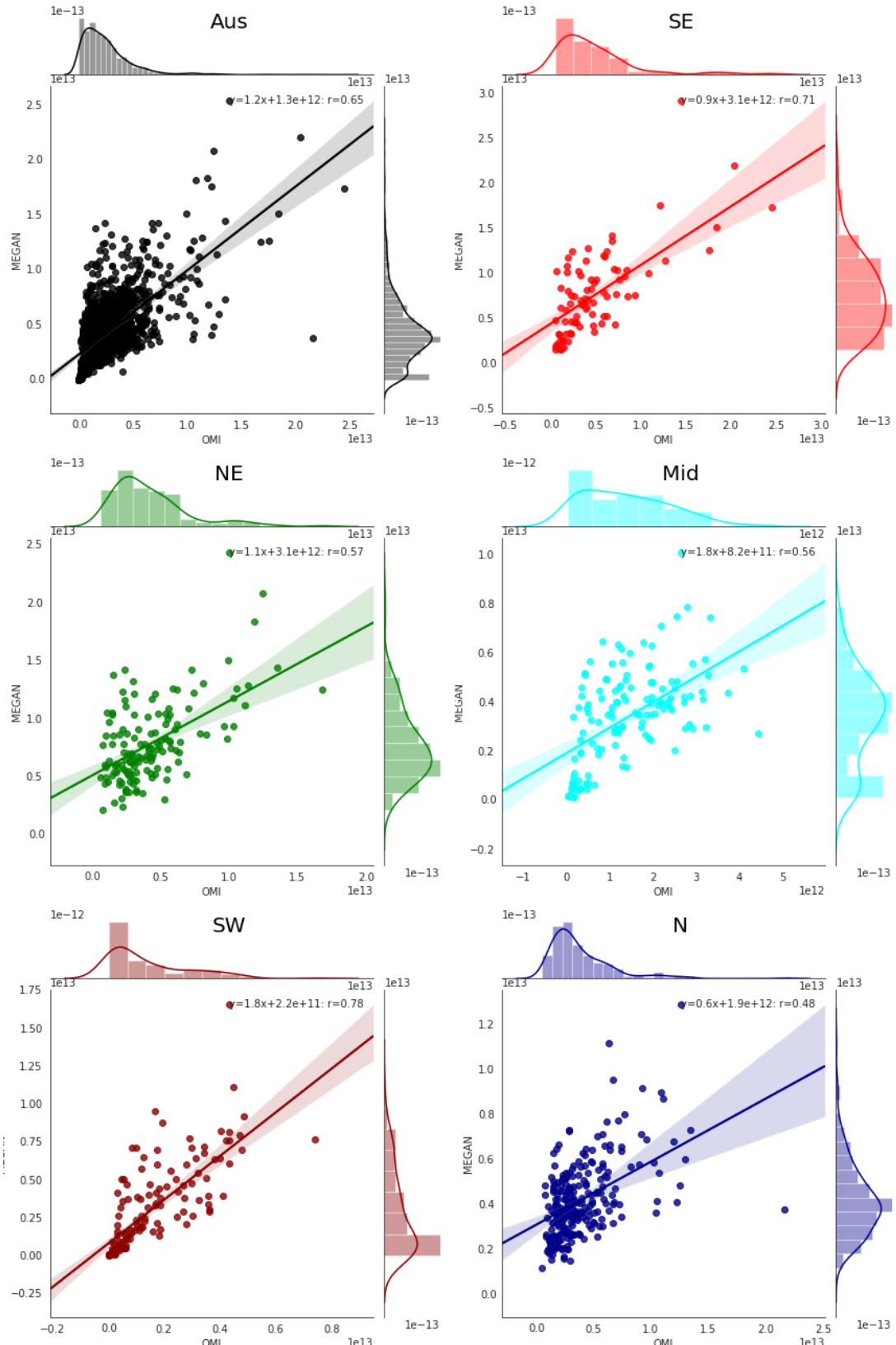


FIGURE 3.17: Scatter plot of MEGAN emissions against top down emissions using monthly averaged grid squares as regression datapoints. Figures use multiple years of summer (DJF) midday values averaged monthly within each region shown. Coloured by regions shown in Figure 3.12.

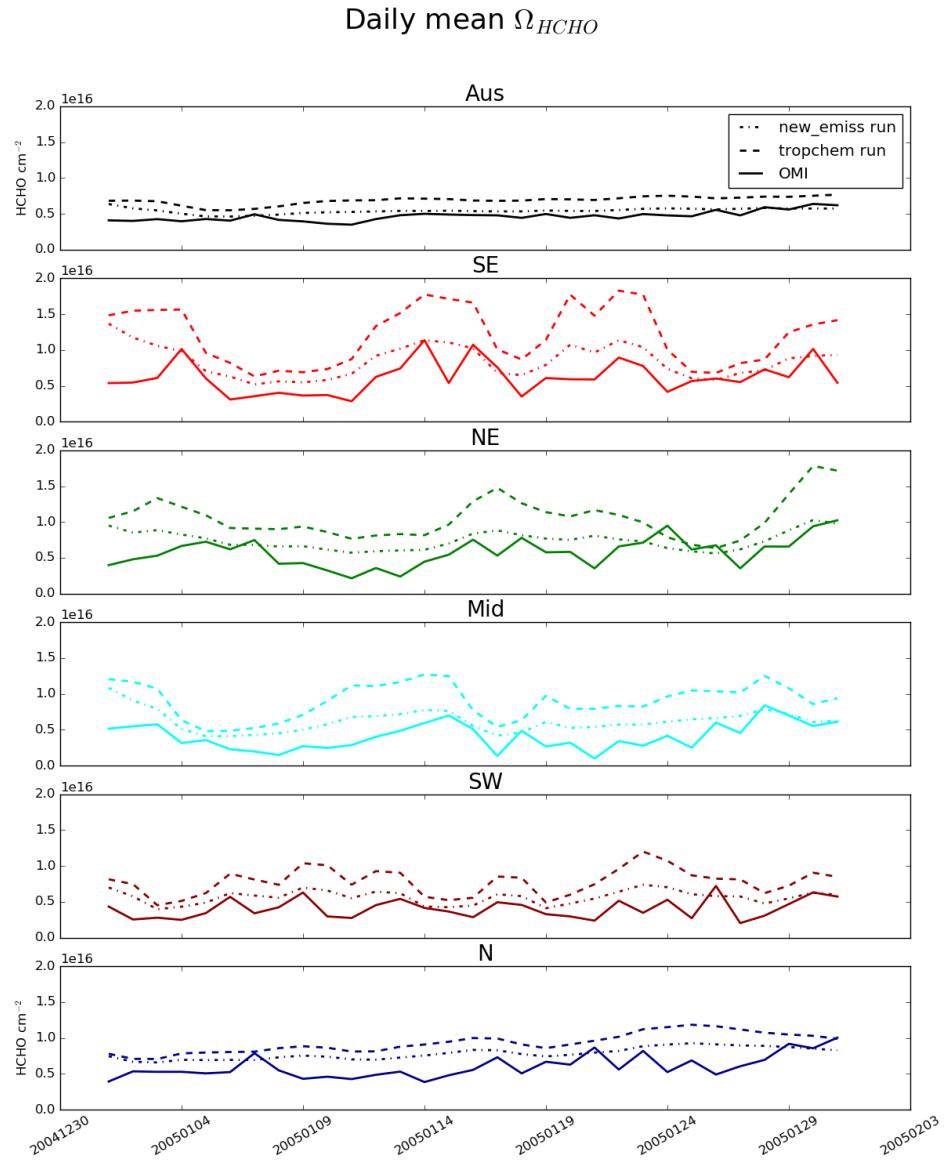


FIGURE 3.18: Daily mean total column HCHO amounts from GEOS-Chem with (new emissions run) and without (tropchem run) a posteriori scaled isoprene emissions, along with the original OMI HCHO columns. Each row shows the average over regions in Figure 3.12.

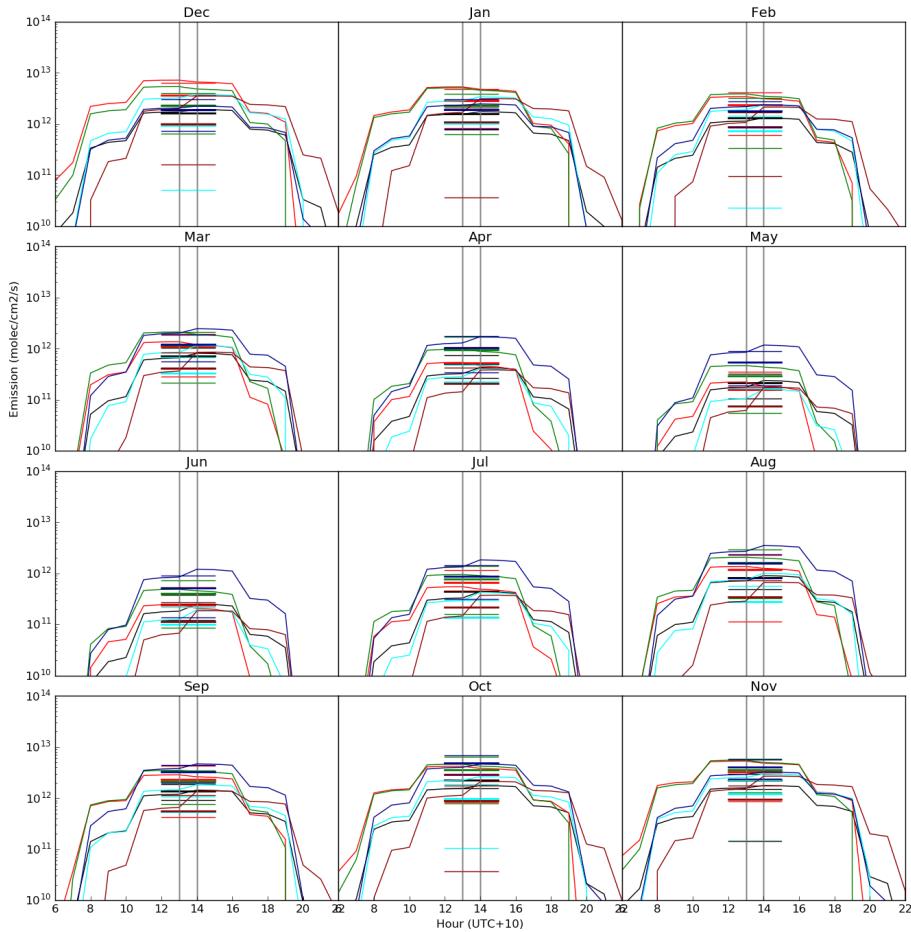


FIGURE 3.19: The diurnal cycle of MEGAN emissions averaged by month over 1, Jan, 2005 to TODO, 2013 are shown with lines, while top-down emissions estimates are shown with plus symbols. MEGAN emissions are estimated hourly per  $2^\circ \times 2.5^\circ$  horizontal grid box, shown here are the area averages (denoted by colour, see Figure 3.12). Top-down estimates are similarly grouped by colour, and shown at the 13:00 LT mark for each month.

TABLE 3.3: Isoprene emissions (Tg/yr)

Australia	notes
43(2)	a) GEOS-Chem: 2005-2010
19(2)	b) Top-down: 2005-2010
26-94	c) Bauwens et al. (2016): 2005-2013
	a: MEGAN diagnostics based on 3-hourly averages
	b: Based on daily peak emissions integrated over a sinusoidal daily curve
	c: Range shown here based on 3 different models and one top-down inversion

concentrations by TODO: XX to YY % in summer, and XX to YY % in winter. The direct correlation between reduced emissions and surface ozone (Figure TODO) is more or less clear monthly, and using multiple years enhances/reduces the relationship. A look at the regression using only summer data reveals a stronger/weaker relationship, analyse per region (TODO).

### 3.3.3 Trends

TODO: Summary of how the VOC NOX Ozone dynamic is changing over time, including how they are linked

TODO: how isoprene emissions are changing over the period Plot time series of isoprene emissions for each region pre and post change subtract seasonal mya cycle and show trend and p value for trend being non-zero.

TODO: how HCHO is changing over the period Plot: time series pre and post isoprene changes as well as OMI for each region. subtract seasonal mya cycle and show trend and p value for trend being non-zero.

TODO: how surface ozone is changing over the period plot: time series of ozone pre and post isop change subtract seasonal mya cycle and show trend and p value for trend being non-zero.

### 3.3.4 Literature comparisons

When comparing the GEOS-Chem (which runs MEGAN) emissions to those calculated using our top-down inversion, we see a decrease of around  $29 \text{ Tg yr}^{-1}$ (66%). Sindelarova et al. (2014) see a  $41 \text{ Tg yr}^{-1}$ decrease in Australia when introducing soil moisture parameterisation. Table 3.3 shows yearly isoprene emissions from this work and some other works for Australia. Figure 3.20 shows how this decrease is distributed spatially, with  $E_{GC}$  and  $E_{OMI}$  in  $\text{Tg yr}^{-1}$ calculated as a multi-year mean. Across all of Australia we see large reductions of total emissions using the new top-down estimate.

### 3.3.5 Comparison with available measurements

TODO: Analyse comparison of gridbox with campaigns of measurements

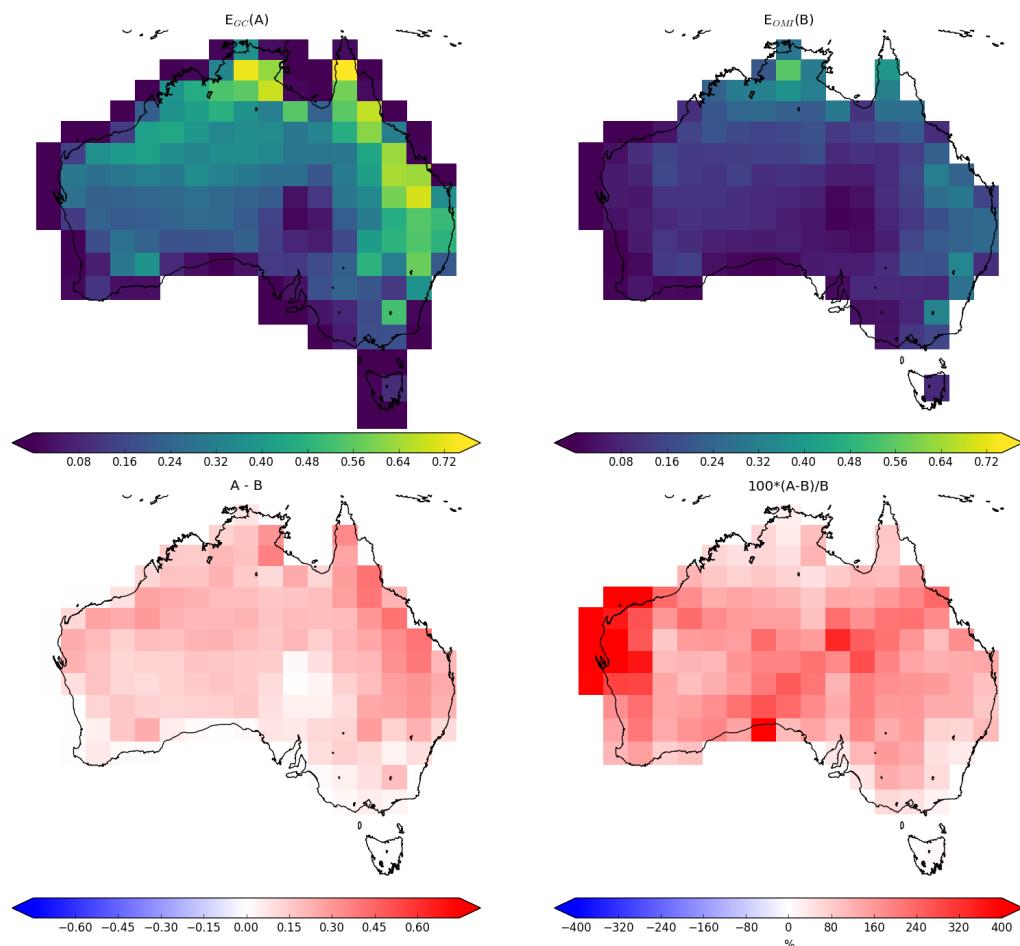


FIGURE 3.20: Top row: multi-year mean emissions in  $\text{Tg yr}^{-1}$  from  $E_{GC}$  (GEOS-Chem; running MEGAN) and  $E_{OMI}$  (top-down emissions) respectively.  $E_{OMI}$  uses an assumed sinusoidal daily cycle, with day-light hours prescribed for each month: see Section 3.2.6). Bottom left and right shows the absolute and relative differences respectively.

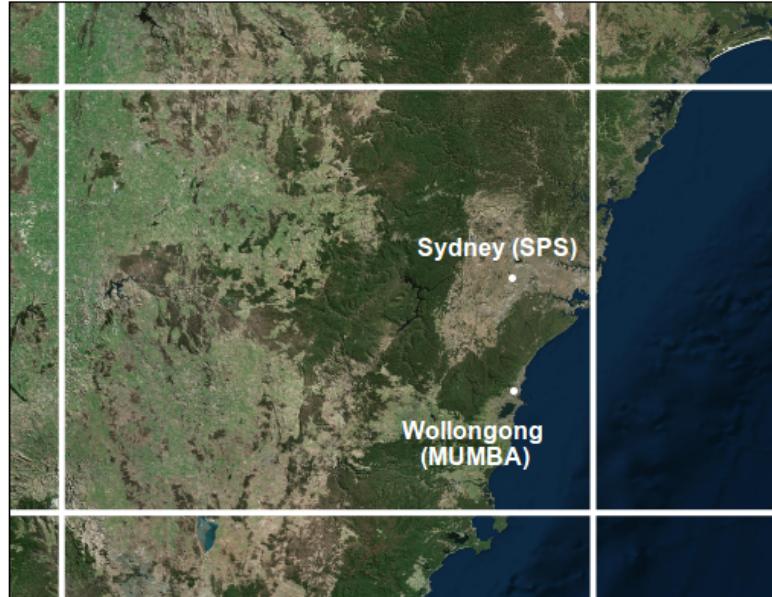


FIGURE 3.21: Gridbox at  $2^\circ \times 2.5^\circ$  containing SPS, and MUMBA campaign data.

Comparison between ground-based measurements and large ( $2^\circ \times 2.5^\circ$ ) averaged grid squares suffers from heavy representational error. Figure 3.21 shows the SPS and MUMBA measurement sites, along with an outline of the  $2^\circ \times 2.5^\circ$  model gridbox. The urban footprint of Sydney and Wollongong are clearly shown, along with some ocean, forest, and rural regions, which are all averaged within calculations made here. Due to high uncertainty in components of the top-down emissions estimate, temporal resolution is also limited. MUMBA, SPS1 and SPS2 each provide only a couple of comparable data points, and these campaigns measured isoprene concentrations while our estimate is of emissions.

Figure 3.22 shows GEOS-Chem output in the grid square containing Sydney overlaid on SPS measurement data. Superficially the comparison is not too bad between these two datasets. This figure shows at a glance how the measurements compare to modelled data, and also the limited nature of the temporal and spatial scale of available measurements. TODO: if possible use midday isoprene multi-year averages, and error bars? then update discussion. 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 since the daily average should not match the peak of the measurements.

Figure TODO shows how isoprene compares against measurements from the MUMBA and SPS campaigns (described in Section 2.2) both before and after scaling isoprene emissions to match the top-down estimation. TODO: discuss results and differences,

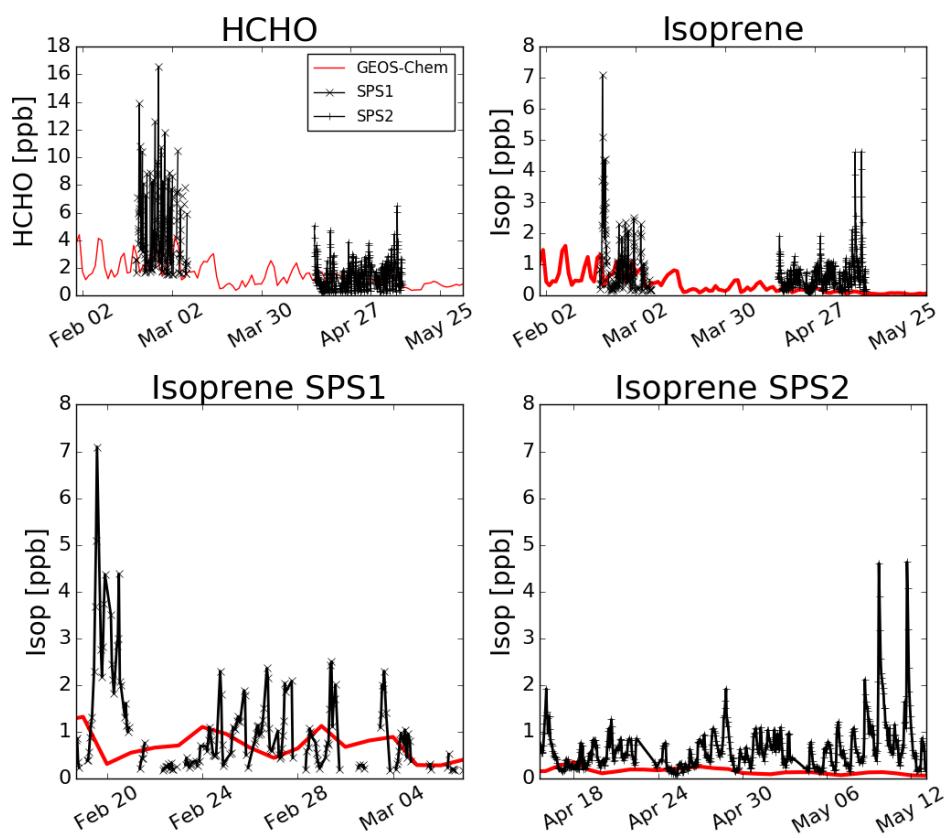


FIGURE 3.22: Comparison between GEOS-Chem HCHO concentrations in the grid square containing Sydney for the duration of the SPS 1 and 2 campaigns

does isoprene improve?

MUMBA data captured concentration at the surface in Wollongong, and is compared in figure TODO against simulated HCHO concentrations at the first level ( $\sim 0$  - 150 metres) over corresponding dates. GEOS-Chem midday and day average concentrations are shown, along with Wollongong bottom-level FTIR measurements to compare against the MUMBA data. TODO Discussion of comparison.

### 3.3.6 HCHO Products and yield

Isoprene reaction chains are diverse, with many branches forming HCHO. HCHO production yields are often classed into two categories: first generation HCHO yield and total (or molar) yield. First generation yield refers to the amount of HCHO produced per unit isoprene consumed by initial oxidation. Total yield refers to time dependent yield of HCHO over multiple oxidation stages (Wolfe et al. 2016). In this work yield ( $Y_{isop}$ ) is not calculated as the slope ( $S = Y_{isop}/k_{HCHO}$ ) can be used instead. This section shows a brief analysis of HCHO lifetimes ( $\tau$ ) at midday.

By assuming yield  $Y_{isop}$  lies between 0.2 and 0.4, we find a range for midday lifetimes of HCHO using equation 3.5:

$$\begin{aligned} S &= \frac{Y_{isop}}{k_{HCHO}} \\ \tau &\equiv \frac{1}{k_{HCHO}} \\ \tau &= \frac{S}{Y_{isop}} \end{aligned}$$

The  $\tau$  is heavily influenced by the assumed yield, and improved methods of estimating yield over Australia are required to improve this estimate. Figure 3.23: shows the GEOS-Chem HCHO lifetime estimated throughout the year. There is a clear seasonal cycle with longer lifetimes in winter months. A clear June (and sometimes March, July and August) increase in HCHO lifetimes is shown, suggesting midday temperature and insolation (which is reduced in winter months) may limit the utility of this top-down emissions estimation technique. The figure is produced using filtered slope information from 2005, and outliers along with low data availability in some months is an issue. Noise in the southwest and middle regions may be indicative of heavy filtering, potentially driven by westerly winds which can lead to both smearing and transported pollution.

## 3.4 Uncertainty

### 3.4.1 Summary

Uncertainties introduced through the inversion process are hard to adequately quantify. We can identify the uncertainties in the linear regression used to relate HCHO to isoprene emissions, as well as in the satellite data product, however these uncertainties lack verification against measurements. Even as this top-down inversion attempts to remedy the lack of measurement over Australia, it suffers from the lack of data-points against which it can be verified.

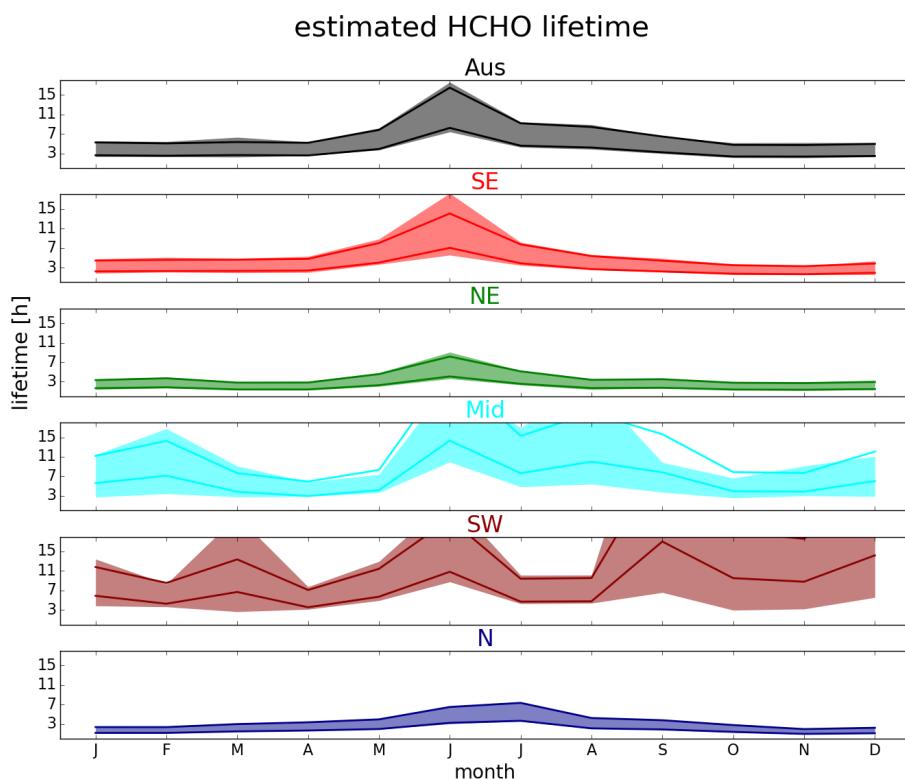


FIGURE 3.23: Monthly area averaged HCHO lifetime ( $\tau$  in hours), with IQR shaded. Solid lines show lifetime assuming yield is 0.2, and 0.4 (higher and lower lines respectively). Coloured by regions shown in Figure 3.12.

This section identifies the overall uncertainties of calculating isoprene emissions using OMHCHO and GEOS-Chem in the top-down method. The major source of uncertainty lies in TODO, and calculations are more or less uncertain in the winter. This limits temporal resolution of isoprene emissions estimates. Table TODO shows each term calculated in this work and the corresponding uncertainty estimate in summer and winter.

### 3.4.2 Top down emissions

There are several factors which need to be considered when looking at the uncertainty in our emissions estimate. 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.

Satellite based chemical concentrations often require both remote and in situ measurements combined with modelled data for validation (Marais et al. 2014). There is less information available from satellite measurements at higher latitudes due to increased error in measurements over the more slanted column paths (De Smedt et al. 2015). Validation is important due to the various uncertainties in the satellite remote sensing process, with a priori 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. They show that a bias in the assumed OMI shape factor leads to a bias between satellite and SEAC<sup>4</sup>RS measurements.

Uncertainty in satellite HCHO, along with top down emissions estimates  $E_{OMI}$  from literature is listed in Table 3.4. The final determination of top-down emissions comes from equation 3.6:  $E_{OMI} = \frac{\Omega - \Omega_0}{S}$ . Assuming each term is independent, we use the following equations to estimate random error in  $E_{OMI}$ :

$$\begin{aligned} z = x + y : \Delta z &= \sqrt{(\Delta x)^2 + (\Delta y)^2} \\ z = x/y : \Delta z &= z \sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2} \\ \Omega - \Omega_0 &= \Phi \\ \Delta \Phi &= \sqrt{(\Delta \Omega)^2 + (\Delta \Omega_0)^2} \\ \Delta E_{OMI} &= E_{OMI} \times \sqrt{\left(\frac{\Delta \Phi}{\Phi}\right)^2 + \left(\frac{\Delta S}{S}\right)^2} \end{aligned}$$

In order to quantify  $\Delta E_{OMI}$  we need to find the uncertainty in the underlying terms:  $\Delta S$ ,  $\Delta \Omega$ , and  $\Delta \Omega_0$ . For  $\Delta S$  ( $\Omega_{GC} = S \times E_{isop} + \Omega_0$  from equation 3.5) we examine related GEOS-Chem output in Section 3.4.3, since  $S$  comes directly from the monthly linear regression of modelled isoprene emissions and column HCHO. Uncertainty in terms  $\Omega$ , and  $\Omega_0$  come from both the OMHCHO slant columns (SC) and the AMFs calculated to transform them into vertical columns ( $VC = SC / AMF$  from equation 2.7). Section ?? describes these calculations.

TABLE 3.4: Uncertainties in literature and here.

product	uncertainty	location	notes
satellite HCHO	40%	North America	(a) mostly due to cloud interference
	X%	where	(b)
top-down $E_{OMI}$	Y%	where	(c)
	X%	Australia	accumulated uncertainty in calculation
	X%	Australia	range found when scaling satellite HCHO

a: Millet et al. (2006) and Palmer et al. (2006)

b:

The summation of these uncertainties through standard quadrature rules provides an estimate of random error in the calculation of  $E_{OMI}$ . In order to calculate the bias or systematic error, an understanding of biases in the underlying terms is required, since there is little in the way of comparable measurements. Known biases:

**OMHCHO** up to 40% underestimated HCHO in the OMI satellite product (pixel bias, (Zhu et al. 2016; De Smedt et al. 2015; Barkley et al. 2013))

**OMHCHO** around 13% overestimation of monthly averaged HCHO (cloud-free bias, (Surl, Palmer, and Abad 2018))

**GEOS-Chem HCHO** TODO under/over estimation of modelled HCHO due to coarse resolution (over prediction of low- $\text{NO}_x$  oxidation pathway TODO:citation)

Figure TODO shows the average summer and winter random uncertainty over Australia, along with a time series of total emissions and the error bars using monthly averaged data and MEGAN emissions for reference.

### 3.4.3 Model Uncertainty

$E_{OMI}$  estimation depends partly on the product it is trying to improve, using modelled yields based on MEGAN. 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. Here GEOS-Chem output is compared against the campaign datasets with the caveat that in situ and point measurements are quite different to modelled (large) area averages.

Uncertainty in modelled yield is estimated through uncertainty in the regression slope (TODO:). Prior works use flight campaigns and in situ data to verify HCHO columns in various locations (TODO: redo cite list from lit review). Figures X to D TODO: make figures comparing campaigns to model HCHO Yield calculations are performed at low resolution ( $2^\circ \times 2.5^\circ$ ) which may lead to overestimation (Yu et al. 2016). TODO: how do I check this? Isoprene to HCHO yield between low and high  $\text{NO}_x$  (0.1 to 1 ppbv respectively) conditions has been estimated through box modelling to be 1.9 to 2.4 mol  $\text{mol}^{-1}$  in Bauwens et al. (2016).

### 3.4.4 Satellite Uncertainty

There are three main sources of error in the satellite 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 gridded average with horizontal resolution of  $0.25^\circ \times 0.3125^\circ$  (latitude by longitude) typically reduces uncertainty by a factor of 1.5 to 4. b) could be determined through comparison of GEOS-Chem output to measured HCHO columns, if they existed. 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  $\sim 30\%$  (calculated in quadrature). Here we use 30% as a rough estimate of error in this term. TODO: Paul palmer calculation and combination for overall Satellite VC uncertainty per pixel and gridded. c) is also determined through a study of GEOS-Chem output, in relation to in situ measurements. Since we expect oceanic background HCHO to be invariant, then variance in remote ocean HCHO can be used as a rough estimate of background uncertainty. 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  $< 0.2$ .

Two simple methods of looking at overall uncertainty from satellite measurements are performed here.

1. using the variance over the remote Pacific ocean to provide relative uncertainty globally (e.g. De Smedt et al. 2012)
2. scaling up HCHO columns by 40% as an upper bound on satellite uncertainty

Analysing variance over the remote Pacific gives a quick measure of uncertainty if we assume HCHO levels within the region are stable. This region should be relatively invariant throughout any month so each month the standard deviation of the midday total column HCHO amounts from  $15^\circ\text{S}$  to  $15^\circ\text{N}$ , and  $180^\circ\text{W}$  to  $120^\circ\text{W}$ . Other literature has found satellite HCHO columns to be up to 40% too low, so scaling them up 40% is another way of quickly analysing how sensitive our calculations are to the satellite HCHO columns.

#### 3.4.4.1 OMI Retrieval

TODO: Calculate remote Pacific variance (also bin by latitudes) and plot over time.

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). Uncertainty introduced through AMF calculation needs to be

additionally determined to give a representation of the confidence in vertical column amounts. 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.

TODO: uncertainty calculation on remote Pacific OMI.

Satellite measured HCHO has been found to be biased low in several studies (e.g. Zhu et al. 2016; De Smedt et al. 2015; 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. Satellite bias is seen to be as high as 40%, which we use as a simplistic method of quantifying potential satellite uncertainty. If satellite HCHO was scaled up by 40% our isoprene emissions estimates would increase by the same, as our estimate is proportional to satellite HCHO.

OMI is scaled up by up to 40% in several papers (cite) we consider HCHO scaled by 1 and 1.4 to be boundaries for modelled yield. If we infer from this that there is 40% bias and random uncertainty remains unchanged, then in the S term, we find TODO increase or decrease in  $\Delta E_{OMI}$  of some amount through changing  $\Delta S =$ .

#### 3.4.4.2 Satellite vertical column recalculation

OMI HCHO vertical columns are recalculated using GEOS-Chem V10.01 a priori HCHO and air density profiles (see Chapter 2). The recalculation is considered assumed to be no more or less uncertain than that used to calculate the default AMF, and uncertainty in the AMF recalculation is not used to alter the pixel uncertainty from the OMH-CHO product. Here we examine the sensitivity of the isoprene emissions estimation technique to the AMF recalculation method. Through looking at how the emissions change based on whether we use the AMF provided ( $AMF_{OMI}$ ), the AMF with shape factor recalculated but the default scattering weights ( $AMF_{GC}$ ), or the fully recalculated AMF ( $AMF_{PP}$ ).

Figure todo shows the emissions over Australia averaged within January 2005. This figure shows estimates from MEGAN (top), and top down estimates using OMH-CHO  $\Omega$ s (row 1: OMI),  $\Omega$  recalculated using GEOS-Chem shape factors (row 2: GC), and using the code from Paul Palmer's group (row 3: PP). Column 2 is emissions without applying anthropogenic or pyrogenic filters, Column 3 is calculated at the lower resolution of  $2^\circ \times 2.5^\circ$ .

Figure todo shows emissions over time from a single grid square, estimated by MEGAN (black) and the three top-down estimates, using  $2^\circ \times 2.5^\circ$  horizontal resolution.

#### 3.4.5 Fire Filtering

Figure 3.24 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 a priori ( $\Omega_{GC}$ ), and those with AMFs recalculated using PP code ( $\Omega_{PP}$ ). The first row shows top-down emissions estimates, 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.

## 3.5 Conclusions and implications

### 3.5.1 Effects from scaling emissions

Using our top-down emissions estimate to recalculate MEGAN isoprene emissions for Australia is detailed in Section 3.2 How these changes affect HCHO and ozone model outputs are discussed in this section.

#### 3.5.1.1 HCHO levels

We examine the effects of scaling 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. 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%

Wollongong FTIR measurements (see Section 2.2.3.4) provide vertical profiles of HCHO which can be converted to total column amounts using modelled air densities. This is the only non-satellite long-term record of vertical profile HCHO available in Australia and we use it here to examine trends and seasonality. The time series for HCHO is shown in figure TODO, along with GEOS-Chem output before and after updating isoprene emissions. TODO: discuss plot here

TODO: Figure showing campaign data against model and recalculated model.

A long term regression of anomalies from the multi-year monthly average shows that TODO. Figure TODO shows the trend for each month along with co-located HCHO from GEOS-Chem outputs. Figure TODO shows how campaign data compares against GEOS-Chem modelled HCHO before and after isoprene scaling is performed. The regressions show improvement/no improvement TODO however there are relatively few data points for comparison as the campaigns only lasted a few months.

Figure todo shows how the longer term FTIR HCHO measurements compare against modelled HCHO before and after changing isoprene emissions

#### 3.5.1.2 Ozone levels

TODO: compare ozone after changing isoprene emissions in GEOS-Chem Changing isoprene emissions over Australia impacts the cities/coasts TODO determine this. This is likely due to the relatively clean atmosphere over the majority of Australia, which is likely  $\text{NO}_x$  limited outside of major population centres.

TODO: Figure shows modelled surface ozone concentrations and their differences between model runs over an average summer (DJF). Figure TODO shows the same

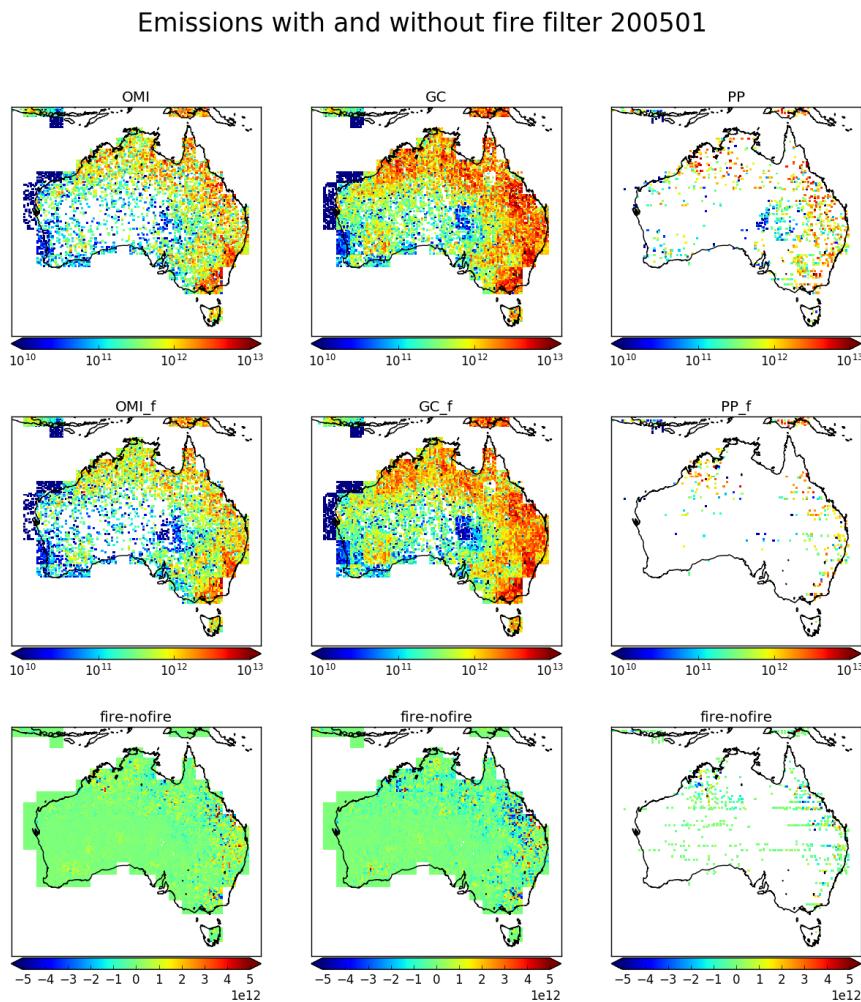


FIGURE 3.24: Emissions estimates using OMI satellite columns (column 1) recalculated with updated shape factor (column 2) and scattering weights (column 3). Turning off the fire and smoke filters gives emissions in row 2, while the difference between row 1 and row 2 is shown in row 3.

over an average winter (JJA). Figure TODO compares the summer differences to surface NO emissions (modelled), and total column NO (satellite).



## Chapter 4

# Stratospheric ozone intrusions

### 4.1 Foreword

This chapter is unchanged (excepting titles, labels, and numbering) from work published by me into the atmospheric chemistry and physics journal: **Greenslade2017**

### 4.2 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) 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).

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 4.3 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 4.4 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 4.5 we analyse how well GEOS-Chem captures the tropospheric ozone seasonality and quantity near our three sites. In Section 4.6 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 4.7 we examine in detail the uncertainties involved in our STT event detection technique and ozone flux estimations.

## 4.3 Data and Methods

### 4.3.1 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).

Ozonesondes are launched approximately weekly from Melbourne (38° S, 145° E), Macquarie Island (55° S, 159° E) and Davis (69° S, 78° 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

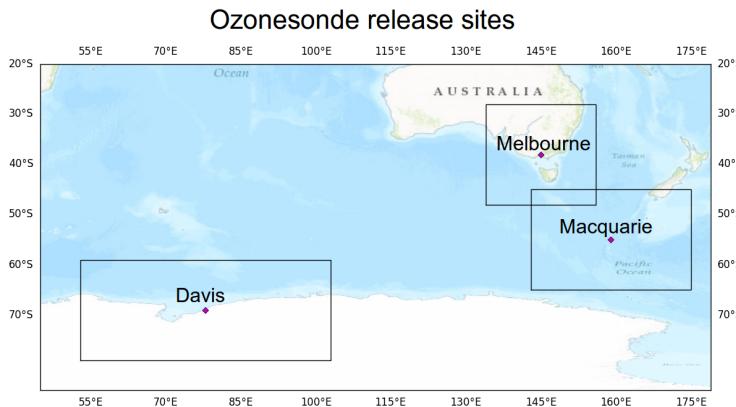


FIGURE 4.1: Ozonesonde release sites and the regions used to examine STT effect on tropospheric ozone levels.

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

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

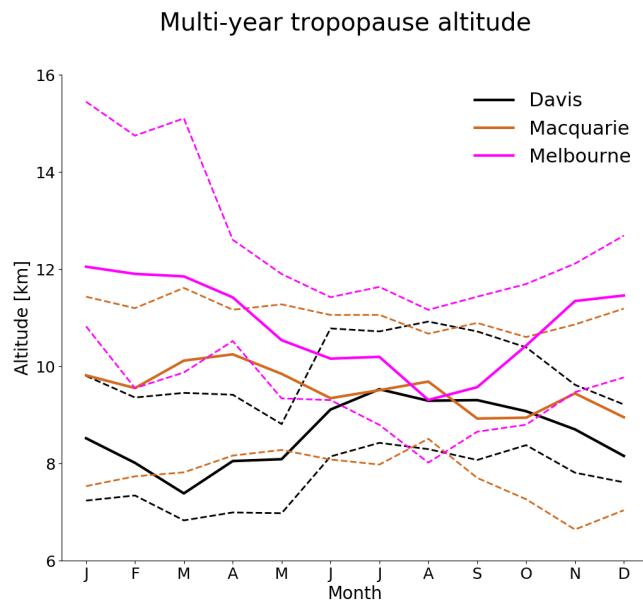


FIGURE 4.2: Multi-year monthly median tropopause altitude (using the ozone defined tropopause) determined from ozonesondes measurements at Davis (2006-2013), Macquarie Island (2004-2013), and Melbourne (2004-2013) (solid lines). Dashed lines show the 10th to the 90th percentile of tropopause altitude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

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  $PV = 2$  PVU ( $1 \text{ PVU} = 10^{-6} \text{ m}^2 \text{ s}^{-1} \text{ K kg}^{-1}$ ) 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

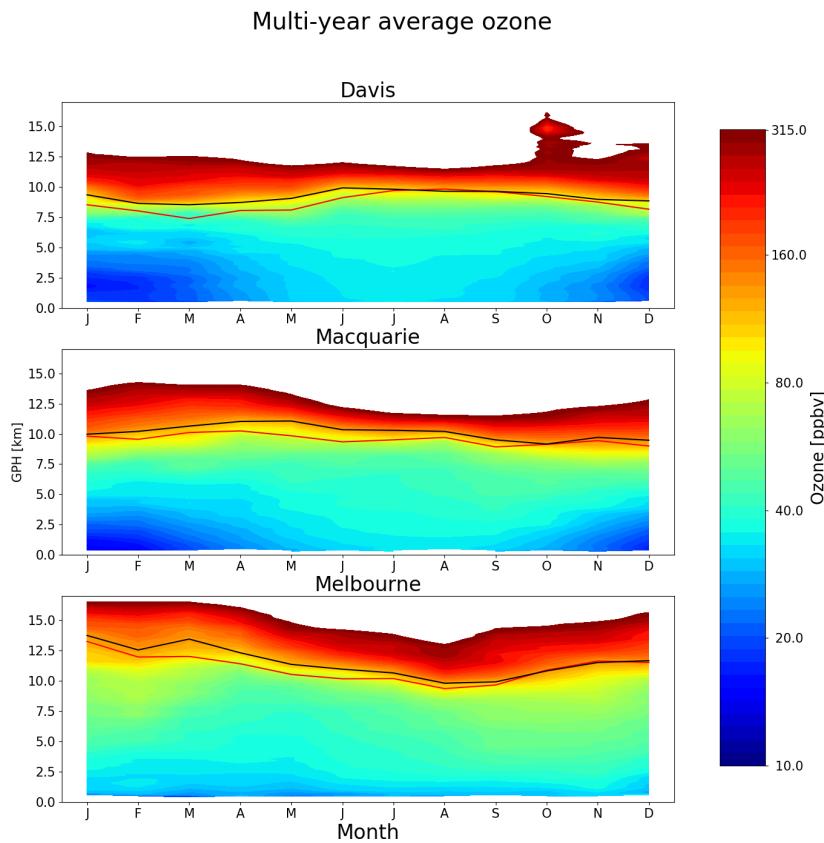


FIGURE 4.3: Multi-year mean seasonal cycle of ozone mixing ratio over Davis, Macquarie Island, and Melbourne as measured by ozonesondes. Measurements were interpolated to every 100 m and then binned monthly. Black and red solid lines show median ozone and lapse-rate defined tropopause altitudes (respectively),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

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 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.

#### 4.3.2 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).

#### 4.3.3 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

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

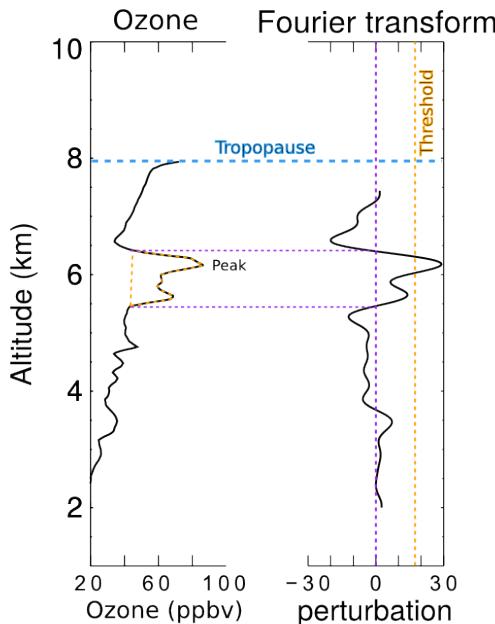


FIGURE 4.4: An example of the STT identification and flux estimation methods used in this work. The left panel shows an ozone profile from Melbourne on 8 January 2004 from 2 km to the tropopause (blue dashed horizontal line). The right panel shows the perturbation profile created from bandpass filtering 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.

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 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.

#### 4.3.4 Biomass burning influence

The STT detection algorithm described in Sect. 4.3.3 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 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.

#### 4.3.5 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. 4.4). 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).

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

## 4.4 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 based on the meteorological classification described in Sect. 4.3.5, 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. 4.3.4) 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

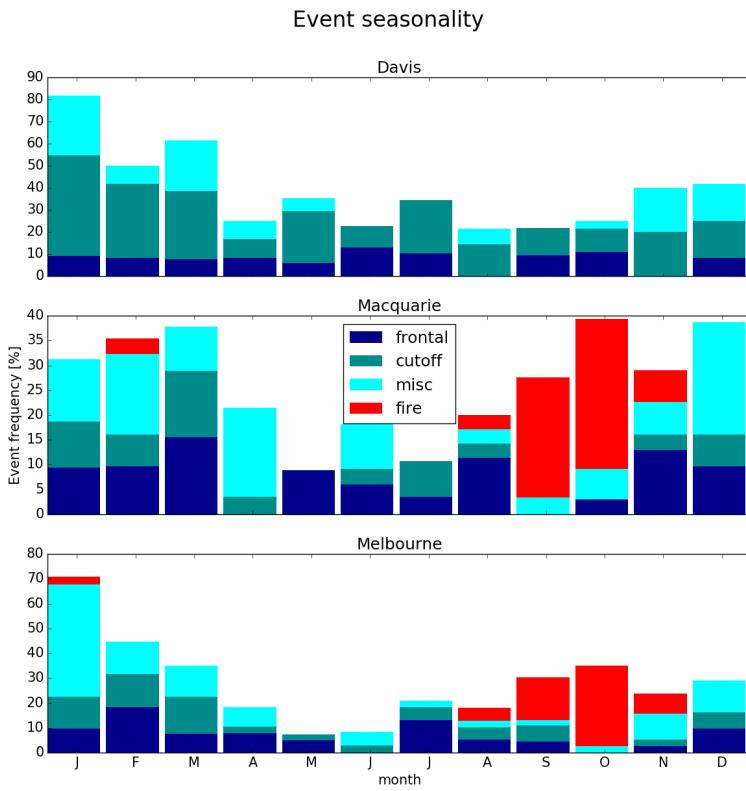


FIGURE 4.5: Seasonal cycle of STT event frequency at Davis (top), Macquarie Island (middle), and Melbourne (bottom). Events are categorised by associated meteorological conditions as described in the text, with low pressure fronts (“frontal”) in dark blue, cut-off low pressure systems (“cutoff”) in teal,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

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 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

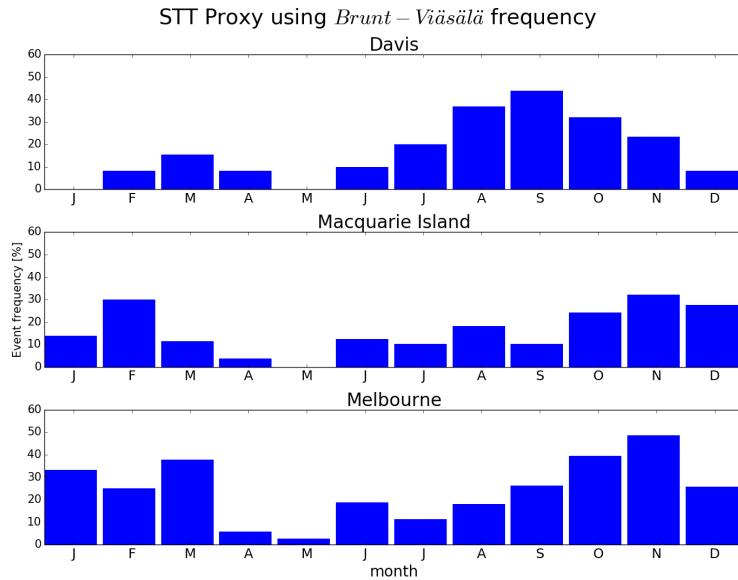


FIGURE 4.6: Seasonal distribution of STT events using the alternative STT proxy, obtained from consideration of the static stability at the ozone and lapse rate tropopauses, for Davis (2006-2013), Macquarie Island (2004-2013), and Melbourne (2004-2013).

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 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

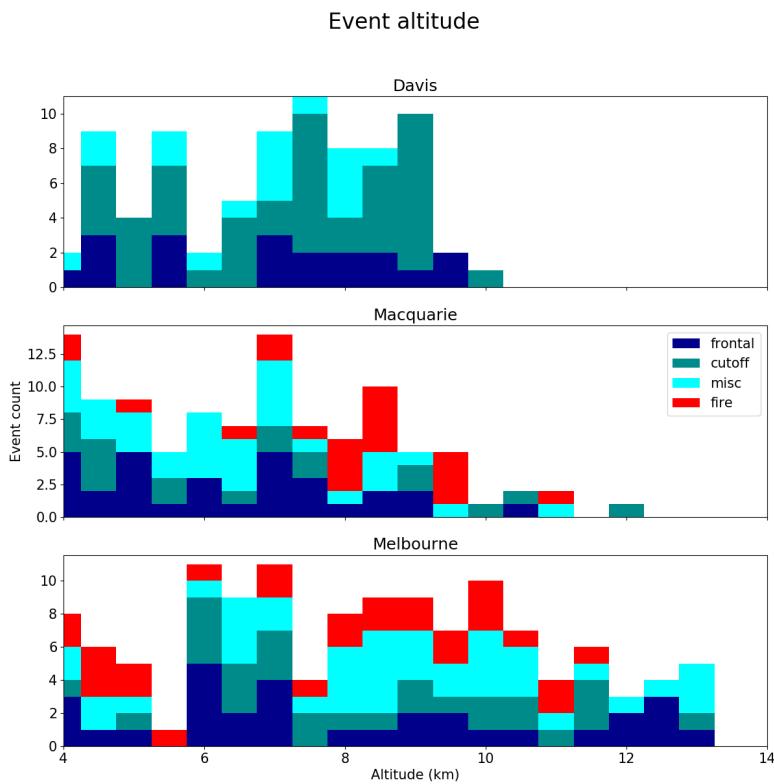


FIGURE 4.7: The distribution of STT events' altitudes at Davis (top), Macquarie Island (middle), and Melbourne (bottom), determined as described in the text. Events are coloured as described in Fig. 4.5.

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

## 4.5 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.

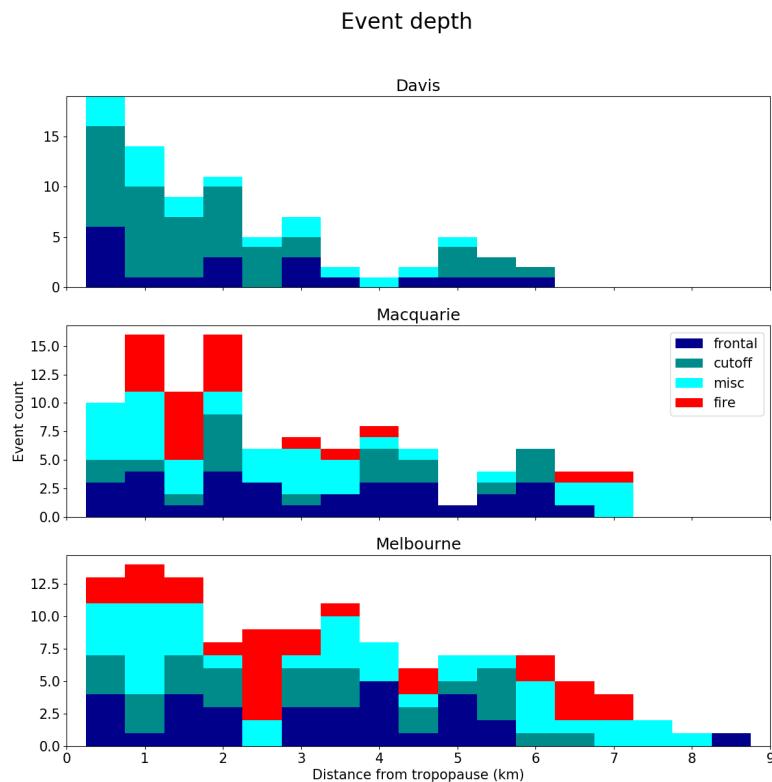


FIGURE 4.8: The distribution of STT events' depths, defined as the distance from the event to the tropopause, at Davis (top), Macquarie Island (middle), and Melbourne (bottom), determined as described in the text.

Events are coloured as described in Fig. 4.5.

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 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

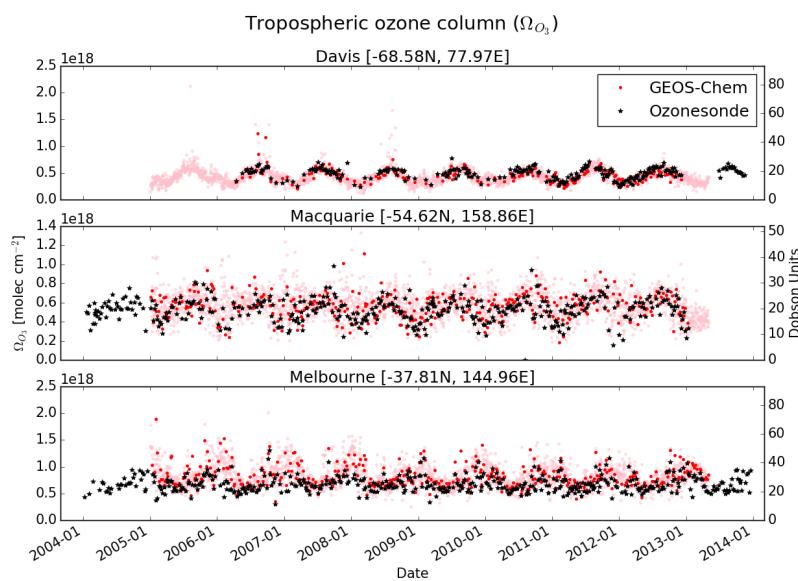


FIGURE 4.9: Comparison between observed (black) and simulated (pink, red) tropospheric ozone columns ( $\Omega_{O_3}$ , in molecules  $\text{cm}^{-2}$ ) from 1 January 2004 to 30 April 2013. For the model, daily output is shown in pink, while output from days with ozonesonde measurements are shown in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

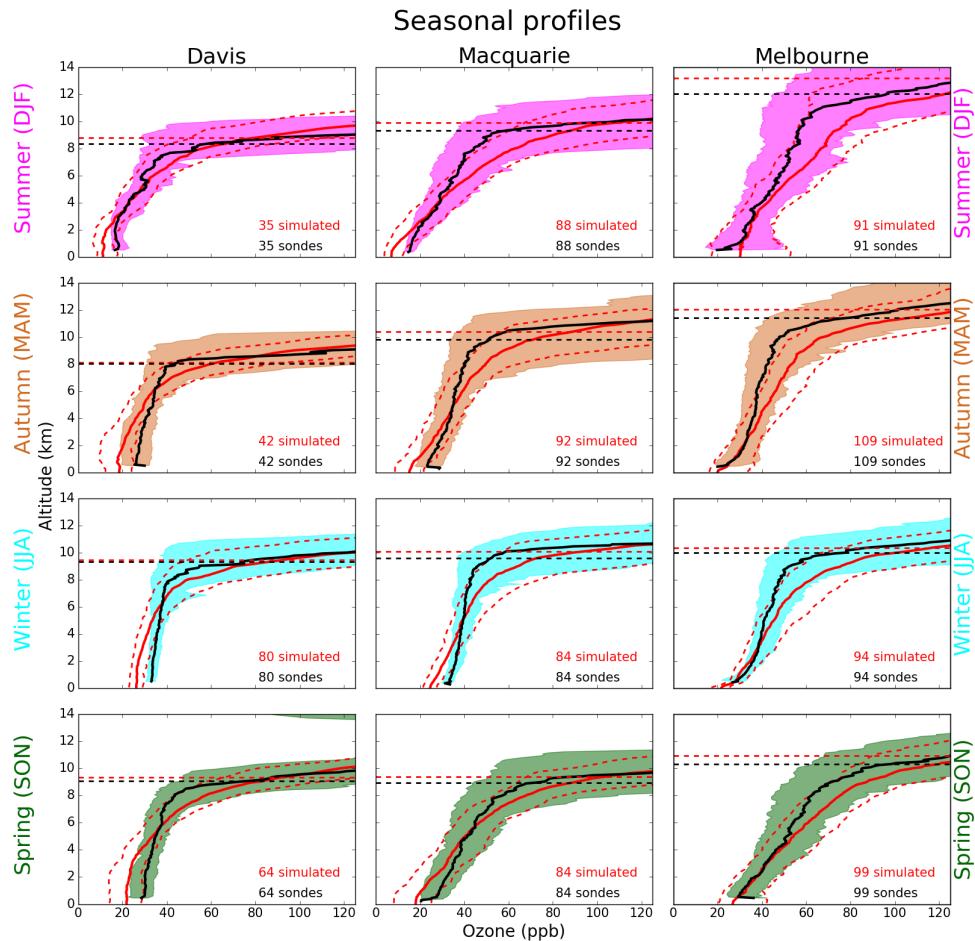


FIGURE 4.10: Observed and simulated tropospheric ozone profiles over Davis, Macquarie Island, and Melbourne, averaged seasonally. Model medians (2005-2013 average) are shown as red solid lines, with red dashed lines showing the 10th and 90th percentiles. Ozonesonde medians (over each season, for all 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).

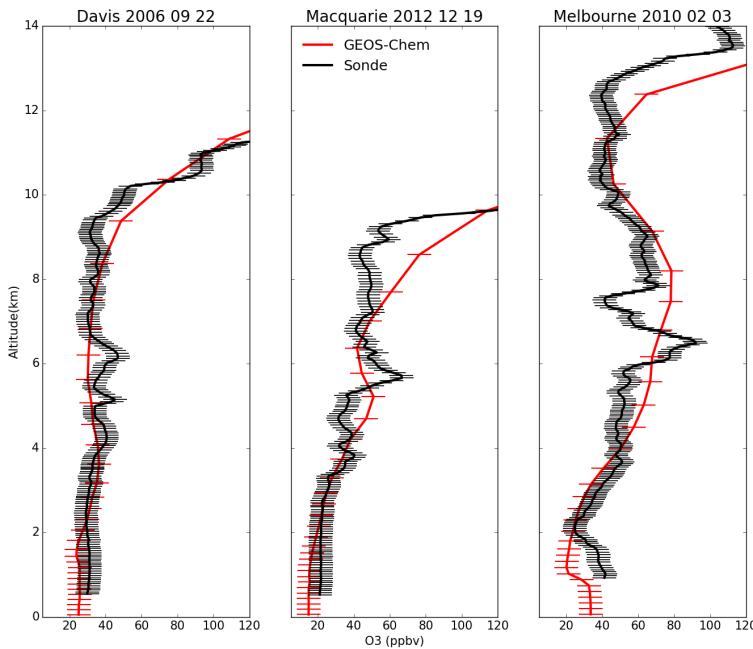


FIGURE 4.11: Example comparisons of ozone profiles from ozonesondes (black) and GEOS-Chem (red) from three different dates during which STT events were detected from the measurements. The dates were picked based on subjective visual analysis. The examples show the best match between model and observations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

(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<sub>3</sub> profiles is steeper in the measurements than the model, at all sites during all seasons. Recently 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.

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

### 4.6.1 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. 4.3.3). 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 flux<sub>*i*</sub> 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.

## 4.6.2 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. 4.7.2). 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  $3.3 \text{ Tg a}^{-1}$ .

## 4.6.3 Comparison to literature

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

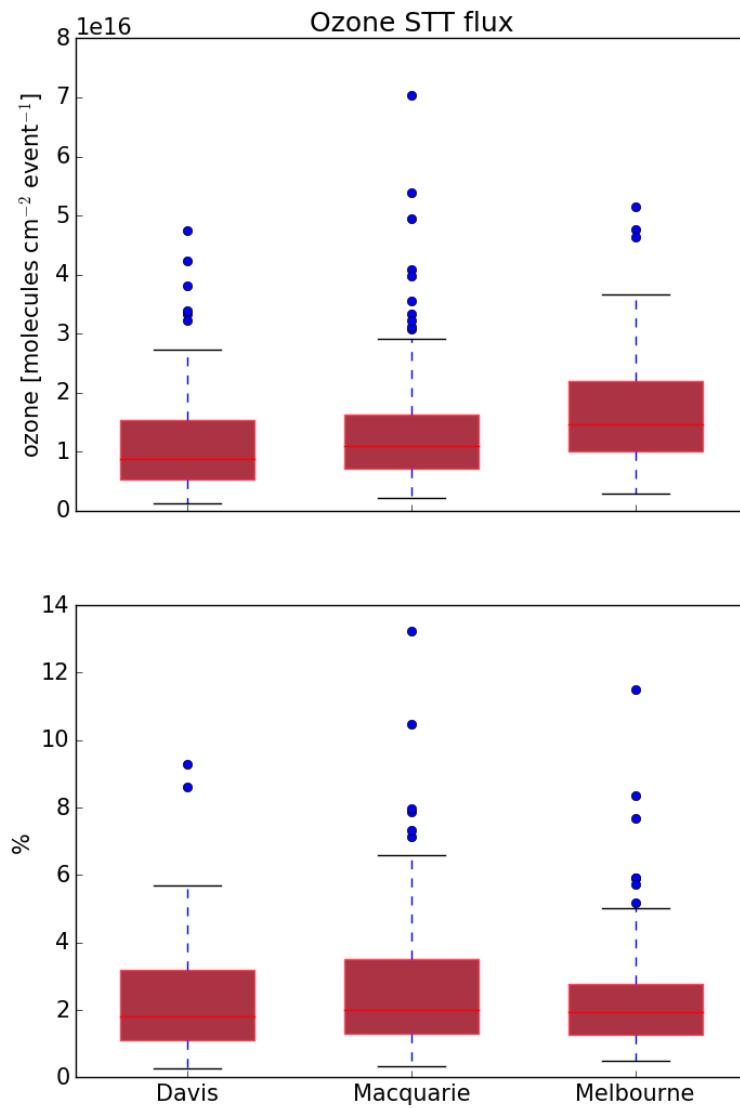


FIGURE 4.12: Top panel: tropospheric ozone attributed to STT events. Bottom panel: percent of total tropospheric column ozone attributed to STT events. Boxes show the inter-quartile range (IQR), with the centre line being the median, whiskers show the minimum and maximum, circles show values which lie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

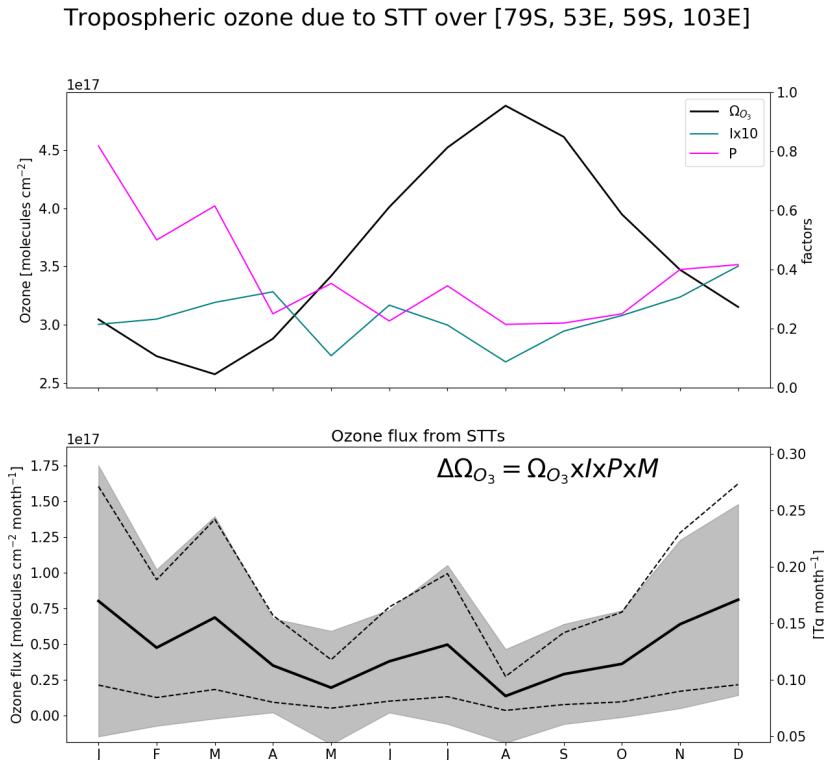


FIGURE 4.13: (Top) Tropospheric ozone, ( $I$ )impact per event, and ( $P$ )probability of event detection per sonde launch, averaged over the region above Davis. The tropospheric ozone column  $\Omega_{O_3}$  (black, left axis) is from GEOS-Chem, while the STT probability  $P$ (magenta, right axis) and 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. 4.7. 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.

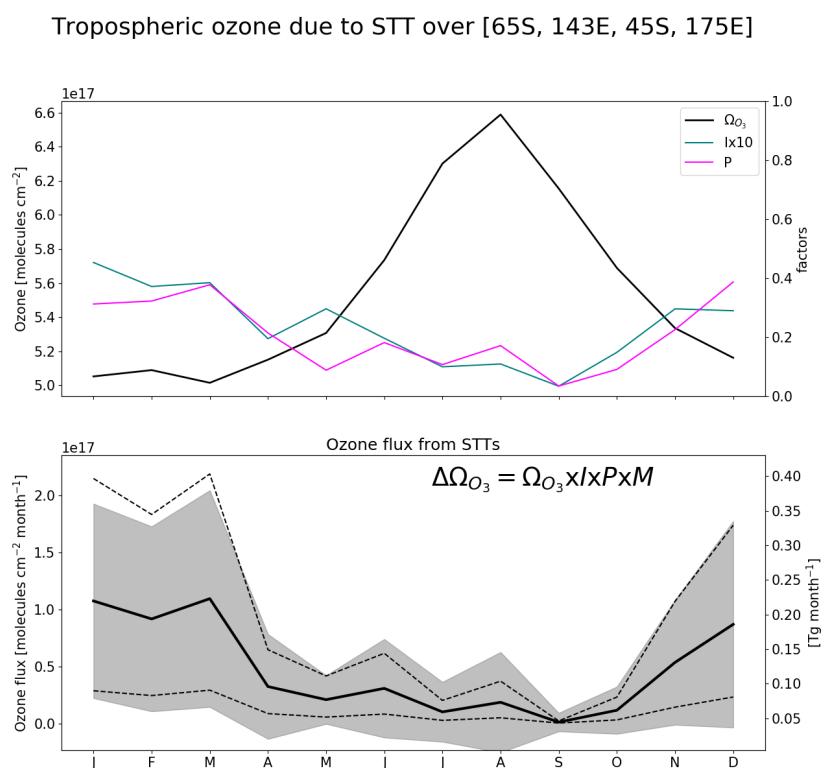


FIGURE 4.14: As described in 4.13, for the region containing Macquarie Island.

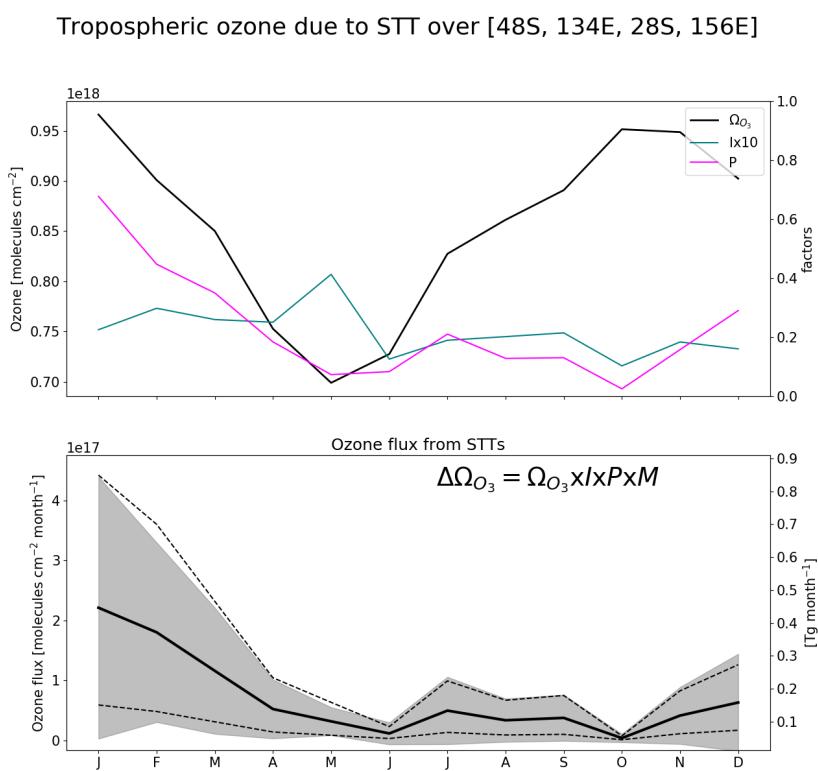


FIGURE 4.15: As described in 4.13, for the region containing Melbourne.

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%)

(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, 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° S and 60° S is 210 Tg yr<sup>-1</sup>, 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° S and 45° 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. 4.4, 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).

## 4.7 Sensitivities and limitations

### 4.7.1 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. 4.3.3, 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.

#### 4.7.2 Flux calculations

Flux is calculated as  $I \times P \times M \times \Omega_{O_3}$ , with each term calculated as described in Sect. 4.6.1. 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. 4.7, 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.

## 4.8 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  $300\text{-}570 \text{ kg km}^{-2} \text{ a}^{-1}$ , 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.

## 4.9 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.



## Chapter 5

# Conclusions

### 5.1 Aims

**In this thesis I aim to improve understanding of natural contributions to ozone over Australia and the southern ocean.** Chapter 1 **Introduction and Literature Review** examines processes leading to ozone creation in the troposphere, primarily isoprene emissions and subsequent chemistry. A summary of how the lack of information available over Australia affects modelling and forecasting is provided. The secondary cause of ozone, stratosphere to troposphere transport (STT), is also outlined with associated processes and causes discussed. Methodologies, tools and data-sets used throughout the thesis are detailed in Chapter 2 **Data and Modelling**. Chapter 3 **Biogenic Isoprene emissions in Australia** shows how models are misrepresenting isoprene emissions by a large margin in Australia, along with the description and implementation of a relatively simple method of improvement. Chapter 4 **Stratospheric ozone intrusions** provides an estimate of stratospheric ozone influx to the troposphere, along with potential classifications and seasonality.

**I aim to recalculate satellite vertical columns of HCHO using updated model a priori information.** HCHO from OMI on board the AURA satellite is examined and recalculated in Chapter 2, where influence of the air mass factor (AMF) is discussed in detail. This AMF was created using an older version of GEOS-Chem with out-of-date HCHO chemistry. Recalculation, binning, and analysis of the satellite HCHO vertical columns is performed using GEOS-Chem v10.01, outlined in Section 2.6.1. This is performed using my own partial AMF recalculation code, along with full recalculation code set up in collaboration with Prof. Palmer and Dr. Surl. Subsequent use of the data makes further use of the recalculation as a basic sensitivity analysis of top-down emissions estimations to a priori satellite information.

**I aim to determine biogenic isoprene emissions in Australia using a top-down inversion of satellite HCHO, through an estimated yield from GEOS-Chem.** In Chapter 3 I determine the linear relationship between total column HCHO and biogenic isoprene emissions over Australia using the global CTM GEOS-Chem. Applying this relationship to satellite based HCHO measurements creates the desired top-down isoprene emissions estimate. This process is described in Section 3.2, and requires intensive filtering of both satellite data (to exclude non-biogenic HCHO sources) and model yield (to minimise spatial smearing). The uncertainty and limitations of top-down estimates due to satellite and model uncertainty, along with temporal and horizontal resolution of available data is examined in Section 3.4. Furthermore, this top-down estimation is used to scale a new simulation of HCHO, and O<sub>3</sub> over Australia,

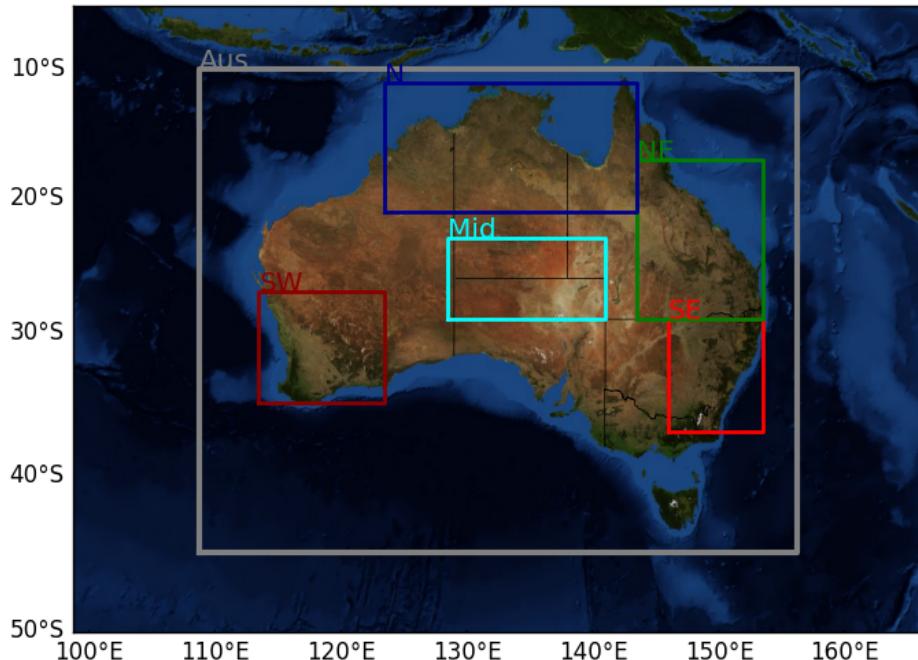


FIGURE 5.1: Sub-regions used in subsequent figures. Australia-wide averages will be black or grey, while results from within the coloured rectangles will match the colours shown here.

described in Section 3.2.8.

**I aim to improve understanding of ozone transported to the troposphere from the stratosphere in Australia and the southern ocean.** In Chapter 4 the seasonal cycle of STT events is characterised, and their contribution to the SH extra-tropical tropospheric ozone budget is quantified using GEOS-Chem to estimate ozone flux extrapolated from three measurement sites. Causal climatology and event seasonality are examined in Section 4.4. STT detection frequencies and modelled tropospheric ozone columns are used to estimate STT ozone flux near three sites in Section 4.6. Findings are compared against relevant literature, and the uncertainties involved in STT event detection and ozone flux estimation are studied.

In the following sections, summaries of results over Australia are often split into regions shown in Figure 5.1.

## 5.2 Isoprene emissions

Being a major driver of continental boundary ozone production, a new isoprene emissions estimate is created using OMI satellite measurements of HCHO. This estimate is and compared against the bottom-up estimate from GEOS-Chem (running MEGAN).

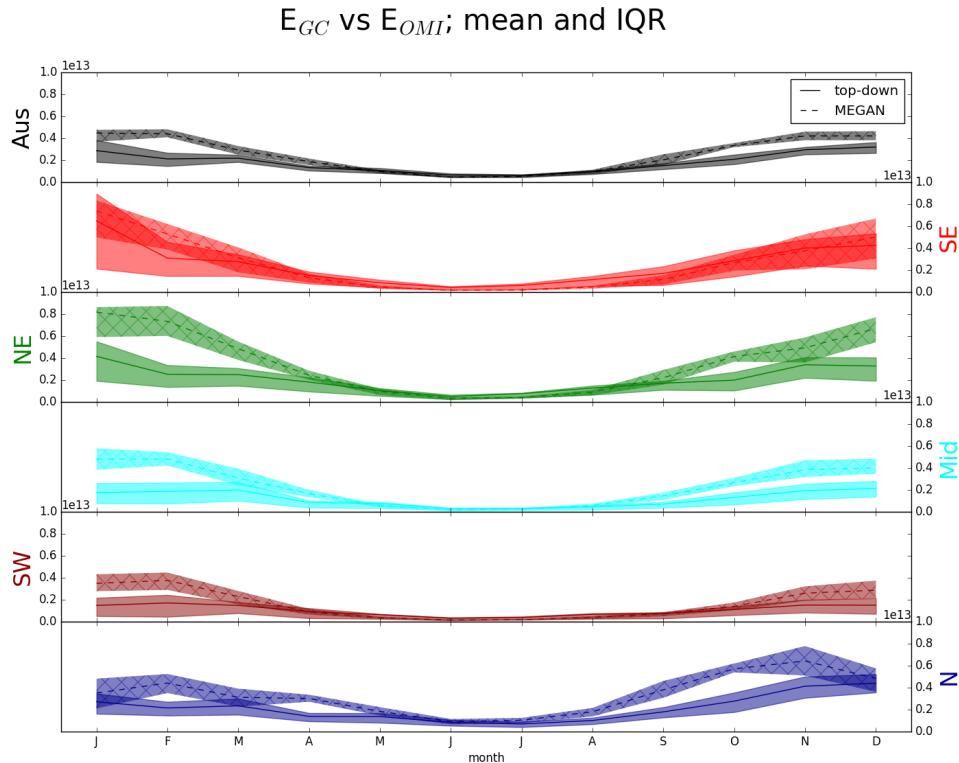


FIGURE 5.2: The multi-year monthly mean (lines) and IQR (shaded) of midday (13:00-14:00 LT) isoprene emissions estimates. Estimates come from MEGAN run by GEOS-Chem ( $E_{GC}$ ), and the OMI top-down technique ( $E_{OMI}$ ). The mean  $E_{GC}$  is shown by the dashed lines and hatched shaded areas show the IQR.  $E_{OMI}$  means are shown using the solid lines, with IQR shown by unhatched shaded areas. Colours denote the region over which the monthly average was taken, as shown in Figure 5.1.

Figure 5.2 shows these emissions compared regionally. Generally months outside of May to August show the a posteriori lower than the a priori, except in the south eastern portion of Australia where only February is notably reduced. MEGAN estimates of isoprene emissions appear to be largely overestimated in Austral summer, suggesting poorly understood emission factors for Australian forests.

### 5.3 Ozone over Australia

Ozone production in the troposphere is a complex process involving various compounds. Tropospheric ozone is enhanced through VOC chemistry, stratospheric transport, and pollution. The first two of these processes are highly uncertain, with few studies performed on either topic within Australia. Of these ozone sources, VOC chemistry uncertainty is dominated by poor understanding of biogenic emissions (mainly) of isoprene. Emissions of isoprene are globally modelled at  $\sim 465\text{-}500 \text{ Tg C yr}^{-1}$  (Guenther

TABLE 5.1: Isoprene emissions from MEGAN and top-down estimation in  $\text{Tg a}^{-1}$ , along with ozone tropospheric column amounts in  $\text{O}_3 \text{ cm}^{-2} \times 10^{17}$ .

Metric	AUS	SEA	NEA	NA	SWA	MID
MEGAN	43(2)	blah				
Ozone	9.70	11.17	11.03	11.19	11.69	9.09
Top-Down	19(2)					
Ozone	9.64	11.11	10.99	11.12	11.63	9.02

Standard deviations shown in parenthesis.

et al. 2006; Messina et al. 2016). However these appear to be overestimated in Australia, as seen here with top-down estimates at  $\sim 19 \text{ Tg C yr}^{-1}$  compared to  $\sim 43 \text{ Tg C yr}^{-1}$  using GEOS-Chem (which implements MEGAN).

TODO: ozone concentrations using emissions from MEGAN/TOP-Down, also how these compare to estimated STT: Tropospheric ozone production is estimated using GEOS-Chem before and after scaling isoprene emissions based on top-down satellite data. Figure 5.3 shows Australian summer surface level ozone with and without scaling isoprene emissions to match the multi-year top-down estimation. Tropospheric column ozone reduces very minimally ( $< 1\%$ ), however surface ozone drops  $\sim 5\%$ . This suggests that impacts from reduced isoprene emission do not propagate vertically. The results are tabulated for each of five regions, and all of Australia, and shown here in Table 5.1. Although isoprene emissions are halved on average over Australia, the reduction has only a small effect on ozone.

Ozone transported from the stratosphere also may have some affect on surface ozone levels, although this is hard to detect with the methods used in this thesis. We only compare the affect of the two disparate influences on the tropospheric ozone column. STT analysis over Melbourne suggests up to  $\sim 10\%$  of the tropospheric column is due to stratospheric influx, with the average increase caused by influx being  $\sim 1$  to  $3.5\%$ . We see that a relatively large drop in isoprene emissions causes  $\sim 1\%$  reduced ozone concentration over summer. Both of these sources become more important in summer, with more frequent STT and stronger isoprene emissions. However the impacts are separated by altitude with most STT occurring in the free troposphere, while isoprene emission impacts are largely seen at the surface.

TODO: how isoprene is linked to ozone in modelling Figure 5.4 shows surface level (up to  $\sim 150 \text{ m}$  altitude) ozone concentration over 2005 before and after scaling modelled isoprene emissions. Reducing isoprene emissions lowers surface ozone concentrations by TODO: XX to YY % in summer, and XX to YY % in winter. The direct correlation between reduced emissions and surface ozone (Figure TODO) is more or less clear monthly, and using multiple years enhances/reduces the relationship.

## 5.4 Outputs and future work

Analysis of STT along with an estimate of ozone flux from the stratosphere over a portion of Australia and the southern ocean is published in Greenslade2017 Ozone

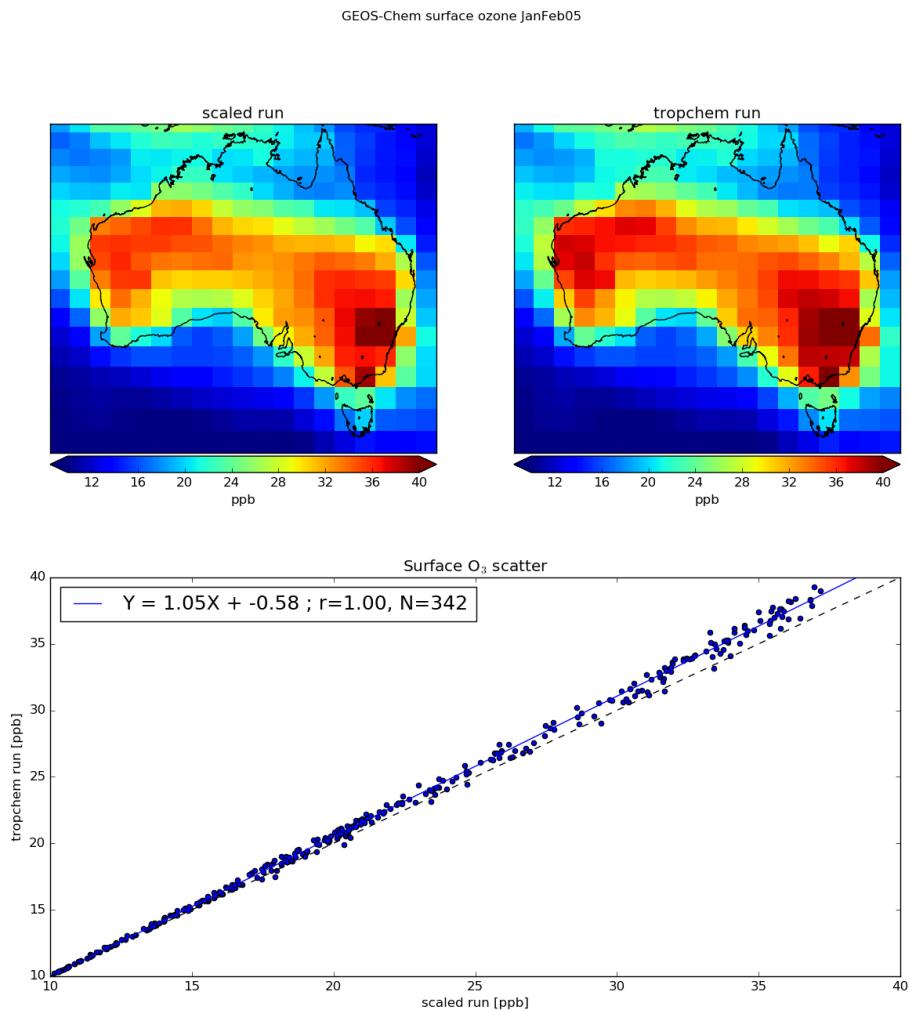


FIGURE 5.3: Ozone maps before (left) and after (right) scaling isoprene emissions in GEOS-Chem for summer. The bottom panel shows the linear regression between the runs along with a black dashed line representing the 1-1 ratio.

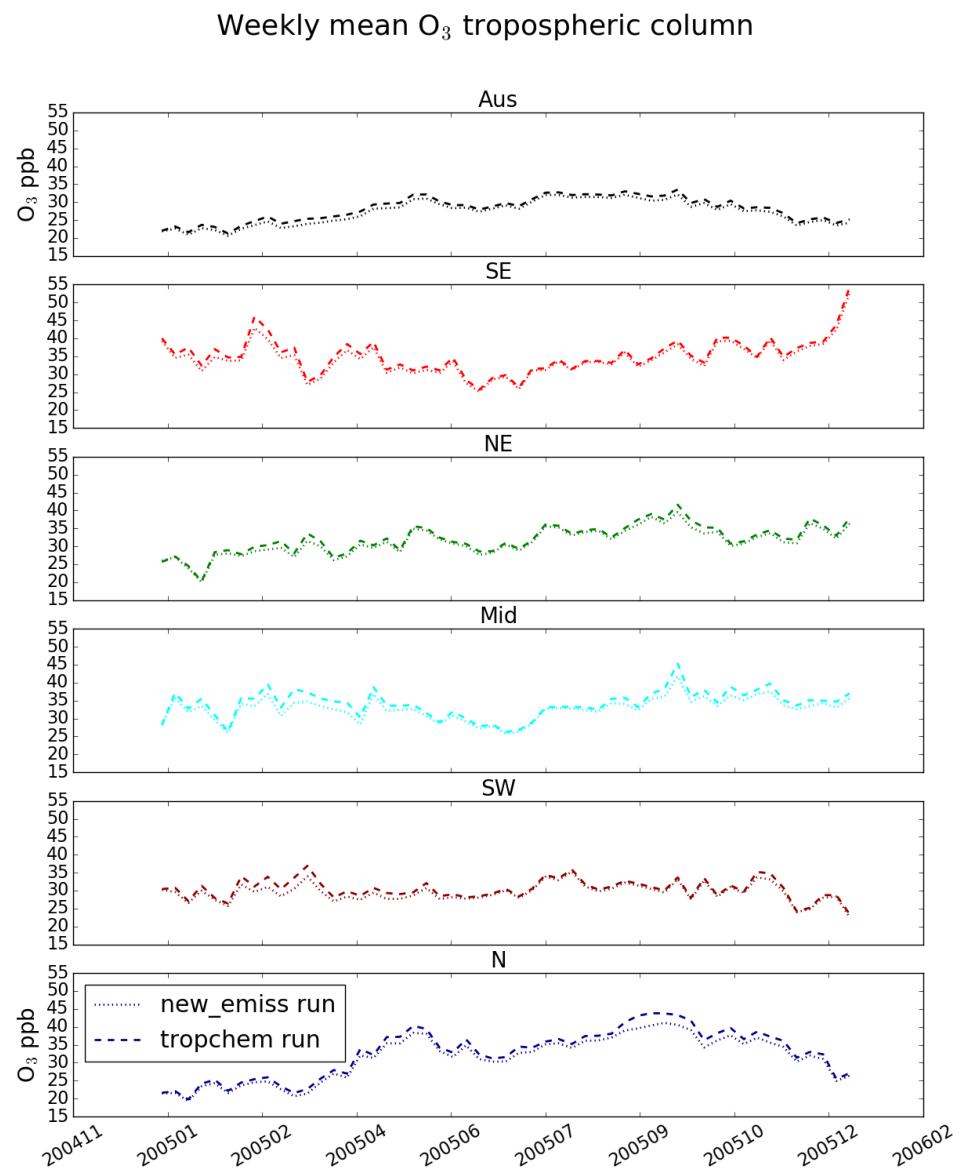


FIGURE 5.4: Surface ozone concentrations (ppb) per region over 2005.

production using GEOS-Chem with updated top-down isoprene emissions is examined in this thesis and may form part of a new publication in the near future with the aim of updating MEGAN isoprene emission factors in Australia.

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). These moisture states are disabled in GEOS-Chem V10.01. Improved isoprene emissions modelling requires this soil moisture problem to be handled. Simply enabling the parameter in its current form is not quite good enough for Australia, due to both the unknown soil moisture and the poorly understood plant responses in this country. TODO: cite the paper which could update MEGAN soil moisture parameterisation.

Using satellite data to improve isoprene emission estimates such as done in this thesis must be used to analyse model improvements, since fully independent measurements are lacking. Measurements of isoprene emissions in Australia would provide a valuable opportunity to verify improvements in emission estimates. Emissions measurement would be greatly valuable however they remain expensive and difficult, especially over the large sparse environment which makes up the Australia outback.



# 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*.

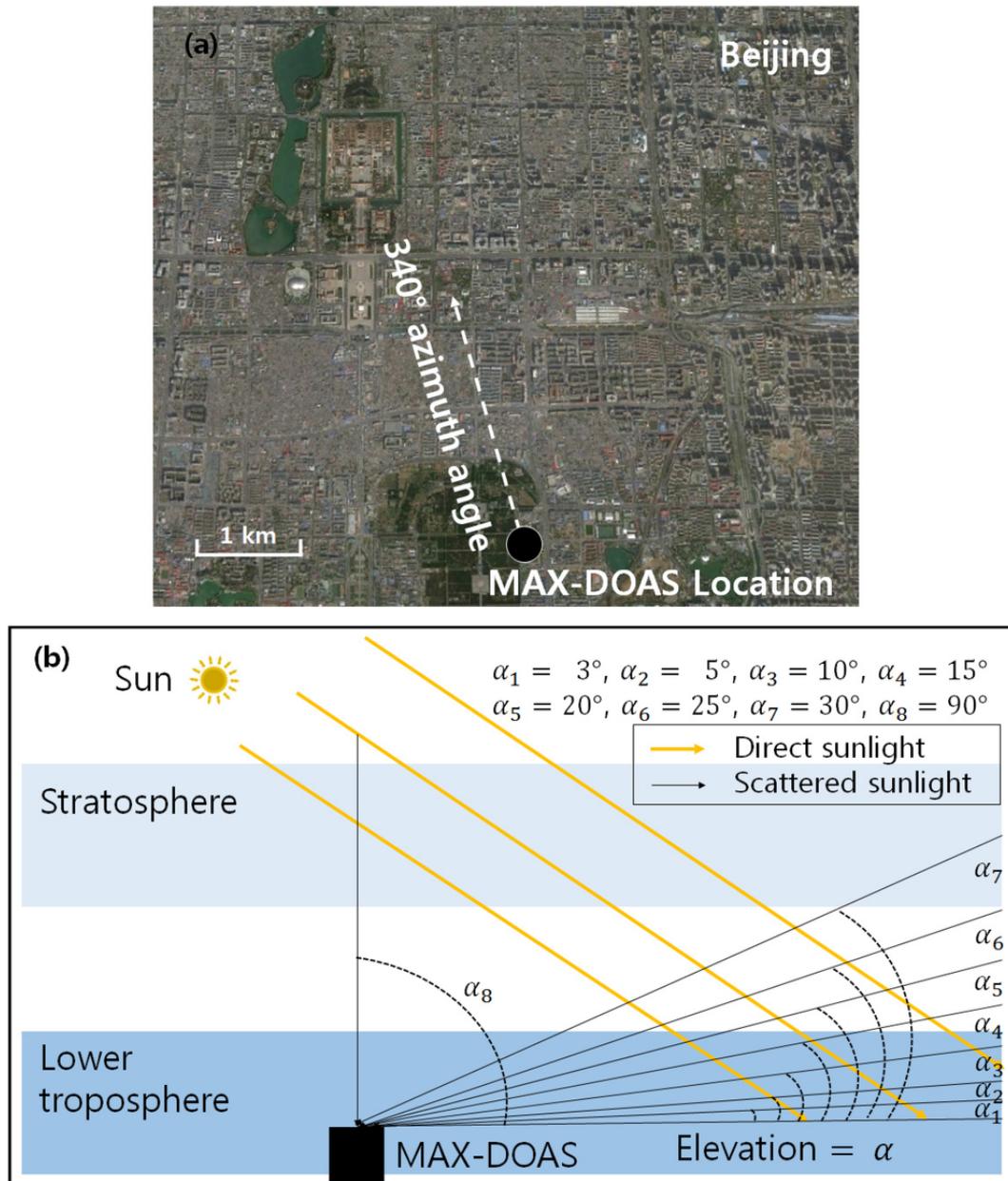


FIGURE A.1: Image from Lee et al. (2015).

### 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

Izci and Hosgün (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 Kanakidou2005) 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.

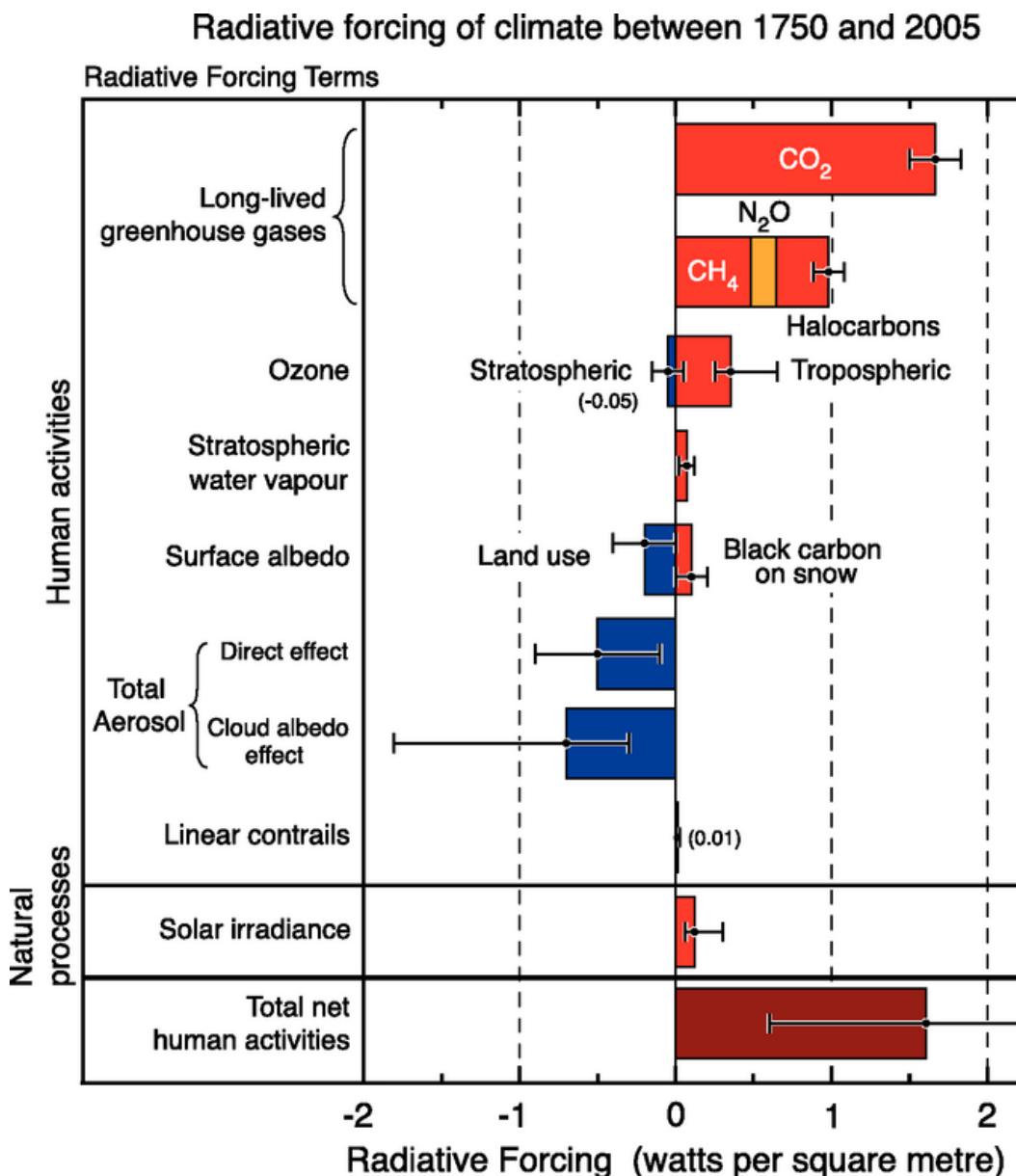


FIGURE A.2: The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from Forster et al. 2007, found at [https://www.ipcc.ch/publications\\_and\\_data/ar4/wg1/en/faq-2-1.html](https://www.ipcc.ch/publications_and_data/ar4/wg1/en/faq-2-1.html).

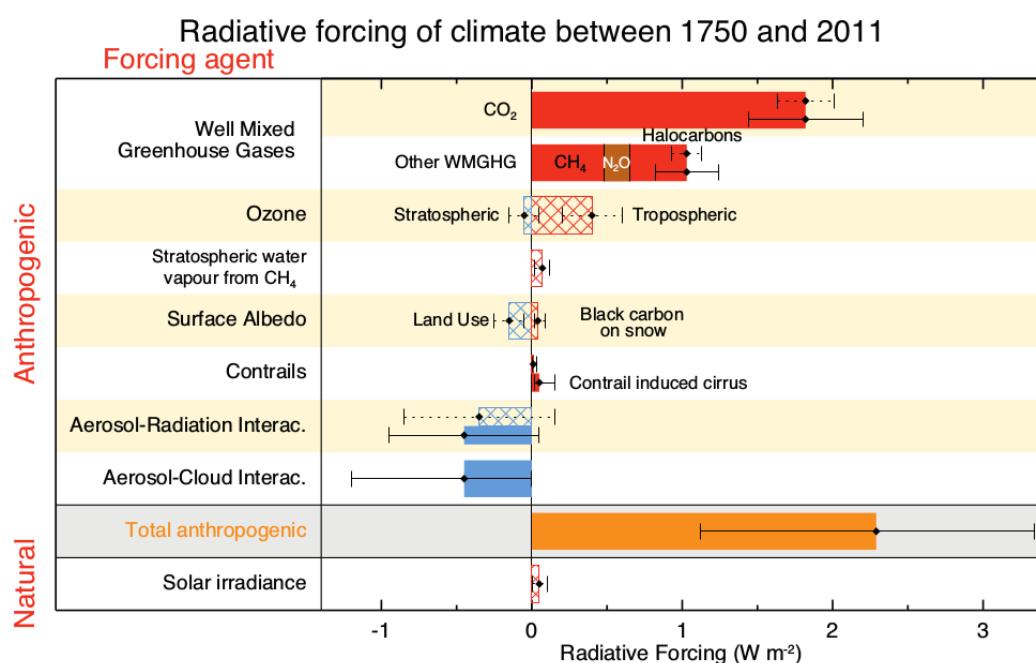


FIGURE A.3: The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 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 (SAPPHO) 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			
NO <sub>X</sub> (molec cm <sup>-3</sup> )	low	low	high
O <sub>3</sub> (molec cm <sup>-3</sup> )	mid	mid	low
HO <sub>X</sub> (molec cm <sup>-3</sup> )			
Influx			
NO <sub>X</sub> (molec cm <sup>-3</sup> s <sup>-1</sup> )	low	low	high
O <sub>3</sub> (molec cm <sup>-3</sup> s <sup>-1</sup> )	mid	mid	low
HO <sub>X</sub> (molec cm <sup>-3</sup> s <sup>-1</sup> )			
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 (molec cm<sup>-2</sup> s<sup>-1</sup>),  $D_{inj}$  is the duration of injection (s),  $H_{PBL}$  is the boundary layer height (cm), and  $N_A$  is the Air number density (molec cm<sup>-3</sup>  $\approx 2.5e19$ ). 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

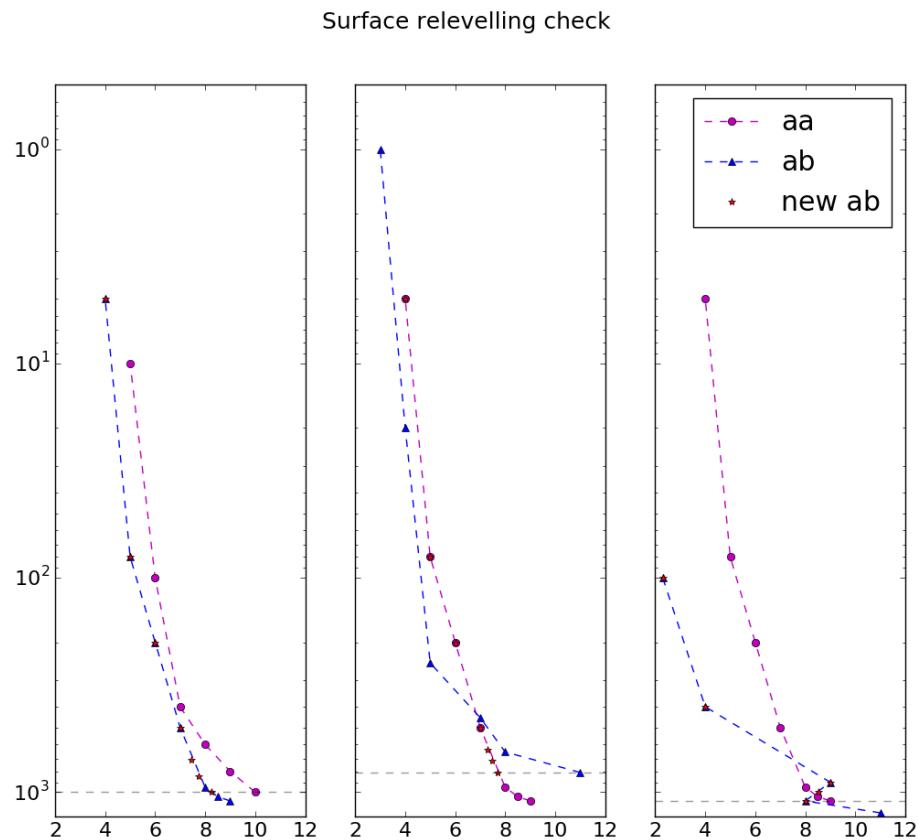


FIGURE A.4: Constructed example of the initial interpolation of OMI's  $\omega$  onto a pressure dimension with mismatched surface pressure.

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.

## A.6 Modelling

### A.6.1 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

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

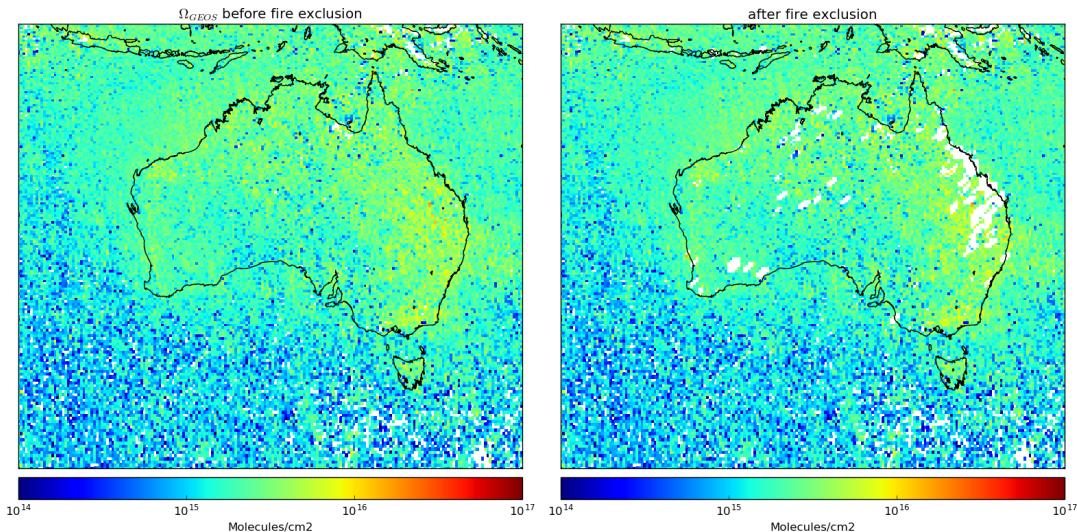


FIGURE A.5: Vertical column HCHO calculated using OMI satellite swaths with GEOS-Chem aprioris, averaged over 1-8 January 2005 with and without fire affected squares removed.

(described in Section ??) TODO: compare scenarios yields and show map of Australia with mapped closest scenario(one colour for each scenario, contourf).

This subsection could analyse nox and LAI correlations to isop,hcho, and ozone over australia?

## A.7 Campaigns

### A.7.1 Daintree

Daintree summary (P. Nelson)

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).

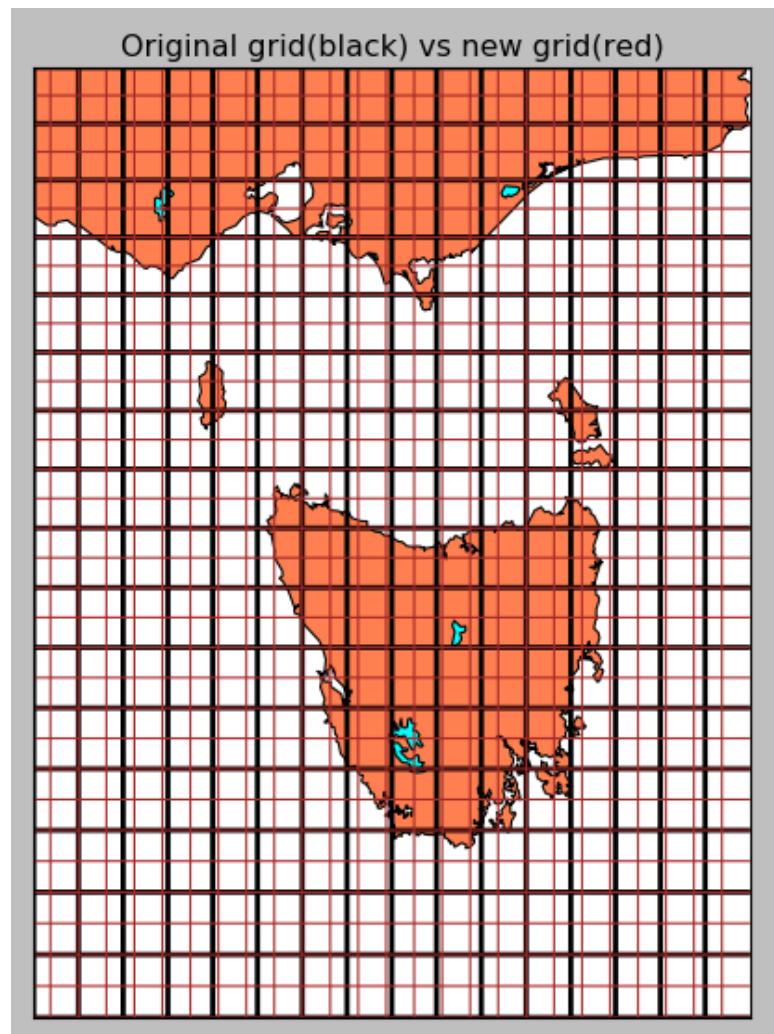


FIGURE A.6: Example of grid space change using  $0.5 \times 0.5$  and  $0.25 \times 0.3125$  latitude by longitude resolution.

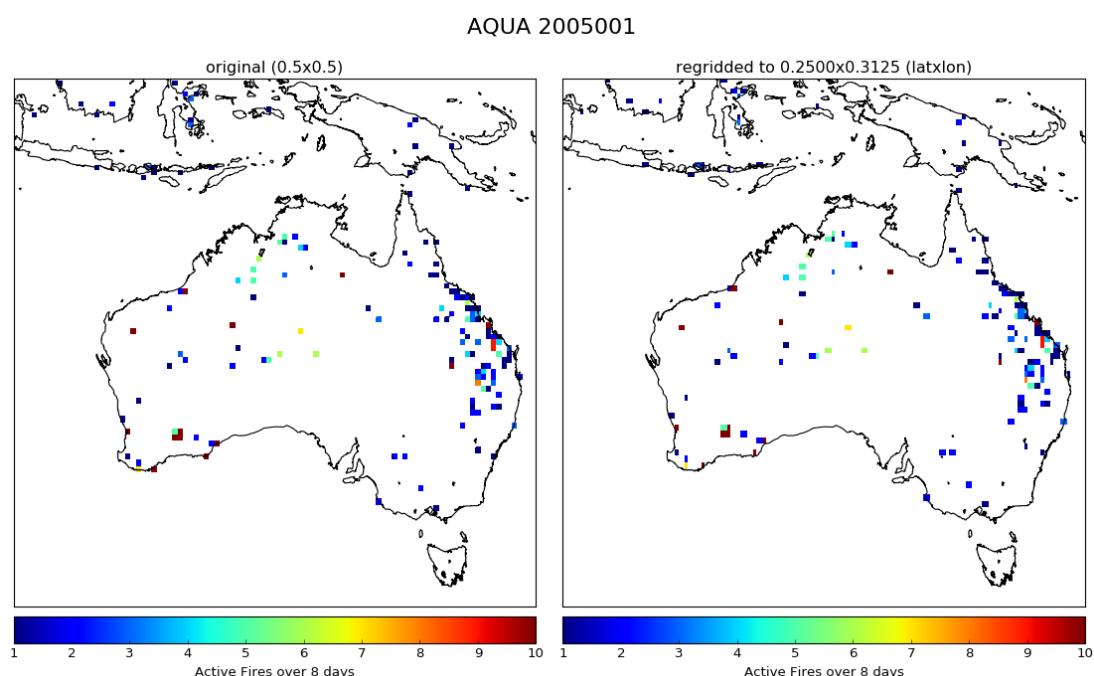


FIGURE A.7: Example of MODIS 8 day grid interpolation from  $0.5 \times 0.5$  to  $0.25 \times 0.3125$  latitude by longitude resolution. This example uses MODIS fire counts for 1-8 January 2005.

## 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).
- 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).
- 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/>.
- 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/>.
- Zeng, Guang et al. (2017). "Attribution of recent ozone changes in the Southern Hemisphere mid-latitudes using statistical analysis and chemistry-climate model simulations". In: *Atmospheric Chemistry and Physics* 17.17, pp. 10495–10513. ISSN: 16807324. DOI: [10.5194/acp-17-10495-2017](https://doi.org/10.5194/acp-17-10495-2017).
- 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://dx.doi.org/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 Maspoch, 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 Himalayas". 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>.
- Christian, Kenneth E. et al. (2018). "Global sensitivity analysis of GEOS-Chem modeled ozone and hydrogen oxides during the INTEX campaigns". In: *Atmospheric Chemistry and Physics* 18.4, pp. 2443–2460. ISSN: 1680-7324. DOI: [10.5194/acp-18-2443-2018](https://doi.org/10.5194/acp-18-2443-2018). URL: <https://www.atmos-chem-phys.net/18/2443/2018/>.
- 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 NO<sub>x</sub> 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](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://www.atmos-chem-phys.net/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://www.elementascience.org/articles/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.2008.02.030](https://doi.org/10.1016/j.atmosenv.2008.02.030).

- 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).
- 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#.VqqEubNM6lM>.
- 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.
- 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).
- 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>.
- 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).
- 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>.

- 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>.
- Jiang, Xiaoyan et al. (2018). "Isoprene emission response to drought and the impact on global atmospheric chemistry". In: *Atmospheric Environment* 183, pp. 69–83. ISSN: 1352-2310. DOI: [10.1016/J.ATMOSENV.2018.01.026](https://doi.org/10.1016/J.ATMOSENV.2018.01.026). URL: <https://www.sciencedirect.com/science/article/pii/S1352231018300402>.
- Izci, Alime and Halit L. Hosgün (2007). "Kinetics of synthesis of isobutyl propionate over amberlyst-15". In: *Turkish Journal of Chemistry* 31.5, pp. 493–499. ISSN: 13000527. DOI: [10.1002/kin](https://doi.org/10.1002/kin).
- 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 hydroperoxyl 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). "HO<sub>x</sub> 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). "HO<sub>x</sub> 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.3d2869ee02fd4f0792f831ac8cbec117>.
- 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).
- 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>.
- Ayers, G. P. et al. (1997). "Formaldehyde production in clean marine air". In: *Geophysical Research Letters* 24.4. DOI: <https://doi.org/10.1029/97GL00123>.
- 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* 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 G. 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>.
- 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/).
- 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).
- 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: <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<sub>2</sub> 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=cafafef03b948fb456696583ed3ab9a5>.
- 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/>.
- 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/>.
- 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.
- Dunne, Erin et al. (2018). "Comparison of VOC measurements made by PTR-MS, adsorbent tubes-GC-FID-MS and DNPH derivatization-HPLC during the Sydney Particle Study, 2012: A contribution to the assessment of uncertainty in routine atmospheric VOC measurements". In: *Atmospheric Measurement Techniques* 11.1, pp. 141–159. ISSN: 18678548. DOI: [10.5194/amt-11-141-2018](https://doi.org/10.5194/amt-11-141-2018). URL: <https://www.atmos-meas-tech.net/11/141/2018/amt-11-141-2018.pdf>.
- 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, HJ and K F Boersma (2003). "Averaging kernels for DOAS total-column satellite retrievals". In: *Atmospheric Chemistry and Physics* 3.1, pp. 1285–1291. ISSN: 1680-7324. DOI: [10.5194/acp-3-1285-2003](https://doi.org/10.5194/acp-3-1285-2003). URL: <http://dx.doi.org/10.5194/acpd-3-895-2003>.
- Lamsal, L N et al. (2014). "Evaluation of OMI operational standard 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/).

- 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. (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).
- 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-2016-306](https://doi.org/10.5194/amt-2016-306). URL: <http://www.atmos-meas-tech-discuss.net/amt-2016-306/>.
- Goldberg, Daniel L. et al. (2017). "A high-resolution and observationally constrained OMI NO<sub>2</sub> satellite retrieval". In: *Atmospheric Chemistry and Physics* 17.18, pp. 11403–11421. ISSN: 16807324. DOI: [10.5194/acp-17-11403-2017](https://doi.org/10.5194/acp-17-11403-2017).
- 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).
- 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).
- 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>.
- Paton-Walsh, C et al. (2017). "The MUMBA campaign: measurements of urban, marine and biogenic air Clare". In: *Earth system science data*, pp. 349–362.
- Guérette, Elise-Andrée et al. (2018). "Emissions of trace gases from Australian temperate forest fires: emission factors and dependence on modified combustion efficiency". In: *Atmospheric Chemistry and Physics* 18.5, pp. 3717–3735. ISSN: 1680-7324. DOI: [10.5194/acp-18-3717-2018](https://doi.org/10.5194/acp-18-3717-2018). URL: <https://www.atmos-chem-phys.net/18/3717/2018/>.
- Cheng, M et al. "Factors controlling volatile organic compounds in dwellings in Melbourne, Australia". In: *Indoor Air* 26.2, pp. 219–230. DOI: [10.1111/ina.12201](https://doi.org/10.1111/ina.12201). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/ina.12201>.
- Lawson, S. J. et al. (2015). "Seasonal in situ observations of glyoxal and methylglyoxal over the temperate oceans of the Southern Hemisphere". In: *Atmospheric Chemistry and Physics* 15.1, pp. 223–240. ISSN: 1680-7324. DOI: [10.5194/acp-15-223-2015](https://doi.org/10.5194/acp-15-223-2015). URL: <http://www.atmos-chem-phys.net/15/223/2015/>.
- ABS (2013). *Melbourne population*. URL: <http://www.abs.gov.au/ausstats/abs@.nsf?Previousproducts/3218.0>MainFeatures52012-13?opendocument&tabname=Summary&prodno=3218.0&issue=2012-13&num=&view=> (visited on 05/10/2017).

- Schenkeveld, V. M. Erik et al. (2017). "In-flight performance of the Ozone Monitoring Instrument". In: *Atmospheric Measurement Techniques* 10.5, pp. 1957–1986. ISSN: 1867-8548. DOI: [10.5194/amt-10-1957-2017](https://doi.org/10.5194/amt-10-1957-2017). URL: <http://www.atmos-meas-tech-discuss.net/amt-2016-420/> <https://www.atmos-meas-tech.net/10/1957/2017/>.
- 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>.
- 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).
- Instrument, O M I (2002). "OMI Algorithm Theoretical Basis Document Volume I". In: I.August, pp. 1–50.
- Spurr, R.J.D., T.P. Kurosu, and K.V. Chance (2001). "A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval". In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 68.6, pp. 689–735. DOI: [10.1016/S0022-4073\(00\)00055-8](https://doi.org/10.1016/S0022-4073(00)00055-8). URL: <https://www.sciencedirect.com/science/article/pii/S0022407300000558>.
- 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>.
- 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>.
- Chen, D. et al. (2009). "Regional CO pollution in China simulated by the high-resolution nested-grid GEOS-Chem model". In: *Atmospheric Chemistry and Physics Discussions*

- 9.2, pp. 5853–5887. ISSN: 1680-7324. DOI: [10.5194/acpd-9-5853-2009](https://doi.org/10.5194/acpd-9-5853-2009). URL: <http://dx.doi.org/10.5194/acp-9-3825-2009>.
- 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).
- Wuebbles, Donald J and Katharine Hayhoe (2002). “Atmospheric methane and global change”. In: *Earth-Science Reviews* 57.3-4, pp. 177–210. ISSN: 0012-8252. DOI: [10.1016/S0012-8252\(01\)00062-9](https://doi.org/10.1016/S0012-8252(01)00062-9). URL: <https://www.sciencedirect.com/science/article/pii/S0012825201000629?via%3Dihub>.
- 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>.
- 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*.
- 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>.
- 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>.
- 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>.
- Carlisle, Wendy (2012). *The coal seam gas rush*. URL: <http://www.abc.net.au/news/2011-11-24/coal-seam-gas-by-the-numbers-map/3664318>.
- Zheng, Y. et al. (2015). "Relationships between photosynthesis and formaldehyde as a probe of isoprene emission". In: *Atmospheric Chemistry and Physics* 15.15, pp. 8559–8576. ISSN: 16807324. DOI: [10.5194/acp-15-8559-2015](https://doi.org/10.5194/acp-15-8559-2015).
- 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>.
- 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).
- 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>.
- 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>.
- Streets, David G et al. (2013). "Emissions estimation from satellite retrievals: A review of current capability". In: *Atmospheric Environment* 77, pp. 1011–1042. DOI: [10.1016/j.atmosenv.2013.05.051](https://doi.org/10.1016/j.atmosenv.2013.05.051). URL: <http://dx.doi.org/10.1016/j.atmosenv.2013.05.051>.
- Kaden DA, Mandin C, Nielsen GD, et al. (2010). *HCHO exposure guidelines*. URL: <https://www.ncbi.nlm.nih.gov/books/NBK138711/> (visited on 10/10/2018).
- 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).
- 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.org/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.org/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).
- 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.org/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>.
- 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>.
- 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).