



## Thesis Title

Jesse Greenslade

Supervisor:  
Dr. Jenny A. Fisher  
Co-supervisor:  
Dr. Clare Paton-Walsh

*This thesis is presented as required for the conferral of the degree:*

Doctor of philosophy

The University of Wollongong  
School of Chemistry

January 1901

## **Declaration**

*I, Jesse Greenslade, declare that this thesis submitted in fulfilment of the requirements for the conferral of the degree Doctor of philosophy, from the University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. This document has not been submitted for qualifications at any other academic institution.*

---

**Jesse Greenslade**

*September 26, 2017*

# **Abstract**

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

# Acknowledgments

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 his noodly appendages.

# Contents

<b>Abstract</b>	iii
<b>1 Outline</b>	1
1.1 Overview . . . . .	1
<b>2 Introduction and Literature Review</b>	3
2.1 Tropospheric ozone and air quality . . . . .	3
2.1.1 Air Quality . . . . .	3
2.1.2 Ozone transported from the stratosphere . . . . .	8
2.1.3 Biogenic precursors to ozone . . . . .	9
2.2 Isoprene and other VOCs . . . . .	10
2.2.1 What are VOCs . . . . .	10
2.2.2 Emissions estimates . . . . .	12
2.2.3 What do they Do? . . . . .	14
2.2.4 Where do they go? . . . . .	16
2.2.5 The isoprene cascade . . . . .	16
2.2.6 How and where do we measure them? . . . . .	22
2.2.7 How do we determine emissions? . . . . .	23
2.3 Hydroxyl (OH) and other radicals (maybe relocate this bit?) . . . . .	26
2.4 Formaldehyde (HCHO) . . . . .	27
2.4.1 How HCHO is measured . . . . .	27
2.4.2 Sources and sinks . . . . .	30
2.4.3 Determining isoprene from HCHO . . . . .	31
2.4.4 Satellite HCHO detection . . . . .	31
2.4.5 Glyoxal TODO: move somewhere fitting? . . . . .	36
2.5 Models . . . . .	37
2.5.1 How can models help . . . . .	37
2.5.2 Relevant model frameworks . . . . .	38
2.5.3 Using models to determine Isoprene emissions . . . . .	41
2.5.4 Uncertainties . . . . .	42
2.6 Aims . . . . .	44

<b>3 Modelling with GEOS-Chem</b>	<b>45</b>
3.0.1 List of runs and outputs used in my work TODO: good place for this? . . . . .	45
3.1 GEOS-Chem framework . . . . .	45
3.2 GEOS-Chem isoprene mechanisms . . . . .	46
3.2.1 Outline . . . . .	46
3.2.2 Emissions from MEGAN . . . . .	48
3.3 Running GEOS-Chem (before isop?) . . . . .	48
3.3.1 Installation and requirements . . . . .	48
3.3.2 Options . . . . .	49
3.3.3 Tropchem mechanisms . . . . .	49
3.3.4 UCX mechanisms . . . . .	49
3.3.5 Run comparisons . . . . .	49
3.4 Analysing output . . . . .	54
<b>4 Stratospheric ozone intrusions</b>	<b>55</b>
4.1 Introduction . . . . .	55
4.2 Data and Methods . . . . .	57
4.2.1 Ozonesonde record in the Southern Ocean . . . . .	57
4.2.2 Model description . . . . .	60
4.2.3 Characterisation of STT events and associated fluxes . . . . .	62
4.2.4 Biomass burning influence . . . . .	64
4.2.5 Classifying synoptic conditions during STT events . . . . .	65
4.3 STT event climatologies . . . . .	66
4.4 Simulated ozone columns . . . . .	69
4.5 Stratosphere-to-troposphere ozone flux from STT events . . . . .	71
4.5.1 Method . . . . .	71
4.5.2 Results . . . . .	75
4.5.3 Comparison to literature . . . . .	76
4.6 Sensitivities and limitations . . . . .	80
4.6.1 Event detection . . . . .	80
4.6.2 Flux calculations . . . . .	81
4.7 Conclusions . . . . .	82
4.8 Contributions and Acknowledgements . . . . .	84
<b>5 Biogenic Isoprene emissions in Australia</b>	<b>85</b>
5.1 Outline TODO: move into intro? . . . . .	85
5.2 Introduction . . . . .	86
5.3 Data . . . . .	87
5.4 Modelling . . . . .	87

5.4.1	GEOS-Chem simulation . . . . .	87
5.5	Methods . . . . .	87
5.5.1	CAABA/MECCA Box model: isoprene source classifications .	87
5.5.2	Campaigns and datasets . . . . .	89
5.5.3	Recalculation of OMI HCHO . . . . .	89
5.5.4	Calculation of Emissions . . . . .	90
5.5.5	Emisions drivers . . . . .	92
5.5.6	HCHO Products and yield . . . . .	92
5.5.7	Accounting for smearing . . . . .	96
5.6	Results . . . . .	98
5.6.1	Emissions affect on GEOS-Chem . . . . .	98
5.6.2	Emissions comparisons . . . . .	98
<b>6</b>	<b>Formaldehyde product over Australia</b>	<b>100</b>
6.1	Australian Biogenic Volatile Organic Compounds (BVOCs) . . . . .	100
6.1.1	Isoprene, Monoterpenes . . . . .	100
6.1.2	Biomass Burning . . . . .	101
6.1.3	MEGAN . . . . .	101
6.2	Satellite HCHO measurements . . . . .	103
6.2.1	Satellite Retrievals . . . . .	103
6.2.2	OMI Algorithm BOAS . . . . .	104
6.2.3	Optical Depth ( $\tau$ ) . . . . .	105
6.2.4	Scattering . . . . .	106
6.2.5	Absorption cross section and number density . . . . .	106
6.2.6	Air Mass Factors . . . . .	106
6.2.7	OMI HCHO data products . . . . .	108
6.2.8	HCHO Vertical Column Calculation . . . . .	108
6.2.9	Uncertainty in OMI total columns . . . . .	111
6.2.10	Reference sector correction for comparison of products to various models . . . . .	111
6.3	Recalculating HCHO from satellite(OMI) data over Australia . . . . .	112
6.3.1	Process Outline . . . . .	112
6.3.2	Quality filtering OMI HCHO slant columns . . . . .	114
6.3.3	Reading OMHCHO daily slant columns . . . . .	116
6.3.4	Regridding to 0.25 by 0.3125 8-day averaged vertical columns	117
6.3.5	Filtering pyrogenic HCHO . . . . .	118
6.3.6	Filtering anthropogenic HCHO . . . . .	119
6.3.7	Recalculating the AMF to create our own vertical HCHO columns . . . . .	120

6.3.8	AMF code from Paul Palmer . . . . .	123
6.3.9	Determination and application of the pacific ocean reference sector normalisation . . . . .	124
6.3.10	Estimation of error or uncertainty . . . . .	127
6.4	Validation and comparisons . . . . .	128
6.4.1	Comparison with standard OMI product . . . . .	128
6.4.2	Comparison with in-situ measurements . . . . .	128
6.4.3	Summary . . . . .	128
6.4.4	Conclusions . . . . .	128
<b>7</b>	<b>Isoprene Emissions in Australia</b>	<b>129</b>
7.1	GEOS-Chem isoprene mechanisms(moved to Modelling chapter) . . .	129
7.1.1	Outline(moved) . . . . .	129
7.1.2	Emissions from MEGAN(moved) . . . . .	129
7.2	Isoprene emissions estimation . . . . .	129
7.2.1	Outline . . . . .	129
7.2.2	HCHO Products and yield . . . . .	130
7.2.3	CAABA/MECCA yield . . . . .	130
7.2.4	Calculation of Emissions . . . . .	130
7.2.5	Calculation of smearing effect . . . . .	130
7.2.6	Calculations of uncertainty . . . . .	130
7.2.7	Extrapolating the circadian cycle . . . . .	131
7.2.8	Comparison with MEGAN . . . . .	131
7.3	New estimates affects on the Australian atmosphere . . . . .	131
<b>Bibliography</b>		<b>132</b>
<b>8</b>	<b>Appendix A</b>	<b>148</b>
8.1	Grid Resolution . . . . .	148

# Chapter 1

## Outline

### 1.1 Overview

In this thesis I will combine satellite and ground based atmospheric measurements with chemical transport modelling to clarify the impact of Australian natural emissions on atmospheric composition and chemistry. This broadly includes O<sub>3</sub>, NO<sub>X</sub>, VOCs, HO<sub>X</sub>, and the various processes which influence their composition and transport. Satellite measurements of HCHO, calculations of VOC emissions, HCHO and VOC atmospheric distributions, variations, and chemistry will be laid out with isoprene calculations being the main theme. The techniques used to analyse and create the products I use in my thesis will be explained. A background of HCHO modelling and MEGAN and GEOS-Chem will be split between here and the introduction to chapter 2.

A review of the literature on several topics is included with a view to touch on all the possible factors affecting my own work. Ozone and ozone precursors, along with the toxic effects and future projections of tropospheric ozone will be discussed. This leads into my **first chapter: “Stratosphere to Troposphere Transport of ozone”**. Here I analyse a dataset of ozonesondes, and calculate how many STTs occur in the high southern latitudes. An overview of my work categorising Stratospheric transport over the southern high latitudes will be the focus of this chapter as well as a paper I hope to submit next month.

An in depth review of how the GEOS-Chem model compares against ozonesondes will also be part of this chapter. GEOS-Chem is used to estimate how much tropospheric ozone is due to these STT events. These processes are some portion of the tropospheric ozone source - ties into another source: VOCs

My **second chapter: “HCHO total columns in Australia”** is focussed on calibrating a grided HCHO product from the OMI satellite measurements over Australia between January 2005 and April 2013. GOES-Chem and MEGAN model

histories will form some part of the background as they have been used several times to undertake similar inversions of VOC emissions. The process follows that of several other publications, however I focus on Australian emissions (largely unmentioned elsewhere) and parameters including precursor yields and grid resolution. Estimation of uncertainties and sensitivities to various factors including model grid resolution, fire, and anthropogenic masking operations is performed. Finally validation is undertaken using in-situ measurements of HCHO from an FTIR instrument on the roof of the chemistry building at the University of Wollongong.

My **third chapter “Isoprene emissions in Australia”** uses the HCHO product developed in chapter 2 along with various modelled parameters to estimate the emissions of isoprene (and possibly other BVOCs) through a simple linear steady state model, along with an analysis of the assumptions required for this model. Ideally the isoprene emissions estimates can be compared with MUMBA isoprene measurements during summer of 2012 - 2013. Again uncertainties and sensitivities will be examined for various parameters. Notably the effects due to smearing and grid resolution will be examined in detail, as these have never been heavily scrutinised for this inversion technique over Australia. A detailed comparison against other emissions estimates for Australia (ie: MEGAN and any other inventories) will be done with improvements or limitations highlighted. The isoprene emissions product will be analysed further by examination of its simulated VOC products.

# Chapter 2

## Introduction and Literature Review

### 2.1 Tropospheric ozone and air quality

#### 2.1.1 Air Quality

Australia is largely covered by environments which are not heavily influenced by human activity. These regions are natural (biogenic) sources of the trace gases (those which make up less than 1% of the earth's atmosphere). Trace gases in the atmosphere can have a large impact on living conditions. They react in complex ways with other elements (anthropogenic and natural), affecting various ecosystems upon which life depends. Biogenic emissions affect surface pollution levels, potentially enhancing particulate matter (PM) and ozone levels.

#### Ozone

Ozone ( $O_3$ ) is mostly located in the stratosphere, where it helpfully prevents much of the shorter wave length solar radiation from reaching the earth's surface (ie UV light). However around 11% of the total column of ozone is located in the troposphere (TODO: cite), where it has several deleterious effects. Ozone in the lower atmosphere is a serious hazard that causes health problems [48], damages agricultural crops worth billions of dollars [87, 124], and increases the rate of climate warming [86]. In the short term, ozone concentrations of  $\sim$ 50-60 ppbv over eight hours or  $\sim$ 80 ppbv over one hour are agreed to constitute a human health hazard [7, 63]. Long term exposure to lower levels cause problems with crop loss and ecosystem damage [4], and both short and long term concentrations may get worse in the future [63, 114]. Further tropospheric ozone enhancements are projected to drive reductions in global crop yields equivalent to losses of up to \$USD<sub>2000</sub> 35 bil-

lion per year by 2030 [87], along with detrimental health outcomes equivalent to  $\sim$ \$USD<sub>2000</sub>11.8 billion per year by 2050 [106]. Recently [124] showed that the net effect of near-surface ozone on is a  $\sim$  14% decrease in net primary productivity (NPP) in China, which could be reduced with drastic measures required to reduce this by 70% by 2030.

The tropospheric ozone concentrations rely on climate and ozone precursor emissions; including NO, NO<sub>2</sub>, CO, VOCs, and HCHO [Marvin2017, 5, 122]. Ozone predictions are uncertain and difficult due to the vagaries of changing climate which affects both 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 [122]. Even with all the work done in the prior decades there remains large uncertainties about ozone precursors in the troposphere [77]. Ozone is also a very important substance for formation of radicals (NO<sub>3</sub>, OH) in the troposphere through photolysis in the presence of water, see Section ?? for more details.

### Particulate matter and SOA

PM in the atmosphere is also a major problem, causing an estimated 2-3 million deaths annually [47, 58, 108, 64]. Aerosols are suspended particulates and liquid compounds in the atmosphere, of which particulate matter (PM) is an important subset. Fine particulate matter (PM<sub>2.5</sub>) 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's known as secondary OA (SOA) [59]. 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 [56]. TODO: POA vs SOA summary, how exactly are these two terms/metrics quantified?

### Factors influencing ozone and PM

One important factor, affecting both PM and ozone concentration, is a group called Volatile Organic Compounds (VOCs). The major source of VOCs in the atmosphere is biogenic, with around 90% of emissions (globally) coming from natural sources [42, 40, 80]. Global non-methane VOC (NMVOC) levels are estimated at 85 %, 13 %, and 3 % from biogenic, anthropogenic, and pyrogenic sources respectively [57]. They also affect radical levels, which drive much of the chemistry in our atmosphere. Due to the lack of in-situ ground based measurements, estimates of VOC emissions are uncertain, with large scale extrapolation required Millet et al.

[80]. VOC emission estimates are based on many factors, including plant type and soil moisture [42], neither of which are well characterised in Australia [109, 11]. In addition the estimates are highly sensitive to them. Changes in parameterisation of soil moisture in MEGAN (one emissions model, described later) lead to massive changes in Australian isoprene emission estimates [109]. This has an compounding effect on the large uncertainties of biogenic VOC emissions [43, 80].

Another important factor is NO<sub>X</sub> (NO, NO<sub>2</sub>) which has a complex non-linear relationship with both ozone and VOCs. TODO: talk about NOx chemistry? 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. 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. Understanding of emissions from these areas is necessary to inform national policy on air pollution levels.

Smoke plumes from biomass burning can carry ozone precursors, creating higher ozone concentrations downwind of the plume's source. Fire emissions include a range of chemicals and each year the affects 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 [92, 38, 26], particularly from July to December [93, 68].

### How do we measure air quality?

Australian air quality is monitored independently within each state, using various metrics. These metrics are measured by varying numbers of monitoring stations in each state. In New South Wales (NSW) the metrics used to determine air quality are: particulate matter (PM), O<sub>3</sub>, CO, NO<sub>2</sub>, SO<sub>2</sub>, and visibility. PM is separated into size bins: with radius < 2.5  $\mu\text{m}$  and < 10  $\mu\text{m}$  being PM<sub>2.5</sub>, and PM<sub>10</sub> respectively. An air quality index equal to the worst of these metrics is used for NSW as shown at <http://www.environment.nsw.gov.au/aqms/aqitable.htm>. Similar methods are used in other states to get an idea of air quality. Measurement stations are generally located in population centres, and don't regularly measure precursor emissions. This is an important omission as naturally emitted precursor gases often get transported into cities where they affect air quality.

Australia is dominated by areas with little anthropogenic influence and no ground based measurements of the natural emissions taking place [117]. One source of information which covers the entirety of Australia is remote sensing data. This comes from instruments on satellites which overpass daily recording reflected solar (and emitted terrestrial) radiation. In conjunction with atmospheric chemistry and radiative models, these measurements can be used to quantify the abundance of

several chemical species in the atmosphere. These measurements can then be used to determine the unmeasured emissions of precursor gases which affect air quality. Satellite data allow us to verify large scale model estimates of natural emissions. Satellite measurements can be performed using spectral fitting followed by conversion to vertical column densities. The use of multiple satellites can even be used to detect intradiel concentrations in trace gas columns, as shown in [111] using OMI and GOME-2 measurements, which have respective overpass times of 1330 and 0930 LT. These various measurements can be used to improve models, which allows prediction of harmful and costly events, and prevention of scenarios in which pollution events are likely.

In Australia most long term air quality measurements are performed in or near large cities. Estimates of atmospheric gas and particulate densities, and their distributions over much of the continent are uncertain, lacking long term in-situ measurements. Although the majority of VOC emissions globally are from biogenic sources, anthropogenic emissions in populated areas must be managed. [81] show that anthropogenic emissions of HCHO in America are mostly negligible, although improved sensitivity from oversampling allowed satellite detection of enhanced HCHO concentrations over Houston and Texas [127]. This is not the case in China, since massive population centres and industrial districts are emitting huge amounts of VOCs into the atmosphere [34]. [34] use GOME measurements over Asia and derive biogenic, anthropogenic, and pyrogenic VOC emissions, and [127] use oversampling of the OMI HCHO measurements to determine anthropogenic highly-reactive VOC emissions. Then with their updated emissions they show how surface ozone is affected, with a seasonal increase of 5-20 ppb simulated by GEOS-Chem.

Many models lack in-situ measurements with which to verify their chemical mechanisms, leading to large discrepancies, as seen in [76]. TODO: briefly talk about Marvin2017a takeaways.

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 [49]. One method of measuring ozone in the troposphere and stratosphere is by releasing weather balloons (with attached ozone detectors) which take readings as they rise up to around 30 km, giving a vertical profile of concentrations. Since 1986, Lauder, New Zealand ( $45^{\circ}\text{S}$ ,  $170^{\circ}\text{E}$ ) has released ozonesondes allowing a multi-decadal analysis of ozone concentrations over the city [13]. Kerguelan Island ( $49.2^{\circ}\text{S}$ ,  $70.1^{\circ}\text{E}$ ), also has a record of ozonesonde profiles, which are directly in the path of biomass burning smoke plumes transported off shore from Africa [8]. SHADOZ is the southern hemispheric additional ozone project, which have released sondes from 15 sites at

different times <http://tropo.gsfc.nasa.gov/shadoz/>.

TODO: get access to Hegglin (10.1038/ngeo604) [45] TODO: Include ozone hole treaty and things put in place for that

Combining satellite data with model outcomes provides a platform for the understanding of natural processes which is especially useful over Australia. Due to the low availability of in-situ data covering most of the Australian continent, a combination of the models with satellite data may provide improved understanding of emissions from Australian landscapes. Improved emissions estimates will in turn improve the accuracy of CTMs, providing better predictions of atmospheric composition and its response to ongoing environmental change.

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. [Marvin2017] use a box model matching conditions in southeast USA to evaluate isoprene mechanisms from several models.

While satellite data is effective at covering huge areas (the entire earth) it only exists at a particular time of day, is subject to cloud cover, and generally does not have fine horizontal or vertical resolution. Concentrations retrieved by satellites have large uncertainties, which arise in the process of transforming spectra to total column measurements, as well as instrument degradation (satellite instruments are hard to tinker with once they are launched). Uncertainty in satellite measurements comes from a range of things, including measurement difficulties introduced by clouds, and instrument sensitivity to particular aerosols [80]. Many products require analysis of cloud and aerosol properties in order to estimate concentration or total column amounts [95, 94, 75, 118]. There are two types of error, arguably the worst of these is systematic error (or bias) which normally indicates a problem in calculation or instrumentation. If the systematic error is known, it can be corrected for by either offsetting data in the opposite direction, or else fixing the cause. A proper fix can only be performed if the sources of error are known and there is a way of correcting or bypassing it. Random error is the other type (often reported as some function of a dataset's variance or else uncertainty), and this can be reduced through averaging either spatially or temporally. By taking the average of several measurements, any random error can be reduced by a factor of one over the square root of the number of measurements. This is done frequently for satellite measurements of trace gases (which are often near to the detection limit over much of the globe). For example: Vigouroux et al. [119] 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. The main source of error in satellite retrievals of HCHO are due to

instrument detection sensitivities, and the vertical multiplication factor (discussed in more detail in Section 6.2.8) [80].

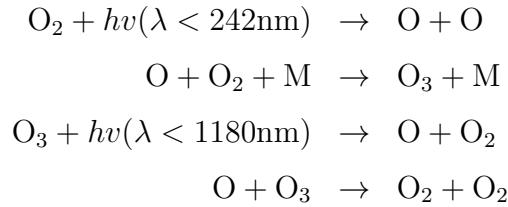
### 2.1.2 Ozone transported from the stratosphere

What drives ozone in the troposphere? Two main contributors are transport from the stratosphere and creation due to biogenic emissions of precursors. At smaller (regional) scales anthropogenic emissions are also important, especially in large cities such as Sydney due to NO<sub>X</sub> emissions from traffic and power production.

TODO: Intro to strat ozone transport first, then mechanism, then other stuff (maybe use subsubsections?)

Historically (in the late 1990's), ozone transported down from the stratosphere was thought to contribute 10-40 ppb to tropospheric ozone levels, matching tropospheric production (shown in equation 2.2) [5, 116]. This number was revised down over the years as measurement and modelling campaigns improved our understanding of global scale transport, mixing, and chemistry. (TODO: cites showing when these estimates were revised) Recently [60] analysed various measurements in the southeast USA and observed STT influence which can be seen to affect surface ozone levels. In their work they use of high spectral resolution lidar (HSRL), ozonesondes, ozone lidar, and airborn in-situ measurements give the structure and temporal evolution of ozone and the low front weather system. An analysis of the Atmospheric Chemistry and Climate Model Inter-comparison Project (ACCMIP) simulations by Young et al. [122] found 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 [83].

In the stratosphere ozone production is generally driven by the Chapman mechanism, as high energy light (with wavelengths  $\lambda < 242 \text{ nm}$ ) photolyses the molecular oxygen (O<sub>2</sub>) in the atmosphere [BrasseurJacob2017]. The Chapman mechanism involves several equations which lead to rough equilibrium of O, O<sub>2</sub>, O<sub>3</sub> and pressure, as follows:



Where  $h\nu$  represents radiation and M is an inert molecule (such as N<sub>2</sub>). The lifetime of O against loss by O<sub>2</sub> is less than a second in the lower atmosphere, and the

produced  $O_3$  quickly returns to  $O$  and  $O_2$ . 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 [BrasseurJacob2017]. This mechanism requires radiation so only takes place during the daytime, during the night there are different processes driving ozone chemistry. TODO: Tie strat and trop ozone to this concept

Ozone transported to the troposphere from the stratosphere can occur through diffusion (slow process (TODO: Cite)) or through mixing, often called Stratosphere to Troposphere Transport events (STT), or intrusions. TODO: compare these two processes and their impacts briefly. Recently global chemical transport models (CTMs) have been used to trace how much ozone is being transported to the troposphere from the stratosphere. There are a few methods of doing this, such as Ojha et al. [91], who use the ECHAM5 CTM with a tracer based on keeping track of ozone formed and transported from the stratosphere. The estimates generally require validation against actual measurements, such as those from ozonesondes or satellites. Hegglin and Shepherd [45] estimate that climate change will lead to increased STT due to an acceleration in the Brewer Dobson circulation. They estimate  $\sim 30$ , and  $\sim 121 \text{ Tg yr}^{-1}$  increases (relative to 1965) in the southern and northern hemispheres respectively

[69] examine southern hemispheric ozone and the processes which control its interannual variability (IAV). IAV is the standard deviation of ozone anomalies (difference 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. While stratospheric transport mostly impacts the upper troposphere, some areas are impacted right down to the surface. Liu et al. [69] look at modelled tropospheric ozone sensitivity to changes in stratospheric ozone, ozone precursor emissions, and lightning over the southern hemisphere from 1992–2011. Their work suggests ozone at 430 hPa is mostly stratospheric in September over 20°S to 60°S at all longitudes. They also see tropospheric ozone sensitivity to emissions from South America (0–20°S, 72.5–37.5°W), southern Africa (5–10°S, 12–38°E), and South and South-east Asia (70–125°E, 10°S–40°N). In the USA recent work by [67] suggests that intrusions during spring are increasing surface ozone levels. Their work also recommends that understanding of frequency and cause of STT needs to be improved to effectively implement air quality standards.

### 2.1.3 Biogenic precursors to ozone

Ozone is formed in the troposphere through oxidation of VOCs in the presence of  $NO_X$ . Net formation or loss of  $O_3$  is determined by interactions between VOCs,  $NO_X$ , and  $HO_X$ , and is a complicated system of positive and negative feedbacks [5].

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 [113].

Methane and isoprene each comprise around a third of the yearly global total emission of VOCs. TODO: update this sentence, methane emissions? or methane as a source of trop ozone?. However, methane is relatively long lived (years) and is well mixed in the atmosphere while isoprene levels are very volatile and spatially diverse due to a life time of around an hour. This means that methane measurements and concentrations are relatively well understood, while NMVOCs are less so.

TODO: isoprene stuff to isop section - this section more about ozone production

TODO: more on ozone formations Ozone precursor concentrations are largely driven by emissions of VOCs and  $\text{NO}_x$

## 2.2 Isoprene and other VOCs

### 2.2.1 What are VOCs

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 ( $\text{CO}_3$ ), and simple oxides of carbon and cyanides. Organic compounds can be categorised based on their vapour pressure, which is the tendency of a liquid or solid to vaporise. Compounds with high vapour pressures at standard temperature are classed as volatile, and have a facility to evaporate at low temperatures. Plants contain tens of thousands of organic compounds, it's likely that fewer than 40 are emitted due to the low volatility of most of them [43].

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. Volatile organic compounds have vapour pressure greater than  $10^{-5} \text{ atm}$ , and are mostly generated naturally by plants, which emit around 1000 Tg per year [42, 37]. Due to their high volatility these compounds are generally seen in the gas phase. Organic compounds with a lower volatility are classed as semi-volatile (SVOCs: vapour pressure between  $10^{-5}$  and  $10^{-11} \text{ atm}$ ) are seen 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 [37].

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. NMVOCs are alkanes, alkenes, aromatic hydrocarbons and isoprene, with isoprene being the most prominent. Estimates put global isoprene emission at roughly  $600 \text{ Tg yr}^{-1}$ , emitted mostly during

the day. [42], and subsequent updates [43, 40], 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. More recently an estimate of global isoprene emissions of around 465 Tg C yr<sup>-1</sup> has been made by [78] using ORCHIDEE. The global emission factors model used to derive these estimates is based on modelling emissions from different plant species (phenotypes), and relatively few Australian species are used when forming in these estimates.

Major emitters are tropical broadleafs (notably eucalypts), and scrubs [40, 3, 90, 83]. Although recent work suggests that some Australian eucalypts may not be as egregious isoprene emitters as once thought Emmerson et al. [27]. Isoprene emissions are often classified as either anthropogenic, biogenic, or pyrogenic. The natural or biogenic sources are roughly ten times higher than the anthropogenic VOC sources [40, 56].

Isoprene affects NO<sub>X</sub> and HO<sub>Y</sub> cycling, and in the presence of NO<sub>X</sub>, forms tropospheric ozone and SOAs [120, 80]. It has a short lifetime during the day, roughly an hour due to OH oxidation [6]). At night when OH concentrations have dropped, isoprene can remain in the atmosphere to be transported. Typically less than half of this night time isoprene is removed through ozonolysis [6], however, in polluted areas where high levels of NO<sub>X</sub> exist, isoprene is consumed by a different radical. During the night time, nitrate radicals (NO<sub>3</sub>) build up, especially in areas with high NO<sub>X</sub> levels. 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 over night [14]. So while night time isoprene is not as highly concentrated, it does have varying biogenic and anthropogenic sinks. At night isoprene has affects on both NO<sub>X</sub> concentrations and ozone levels, and can form harmful SOAs [14, 73]. 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 [7].

Land use changes could drastically affect isoprene sources, for instance in the tropics where large scale deforestation has occurred, converting forest into crop lands [56].

Bottom up inventories of VOCs remain largely uncertain due to extensive extrapolation over plant functional types, changing land cover, and parameterised environmental stressors [43, 56]. This problem is even more pronounced in Australia due to poorly studied paramaterised PFT and soil moisture. [85] show how isoprene is poorly captured by the MEGAN model and analyse the affect of changing the soil moisture parameter, which can reduce the overall bias for Australia. TODO: more on Muller2008 [27] shows that isoprene emissions modelled by MEGAN in south-eastern Australia may be 6 times too high. They compare emissions estimates from MEGAN against data from several field campaigns and see overestimated isoprene

emissions, as well as underestimated monoterpene emissions. There are many uncertainties in estimates of emissions in Australia due to missing or extrapolated data. For instance Emissions in MEGAN are based on plant functional types, which can vary heavily even within species. Many plant emissions have not been published, such as those for any Australian acacias. And soil moisture is not well quantified which has a large effect on emissions. TODO: more on K. Emmerson

### 2.2.2 Emissions estimates

There are two commonly used ways of estimating isoprene emissions, top-down or bottom-up. Bottom-up emission estimates generally model the flora and events which emit isoprene, like Eucalypts, factories, shrubs, leaf areas under sunlight, etc. Understanding how much isoprene is emitted, when and by what is more complicated than it sounds, and since little data exists with which to verify these bottom-up emission inventories they are uncertain on a large scale. 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. For isoprene this is done by looking at atmospheric HCHO enhancement, which can be largely attributed to isoprene emissions as long NO<sub>X</sub> and transport effects are accounted for.

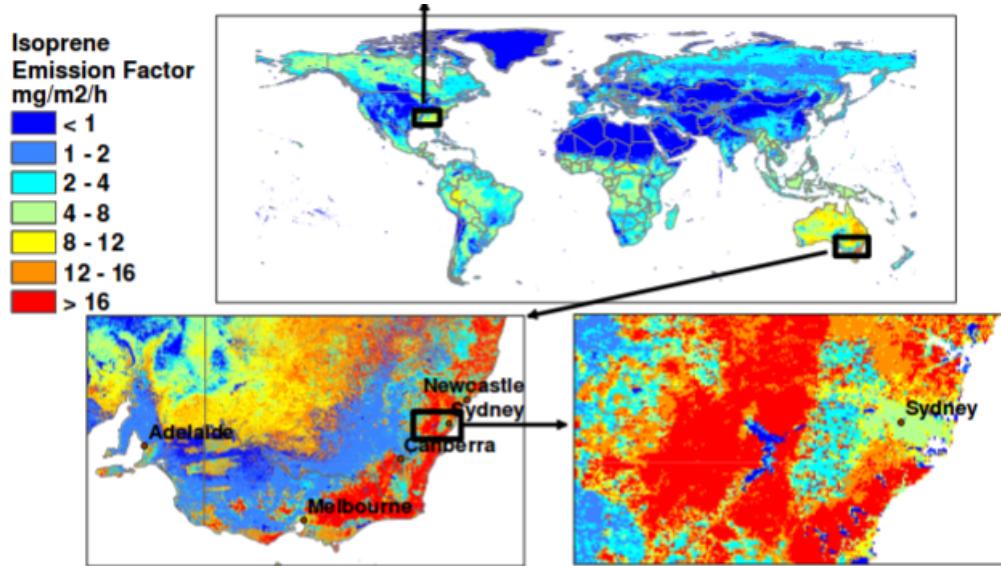
It used to be thought that emissions of anthropogenic and biogenic VOCs (BVOCs) were roughly similar (TODO abstract of [Mueller1992], and more cites). In the 1990's, the World Meteorological Organisation (WMO) estimated that we are emitting 360 Mt yr<sup>-1</sup> of methane (CH<sub>4</sub>), one of the more abundant and potent VOCs, while biogenic emissions were around 200 Mt yr<sup>-1</sup> [5]. At that point in time, emissions of other VOCs were estimated at 1150 Mt yr<sup>-1</sup> (of carbon) from biogenic sources, and 100 Mt yr<sup>-1</sup> from anthropogenic sources [42, 5]. These estimates were based on the Model of Emissions of Gases and Aerosols from Nature (MEGAN, Guenther et al. [42]). It's remains clear that biogenic VOC (BVOC) emissions are far greater than anthropogenic emissions of VOCs, making up 87% of NMVOC emissions [56, 57]. Non methane BVOC emissions are estimated to be  $\sim 1150 \text{ TgCyr}^{-1}$ , of which isoprene (44%) and monoterpenes (11%) are the main single contributors. [43, 57].

Guenther et al. [40] estimated that the Australian outback is among the world's strongest isoprene emitters with forests in SE Australia having emission factors greater than 16 mg m<sup>-2</sup> h<sup>-1</sup> (see figure 2.1). Measurement campaigns in SE Australia have since cast doubt on the emission factors used by MEGAN, as the Eucalyptus trees and soil moisture were poorly studied [27]. These emissions factor estimates are not well verified as there is little coverage of isoprene (or other BVOC) emissions measurements over Australia. However, comprehensive coverage of one

high yield product (HCHO) in the atmosphere over Australia exists in the form of satellite measurements.

The estimates are still fairly uncertain, as global measurements are difficult and regional emissions can be very different. The global uncertainty of isoprene emission was estimated to be a factor of 3 ( $250\text{-}750 \text{ Tga}^{-1}$ ) [56]. The improvement over the years have been incremental, with updated estimates of TODO: find recent uncertainty estimate. The lack of accuracy in BVOC emissions estimates has a large effect on determining with confidence the sources and distribution of pollutants including ozone and organic aerosols. Most of the tropospheric SOA comes from biogenic precursors, the evidence for this has grown over the last two decades [42, 56, 41]. Accuracy in VOC measurements is important: it has been shown that even the diurnal pattern of isoprene emissions has an effect on modelling ground level ozone [46, 30]. These uncertainties could explain why models of HCHO over Australia are poor at reproducing satellite measurements [110]. Over Australia specifically MEGAN has problems involving unpublished plant functional types and their emissions, as well as poorly optimised soil moisture parameterisation [27]. Australia also lacks a clear estimate of emitted monoterpenes. [27] suggest that monoterpenes may be emitted in similar quantities to isoprene, with more measurements required to determine if this is so. 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.

TODO: some more on Kefauver2014 review [57] reviews remote sensing of BVOCs, which are on the rise, examining the last 20 years of data and analysis of the satellite products. Their review encompasses the latest reports up to 2014 however the modelled isoprene and BVOC emissions from MEGAN [43] of  $500$  and  $1150 \text{ Tga}^{-1}$  respectively are still the global go to estimates. The review reinforces the message that NMVOCs affect the oxidative capacity of the atmosphere and are largely driven by and sensitive to vegetation. The tropospheric affects from NMVOCs on the hydroxyl radical ( $\text{OH}$ ), ozone ( $\text{O}_3$ ), SOAs, and methane longevity, all interconnect to form a very complex system which still suffers from relatively large uncertainties in both measurement and chemistry mechanisms. One focus of Dr. Kefauver et al.'s review is HCHO, which is the dominant product of most BVOCs which is measurable by remote sensing. The main datasets of HCHO are from four satellite instruments: GOME on ERS-2, SCIAMACHY on ENVI-SAT, OMI on EOS AURA, and GOME2 on MetOp-A. These satellites have slightly different spectral and spatial resolutions, as well as using varied processes to estimate HCHO from detected radiances. This can lead to different estimates between instruments or methodologies as described in [Lorent2017], which means validation and comparison is more important when using these remotely sensed data.



**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.

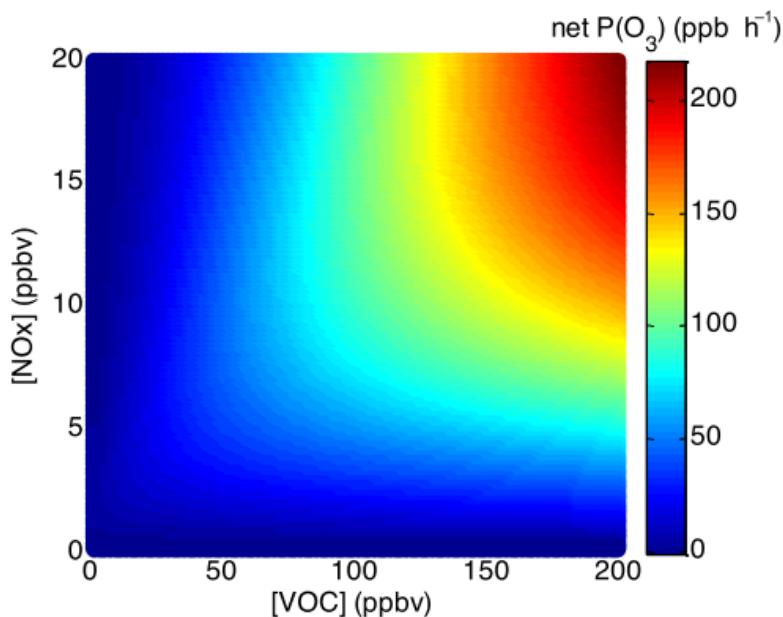
**Figure 2.1:** Part of a figure from Guenther et al. [40] showing global isoprene emission factors.

It is important to note that many estimates of isoprene emission are based on a few algorithms which can depend greatly on input parameters [90]. Yue, Unger, and Zheng [123] has shown that this is still a problem by looking at land carbon fluxes and modelling the sensitivity to VOC emissions estimates using two independent models of VOC emission. One model is photosynthesis based and estimates isoprene emissions using electron transfer energies and leaf physiology [89], while the other (MEGAN) uses the light and canopy temperature ([42] TODO: Arneth et al., 2007; Unger et al., 2013). Both are sensitive to light and temperature parameterisations.

### 2.2.3 What do they Do?

VOC emissions result in radical cycling, acid deposition, and the production of tropospheric ozone, and secondary organic aerosols (SOAs) [5, 56]. These have impacts on climate (through radiative forcing) and air quality, affecting both human health and crop yields [31, 87, 64]. Understanding the drivers of trends in biogenic VOC emissions (BVOCs) is needed in order to estimate future carbon fluxes, changes in the water cycle, air quality, and other climate responses [123].

There is a complex relationship between  $\text{NO}_x$ , VOCs, and ozone, figure 2.2 shows this relationship over Houston, as modelled in [77]. 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 depending on location various other factors [77].



**Figure 1.** Ozone production empirical kinetic modeling approach (EKMA) diagram using box model results with NO<sub>x</sub> levels varying from 0 to 20 ppbv and VOC levels from 0 to 200 ppbv. The mean concentrations of other species and the speciation of NO<sub>x</sub> and VOCs observed during DISCOVER-AQ in Houston in 2013 were used to constrain the box model. This diagram clearly shows the sensitivity of ozone production to NO<sub>x</sub> and VOCs in Houston.

**Figure 2.2:** Ozone production figure copied from Mazzuca et al. [77].

### 2.2.4 Where do they go?

Isoprene forms many products with various lifetimes, here I will present an overview of some important mechanisms which affect oxidation capacity, ozone and aerosol production. Photolysis and oxidation of many VOCs initially form alkyl radicals ( $\dot{R}$ ), and reactions with ozone (with alkenes or VOCs containing a double bonded carbon) lead to organic peroxy radicals ( $\text{RO}\dot{\text{O}}$ ). These go on to form many products and lead to (amongst other things) aerosol, formaldehyde, and ozone formation, depending on various other factors such as sunlight and NO pollution [5]. The first step in oxidising alkenes (including isoprene) is a replacement by OH of a double bond, as summarised by the equation from [97]:  $\text{R}-\text{CH}=\text{CH}-\text{R}' + \text{OH} \longrightarrow \text{R}-\text{CH}(\text{OH})\text{CH}-\text{R}'$  where R and R' represent hydrocarbons. This OH adduct then reacts with  $\text{O}_2$  to produce a hydroxyperoxy radical ( $\text{RO}\dot{\text{O}}$ ), which for isoprene can be any of six different isomers [97]. This is where the isoprene oxidation path splits depending on the  $\text{NO}_x$  concentration, TODO: find recent cite and talk about ROO to products path. Reaction with NO leads to ozone production, while reaction with  $\text{NO}_2$  forms hydroxynitrate ( $\text{RONO}_2$ ). [97] examine this branching step, and for specific conditions they determine the reaction rates for each branch, the most frequent pathway is the formation of ROO (99.3%). TODO: is  $\text{RONO}_2$  PAN? Although the PAN formation branch is quite minor, it can be very important as it can transport  $\text{NO}_x$  to otherwise clean regions. PAN has a relatively long lifetime (against OH, order of 1 day) and is able to transport and release the  $\text{NO}_x$  in environments which are quite far from any emissions. This transport effect can be exacerbated by fast winds and low OH concentrations (TODO: cites), which makes PAN another important factor in modelling atmospheric chemistry.

In the presence of  $\text{NO}_x$ , the ISOPOO radicals form organic nitrates after reacting with NO. Any organic nitrates which are formed affect levels of both  $\text{HO}_x$  (H, OH, peroxy radicals) and  $\text{NO}_x$ , acting as a sink ([73] and references therein). VOCs are removed by wet and dry deposition, or transformed by reaction with OH,  $\text{NO}_3$ , or  $\text{O}_3$ . The process of deposition only accounts for a small fraction of the VOC loss, with the possible exception of the long lived methane compound [6]. Primary reductions occur through photolysis, OH oxidation, ozonolysis, and at night time in polluted areas: loss by reaction with  $\text{NO}_3$  [6, 14]. In the presence of  $\text{NO}_x$  ( $\equiv \text{NO} + \text{NO}_2$ ), non-methane organic compounds (NMOCs) and NMVOCs end up forming tropospheric ozone. This is achieved through photolysis of  $\text{NO}_2$ , concentrations of which are increased by NMOC and NO reactions [6].

### 2.2.5 The isoprene cascade

TODO: Organise this section, remove repeated bits and improve order.

Isoprene is emitted and enters the atmosphere in the gas phase, where it reacts with various chemicals, forming many new chemicals and reactions at various time scales. The estimated yield of HCHO is but one aspect of the many processes going on in this space. The primary first step for atmospheric isoprene is photooxidation, reacting with OH to form isoprene hydroxyperoxy radicals (ISOPOO, or ROO, or RO $\cdot$ ) [Marvin2017, Patchen2017, 121].

The oxidation products of isoprene through addition of OH (forming ISOPOO) followed by addition of O<sub>2</sub> produces various isomers of alkylperoxyl radicals (organic peroxy radicals, or RO<sub>2</sub>), which react with HO<sub>2</sub> or NO and produce stable products (often called oxidised VOCs or OVOCs) [Nguyen2014]. The ISOPOO radicals are eventually destroyed by NO, HO<sub>2</sub> and other RO<sub>2</sub>, with most pathways potentially producing HCHO [121]. Oxidation reactions are important and quickly stabilise the ratio of NO to NO<sub>2</sub>. NO<sub>x</sub> removed primarily by conversion to nitric acid (HNO<sub>3</sub>) followed by wet or dry deposition [7]. There is still large uncertainty around the fate of various RO<sub>2</sub> radicals, which limits understanding of the relative importance of some chemical processes [18].

### Isoprene oxidation TODO: organise into subsubsections?

Photolysis and oxidation of many VOCs initially form alkyl radicals ( $\dot{R}$ ), and reactions with ozone (with alkenes or VOCs containing a double bonded carbon) lead to organic peroxy radicals (RO $\cdot$ ). These go on to form many products and lead to (amongst other things) aerosol, formaldehyde, and ozone formation, depending on various other factors such as sunlight and NO pollution [5]. The fate of ISOPOO depends upon concentrations of NO<sub>x</sub>, reacting with NO to form HCHO, MVK, MACR, and to a lesser extent organic nitrates (ISOPN).

The first step in oxidising alkenes (including isoprene) is a replacement by OH of a double bond, as summarised by the equation from [97]: R–CH=CH–R' + OH → R–CH(OH)CH–R' where R and R' represent hydrocarbons. This OH adduct then reacts with O<sub>2</sub> to produce a hydroxyperoxy radical (RO $\cdot$ ), which for isoprene can be any of six different isomers [97]. There is still uncertainty about which pathways are most important following ISOPOO production: HO<sub>2</sub> reactions predominantly produce hydroxyhydroperoxides (ISOPOOH), NO reactions largely produce methyl vinyl ketone (MVK) and methacrolein (MCR), and RO<sub>2</sub> reactions are also possible [70].

This is where the isoprene oxidation path splits depending on the NO<sub>x</sub> concentration, TODO: find recent cite and talk about ROO to products path. Reaction with NO leads to ozone production, while reaction with NO<sub>2</sub> forms hydroxynitrate (TODO: PAN?) The first generation of organic nitrates produced by isoprene oxidation range from 7% to 12%, shown in laboratory experiments [98, 73], A portion

of isoprene nitrates are recycled back to  $\text{NO}_X$ , so may serve as a reservoir of nitrogen and allow its transport to the boundary layer of remote regions (TODO: as prior todo)[97, 98]. PAN can then transport  $\text{NO}_X$  as it has a relatively long lifetime (against OH, order of 1 day) to release the  $\text{NO}_X$  in a new environment. This transport affect can be exacerbated by fast winds and low OH concentrations (TODO: cites), which makes PAN another important factor in modelling atmospheric chemistry.

Isoprene oxidation by OH is less well understood when lower concentrations of NO are present in the atmosphere. Initially isoprene was thought to be a sink for atmospheric oxidants [e.g. 43]. It was thought that in low NO environments, like those far from anthropogenic pollution and fires, oxidation of isoprene would create hydroxyhydroperoxides (ISOPOOH) and lead to low concentrations of OH and  $\text{HO}_2$  (together known as  $\text{HO}_X$ ) Paulot et al. [99]. In Paulot et al. [99], the  $\text{HO}_X$  levels are shown to be largely unaffected by isoprene concentrations. They show that ISOPOOH is formed in yields  $> 70\%$ , and MACR and MVK is formed with yields  $< 30\%$ . The formation of MACR and MVK produces some  $\text{HO}_X$ , although not enough to close the gap. Paulot et al. [99] goes on to suggest (and provide experimental evidence) that dihydroxyperoxides (IEPOX) are formed from oxidation of the ISOPOOH, which form precursors for SOAs as well as closing the  $\text{HO}_X$  concentration gap. They then use GEOS-Chem, modified to include IEPOX formation, to estimate that one third of isoprene peroxy radicals react with  $\text{HO}_2$ , and two thirds react with NO. Their work showed another pathway for isoprene based SOA creation, and additionally estimated  $95 \pm 45 \text{ TgC yr}^{-1}$  IEPOX being created in the atmosphere without any inclusion in CTMs at that time. Peeters and Müller [100] suggested that the work of Paulot et al. [99] only partially bridges the gap between clean air HO 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 HPALDs, and resulting photolabile peroxy-acid-aldehydes (PACALDs). These reactions are highly non-linear and subject to large uncertainty, however when compared against several campaigns they were shown to improve one particular models (IMAGES)  $\text{HO}_X$  concentrations.

In [17], MACR products are examined in various conditions and hydroxy recycling is also observed in low NO conditions.

Although understanding of OH production/recycling in these low NO conditions has been improved, many observations of OH are still quite under-predicted in models [72]. It was shown in Mao et al. [72], for a remote forest in California, that the traditional method of OH measurement may be affected by instrument internally generated OH from VOC oxidation. This lends more credence to the current under-

standing of VOC oxidation as it closed the gap between measurements and model predictions.

Dillon et al TODO: cite (T. J. Dillon, J. N. Crowley, H. Harder, D. Kubistin and M. Martinez-Harder, Experimental evidence for HO<sub>2</sub> production from OH + isoprene in clean air, Book of Abstracts, 21st International Symposium on Gas Kinetics, KULeuven, Leuven, 2010, ). Showed that HO<sub>2</sub> is produced at near unity yields following isoprene oxidation initiated by HO.

In [Nguyen2014], many chamber studies, scientists, and groups worked together to improve understanding of ambient atmospheric oxidation mechanisms of biogenic hydrocarbons (such as isoprene). This work discussed how large uncertainties persist in isoprene oxidation, which carries through to uncertainties in predictions by atmospheric models. Nguyen2014 show preliminary estimates of low-NO yields of MVK and MCR to be 6±3% and 4±2% respectively, consistent with TODO:Liu2013, but only when cold-trapping methods are employed. These yields each increase (due to interference by OVOCs) to greater than 40% when directly sampled by GC-FID.

During the night isoprene is oxidised by NO<sub>3</sub> radicals, which joins to one of the double bonds and produces organic nitrates in high yield (65% to 85%) [73]. (todo: read mao2013 para 3 cites for) These organic nitrates go on to produce further SOAs, largely due to NO<sub>3</sub> reacting with first generation isoprene oxidation products [Rollins2009]. [Rollins2009] 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.

Even with the recent boom in isoprene analysis, uncertainties remain in the isoprene oxidation mechanisms. Examples (taken from Nguyen2014) include isoprene nitrate yields, which range from 4-15% [98], 90% disagreements in MAC and MVK yields TODO:[Liu2013], various possible sources for SOA TODO:[Chan2010, Surratt2010, Lin2013], unknown HPALD fates, incomplete O<sub>2</sub> incorporation TODO:[Peeters2009, 18], and under-characterized RO<sub>2</sub> lifetime impacts TODO:[Wolfe2012]. TODO: get those citations and read abstracts.

## Ozononlysis

Ozononlysis is the splitting of carbon chains by ozone molecules, and is among the primary oxidation pathway for volatile alkenes [88]. Criegee intermediates are formed when isoprene reacts with O<sub>3</sub>, with C<sub>1</sub> and C<sub>4</sub> Criegees discussed in detail, with proposed mechanisms, in [88]. The C<sub>1</sub> stabilised Criegee (CH<sub>2</sub>OO, ~ 61%) is therein proposed to react with water yielding 73% hydroxymethyl hydroperoxide

(HMHP), 6% HCHO + H<sub>2</sub>O<sub>2</sub>, and formic acids + H<sub>2</sub>O, and the same products with yields of 40, 6, and 54% respectively when this Criegee reacts with (H<sub>2</sub>O)<sub>2</sub>. TODO: more on Nguyen2016

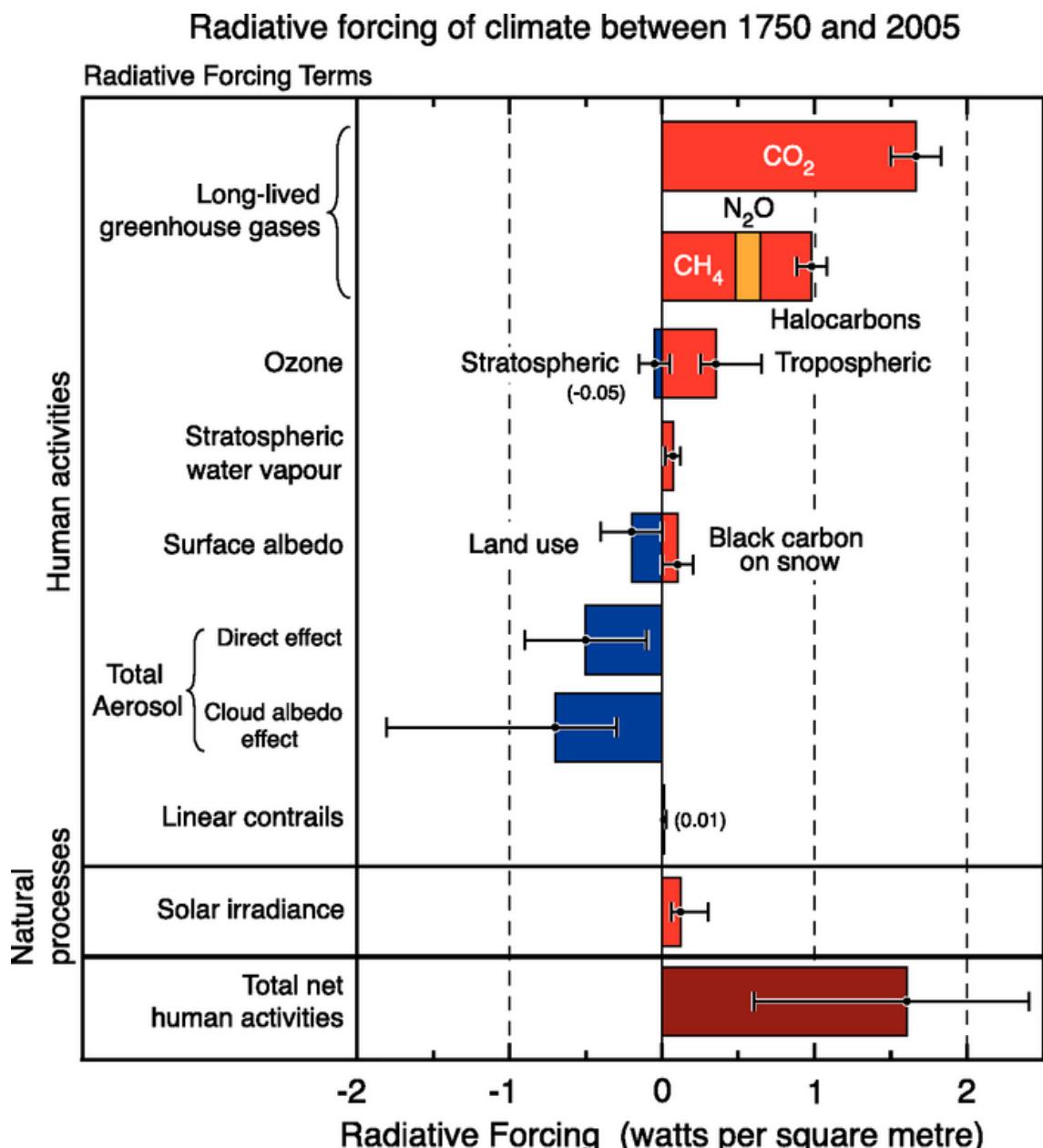
## SOA

Gas phase emissions with higher vapor pressures can be oxidised into lower vapor 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 [56]. In the [56] 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 (~ 1%) of the resulting PM [59, 12]. 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 as a full profile of aerosols is required for remote sensing techniques [56].

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 ~ 10 years ago, as chamber studies did not extend over a large enough parameter range and the importance of heterogeneous aerosol chemistry on SOA formation was unquantified [56].

One of the large uncertainties with OA is the total effect on radiative forcing, ten 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 [56]. 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 [51], 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 [31]. Figure 2.3 shows the radiative forcings (RF) of various atmospheric constituents, it's clear that OA uncertainty dominates. Figure 2.4 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's play an indirect and complex role in cloud properties, with a net cooling effect [115, Chapter 7,8]

(TODO: read more of Kanakidou2005) The emissions of precursors to SOA was



**Figure 2.3:** The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from [31], 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).

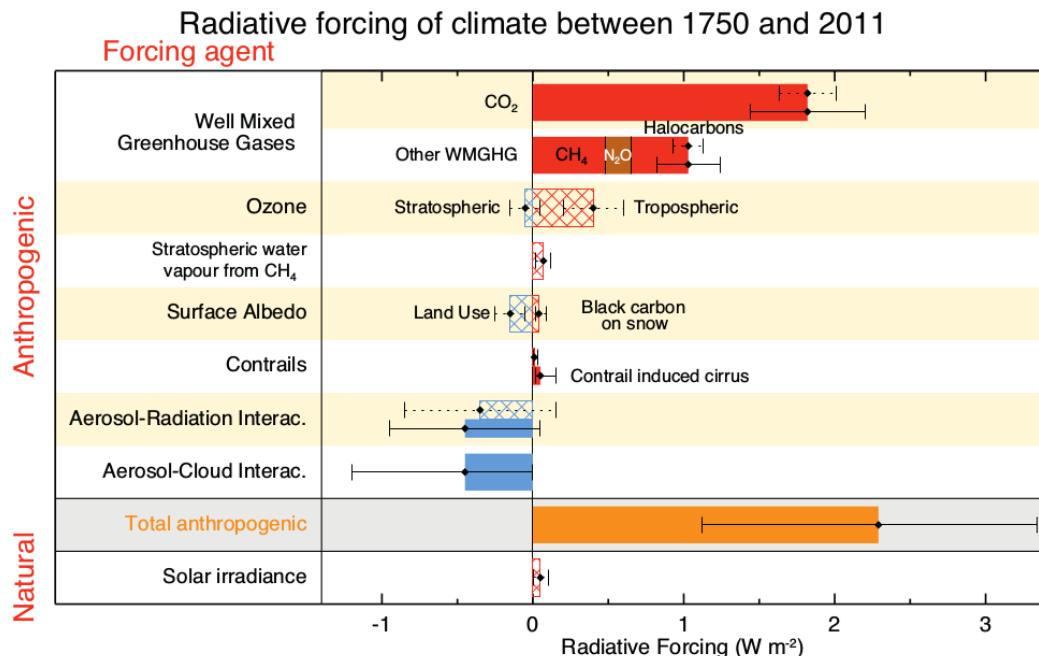


Figure 8.15 | Bar chart for RF (hatched) and ERF (solid) for the period 1750–2011, where the total ERF is derived from Figure 8.16. Uncertainties (5 to 95% confidence range) are given for RF (dotted lines) and ERF (solid lines).

**Figure 2.4:** The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from [115], chapter 8.

and is quite uncertain, in [56] 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 [56], 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. [56] concluded that it is very likely that organics contribute to particle growth and formation rates.

## 2.2.6 How and where do we measure them?

[56] summarised the difficulty of chamber experiments used to measure isoprene reactions and the possible unsuitability of chamber study yields in the natural atmosphere. This is due to the complex relationship between NO<sub>x</sub>, NO<sub>3</sub>, OH, O<sub>3</sub>, and the formation of aerosols being hard to attribute any single precursor.

There are relatively few measurements of isoprene in the southern hemisphere,

including MUMBA(TODO CITE), other campaigns?, and very recently that girl from Macquarie University with an instrument in the daintree rainforest(TODO CITE, DESCRIBE?).

### 2.2.7 How do we determine emissions?

Emissions inventories such as MEGAN use models of emissions based on tree types, weather, etc...

Another way infers emissions based on the longer lived products like formaldehyde (HCHO). Recently [111] used satellite HCHO measurements to constrain anthropogenic sources of isoprene and found good global agreement with the bottom up estimates, although regions had sources differ by up to 25-40%. This study used the RETRO 2000 database for anthropogenic emission aprioris except for Asia in 2008 where REASv2 was used. Since 1997, when GOME first measured HCHO over Asia [Thomas1998], satellites have been able to provide a total column measurement of one of the primary products of isoprene.

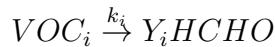
#### Satellite inversion

Satellites recording reflected solar spectra use Differential Optical Absorption Spectroscopy (DOAS) to measure various trace gases in the atmosphere, including formaldehyde. Formaldehyde levels in the continental boundary layer are generally dominated by chemical formation due to VOC (largely isoprene) emissions [57]. While satellite measurements can only be used during daytime hours, HCHO lifetimes are sufficiently short that any nighttime chemistry will not affect midday observations [121]. Satellites can be used to measure the seasonal and interannual variability of HCHO over Australia. These records can be compared with modeled estimates of HCHO and used as a proxy to estimate isoprene emissions. This has been done in North America [94, 80], South America, Africa, China, Europe [25], and recently globally [32, 11]. Often these works use two forms of measurement such as satellite and aircraft data combined for validation [74]. There is less information available from satellite measurements at higher latitudes due to increased errors [23].

Using HCHO to determine emissions of isoprene was initially performed by [95, 94], who used in-situ summertime HCHO measurements over North America as model validation. Isoprene emissions fluxes were derived using the Global Ozone Monitoring Experiment (GOME) satellite instrument. Palmer's method improved biogenic isoprene emissions estimates (compared with in-situ measurements) over two available inventories: the U.S. EPA Biogenic Emissions Inventory System (BEIS2) and the Global Emissions Inventory Activity (GEIA). This showed an inversion technique which could be used to improve large scale emissions estimates

without further expensive measurement campaigns.

Initially studies assumed a simple linear steady-state relationship between HCHO and its precursors [94, 96, 80]. This allowed a simple calculation of isoprene using the measured HCHO, with estimated reaction rates and yields. The methodology for calculating VOCs from HCHO is laid out in Palmer [94], and takes into account the expected lifetime and reaction rates of the precursor VOCs and HCHO. Assuming HCHO is produced quickly from short-lived intermediates, and the column is at steady state:



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

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

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

Estimates of  $Y_i$  can be attained from a model as shown in Millet et al. [80]. This involves a reduced major axis (RMA) correlation calculation between modelled HCHO and isoprene columns, multiplied by their loss rates (to photolysis and oxidation) (as a normalising factor). In high NOx environments where HCHO has a lifetime on the order of 30 minutes, it can be used to map isoprene emissions with spatial resolution from 10-100 kms. Horizontal transport 'smears' the HCHO signal so that source location would need to be calculated using windspeeds and loss rates [95, 94]. For more details on this see section ??.

Another method of correcting isoprene emissions using observed HCHO total column involves a Bayesian inversion. Shim et al. [107] work with GOME HCHO observations and GEOS-Chem, looking at areas with high signal to noise ratio (higher HCHO concentrations). They show that the model underestimates isoprene emissions and HCHO concentrations by 14-46%, with the corrected VOC emissions reducing the model biases to 3-25%.

The Bayesian inversion is also used in Curci et al. [21], where a regional CTM (CHIMERE) simulates HCHO, which is compared against OMI observed HCHO and shown to be regionally biased. This bias is expected to be caused by errors in MEGAN's natural isoprene emissions. The CHIMERE model is used to derive yields

of HCHO from the various local VOCs and these are then used in estimating local emissions. The model is run initially with emissions of BVOCs and reactive anthropogenic VOCs (RAVOCs) turned off in order to work out the background ( $b$ ) values of these compounds. The Bayesian inversion is used to correct regionally biased biogenic isoprene emissions by optimising these parameters in order to simulate HCHO closest to the observed HCHO levels. [21] uses CHIMERE as the forward model to determine the relationship between HCHO ( $y$ ), isoprene and reactive anthropogenic VOCs ( $x$ ), using

$$y = \mathbf{K}x + b + \epsilon \quad (2.1)$$

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

TODO: Read through this list of sources on the hcho to isop process : taken from Wolfe2015 Such techniques have informed isoprene emission inventories in North America (Abbot et al., 2003; Millet et al., 2008 [94, 80, 96]), South America ([9], 2008), Europe [21, 25], Africa [75], Asia (Fu et al., 2007; Stavrakou et al., 2014), and globally (Fortems-Cheiney et al., 2012; [107]; Stavrakou et al., 2009).

More recently, full inversions that better account for transport, source attribution, and chemical schemes have been implemented [32]. TODO: full description of this better inversion technique going through FortemsCheiney2012.

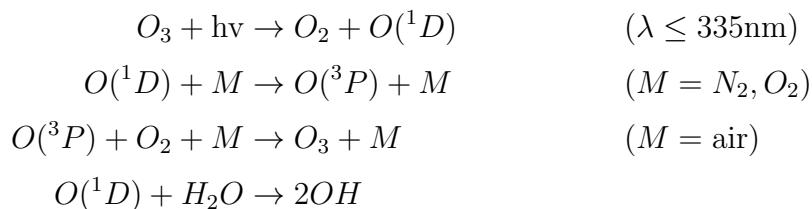
Validation is important due to the various uncertainties in the satellite remote sensing process, with apriori assumptions having the greatest effect on structural uncertainty between measurements techniques Lorente et al. [71]. [128] use SEAC<sup>4</sup>RS aircraft HCHO measurements over the southeastern US as model validation, and show a bias in the assumed OMI shape factor that leads to a bias between satellite and SEAC<sup>4</sup>RS measurements. [74] compare OMI based isoprene emission estimates against relaxed eddy accumulation measurements from African field campaigns, as well as MEGAN and GEOS-5 inventories.

Dufour et al. [25] use HCHO from SCIAMACHY, and examine Europe using CHIMERE as the chemical model. In their work they show that satellite measurements can reduce source emission uncertainty by a factor of two, where emissions are relatively large.

## 2.3 Hydroxyl (OH) and other radicals (maybe relocate this bit?)

The OH radical drives many processes in the atmosphere, especially during the day when photolysis of ozone drives OH concentrations [5]. OH is a key species which reacts with nearly all the organic compounds in the troposphere. The exceptions are chlorofluorocarbons (CFCs), and Halons not containing H atoms [5]. OH and HO<sub>2</sub> concentrations largely determine the oxidative capacity of the atmosphere. Oxidation and photolysis are the two main processes through which VOCs are broken down into HCHO, O<sub>3</sub>, CO<sub>2</sub> and various other species. Over land, isoprene and monoterpenes (comprised of two isoprene units) account for 50% and 30% of the OH reactivity respectively [35].

Ozone is an important precursor to HO, as excited oxygen atoms (O(<sup>1</sup>D)) are created through photolysis, which then go on to mix with water and form OH, as shown in this equation taken from Atkinson [5]:



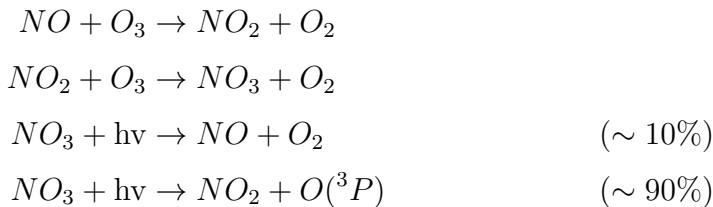
This shows how some of the O(<sup>1</sup>D) recycles back to Ozone, while some forms OH.

NB: The wavelength was updated to 350 nm in Atkinson and Arey [6].

In the late 90's it was understood that OH radicals are formed exclusively from photolysis of O<sub>3</sub>, HONO, HCHO, and other carbonyls (R<sub>2</sub>C=O) Atkinson [5]. Isoprene was thought to be a sink of OH until it was shown by [99] that the radicals are recycled. This recycling process is discussed in more detail in section ???. Monoterpene oxidation by O<sub>3</sub>, OH and NO<sub>3</sub> radicals may also form aerosols, with ozone forming the most particles [56].

Nitrate radicals NO<sub>3</sub> are also largely formed through ozone reactions. They are photolysed very rapidly during the day, with a lifetime of about 5 s [5]. If NO and

$O_3$  are both in the atmosphere, the following reactions [5] occur:



A build up of  $NO_3$  radicals can be seen at night, when the quick photolysis is not occurring [5, 14].

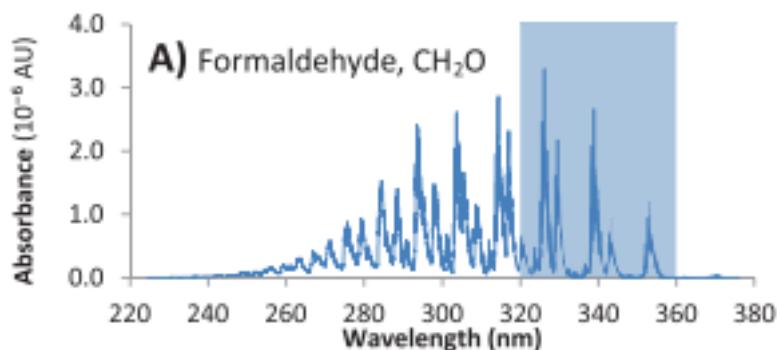
## 2.4 Formaldehyde (HCHO)

In the remote troposphere HCHO production is dominated by methane oxidation, while in the continental boundary layer (CBL) production is largely due to NMVOCs [2, 57].

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. It is dangerous at low levels, with WHO guidelines for prolonged exposure at 80ppb. In the continental boundary layer, HCHO enhancement is generally driven by short lived (< 1 hr) precursors driven by isoprene and other VOC emissions, with a lifetime of a few hours [57]. HCHO is used as an adhesive in plywood, carpeting, and in the creation of paints and wallpapers. Emissions in enclosed spaces can build up to dangerous levels, especially if new furnishings are installed [22]. At global scales HCHO in furniture is less important, as concentrations are driven by photochemical reactions with methane and other VOCs.

### 2.4.1 How HCHO is measured

There are a few ways to measure HCHO, including Fourier Transform Infra-Red Spectrometry (FTIR) and Differential Optical Absorption Spectroscopy (DOAS). 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 works for both in the home HCHO detection and global measurements from in-situ and remote sensing instruments [42, 39, 22]. As a trace gas HCHO interferes with light over a few wavelength bands, which allows instruments to detect concentrations along the path the light has travelled. Figure 2.5 shows the interference spectrum of HCHO as well as a typical band used to examine interference in the DOAS technique. One



**Figure 2.5:** HCHO spectrum, with a typical band of wavelengths used for DOAS path measurements. This is a portion of an image from Davenport et al. [22].

difficulty is that this interference is relatively small (HCHO is optically thin) and other compounds absorb light at similar wavelengths [22]. FTIR measurements can have a range of uncertainties, including systematic and random measurement errors and uncertainties in apriori shape factors and water profiles (eg: Franco et al. [33]). Multiple axis DOAS (MAX-DOAS) also examines the infra-red light interference. In Franco et al. [33], 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. Generally satellites use a DOAS based technique, and then chemical transport, and radiative transfer models are used to transform the non-vertical light path interference into vertical column amounts.

MAX-DOAS is a remote sensing technique which requires many DOAS measurements taken simultaneously. 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 2.6, which was taken from Lee et al. [62]. Recently MAX-DOAS has been used to examine HCHO profiles in the clean free troposphere [33, 105] as well as in polluted city air [62]. 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 [33, 62, eg.].

Other measurement techniques include chromatographic and fluorimetric methods, both of which differ widely from each other and the spectroscopic methods [44]). [44] examine a single air mass with 8 instruments using the four techniques (MAX-DOAS, FTIR, chromatographic, and fluorimetric), and show that reasonable agreements can be achieved. Generally the measurements were somewhat close, the five Hantzsch instruments agreeing to within 11% (after removing two perhaps faulty measurements), although different calibration standards were used. Titra-

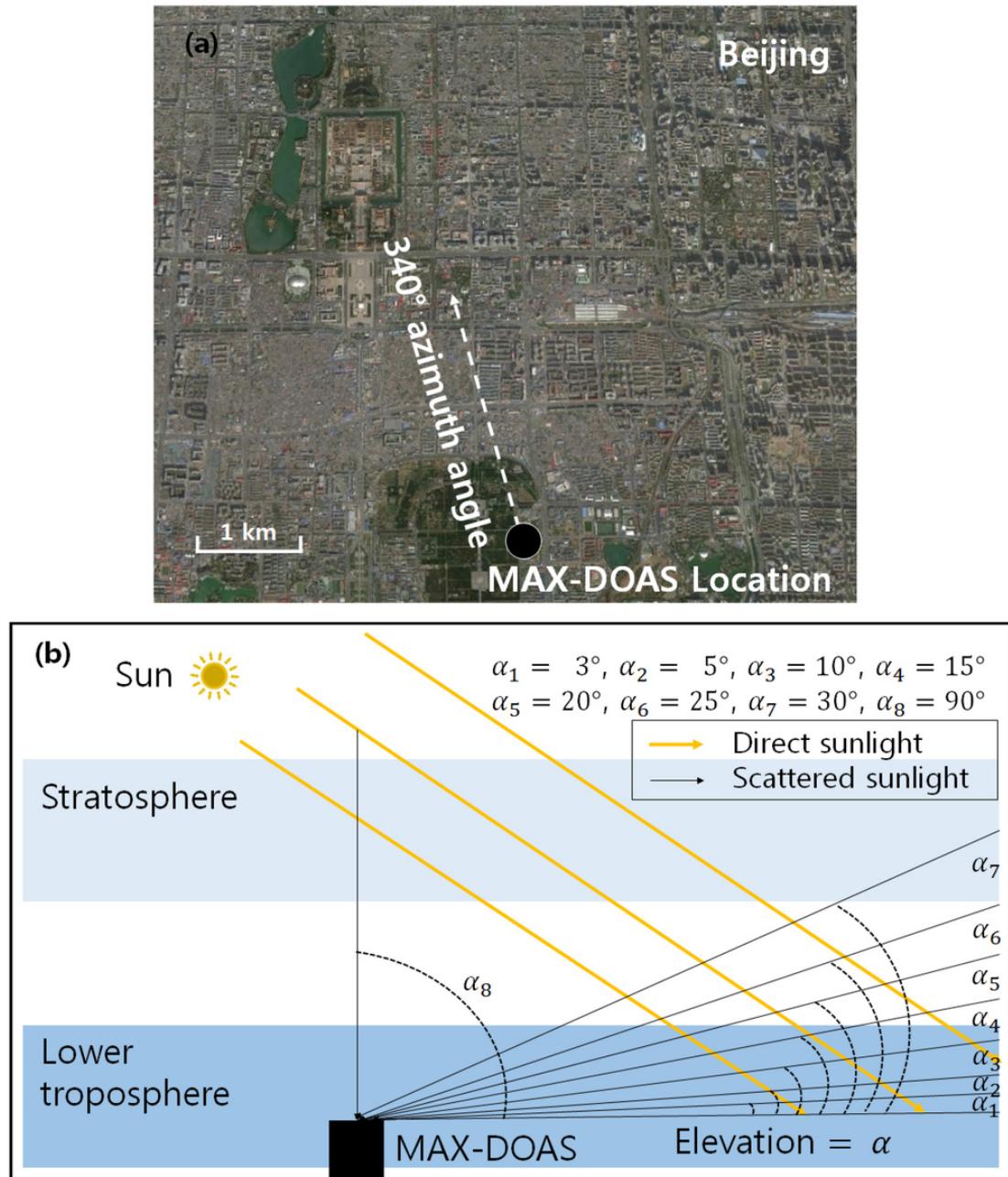


Figure 2.6: Image from Lee et al. [62].

tion for the different calibration solutions could not be resolved, and may account for absolute offsets up to 30%. These differences and non-uniformities even among identical instruments is part of the reason HCHO does not have a consistent network for global measurements like those for GHGs or Ozone [32].

### 2.4.2 Sources and sinks

In the atmosphere HCHO is primarily produced through the oxidation of methane ( $\text{CH}_4$ ) by the hydroxyl radical ( $\text{OH}$ ).  $\text{CH}_4$  concentrations are thought to be well constrained in models, with the ACCMIP comparison showing only  $\sim 3\%$  IQR [122]. Within the continental boundary layer, the major source of HCHO enhancement is VOC emissions reacting with OH radicals in the presence of  $\text{NO}_X$  [120, 80, 57]. There is a complex relationship between VOCs,  $\text{HO}_X$ , and  $\text{NO}_X$ , and with higher levels of  $\text{NO}_X$  the speed that VOCs are converted into HCHO increases, as does the HCHO concentration [121]. Isoprene is the main VOC precursor of HCHO in the continental boundary layer, except near fires or anthropogenic sources of HCHO and precursors [42, 57, 121].

Biomass burning (BB) can be a source of HCHO, and various other pollutants, precursors, and aerosols TODO: best early cites for this? [42]. Additionally HCHO is emitted into the atmosphere directly through fossil fuel combustion, natural gas flaring, ethanol refining, and agricultural activity [121]. Background levels of HCHO are due to methane oxidation, while enhancements to regional and continental HCHO are largely driven by isoprene emissions [42, 94, 107, 57]. Atkinson [5] 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.

HCHO has two major sinks, one being reactions with OH (oxidation), the other being photolysis [20, 120, 66, 57]. These reactions lead to a daytime lifetime of a few hours [5, 80]. Both these loss processes (photolysis, oxidation) form CO and hydroperoxy radicals ( $\text{HO}_2$ ), and have global significance to radiative forcing and oxidative capacity [33]. The other notable sinks are wet and dry deposition, although these are not as significant [5] (todo add more cites here).

In the past, HCHO levels were underestimated by models, often with large discrepancies, due to the poor understanding of methyl peroxy radical ( $\text{CH}_3\text{OO}$ ) chemistry [120]. [Marvin2017] found that discrepancies in modelled HCHO concentrations are primarily due to second and later generation isoprene oxidation chemistry.

### 2.4.3 Determining isoprene from HCHO

Isoprene is hard to measure directly due to its short lifetime and weak spectral absorption, instead formaldehyde is often used as a proxy [80, 34, 25, 75, 10, 57, 11]. Formaldehyde formed in the troposphere is mostly due to VOC (roughly one third each: methane, isoprene, others) oxidation. We can model this oxidation process in order to work out how much VOC is present based on the total HCHO. This requires among other things an idea of which VOCs are present and their yields of HCHO.

Formaldehyde (HCHO) is often used as a proxy to estimate isoprene emissions [75, 10]. Satellite based chemical concentrations rely on ground-based measurements and modelled data for validation. They provide various readings with daily global coverage which is not otherwise feasable.

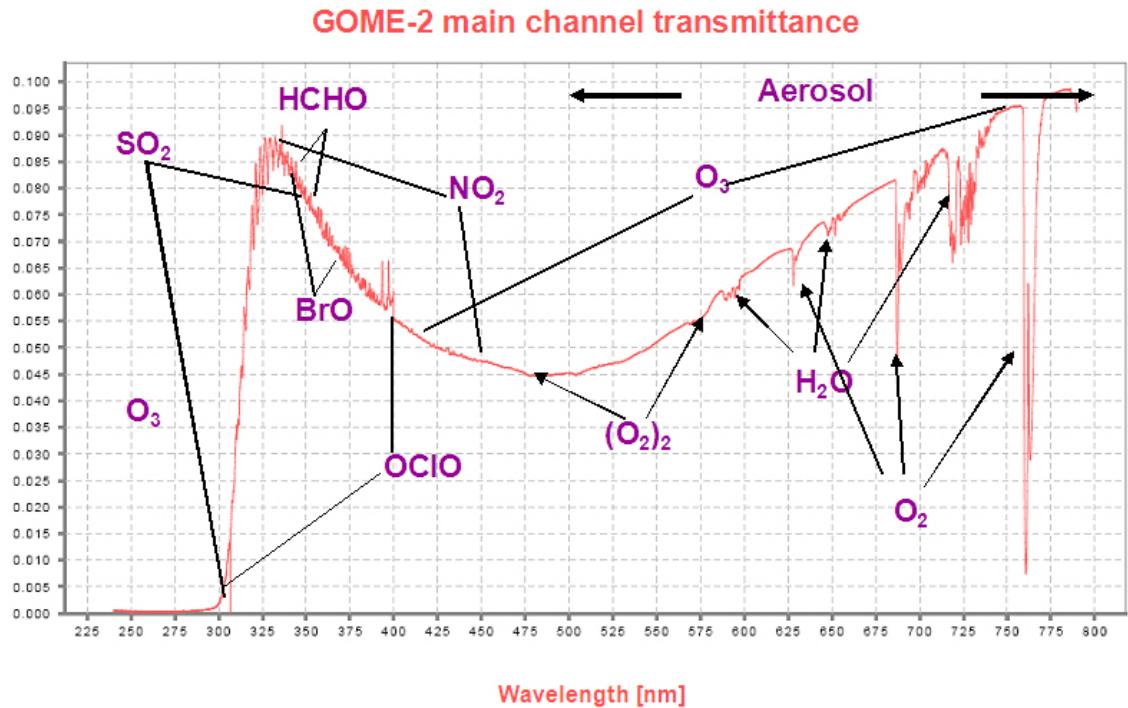
### 2.4.4 Satellite HCHO detection

TODO: Refactor this section so it's readable

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 can 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 2.7. Measurements done using DOAS often apply a forward radiative transfer model (RTM) such as LIDORT in order to determine a trace gas's radiative properties at various altitudes.

Satellites record near nadir (nearly vertical) reflected spectra between around 250-700 nm split into spectral components at around 0.3 nm in order to calculate trace gases including O<sub>3</sub>, NO<sub>2</sub>, and HCHO (eg: Leue et al. [65]). Several public data servers are available which include products from the satellites just mentioned, including NASA's Mirador (<http://mirador.gsfc.nasa.gov/>) and the Belgian Institute for Space Aeronomy (IASB-BIRA) Aeronomie site (<http://h2co.aeronomie.be/>).

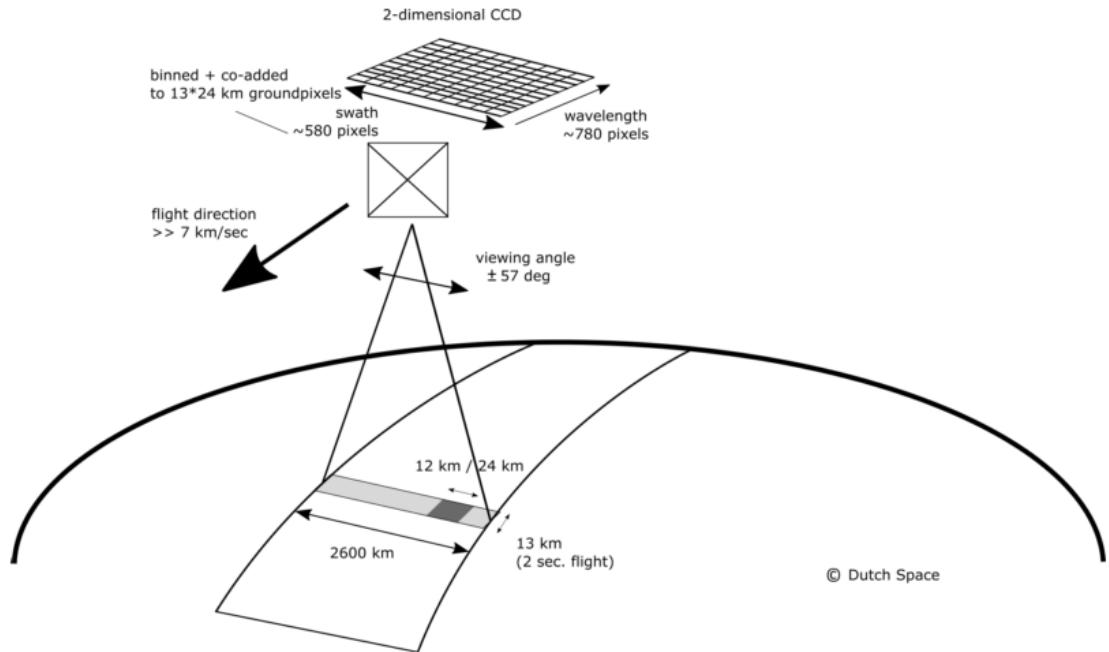
Instruments including MODIS on board the AQUA and TERRA satellites are able to determine aerosol optical depth (AOD), a measure of atmospheric scatter and absorbance. An AOD of under 0.05 indicates a clear sky, while values of 1 or greater indicate increasingly hazy conditions. This is an important atmospheric property allowing us to track dust storms and pollution events as well as determine



**Figure 2.7:** An example spectrum showing interferences used for species concentration measurements by GOME-2. Image by EUMETSAT and ESA [29].

where measurements from other instruments may be compromised by high interference. Satellite measured AOD requires validation by more accurate ground based instruments like those of AERONET which uses more than 200 sun photometers scattered globally.

The OMI instrument on board AURA has been active since July 2005, it records spectra from 264-504 nm using an array of 60 detectors with mid-resolution (0.4-0.6 nm). This band of wavelengths allows measurements of trace gases including O<sub>3</sub>, NO<sub>2</sub>, SO<sub>2</sub>, HCHO, and various other quantities like surface UV radiation. Recently [104] analysed the performance over time of the instrument and found irradiance degradation of 3-8%, changed radiances of 1-2%, and a stable wavelength calibration within 0.005-0.020 nm. They also provide a very nice summary of the OMI instrument with their Fig. 1 and Table 1, which I've copied to Fig. 2.8, as it shows the instruments spectral, temporal, and spatial resolutions. These changes are measured excluding the row anomaly (RA) effect, which is relatively stable since 2011, although it is still growing and remains the most serious concern. An analysis of the row anomaly by [49] state that OMI ozone columns remain suitable for scientific use, with recommendation for further evaluation. And analysis of OMI output by [104] concludes that data is still of high quality and will deliver useful information for 5-10 more years, with radiances only changing by 1 – 2% outside of RA impacted areas. The RA began in June 2007, with some cross-track rows seemingly blocked. The most likely cause is some instrument insulation partially obscuring the radiance



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 2.8:** Figure 1 and Table 1 from [104], 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 7km s<sup>-1</sup> and an exposure time of 2 s lead to a ground pixel size of 13km in the flight direction. The viewing angle of 114 leads to a swath width on the ground of 2600 km.” The table shows the optical properties for OMIs three channels.

port [104].

Soon even more HCHO data will be available in the form of geostationary satellite measurements [61]. [61] examine simulated geostationary measurements against GEOS-Chem column simulations to determine the most important instrument sensitivities. Geostationary satellites can provide temporally rich measurements over an area, as they are not sweeping around the earth but fixed relative to one latitude and longitude.

## DOAS

TODO: some of this is repeated in isoprene chapter satellite section.

The DOAS technique uses solar radiation absorption spectra to measure trace gases through paths of light. The RTM used in DOAS techniques is based on Beer’s law relating the attenuation of light to the properties of the medium it travels

through. Beer's law states that  $T = I/I_0 = e^{-\tau}$  with T being transmittance,  $\tau$  being optical depth, and I,  $I_0$  being radiant flux received at instrument and emitted at source respectively. Using  $\tau_i = \int \rho_i \beta_i ds$  gives us:

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

Where i represents a chemical species index,  $\rho$  is a species density(molecules per  $\text{cm}^3$ ),  $\beta$  is the scattering and absorption cross section area ( $\text{cm}^2$ ), and the integral over ds represents integration over the path from light source to instrument. The forward RTM used for satellite data products also involves functions representing extinction from Mie and Rayleigh scattering, and the efficiency of these on intensities from the trace gas under inspection, as well as accounting for various atmospheric parameters which may or may not be estimated (e.g. albedo).

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 should also account for instrument sensitivities to various wavelengths at various altitudes, and is unique for each trace gas under consideration.

## AMF

One important aspect of satellite retrievals of trace gas vertical columns is the Air Mass Factor (AMF), which characterises measurement sensitivity to a trace gas at various altitudes [95, e.g.]. These AMFs are measures of how radiance at the top of the atmosphere (TOA) changes with trace gas optical depths at specific altitudes [71]. Calculation of the AMF is important as it is multiplied against the estimated slant columns in order to give vertical column amounts.

Related to the AMF is the averaging kernal (AK), which is used to handle instrument measurements which are more or less sensitive to concentrations at different altitudes in the atmosphere. DOAS methods can be heavily influenced by the initial estimates of a trace gas profile (the *apriori*) 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 unrealistic *apriori* assumptions. One way to remove these *apriori* influences is through the satellite's AK (or by using AMFs), which takes into account the vertical profile of the modelled trace gas and instrument sensitivity to the trace gas [28, 95]. An measurement should be compared to other measurements or models only when the affects of these influences can be accounted for.

Lorente et al. [71] show that AMF calculations can still be the largest source of uncertainty in satellite measurements.

### Satellite uncertainties

Satellite measurements of HCHO are relatively uncertain, however this can be improved by averaging over larger grid boxes or longer time scales. An example of this can be seen in Dufour et al. [25], where monthly averaging is used to decrease the measurements uncertainty. They examine HCHO in Europe, which is low; near the detection limit of satellite measurements. Taking monthly averages allows enough certainty that useful inversions can be determined to estimate the source emissions of HCHO. The finer nadir resolution of OMI (13 by 24 km<sup>2</sup>) compared to other satellites reduces cloud influence [80, 81]. Although the uncertainty in each pixel is  $\sim 2 \times 10^{16}$ , which is 5× higher than GOME, there are  $\sim 100 - 200\times$  as many measurements due to the smaller footprint and better temporal resolution of OMI, which allows a greater reduction of uncertainty with averaging [50, 81].

In cloudy, hazy or polluted areas measurements are more difficult to analyse [e.g. 94, 74]. Recent work by [118] showed that updating how the surface reflectivity is incorporated into satellite measurements can change the retrievals by 50 % in polluted areas.

In satellite HCHO products, concentrations over the remote pacific ocean are sometimes used to analyse faulty instrument readings. This is due to the expected invariance of HCHO over this region. For instance GOME (an instrument which measures trace gases on board the ERS-2) corrects for an instrument artifact using modelled HCHO over the remote pacific [**Shim2015**]. OMI HCHO products use a similar technique to account for sensor plate drift and changing bromine sensitivity [39]

For many places the tropospheric column HCHO measured by satellite is biased low, Zhu et al. [128] 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. [23] also found a low bias from 20 - 40% when comparing OMI and GOME2 observations against ground based vertical profiles, and Barkley et al. [9] determine OMI to be 37% low compared with aircraft measurements over Guyana. These bias can be corrected by improving the assumed apriori HCHO profiles which are used to calculate the AMFs of the satellite columns. [80] examine OMI HCHO columns over North America and determine overall uncertainty to be 40%, with most of this coming from cloud interference. [81] 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.

Uncertainty in the OMI satellite instrument is calculated by the Smithsonian Astrophysical Observatory (SAO) group using the uncertainty in backscattered radiation retrievals [39, 1]. Another method of calculating the uncertainty is used by the Belgian Institute for Space Aeronomy (BIRA) group, who determine uncertainty from the standard deviation of HCHO over the remote pacific ocean (TODO: use both these methods for HCHO section)[24, 23].

A full analysis of the AMF uncertainty in OMI measurements, as well as the structural uncertainty (between different systems of calculations applied to the same data) is performed by [71]. They show that in scenarios where the gas is enhanced in the lower troposphere, AMF calculation is the largest uncertainty in satellite measurements. In polluted environments the structural uncertainty is estimated at 42 %, or 31 % over unpolluted environments. The importance of apriori and ancillary data (such as surface albedo and cloud top height) is also shown, as it sharply affects the structural uncertainty. [71] determine the structural uncertainty using ensemble techniques on seven AMF calculation approaches used by different retrieval groups.

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  $\text{cm}^{-2}$  [15, 80]. The conversion factor for slant to vertical columns (AMF) calculation also suffers from errors; primarily from surface albedo, HCHO vertical profile apriori, aerosol, and cloud influence [80]. AMF uncertainties for GOME are calculated to be 1 to  $1.3 \times 10^{15}$  molecules  $\text{cm}^{-2}$  by Shim et al. [107].

#### 2.4.5 Glyoxal TODO: move somewhere fitting?

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. Isoprene has been used to estimate isoprene emissions (see 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) [79]. Recently [79] updated GEOS-Chem to include the prompt formation of glyoxyl and compared this with satellite and airplane measurements over the USA. The glyoxyl formed by isoprene oxidation has a prompt yield mechanism in low  $\text{NO}_x$  conditions which is lacking in HCHO formation. With coming geostationary satellites, which provide greater time resolved measurements of HCHO and  $\text{CHOCHO}$ , this mechanism could be used to

clearly show when low NO<sub>X</sub> isoprene chemistry is being undertaken [79].

Under high NO<sub>X</sub> conditions, glyoxyl forms rapidly, similarly to HCHO. However, glyoxyl also forms in low NO<sub>X</sub> environments both slowly (through isoprene epoxy-diols), and rapidly (through di-hydroperoxide dicarbonyl compound photolysation [18]. This process is similar to the proposed mechanisms for hydroperoxyaldehydes by Jozef et al. [55] and carbonyl nitrates [84].

## 2.5 Models

### 2.5.1 How can models help

Used to forecast weather, pollution, and fires. ...

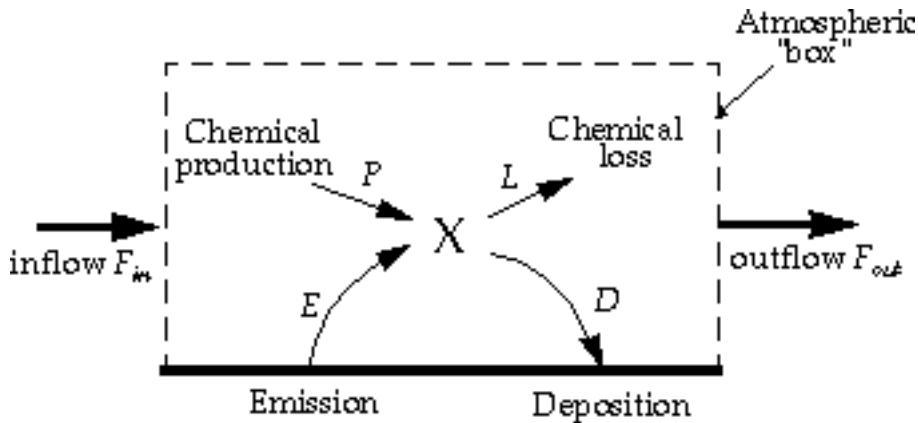
Models can fill the gaps between where we have measurements, as well as predict and avoid hazardous scenarios. They are used ideally to steer us away from unsustainable pollution and help complete our understanding from small to large scales. They can be used to increase measurement accuracy (for instance in satellite measurements) and determine where we lack information, as well as check 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. Currently, improved isoprene understanding is critical for effective air quality measuring [**Marvin2017**].

Chemical Transport Models (CTMs) simulate production, loss, and transport of chemical species. This is generally calculated using one or both of the Eulerian (box) or Lagrangian (puff) frames of reference. CTMs normally solve the continuity equations simultaneously with chemical production and loss for chemicals under inspection. The continuity equations describe transport of a conserved quantity such as mass, which, solved together with production and loss of a chemical forms the basis for a CTM. This basis enables a record of the chemical densities and transport over time as a model runs. 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 (the amount of  $q$  per unit area per unit time entering or leaving the field), and  $\sigma$  is the generation of  $q$  per unit volume per unit time. Note that  $\sigma$  can be positive or negative due to sources and sinks.

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



**Figure 2.9:** Standard box model parameters, image taken from Jacob [52].

and transport calculated and stored for arbitrary sequential steps in time at each location. The mass balance equation must be satisfied in any realistic long term box model and is as follows:

$$\frac{dm}{dt} = \sum \text{sources} - \sum \text{sinks}$$

$$= F_{in} + E + P - F_{out} - L - D$$

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 2.9. 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 can be complex and time consuming.

In many CTMs the isoprene emissions are calculated elsewhere with their own models (EG: Guenther et al. [40]). These estimates can then be used as boundary conditions. Trace gases with short lifetimes and complex chemistry such as isoprene are often hard to measure which makes verifying model estimates difficult.

### 2.5.2 Relevant model frameworks

Atmospheric chemistry models (ACMs) require various inputs and can be sensitive to ozone and oxidative parameterisations. TODO: read more Christian 2017, TODO: put some more generic ACM info here.

#### Box models

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. 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 [102]. 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, aqueus phase, and heterogenous reactions. [102] For the numerical integration, MECCA uses the KPP software [103], 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. CAABA/MECCA been implemented for various calculations including ozone chemistry throughout the atmosphere in [125].

MECCA could also be used as the chemistry mechanism for a more complex, 3-dimensional model [e.g. 53]. The connection is established via the MESSy interface (<http://www.messy-interface.org>) developed by [54] 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).

TODO: continued description of CAABA MECCA box model

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 generally requires a meteorological model to provide winds, and other transport mechanisms.

### Chemical transport

GEOS-Chem has a meteorological model coupled to a chemical model, which splits the world into a three dimensional grid (connected boxes). GEOS-Chem is a well supported global, Eulerian CTM with a state of the science chemical mechanism, with transport driven by meteorological input from the Goddard Earth Observing System (GEOS) of the NASA Global Modeling and Assimilation Office (GMAO). GEOS-Chem simulates more than 100 chemical species from the earth's surface up to the edge of space (0.01 hPa) and can be used in combination with remote and in-situ sensing data to give a verifiable estimate of atmospheric gases and aerosols. It was developed, and is maintained, by Harvard University staff as well as users and researchers worldwide. Several driving meteorological fields exist with different resolutions, the finest at 0.25 by 0.3125° horizontally at 5 minute time steps with 72 vertical levels.

Global CTMs are often run using one or several emission models (or the output

from them) to determine their boundary conditions. GEOS-Chem has boundary conditions based on several meteorological and emissions inventories, the following are the versions of theses used by GEOS-Chem v 10.01. Meteorological fields can be driven by NASA's GEOS-5 data ( $0.5^\circ \times 0.666^\circ$ ) [Chen2009], which exists up to 2013, or GEOS-FP data ( $0.25^\circ \times 0.3125^\circ$ ). Fire emissions come from the GFED4 product [36]. Anthropogenic VOC emissions come from the EDGAR inventory, while biogenic VOC emissions are coupled to the MEGAN model TODO:cites. The estimated biogenic VOC emissions are important to the work done in this thesis and are discussed in somewhat more detail in section ??.

The Model of Emissions of Gases and Aerosols in Nature (MEGAN) is one of the more commonly used natural emission models [83]. (TODO: more cites which say this/use MEGAN)

### Land based emissions

MEGAN can provide an estimate of biogenic emissions of various chemicals including isoprene and monoterpenes. It “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.” [41]. It allows parameterisation of various BVOC emissions, with descriptions given in [41]. Instructions to run version 2.1 are available at [http://lar.wsu.edu/megan/docs/MEGAN2.1\\_User\\_GuideWSU.pdf](http://lar.wsu.edu/megan/docs/MEGAN2.1_User_GuideWSU.pdf), and a version using the Community Land Model (CLM) is available at <http://www.cesm.ucar.edu>. It uses meteorological fields from the Weather Research and Forecasting (WRF) modelling system. Version 2.1 (updated from 2.0 [40]) includes 147 species, in 19 BVOC classes, which can be lumped together to provide appropriate output for mechanisms in various chemical models.

MEGAN was developed as a replacement for two earlier canopy-environment emission models (BIES and GEIA), and initially included a simple canopy radiative transfer model, which parameterised sun-lit and shaded conditions through a canopy. Early models didn't account for abiotic stresses, such as drought, prior rainfall and development processes, although these influenced species specific emissions by more than an order of magnitude [89]. MEGAN includes global measurements of leaf area index, plant functional type, and photosynthetic photon flux density, from remote sensing databases [57]. Isoprene emissions were based on temperature, leaf area, and light, but have since been updated to include leaf age activity [43], and a leaf energy balance model [40] 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 [109]. Soil moisture effects on isoprene emission are very important, and can drastically affect estimates.

MEGAN generally overpredicts isoprene and underpredicts monoterpene emissions in southeast Australia, with peaks and troughs captured but not at the right magnitude [27].

MEGAN has recently been analysed using 30 years of meteorological reanalysis information by Sindelarova et al. [109]. They estimate emissions of Biogenic VOCs (BVOCs) to be  $760 \text{ Tg(C)yr}^{-1}$ , 70% ( $532 \text{ Tg(C)yr}^{-1}$ ) of which is isoprene. This is similar to isoprene emission estimates from MEGAN itself, of  $400\text{-}600 \text{ Tg(C)yr}^{-1}$  [40].

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. [107] estimated global isoprene emissions to be  $\sim 566 \text{ TgC yr}^{-1}$ . This estimate is greater than initially thought and leads to decreased ( $\sim 10\%$ ) simulated OH concentrations to  $9.5\text{e}5 \text{ molec cm}^{-3}$ .

### Radiative transfer

#### 2.5.3 Using models to determine Isoprene emissions

TODO: Move stuff to here? Another chemical retrievable from satellite observation is Glyoxal, which can be used to further determine what sort of precursors to HCHO are being emitted [Miller2014, 79]. TODO: go through 2014 paper and see if it's easy to retrieve, then email Dr. Chris Miller.

### Factors affecting isoprene emissions estimates

Marais et al. [74] examine factors affecting isoprene emissions, showing how emissions are sensitive to various environmental factors. Their work used MEGAN [42] and GEOS-Chem to look at how these factors affect surface ozone and particulate matter in Africa. One of the important uncertainties seen in MEGAN within this work is the isoprene emissions due to plant type. Canopy level isoprene measurements are made using relaxed eddy accumulation (REA) at several sites in Africa. One plant type near a measurement site emits more than other species and it's actual distribution on a larger scale is completely unknown - leading to possible overestimations in MEGAN. Current emissions estimates require more validation against observations, and recently a comparison of two major VOC models (MEGAN and ORCHIDEE) was undertaken by [78] reiterating this requirement. In their work they examine model sensitivities and show that the important parameters are leaf area index (LAI), emission factors (EF), plant functional type (PFT), and light density fraction (LDF). There is high uncertainty in LAI and EF, which require more or improved measurements at the global scale. LDF parameterisation needs improvement and

these models require more PFTs. [27] analyse EF sensitivity of a high resolution model of atmospheric chemistry over southeast Australia, comparing isoprene and monoterene emissions against 4 separate campaigns. They show that the effect on total emissions is roughly linear and that no blanket EF changes are appropriate for all regions/seasons. They also mention that Australian eucalypt emissions are based on samples from young trees, which may emit more isoprene than older trees.

Stavrakou et al. [112] 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 isoprene emissions by a factor of two over the region. The Asian region is also shown to have a strong correlation with the Oceanic Nio Index (ONI), with positive anomalies associated with El Nio. 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 [112, 61].

### 2.5.4 Uncertainties

Here I will attempt to list and partially explain the major uncertainties models have in relation to VOCs, SOAs, and ozone.

#### Emissions Inventories

Using different emissions inventories in an ACM 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. [126] examine the affects on CO and HCHO when running simulations with two different inventories. TODO: find where I took notes about Zeng2015 and put them here.

#### Resolution

GEOS-Chem simulations are somewhat sensitive to the resolution at which you run. For example: TODO: read abstract/cite Wild and Prather 2006 show that reduced resolution increases OH concentrations and ozone production rates, [Wild2006]. [16] 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, however they continue at lower resolution to save computational time.

For many global scale analyses, errors from resolution are less important than those from chemistry, meteorology, and emissions [16].

### Chemistry mechanisms

There is still much work to be done in models to correctly simulate the various precursors to HCHO. Often HCHO is used as a way of checking if these precursors are correctly modelled since HCHO measurements are more readily available (for instance from satellites). GEOS-Chem (an ACM) has recently been analysed for sensitivity for ozone along with oxidants ( $\text{OH}$  and  $\text{HO}_2$ ) [16]. [16] found that GEOS-Chem ozone was most sensitive to  $\text{NO}_2$  photolysis, the  $\text{NO}_2 + \text{OH}$  reaction rate, and various emissions. They used GEOS-Chem v9-02, with  $4^\circ \times 5^\circ$  resolution, and while the low resolution adds errors in OH concentrations and  $\text{O}_3$  production rates, the errors from chemistry, meteorology, and emissions are much larger.

**Marvin2017** suggest that isoprene mechanisms in several contemporary models (including GEOS-Chem) are inadequate. They show that for a specific measurement campaign, the HCHO concentrations are underestimated in a way that can not be easily fixed through rate constant changes. Recently [**Marvin2017**] compared five global ACMs isoprene mechanisms by evaluating simulated HCHO mixing ratios compared to in situ measurements from the Southeast Nexus (SENEX) aircraft campaign (in southeastern USA). They compared five models (GEOS-Chem, CB05, CB6r2, MCMv3.2, and MCMv3.3.1) and found all of them underestimated the HCHO concentrations (by 15 – 30%).

### 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  $\sim 10\%$  of them being rain-clouds [56]. Wet scavenging performed in clouds not only depends on large scale cloud processes, but also on the microphysics of aerosols being scavenged, differing between aerosol sizes and hygroscopic properties.

### Soil Moisture

Australia has a unique climate, along with soil moisture, clay content and other important properties which affect VOC emissions. These properties are poorly understood in Australia due to the continents size and the relative sparsity of population centres, which make many areas very difficult or expensive to reach. Soil moisture plays an important role in VOC emissions, as trees under stress may stop emitting various chemicals. This is especially true for Australia due to frequent droughts and wildfires. The argument for improved understanding of land surface properties, specifically soil moisture, is an old one[**Chen2001**, 82, 101]. [101] show how quickly

soil moisture anomalies affect rainfall and other weather systems, while [Chen2001] specifically show how important fine scale soil moisture information is when modelling land surface heat flux, and energy balances. Modelled emissions are sensitive to soil moisture, especially near the soil moisture threshold (or wilting point), below which trees stop emitting isoprene and other VOCs completely as they can no longer draw water [11]. MEGAN accounts for soil moisture by applying it as an emission factor (EF) which scales the emission rate of various species. Sindelarova et al. [109] show reductions in modelled Australian isoprene emissions of 50% by accounting for soil moisture.

## 2.6 Aims

# Chapter 3

## Modelling with GEOS-Chem

### 3.0.1 List of runs and outputs used in my work TODO: good place for this?

TODO: Go through work process and clarify these items

1. UCX Satellite output (1300-1400LT):
  - (a) Create shape factors for AMF recalculation in OMI
  - (b) TODO: should I use this one for yields also?
2. Tropchem satellite output, modified to include tropopause height:
  - (a) recreate the AMFs for OMI when running code from Dr. Paul Palmer, modified by Dr. Luke Surl.
  - (b) in conjunction with an identical run where isoprene emissions are halved in order to determine smearing.
3. Tropchem(isoprene emissions halved) satellite output:
  - (a) used to determine smearing.
4. Tropchem daily tracer averages
  - (a) Currently used to determine yield and new emissions estimates (may switch to UCX satellite or trac avg in future?)

### 3.1 GEOS-Chem framework

GEOS-Chem is an atmospheric chemical model (ACM), using a 3-D grid of boxes with transport driven by the GEOS meteorological model and chemistry calculated in each box independently. Many of these terms are described in Section 2.5.2.

## 3.2 GEOS-Chem isoprene mechanisms

### 3.2.1 Outline

The isoprene reactions simulated by GEOS-Chem were originally based on [Horowitz1998]. This involved simulating  $\text{NO}_x$ ,  $\text{O}_3$ , and NMHC chemistry in the troposphere at continental scale in three dimensions, with detailed NMHC chemistry with isoprene reactions and products. The mechanism was subsequently updated by Mao et al. [73], who change the isoprene nitrates yields and add products based on current understanding as laid out in Paulot et al. [98] and Paulot et al. [99]. Further mechanistic properties, like isomerisation rates, are based on results from four publications: citeCrounse2011,Crounse2012,Peeters2010,Peeters2011. (TODO: check abstracts Peeters papers). [19] examines the isomerisations associated with the oxidation of isoprene to six different isomers ( $\text{ISO}_2$ ) formed in the presence of oxygen: isoprene +  $\text{OH} \rightarrow \text{O}_2$   $\text{ISO}_2$ . They determine rates and uncertainties involved in these reactions, and study the rate of formation of  $\text{C}_5$ -hydroperoxyaldehydes (HPALDs) by isomerisation. [17] examine the fate of methacrolein (MACR), one of the products of isoprene oxidation. Prior to this work MACR oxidation chamber studies were performed in high NO or  $\text{HO}_2$  concentrations, giving peroxy lifetimes of less than 0.1 s. In most environments this is not the case, and MACR products over various NO concentrations and peroxy radical lifetimes are determined in their work. [100] examine photolysis of hydroperoxy-methyl-buteneals (HPALDs, produced by isoprene isomerisation), which regenerates OH levels in areas with high isoprene emissions. Additionally, photolysis of photolabile peroxy-acid-aldehydes creates OH and improved model agreement with continental observations. The OH and HPALD interactions are central to maintaining the OH levels in pristine and moderately polluted environments, which makes isoprene both a source and a sink of OH TODO: cite and DL;<http://www.nature.com/ngeo/journal/v5/n3/full/ngeo1405.html>.

Formation of isoprene nitrates have an effect on ozone levels through  $\text{NO}_x$  sequestration, and the yields and destinies of these nitrates is analysed in Paulot et al. [98]. They use anion chemical ionization mass spectrometry (CIMS) to determine products of isoprene photooxidation. In a chamber with clean air and high NO concentrations, isoprene photooxidation is initially driven by OH addition, followed by  $\text{NO}_x$  chemistry (150 min - 600 min), and finally  $\text{HO}_x$  dominated chemistry. The yields of various positional isomers of isoprene nitrates is estimated, and pathways of their oxidation products is shown and used in the GEOS-Chem isoprene mechanism [98, 73].

In low  $\text{NO}_x$  conditions, isoprene oxidises to yield 70% hydroxyhydroperoxides (ISOPOOH), which then oxidises to create dihydroxyperoxides (IEPOX) with OH recycling maintaining the OH levels in the atmosphere [99]. In older models isoprene

produced ISOPOOH which then titrated OH, however, the loss of OH has not been seen in measurements [99, 73]. The isoprene mechanism in GEOS-Chem has been updated to include OH regeneration from oxidation of epoxydiols and slow isomerisation of ISOPO<sub>2</sub> [73].

ISOPN can be oxidised (by OH) to form nitrated organic products [98]. In low NO<sub>X</sub> ISOPOO reacts with HO<sub>2</sub> (producing hydroxy hydroperoxides, ISOPOOH), RO<sub>2</sub> (producing mainly MACR, MVK, and HCHO), or isomerises (1,5-H shift producing MACR, MVK, HCHO, or 1,6-H shift producing hydroperoxyenals HPALDs). ISOPOOH can be oxidised (by OH) to produce epoxydiols (IEPOX), precursors to SOA [99]. HPALDs can photolyse to regenerate OH and small VOCs [Wolfe2012, 19, 55] TODO: Check out crounse2011. See section 2.2.5 for more information.

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

Under low NO<sub>X</sub> conditions, ISOPO<sub>2</sub> may react with HO<sub>2</sub> to form ISOPOOH. In this case there is also production of HCHO (4.7%), MVK(7.3%), and MACR (12%). As stated in earlier; most ISOPOOH will form IEPOX (epoxydiols) after reacting with OH and lead to OH regeneration. The other mechanism in low NO<sub>X</sub> environments is unimolecular isomerisation of ISOPO<sub>2</sub>. This leads to production of hydroperoxyaldehydes (HPALDS), which generally photolyse and have an OH yield of 100%. Mao et al. [73] show that a lower (factor of 50) rate constant for ISOPO<sub>2</sub> isomerisation leads to better organic nitrate agreements with ICARTT.

This update leads to more accurate modelling of OH concentrations, especially in low NO<sub>X</sub> conditions common in remote forests. Prior to Mao et al. [72], measurements of OH in high VOC regions may have been up to double the real atmospheric OH levels, due to formation of OH inside the instrument. Mao et al. [72] examine an upgraded method of measurement, and compare these against a regional atmospheric chemistry model (RACM2), with the OH recycling updates from Paulot et al. [99] as discussed in prior paragraphs.

The updates to isoprene chemistry by Mao et al. [73], and those shown in [19, 17] are the last before version 11, which was not used in this work.

The full current mechanism is described online at [http://wiki.seas.harvard.edu/geos-chem/index.php/New\\_isoprene\\_scheme](http://wiki.seas.harvard.edu/geos-chem/index.php/New_isoprene_scheme).

### 3.2.2 Emissions from MEGAN

MEGAN simulates biogenic emissions of various gases including isoprene, based on various meteorological, land cover, and plant type parameterisations.

One of the important parameters in Australia is the soil moisture activity factor ( $\gamma_{SM}$ ), which can have large regional affects on the isoprene emissions [109, 11]. Generally if soil moisture is too low, isoprene emissions stop [**Pegoraro2004**, 90], 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 ( $\theta$ ):

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

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 **Chen2001**, while  $\theta_1$  is set globally based on **Pegoraro2004**.

In GEOS-Chem the isoprene emissions can be globally multiplied by a constant factor. By running the model two extra times, with the biogenic isoprene emissions turned off in one run and halved in another, while other parameters remain unchanged. These modified runs allow an estimate of model sensitivity to isoprene emissions and smearing impact as described in Section 5.5.7.

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

### 3.3.1 Installation and requirements

GEOS-Chem instructions for download, compilation, and running can be found in the user guide provided by Harvard: <http://acmg.seas.harvard.edu/geos/doc/man/>. In order to build and run GEOS-Chem a high-speed computing system is optimal, as globally gridded chemical calculations can take a long time to perform. I installed GEOS-Chem onto a suitably configured workspace on the National Computational Infrastructure (NCI, <http://nci.org.au/>). This workspace included access to compilers and libraries which are needed to build the Fortran based GEOS-Chem source code, and IDL, Python, and various editors and scripting languages to read, run, edit, and analyse both GEOS-Chem and its output.

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

### 3.3.2 Options

### 3.3.3 Tropchem mechanisms

### 3.3.4 UCX mechanisms

### 3.3.5 Run comparisons

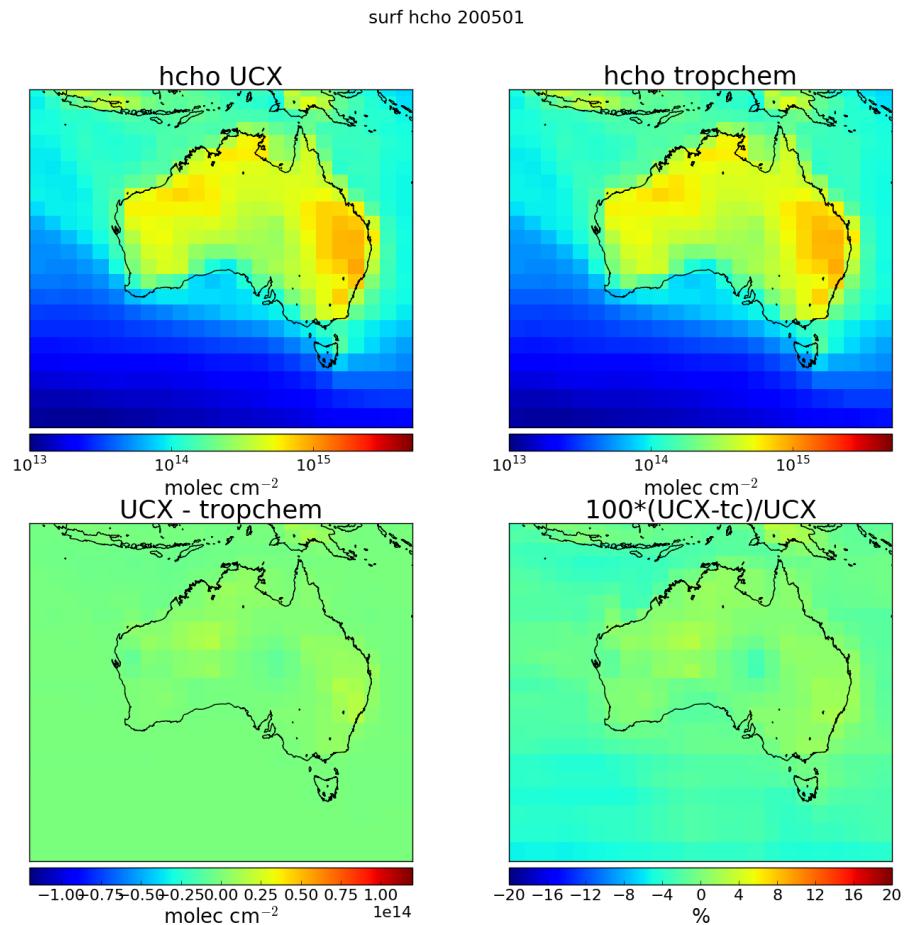
There are many options available when running GEOS-Chem depending on the desired chemistry, resolution, meteorology, and boundary conditions. Here we compare the model output with and without enabling the Universal tropospheric-stratospheric Chemistry eXtension (UCX). Both runs use  $2^{\circ}$  latitude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

At the surface level, GEOS-Chem output of HCHO does not differ much between runs with or without the Unified Chemistry eXchange (UCX). Figure 3.1 shows an example of the surface level HCHO amounts with and without UCX turned on. The differences do not exceed 3% over Australia for the averaged month of January, 2005.

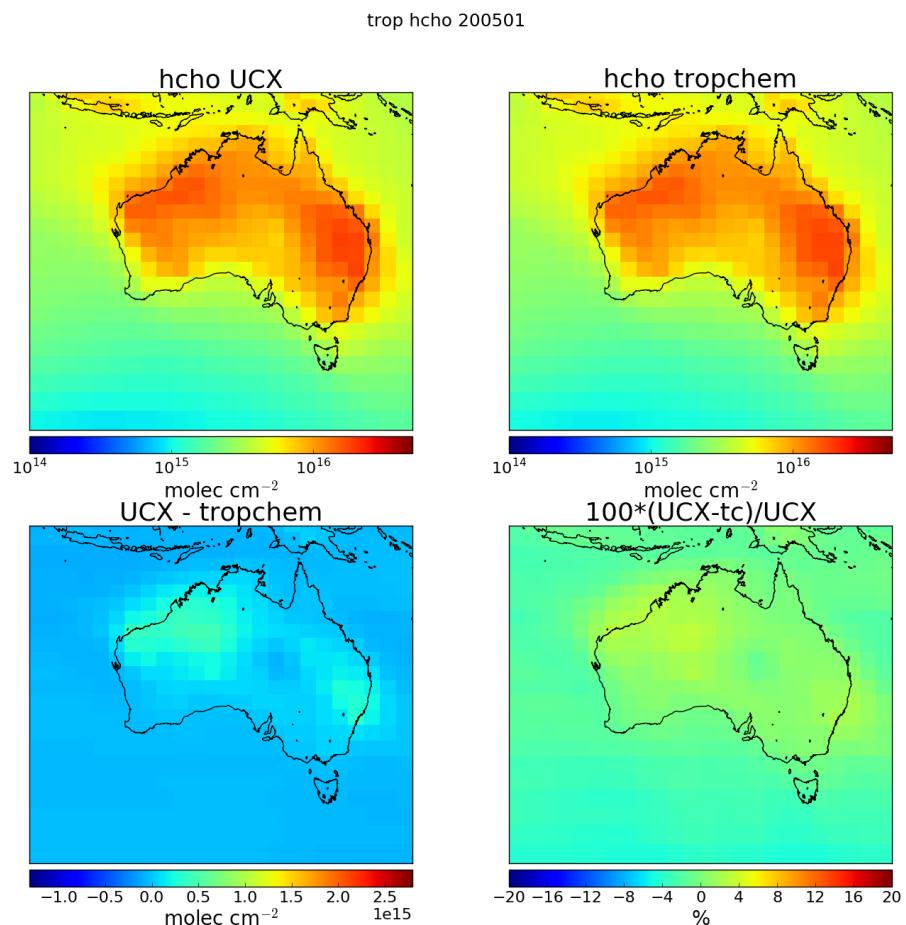
Figure 3.2 shows very similar results extended throughout the troposphere, with relative tropospheric column differences not exceeding 5% over Australia in January of 2005.

Figure 3.3 shows the differences in isoprene tropospheric columns over Australia. Here we start to see a higher relative difference in concentrations, although this is generally over the areas with less absolute concentrations. Very little isoprene is seen away from the continent (4-5 orders of magnitude less), due to the short lifetime of isoprene, and lack of emissions over the oceans. Generally isoprene is 0-30% higher over Australia when the UCX mechanism is turned on. This enhancement can be seen throughout the entire tropospheric column as shown by Figure 3.4.

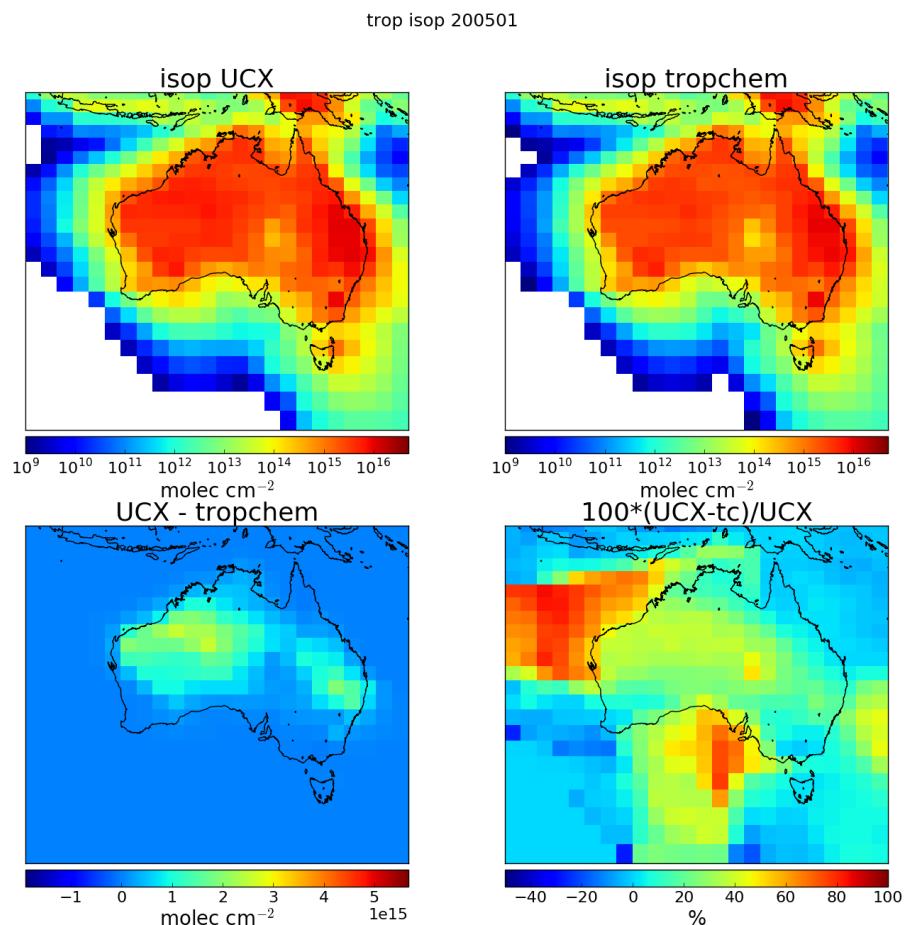
TODO: The difference in isoprene between UCX and tropchem is caused by WHAT? Comparing biogenic emissions of isoprene between runs should be very similar, as they are both estimated by MEGAN using the same input files and emission factor maps. Figure 3.5 shows these emissions and the differences between runs, which is unexpectedly high. The differences are most prominent over Australia and North America, with emissions averaging TODO over Australia for UCX and



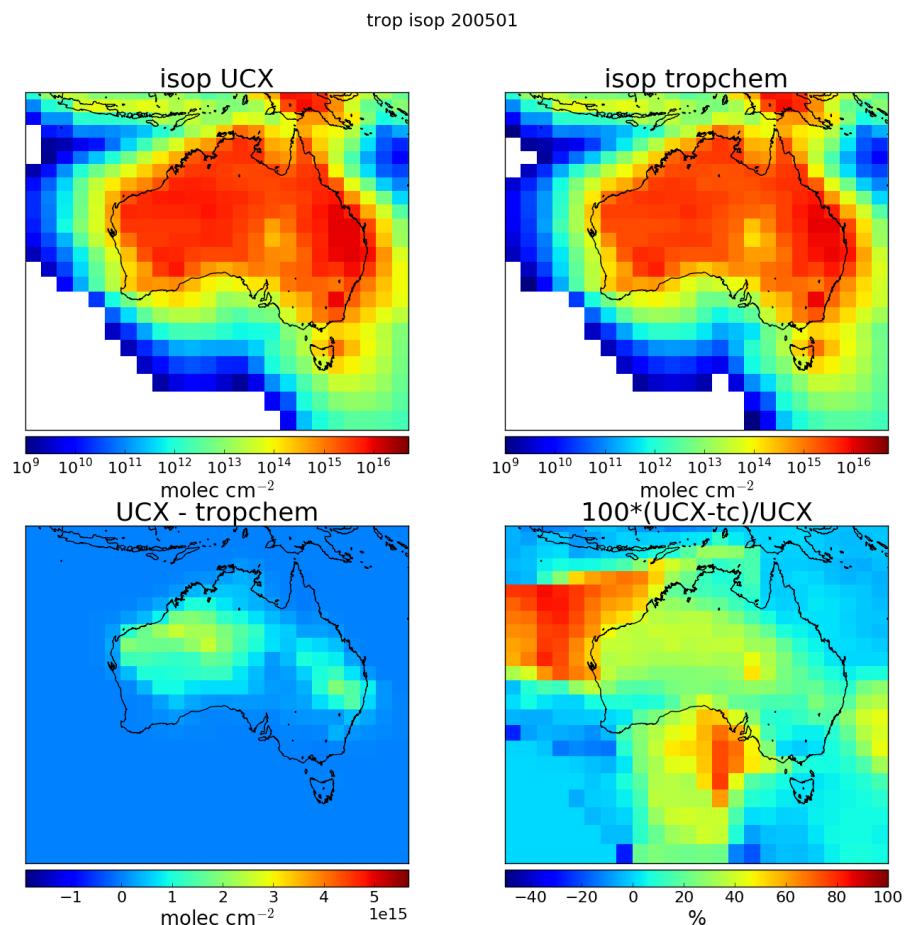
**Figure 3.1:** Surface HCHO simulated by GEOS-Chem with UCX (top left), and without UCX (top right), along with their absolute and relative differences (bottom left, right respectively). Amounts simulated by GEOS-Chem for the 1st of January, 2005.



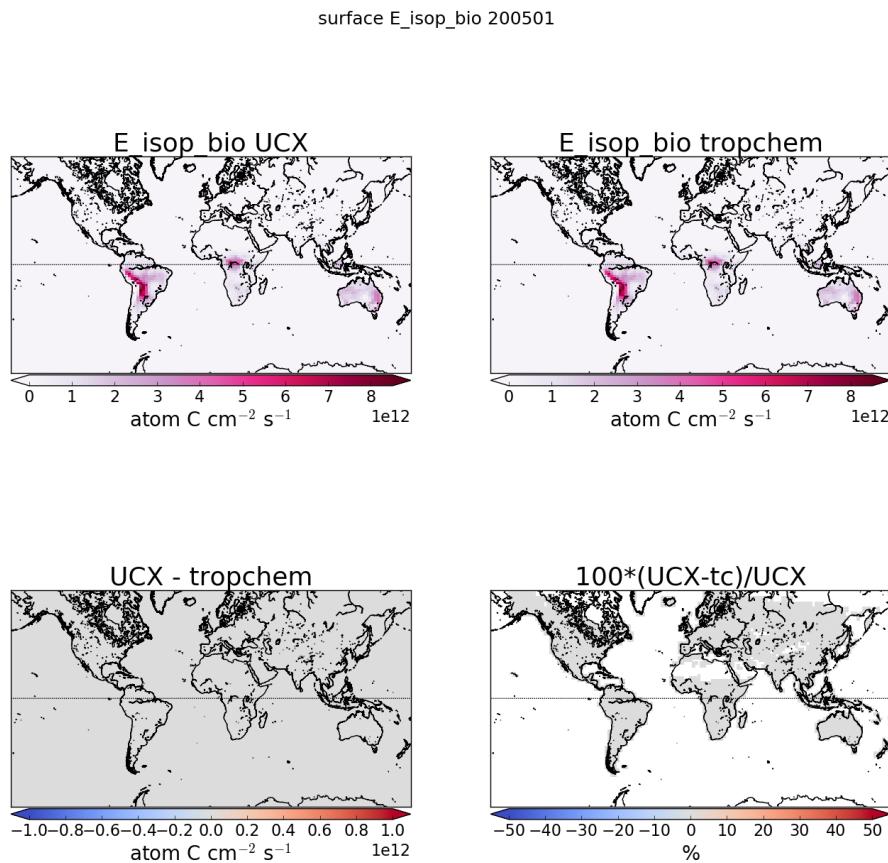
**Figure 3.2:** As figure 3.1, except looking at the tropospheric columns of formaldehyde.



**Figure 3.3:** As figure 3.1, except looking at surface ( $<\sim 120$  m) amounts of isoprene.



**Figure 3.4:** As figure 3.1, except looking at tropospheric columns of isoprene.



**Figure 3.5:** Top left, right: Modelled isoprene emissions with and without the UCX mechanism respectively, averaged over January 2005. Bottom left, right: absolute, relative differences respectively.

non UCX runs respectively. One possible explanation is the radiative affects of different ozone concentrations in the stratosphere calculated by the UCX mechanism. With higher ozone levels, less radiation would reach the surface leading to lower isoprene emissions (TODO: is this correct?). Figure TODO: shows the total column ozone between UCX and non-UCX runs, we can see that UCX has TODO: less or more ozone over Australia/USA in January. TODO: determine if that's the cause

### 3.4 Analysing output

# Chapter 4

## Stratospheric ozone intrusions

### 4.1 Introduction

Tropospheric ozone constitutes only 10% of the total ozone column but is an important oxidant and greenhouse gas which is toxic to life, harming natural ecosystems and reducing agricultural productivity. Over the industrial period, increasing tropospheric ozone has been estimated to exert a radiative forcing (RF) of  $365 \text{ mWm}^{-2}$  [114], equivalent to a quarter of the CO<sub>2</sub> forcing [31]. 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 [**Jacobson2000**]. The contribution of this source to overall tropospheric ozone budgets remains uncertain [**Skerlak2014**], especially in the southern hemisphere (SH). Models show that stratospheric ozone depletion has propagated to the upper troposphere [114]. However, work based on the Southern Hemisphere Additional OZonesonde (SHADOZ) network suggests stratospheric mixing may be increasing upper tropospheric ozone near southern Africa [**Liu2015, Thompson2014**]. 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 [**Danielson1968, Lefohn2011, Langford2012, Zhang2014**, 63, 67]. 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 [**Lin2012**, 67]. 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 [63]. Estimates of the overall contribution of STT to tropospheric ozone vary widely [**Galani2003**, **Lefohn2011**, e.g. 116, 113]. 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 [116]. In contrast, a more recent analysis of the Atmospheric Chemistry and Climate Model Inter-comparison Project (ACCMIP) simulations by Young et al. [122] 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 [83]. This wide range in model estimates exists in part because STT is challenging to be accurately represented, 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 [**Terao2008**]. Extra-tropical STT events most commonly occur during synoptic-scale tropopause folds [**Sprenger2003**, **Tang2012**, **Frey2015**] 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 [**Frey2015**], tropical cyclones [**Das2016**] and mid-latitude synoptic scale disturbances [**Mihalikova2012**, e.g. 116]. STT events have been observed in tropopause folds around both the polar front jet [**Vaughan1994**, **Beekmann1997**] and the subtropical jet [**Baray2000**]. 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 [125] and the tropopause fold activity [**Akritidis2016**] 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 [**Tyrlis2014**]. They are also observed near cut-off lows [**Price1993**, **Wirth1995**], 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 [**Trickl2014**]. The long range transport of enhanced ozone can be facilitated by

upper tropospheric winds, with remarkably little convective mixing, as shown by [Trickl2014] who measure STT air masses two days and thousands of kilometres from their source. Cooper2004 also shows how STT advection can transport stratospheric air over long distances, with a modelled STT event spreading from the northern Pacific to the East coast of the USA over a few days.

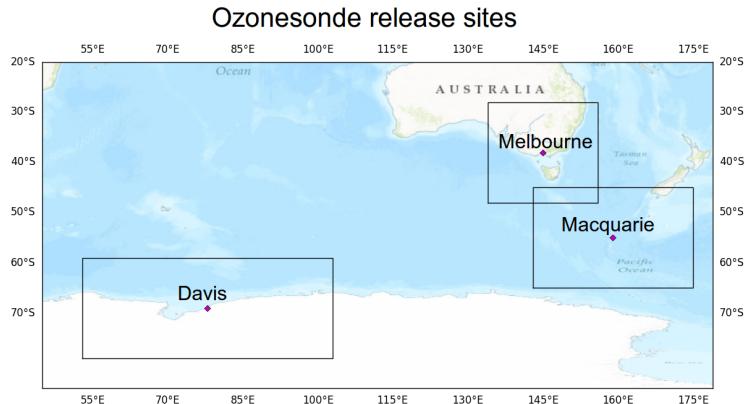
The strength (ozone enhancement above background levels), horizontal scale, vertical depth, and longevity of these intruding ozone tongues vary with wind direction and strength, topography, and season. While the frequency, seasonality, and impacts of STT events have been well described in the tropics and northern hemisphere (NH), observational estimates from the SH extra-tropics are noticeably absent in the literature. The role of STT in the SH remains highly uncertain due to the more limited data availability compared to the NH and the temporal sparsity of these datasets [Mze2010, Thompson2014, Liu2015].

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

## 4.2 Data and Methods

### 4.2.1 Ozonesonde record in the Southern Ocean

Ozonesondes provide a high vertical resolution profile of ozone, temperature, pressure, and humidity from the surface and up to 35 km. In the troposphere, the ozonesondes generally perform 150-300 measurements. Ozone mixing ratio is quantified with an electrochemical concentration cell, using standardised procedures when constructing, transporting, and releasing the ozonesondes (<http://www.ndsc.ncep.gov/>).



**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

[noaa.gov/organize/protocols/appendix5/](http://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 [Smit2007].

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 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 [Alexander2013]. 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.

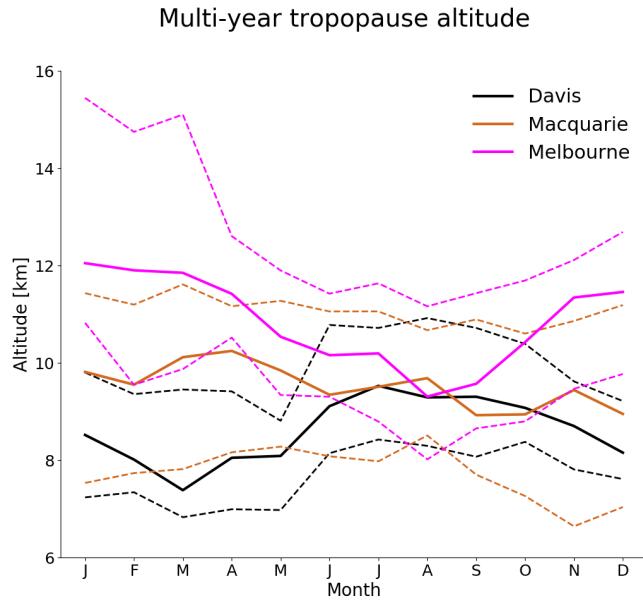
Two common tropopause height definitions are the standard lapse rate tropopause [WMO1957] and the ozone tropopause [Bethan1996]. The lapse rate tropopause is defined as the lowest altitude where the lapse rate (vertical gradient of temperature) is less than  $2^{\circ}\text{C km}^{-1}$ , provided the 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) [Bethan1996]:

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

The ozone tropopause may misdiagnose the real tropopause altitude during stratosphere-troposphere exchange; however, it is useful at polar latitudes in winter, where the lapse-rate definition may result in artificially high values for tropopause height [Bethan1996, Tomikawa2009, Alexander2013]. We require lapse rate defined tropopauses to be at a minimum of 4 km altitude. Another commonly used tropopause definition is determined with the use of PV (dynamical tropopause). In the extra-tropics the isosurface where  $\text{PV} = 2 \text{ PVU}$  ( $1 \text{ PVU} = 10^{-6} \text{ m}^2 \text{ s}^{-1} \text{ K kg}^{-1}$ ) is often used to define the tropopause, allowing the 3D representation of tropopause folds and other tropopause features in a sufficiently resolved model [Skerlak2014, Tyrlis2014]. 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 ( $38^{\circ}$  S) to 7-9 km at Davis ( $69^{\circ}$  S).

Figure 4.3 shows multi-year averaged ozone mixing ratios measured by ozonesonde over the three stations. Over Melbourne, increased ozone extending down through the troposphere is apparent from December to March and from September to November. The increased tropospheric ozone in these months is due to STT (in summer), and possible biomass burning influence (in spring), both discussed in more detail in the following sections. Over Davis and Macquarie Island, tropospheric ozone is higher between March and October, although the seasonal differences are small

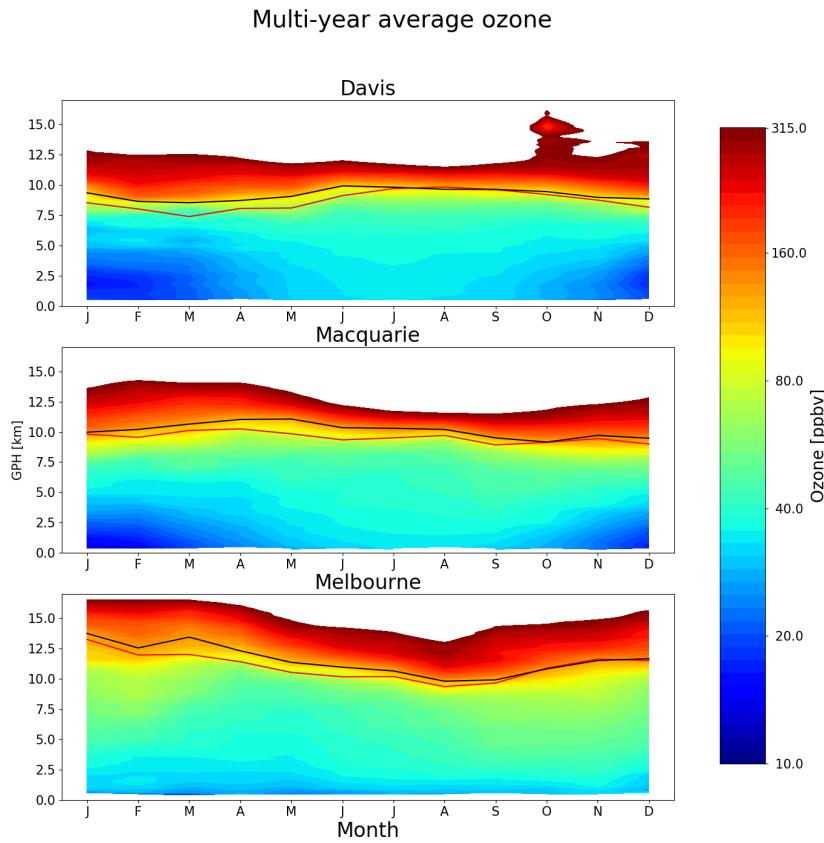


**Figure 4.2:** Multi-year monthly median tropopause altitude (using the ozone defined tropopause) determined from ozonesondes measurements at Davis (2006-2013), Macquarie Island (2004-2013), and Melbourne (2004-2013) (solid lines). Dashed lines show the 10th to the 90th percentile of tropopause altitude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

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 [**Lelieveld2000**]. NO<sub>2</sub> stratospheric observations have been conducted in the Southern hemisphere at Lauder, Macquarie Island and Arrival Heights [**Struthers2004**] which displays a winter minima in seasonality consistent with an ozone maxima. Influence from the ozone hole can be seen over Davis in October, with relatively low ozone levels extending up 5-6 km into the stratosphere.

#### 4.2.2 Model description

To provide regional and global context to the ozonesonde observations, we use the GEOS-Chem version 10-01 global chemical transport model [**Bey2001**], 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) [**Eastham2014**]. 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 [41] for biogenic emissions, the Emissions Database for Global Atmospheric Research (EDGAR)



**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.

version 4.2 for anthropogenic emissions, and the Global Fire Emissions Database (GFED4) inventory [36] 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 [Rienecker2008].

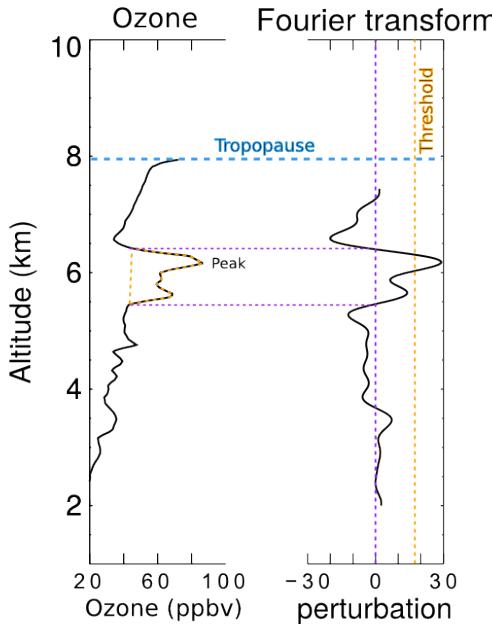
#### 4.2.3 Characterisation of STT events and associated fluxes

We characterise STT events using the ozonesonde vertical profiles to identify tropospheric ozone enhancements above a local background (in moles per billion moles of dry air, referred to here as ppb). The process is illustrated in Figure 4.4 on an example ozone profile. First, the ozone vertical profiles are linearly interpolated to a regular grid with 20 m resolution from the surface to 14 km altitude. Small vertical-scale fluctuations in ozone, which are captured by the high-resolution ozonesondes, can be regarded as sinusoidal waves superimposed on the large vertical scale background tropospheric ozone. As such, the interpolated profiles are bandpass-filtered using a fast Fourier transform [Press1992] to retain these small vertical scales, between 0.5 km and 5 km (removing low and high frequency perturbations). The high frequency perturbations are removed as they may represent noise in the measurements. The perturbations with scales longer than 5 km represent the vertical gradient of ozone concentration from the surface to the stratosphere. In what follows, these filtered vertical profiles are referred to as perturbation profiles.

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

### 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.

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 [Sprenger2003], 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 **Tang2010** to detect STT events from ozonesonde measurements. Their definition is based on subjective analysis of sondes released from 20 stations ranging in latitude from 35° S to 40° N. They identify an STT event if, starting from 5 km altitude, ozone exceeds 80 ppb and then within 3 km decreases by 20 ppb or more to a value less than 120 ppb. Their technique would miss many events due to the lower ozone concentrations found in the cleaner Southern Hemisphere.

#### 4.2.4 Biomass burning influence

The STT detection algorithm described in Sect. 4.2.3 assumes all ozone enhancements are caused by stratospheric intrusions. In some cases, however, these perturbations may in fact reflect ozone production in lofted smoke plumes. Biomass burning in southern Africa and South America has previously been shown to have a major influence on atmospheric composition in the vicinity of our measurement sites [92, 38, 26], particularly from July to December [93, 69]. 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 [**Jaffe2012**]. 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 [**AIRS3STD**]. CO is emitted during incomplete combustion and is an effective tracer of long-range transport due to its long lifetime [**Edwards2003**, 26]. In the Southern Hemisphere, biomass burning is the primary source of CO, making CO a good proxy for fire plumes [**Sinha2004**, **Mari2008**]. 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.

#### 4.2.5 Classifying synoptic conditions during STT events

Synoptic scale weather patterns are examined using data from the European Centre for Medium-range Weather Forecasts (ECMWF) Interim Reanalysis (ERA-I) [Dee2011]. This is done using the ERA-I data products over the three sites on dates matching the detected STT events. We use the ERA-I 500 hPa data to subjectively classify the events based on their likely meteorological cause, by visually examining each date where an event was detected. During STT occurrence, the upper troposphere is typically characterised by nearby cyclones, cut-off lows, or cold fronts. Over Melbourne and Macquarie Island, we find that frontal and low pressure activity are prevalent during STT events (see Sect. 4.3). Over Davis, the weather systems are often less clear, however we see a higher portion of probable cut-off lows. The stratospheric polar vortex may create tropopause folds without other sources of upper tropospheric turbulence such as low pressure fronts or cyclones [Baray2000, Sprenger2003, Tyrlis2014]. 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 [Sprenger2003].

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

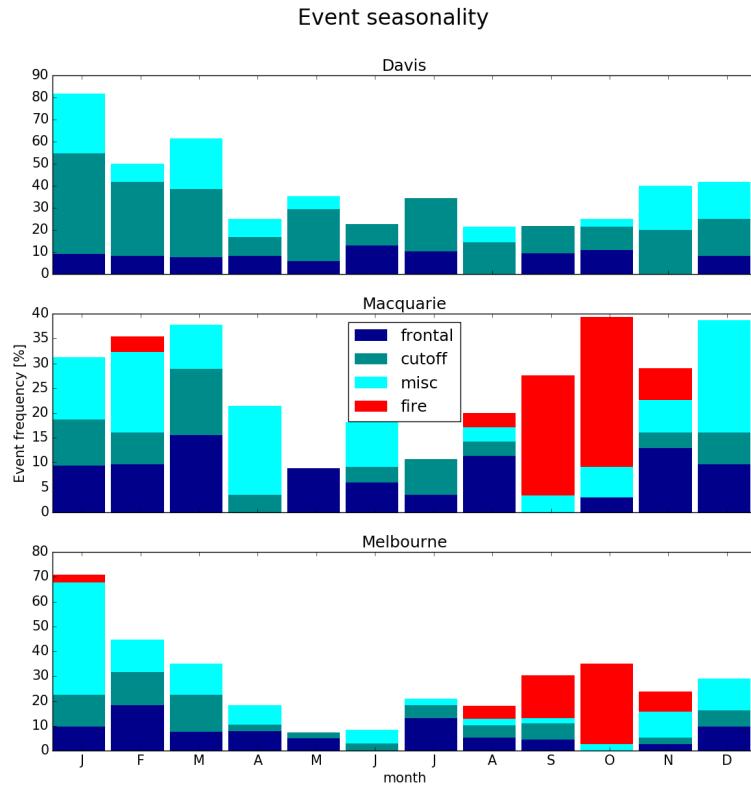
Site	Events	Cut-offs	Frontals	Misc	Fire
Davis	80	44	19	17	0
Macquarie Island	105	19	31	34	21
Melbourne	127	28	31	41	27

### 4.3 STT event climatologies

Figure 4.5 shows the seasonal cycles of STT frequency at Davis, Macquarie Island, and Melbourne. Frequency is determined as detected event count divided by total launched ozonesondes for each month. STT events in Figures 4.5-4.8 are coloured based on the meteorological classification described in Sect. 4.2.5, with events classified as either low pressure fronts (frontal, dark blue), cut-off low pressure systems (cutoff, teal), or indeterminate (misc, cyan). Events that may have been influenced by transported smoke plumes (Sect. 4.2.4) are shown in red. Ozonesonde releases are summarised in Table 4.1 and detected event counts are summarised in Table 4.2.

There is an annual cycle in the frequency of STT events (Fig. 4.5) with a summertime peak at all three sites. This summertime peak is due to a prevalence of summer low-pressure storms and fronts, which increase turbulence and lower the tropopause [Reutter2015]. 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 Tyrlis2014.

The SH summer maximum we see for STT ozone flux can also be seen in Fig. 16 of [Skerlak2014], 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 Skerlak2015, 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 [Skerlak2015], 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. [Wauben1998] 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



**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).

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. [Sprenger2003] examine modelled STT folds using ECMWF output over March 2000 - April 2001, and show that for this year there is a clear austral winter maximum, again over the  $20^{\circ}$  S to  $40^{\circ}$  S band. The winter maximum does not include Melbourne, or the southern ocean, which explains why we see a seasonality not readily evident in these global-scale studies.

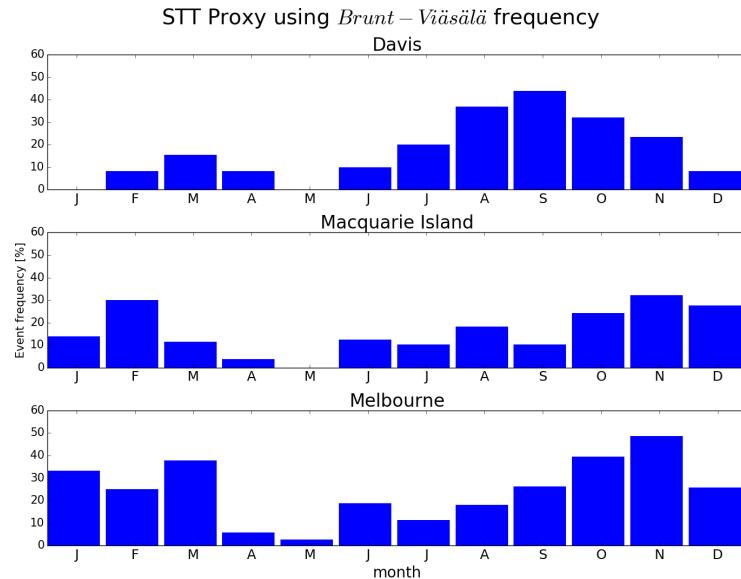
The measurement sites are not in the regions which have a clear winter maximum seen in [Sprenger2003], and the large scale winter maximum shown by all three studies seems to be dominated by the system in that region. The seasonality of our three sites is not driven by the larger STT system seen over 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 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 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.



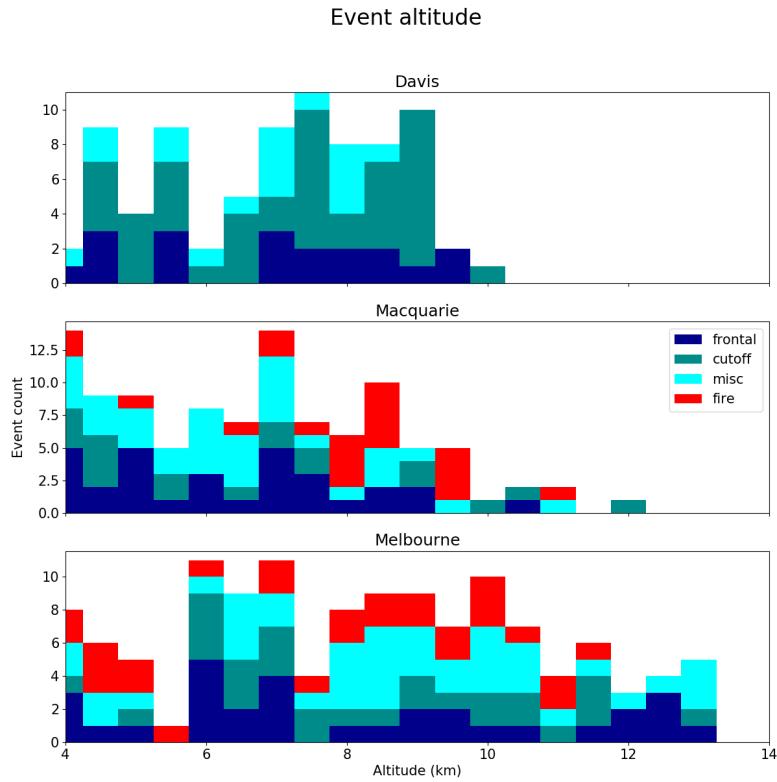
**Figure 4.6:** Seasonal distribution of STT events using the alternative STT proxy, obtained from consideration of the static stability at the ozone and lapse rate tropopauses, for Davis (2006-2013), Macquarie Island (2004-2013), and Melbourne (2004-2013).

## 4.4 Simulated ozone columns

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

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

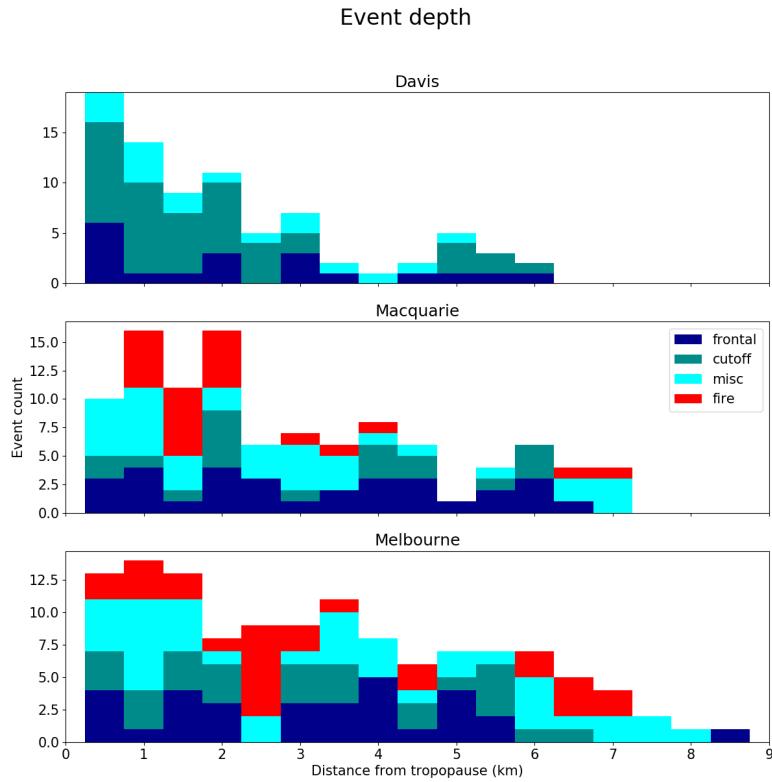
Figure 4.10 shows the observed and simulated ozone profiles at all sites, averaged seasonally. The model generally underestimates ozone in the lower troposphere (up to 6 km) over Davis, although this bias is less pronounced during summer.



**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.

Over Melbourne, ozone in the lower troposphere is well represented, but the model overestimates ozone from around 4 km to the tropopause. Over Macquarie Island we see model overestimation of ozone above 4 km, as well as underestimated ozone in the lower troposphere, suggesting that this region is influenced by processes seen at both of our other sites. Also shown is the mean tropopause height simulated by the model (horizontal dashed red line), which is always higher than the observed average, although this difference is not statistically significant. The effect of local pollution over Melbourne during austral summer (DJF) can be seen from the increased mean mixing ratios and enhanced variance near the surface. The gradient of the  $O_3$  profiles is steeper in the measurements than the model, at all sites during all seasons. Recently [Hu2017] 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



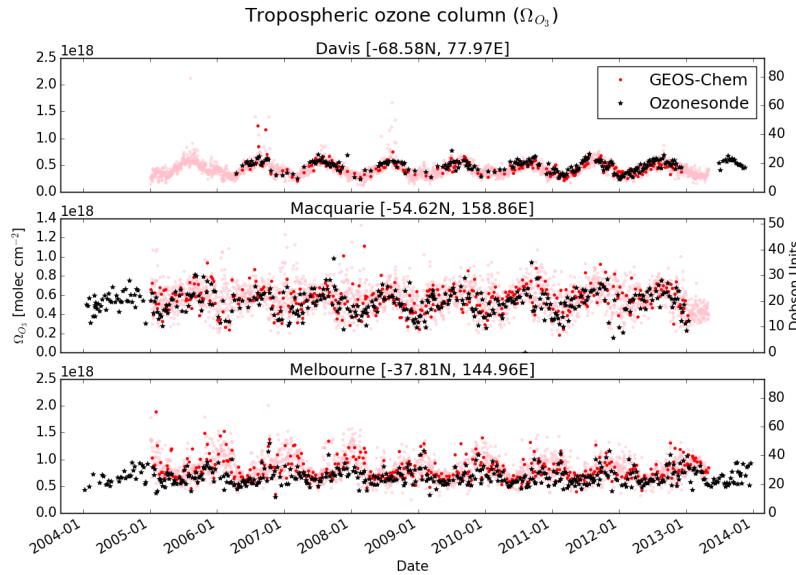
**Figure 4.8:** The distribution of STT events’ depths, defined as the distance from the event to the tropopause, at Davis (top), Macquarie Island (middle), and Melbourne (bottom), determined as described in the text. Events are coloured as described in Fig. 4.5.

the upper troposphere is too low to consistently allow detection of STTs, although in a few cases (e.g., Melbourne in Fig. 4.11) it appears that the event was large enough to be visible in the model output.

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

### 4.5.1 Method

We quantify the mean stratosphere-to-troposphere ozone flux due to STT events at each site based on the integrated ozone amount associated with each STT event (see Sect. 4.2.3). Events that may have been influenced by transported biomass burning are excluded from this calculation. Our estimate provides a preliminary estimate of how much ozone is transported from the stratosphere by the events detected by our method. The estimate is conservative for several reasons: it ignores secondary ozone peaks which may also be transported from the stratosphere, it



**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.

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.

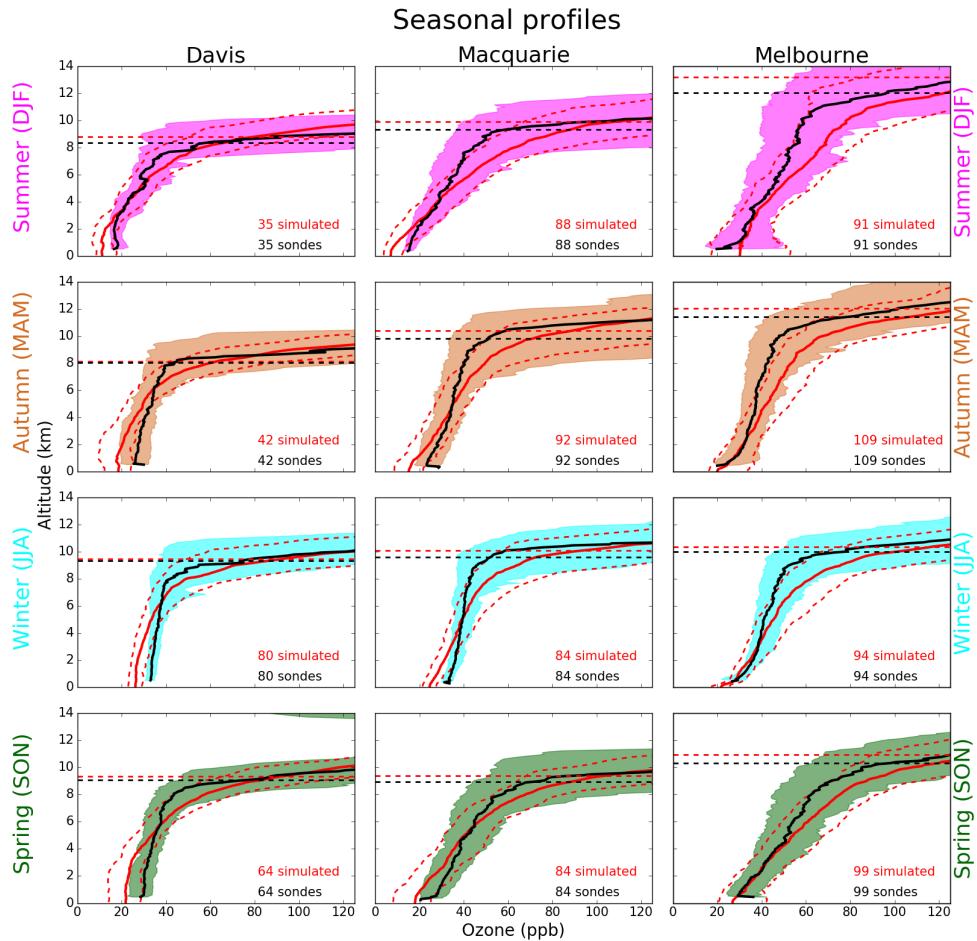
Observed tropospheric columns are calculated from the ozonesondes by calculating the ozone number density ( $\text{molecules cm}^{-3}$ ) using measured ozone partial pressure ( $P_{O_3}$ ) and integrating vertically up to the tropopause:

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

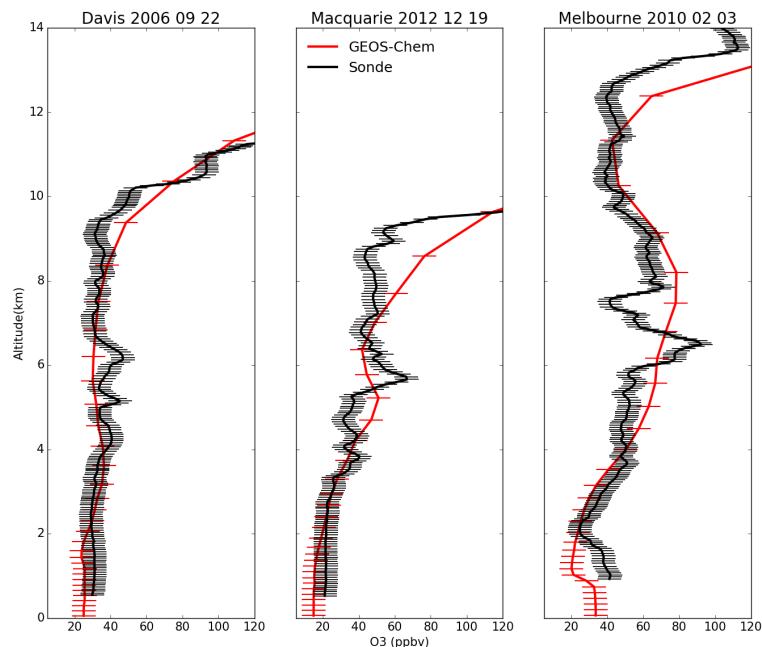
where  $z$  is the altitude (GPH),  $TP$  is the altitude at the tropopause,  $T$  is the temperature, and  $k_B$  is the Boltzmann constant.

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 strato-



**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.

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

The longevity of ozone events is very difficult to determine, and we have chosen 2 days as a representative number based on several examples in [Lin2012] where intrusions were seen to last from 1-3 days (occasionally longer) and an analysis of one large event by [Cooper2004] showing that most of the ozone had dispersed after 48 hours. Worth noting is the recent work of [Trickl2014], where intrusions are detected  $> 2$  days and thousands of kilometres away from their initial descent into the troposphere over Greenland or the Arctic. In those regions with high wind shear, mixing appears to be slower, which allows ozone intrusions to be transported further without dissipating into the troposphere. Relative uncertainty in our *M* term is set to 50%, as we assume these synoptic events to generally last from 1-3 days.

### 4.5.2 Results

The top panel of Fig. 4.12 shows the STT ozone enhancements, based on a vertical integration of the ozone above baseline levels for each ozonesonde where an event was detected. The area considered to be 'enhanced' ozone is outlined with yellow dashes on the left panel of Fig. 4.4. We find that the mean ozone flux associated with STT events is  $\sim 0.5\text{--}2.0 \times 10^{16}$  molecules cm<sup>-2</sup>. 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

**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%)

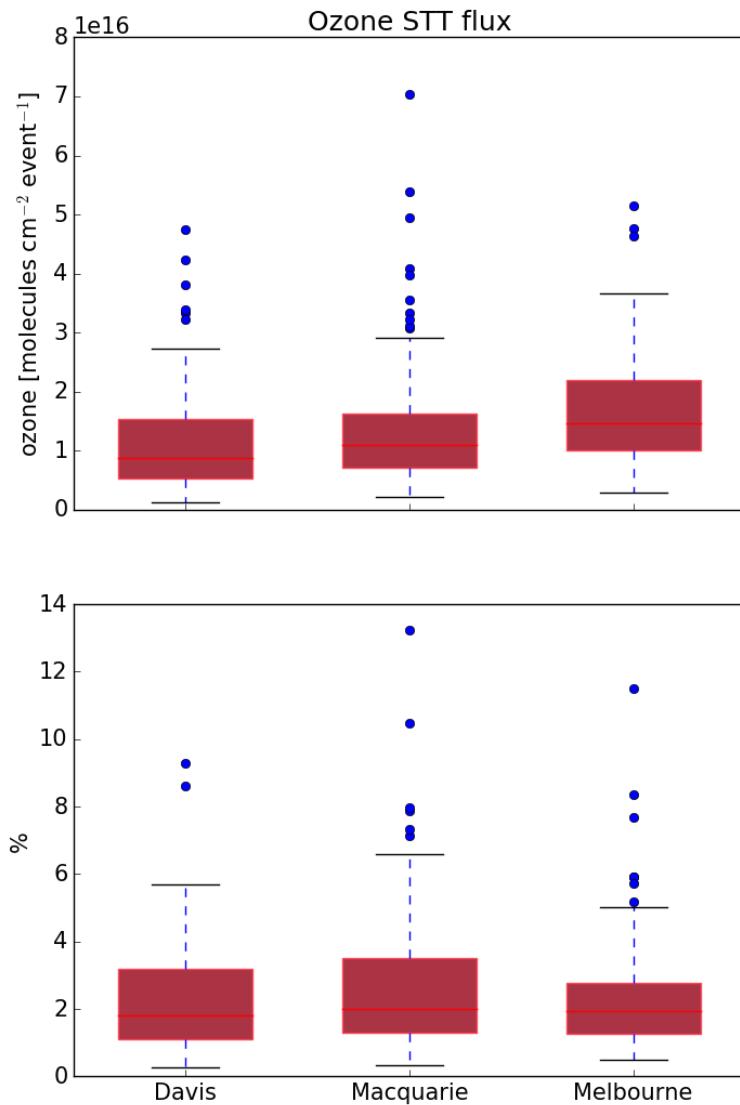
enhanced ozone column amount is used to determine the relative increase. At all sites, the mean fraction of tropospheric ozone attributed to STT events is  $\sim 1.0\text{--}3.5\%$ . On three separate days over Macquarie and Melbourne, this value exceeds 10%.

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

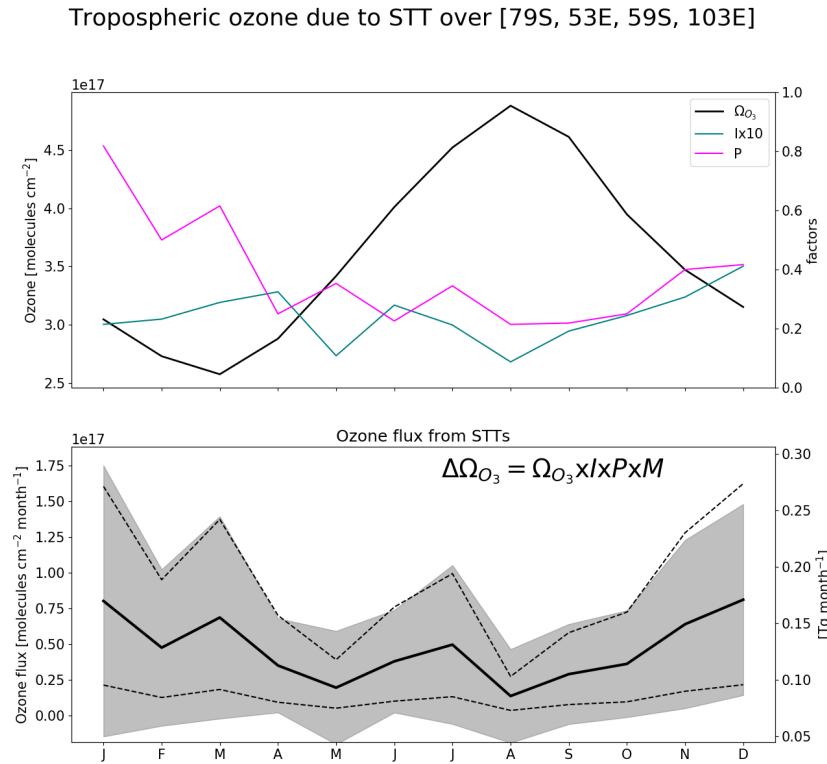
### 4.5.3 Comparison to literature

[Skerlak2014] show an estimate of roughly  $40$  to  $150 \text{ kg km}^{-2} \text{ month}^{-1}$  in these regions, over all seasons (see Fig. 16, 17 in their publication) while we estimate from  $0$  to  $180 \text{ kg km}^{-2} \text{ month}^{-1}$  STT impact, following a seasonal cycle with the maximum in austral summer. We estimate higher maximum flux over Melbourne, ( $178$ , and  $150 \text{ kg km}^{-2} \text{ month}^{-1}$  in January and February) than in either Davis ( $89 \text{ kg km}^{-2} \text{ month}^{-1}$  in March) or Macquarie Island ( $68 \text{ kg km}^{-2} \text{ month}^{-1}$  in January). Our calculated seasonal contributions, along with total uncertainty are shown in Table 4.3.

This result disagrees with several other studies which have found STT ozone fluxes in the SH extra-tropics are largest from autumn or winter to early spring. [Roelofs1997] 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

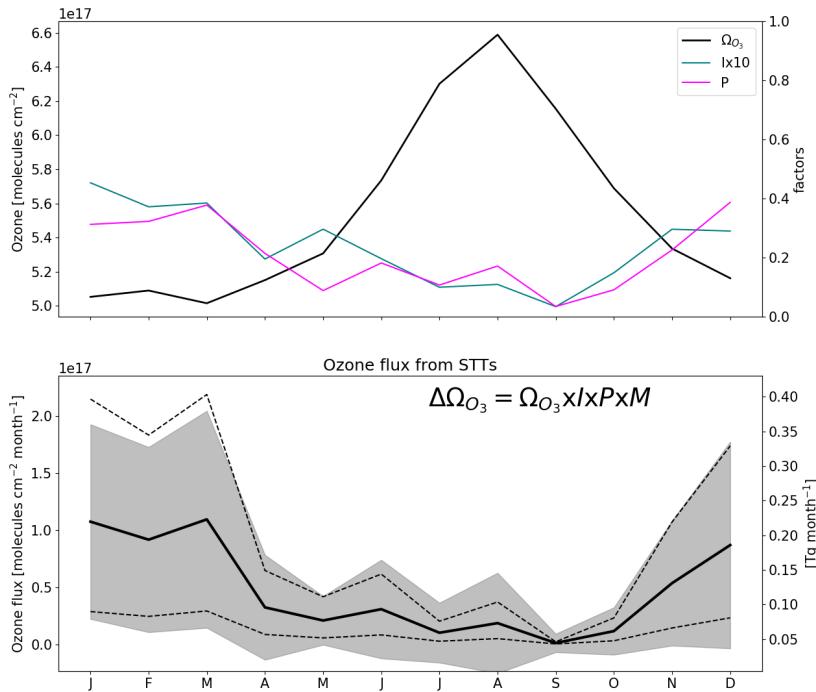


**Figure 4.12:** Top panel: tropospheric ozone attributed to STT events. Bottom panel: percent of total tropospheric column ozone attributed to STT events. Boxes show the inter-quartile range (IQR), with the centre line being the median, whiskers show the minimum and maximum, circles show values which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

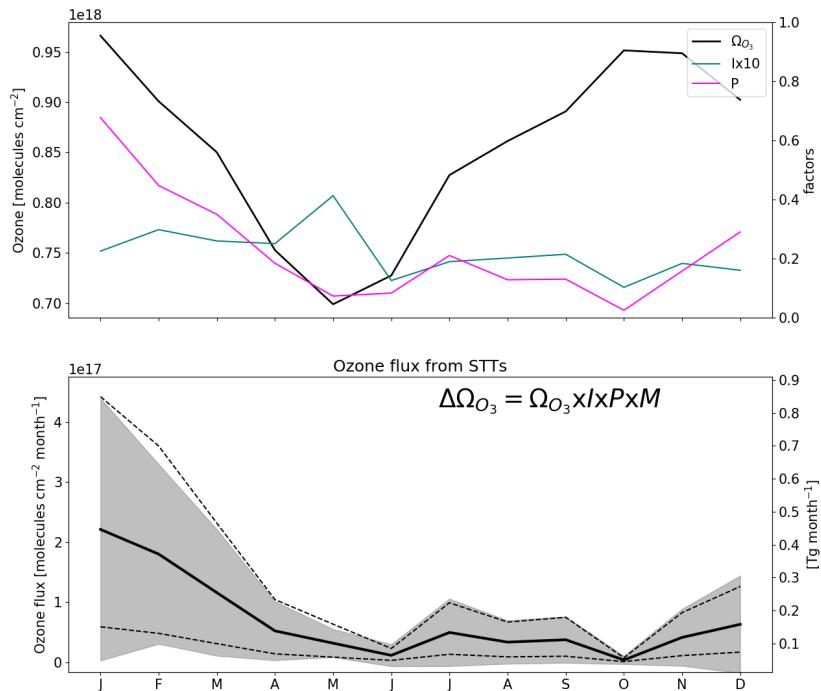


**Figure 4.13:** (Top) Tropospheric ozone, ( $I$ )mpact per event, and ( $P$ )robability of event detection per sonde launch, averaged over the region above Davis. The tropospheric ozone column  $\Omega_{O_3}$  (black, left axis) is from GEOS-Chem, while the STT probability  $P$  (magenta, right axis)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

Tropospheric ozone due to STT over [65S, 143E, 45S, 175E]

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

Tropospheric ozone due to STT over [48S, 134E, 28S, 156E]

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

winter, however flux is maximised in summer due to our detected event frequencies. [Elbern1998] 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. Olsen2003 used PV and winds from the GEOS reanalysis combined with ozone measurements from the TOMS satellite to estimate that the ozone flux between  $30^{\circ}$  S and  $60^{\circ}$  S is  $210 \text{ Tg yr}^{-1}$ , with the maximum occurring over SH winter. Liu et al. [69] model the upper tropospheric ozone and its source (emissions/lightning/stratospheric) over the Atlantic ocean between  $30^{\circ}$  S and  $45^{\circ}$  S, and suggest that most of this is transported from the stratosphere from March to September, which is when the subtropical jet system is strongest.

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

## 4.6 Sensitivities and limitations

### 4.6.1 Event detection

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

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

tropopause.

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

#### 4.6.2 Flux calculations

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

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

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

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

Considering the  $I$  factor, as discussed in here and in Sect. 4.6, there are several uncertainties in our method that are likely to lead to a low bias, such as the conservative estimate of flux within each event. Although there is little available data on SH ozone events for us to compare against, consider **Terao2008**, 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 [Hu2017] suggest that in general GEOS-Chem (with GEOS-5 met. fields) underestimates STT, with  $\sim 360 \text{ Tg a}^{-1}$  simulated globally, compared to  $\sim 550 \text{ Tg a}^{-1}$  observationally constrained. Transport uncertainty is very difficult to estimate with the disparate point measurements; it's possible that detected events are (at least partially) advected out of the analysis regions, which would mean we overestimate the influx into the region, and it is also possible that we are influenced by STT events outside the regions of analysis. Uncertainty in event longevity is set to 50%, however this implies a very simplistic model of event lifetimes. A great deal of work could be done to properly model the regional event lifetimes, however this is beyond the scope of our work.

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

## 4.7 Conclusions

Stratosphere-to-troposphere transport (STT) can be a major source of ozone to the remote free troposphere, but the occurrence and influence of STT events remains poorly quantified in the southern extra-tropics. Ozonesonde observations in the SH provide a satellite-independent quantification of the frequency of STT events, as well as an estimate of their impact and source. Using almost ten years of ozonesonde pro-

files 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.

## 4.8 Contributions and Acknowledgements

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

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

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

# Chapter 5

## Biogenic Isoprene emissions in Australia

### 5.1 Outline TODO: move into intro?

Currently this is an outline, TODO: Remove or put into aims.

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

With the vertical columns of biogenic HCHO we can infer the local (grid space) isoprene emissions using effective molar formaldehyde yield (In other continents around 2-3, or 1 in low NO<sub>X</sub> conditions) [94, 75, 11]. If we assume there is fast HCHO yield, so that the effect of chemical transport is minimal, and that HCHO and isoprene are at steady states, then we can calculate local yield from our CTM. Yield is calculated from the modelled slope between isoprene emissions and HCHO total column within each gridbox over Australia, as performed in [94], using modelled values between 1200-1400 LT which is around the overpass time of the OMI. This modelled yield is then used in conjunction with the recalculated OMI measurements in order to estimate isoprene emissions. To calculate emissions we use a reduced major axis (RMA) regression between modelled average values of the loss rates and total columns, an example is shown in figure TODO: figure with RMA of these over whatever time and space I end up using.

## 5.2 Introduction

One of the most popular emissions inventories for biogenic isoprene, the Model of Emissions of Gases and Aerosols from Nature (MEGAN). Global atmospheric studies often use MEGAN along with a chemical transport model (CTM) to examine transport, deposition, and various chemical processes in the atmosphere. Emissions of 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.

(MEGAN) is poorly calibrated for Australian conditions, their emissions of isoprene ( $C_5H_8$ ) may be overestimated, especially in the southeast. [85] compared MEGAN against emissions calculated using top down estimates from the GOME2 satellite measurements of formaldehyde. [111] showed that this overestimate may be a factor of 2-3 in January. [109] show how 50% of the isoprene emissions could be reduced by accounting for lower soil moisture. [27] 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. Satellite based emissions estimates may allow us to improve the models without requiring lots of hard work on calibrating MEGAN to the large data sparse continent of Australia. Emissions of monoterpenes ( $C_{10}H_{16}$ , two units of isoprene) may also be underestimated in southeastern Australia, which could lead to the unique scenario of neither type of emission dominating VOC chemistry over the forests [27].

Isoprene is emitted and enters the atmosphere in the gas phase, where it reacts with various chemicals, forming many new chemicals and reactions at various time scales. The primary first step for atmospheric isoprene is its oxidation by OH radicals leading to isoprene hydroxy peroxy radicals (ISOPOO) [Marvin2017, 121]. The fate of ISOPOO depends upon concentrations of  $NO_x$ , reacting with NO to form HCHO, MVK, MACR, and to a lesser extent organic nitrates (ISOPN). ISOPN can be oxidised (by OH) to form nitrated organic products [98]. In low  $NO_x$  ISOPOO reacts with  $HO_2$  (producing hydroxy hydroperoxides, ISOPOOH),  $RO_2$  (producing mainly MACR, MVK, and HCHO), or isomerises (1,5-H shift producing MACR, MVK, HCHO, or 1,6-H shift producing hydroperoxyenals HPALDs). ISOPOOH can be oxidised (by OH) to produce epoxydiols (IEPOX), precursors to SOA [99]. HPALDs can photolyse to regenerate OH and small VOCs [Wolfe2012, 19, 55] TODO: Check out crounse2011. Formaldehyde is produced with high yields in many of the isoprene reactions, and is often used as a check on how well these reactions are simulated by comparison against in situ measurements [Marvin2017]. TODO: less

details about isoprene cascade? Image showing pathway and all the HCHO yields?

Here we introduce how uncertain isoprene emissions are over Australia, and discuss literature which shows how the estimates may be too high. Section ?? describes the methods, models, and campaign data we use to determine and analyse isoprene emissions. The OMI measurements used in this research are recalculated using an updated estimate of HCHO profiles and validated against Wollongong total column measurements.

## 5.3 Data

## 5.4 Modelling

### 5.4.1 GEOS-Chem simulation

The GEOS-Chem global atmospheric chemistry model (V10.01) simulates and records up to 66 chemical species (tracers) in the standard run, at 2 by 2.5° horizontal resolution, with 47 levels up to the top of the atmosphere (TOA at 0.01 hPa). GEOS-5 meteorological fields from NASA's ... (TODO: ref and note) are used to drive transport and coupled with the chemical module of GEOS-Chem. MEGAN is used to determine biogenic emissions for our default GEOS-Chem simulation, with subsequent modifications based on top-down estimates made herein.

Output for an area averaged over 1200 - 1400 local time is saved for comparison and recalculation with satellite overpass records. These averages are used to calculate both the GEOS-Chem based AMF, and the modelled background HCHO over the remote pacific which is used in the reference sector correction for OMI column retrievals.

## 5.5 Methods

### 5.5.1 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.

TODO: Description of CAABA/MECCA and how it's used.

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. 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 [102]. For our purposes it's worth noting that MECCAs chemical mechanism includes basic  $O_3$ ,  $CH_4$ ,  $NO_X$ , and  $HO_X$  chemistry, as well as non methane hydrocarbon (NMHC) chemistry, considering gas phase, aqueus phase, and heterogenous reactions. [102] For the numerical integration, MECCA uses the KPP software [103], 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. CAABA/MECCA been implemented for various calculations including ozone chemistry throughout the atmosphere in [125]. 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).

TODO: Description of how scenario yields can be calculated. We prescribe parameters values approximating three separate (and relatively broad) scenarios: forest, urban, and scrubland. These parameters are shown in Table 5.1! The initial concentrations, the influx, (TODO: check this) and the deposition rates of various chemical species or families is prescribed, 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 its 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:

$$Y_{100} = 10^9 \times \frac{C_{PM} E_{inj} D_{inj}}{(N_A H_{PBL})} \quad (5.1)$$

**Table 5.1:** 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?			

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:

$$Y_{HCHO} = \frac{\Delta HCHO_{\text{Accumulated}}}{Y_{100}} \quad (5.2)$$

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.

### 5.5.2 Campaigns and datasets

TODO: these summaries. OMI summary

Daintree summary (P. Nelson)

MUMBA summary

SPS1,2 summary

### 5.5.3 Recalculation of OMI HCHO

The AMFs for OMI are recalculated using shape factors based on GEOS-Chem HCHO profiles, averaged between 1200 and 1400 local time (LT). The method used here largely follows that of Palmer et al. [95]. When comparing satellite observations to a chemical model, recalculation of the satellite AMF using modelled vertical gas profiles removes any bias introduced by differences from the a-priori shape factor to the model. The AMF is needed to transform the slant column, as viewed by the

satellite, into a vertical column:

$$AMF = \frac{\Omega_s}{\Omega_v} \quad (5.3)$$

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

The vertical shape factor  $S_z(z)$  is defined as a normalized vertical number density profile  $S_z(z) = \frac{\eta(z)}{\Omega_v}$  where  $\eta(z)$  is the number density in molecules  $\text{m}^{-3}$ . The AMF can be expressed as

$$AMF = AMF_G \int_0^\infty w(z) S_z(z) dz \quad (5.4)$$

Where  $w(z)$  is the scattering weights describing the sensitivity of the backscattered spectrum to the abundance of an absorber at altitude  $z$ . It's worth noting that in the OMI satellite product, the provided  $\omega(z)$  incorporates the  $AMF_G$  term and the equation 5.4 should be implemented without this term if using the satellite  $\omega$ . 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.

Two HCHO products are created, both using GEOS-Chem output at global 2 by  $2.5^\circ$  horizontal resolution. One uses the OMI product's  $\omega_z$  and equation 5.4 in order to calculate an AMF. While the other uses code provided by Dr. Paul Palmer, with alterations by Dr. Randal Martin, and Dr. Luke Surl to run LIDORT on the satellite slant columns and the GEOS-Chem output in order to calculate an AMF. These two calculations are compared over Australia in figure(s) TODO: Map comparison, regression, and time series once AMFpp is working properly. The effect of not recalculating the  $\omega_z$  is TODO: summarise changes between these calculations.

The AMF calculated using Dr. Palmer's code uses a more strict series of filters, leading to fewer satellite based HCHO columns and reduced coverage over Australia. Stricter filtering must be balanced against both coverage and the sensitivity of the AMF determination to recalculating  $\omega_z$ .

### 5.5.4 Calculation of Emissions

As is done in Palmer [94], Millet et al. [80], and Bauwens et al. [11], we assume that HCHO, and Isoprene columns are in a steady state, with no horizontal transport. Emissions of precursors are easy to calculate as long as we know the molar HCHO yields ( $Y_i$ ) and effective chemical loss rates ( $k_i$ ):

$$\Omega_{HCHO} = \frac{1}{k_{HCHO}} \sum_i k_i Y_i \Omega_i = \frac{1}{k_{HCHO}} \sum_i Y_i E_i \quad (5.5)$$

This works if there is fast HCHO yield, so that the effect of chemical transport is minimal. The background HCHO is calculated using measurements in the remote pacific at the same time and latitude. Table 5.2 shows the average yield calculated for Australia. (TODO: this table and some notes)

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

$$\Omega_{HCHO} = S \times E_{isop} + B \quad (5.6)$$

Where  $B$  is the background HCHO, and  $S = Y_{isop}/k_{HCHO}$  is determined monthly as the regression between  $\Omega_{HCHO}$  and  $E_{isop}$  on daily saved outputs from GEOS-Chem over Australia using 2 by  $2.5^\circ$  horizontal resolution. Modeled background emissions can be ignored here as they do not affect the slope calculation. Once we have calculated this slope, we use the same formula (Eqn. 5.6) to determine the isoprene emissions. By replacing  $\Omega_{HCHO}$  and  $B$  with OMI based values,  $E_{isop}$  is the only unknown.

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

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

ground (as  $E_{isop} = \frac{\Omega_{HCHO}-B}{S}$ ). These are set to zero, which increases the average by around TODO: X%. The measured background HCHO is the average concentration measured in the remote pacific at the same time.

Figure 5.1 shows the modelled isoprene emissions and column HCHO concentrations along with the RMA regression line, sampled from grid boxes over Australia for January 2005. Some affects from the low emissions in grid boxes which are largely oceanic can be seen and are handled by TODO: handle these and document here. Due to the low horizontal resolution of GEOS-Chem (2 by  $2.5^\circ$ , latitude by longitude), calculations from grid boxes on the coast which are largely oceanic need to be discarded as the change in HCHO is not dominated by emissions of isoprene, as is assumed for Eqn 5.6. A nested version of GEOS-Chem allows a much better analysis of coastal regions, at  $0.25$  by  $0.3125^\circ$  resolution.

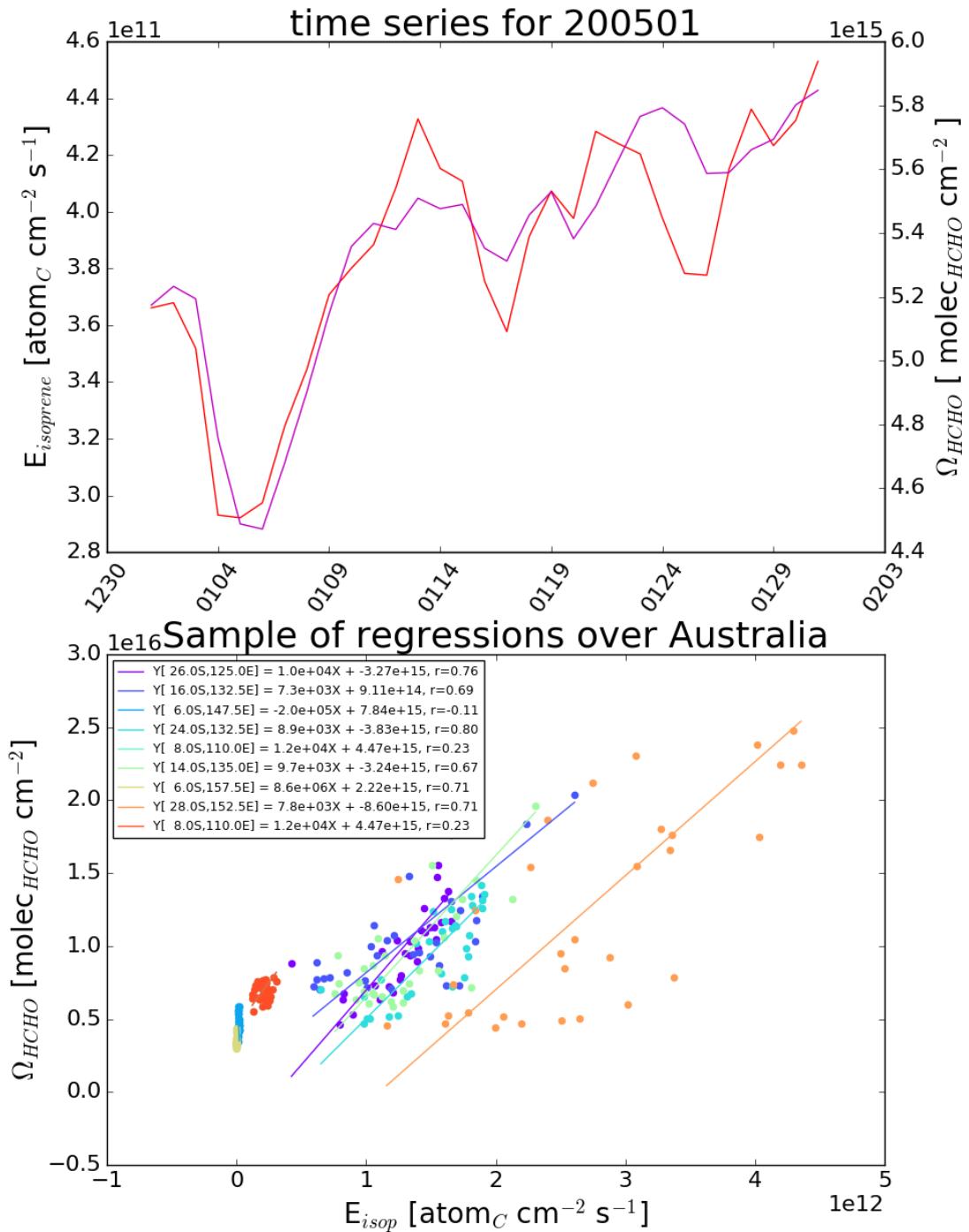
TODO: put this into results Using this modelled slope at  $2^\circ$  by  $2.5^\circ$  and applying it to equation 5.6 with B and  $\Omega_{HCHO}$  calculated using OMI satellite measurements provides a new estimate of isoprene emissions. Figure 5.2 shows the emissions calculated this way along with the Emissions output by GEOS-Chem averaged over January, 2005.

### 5.5.5 Emisions drivers

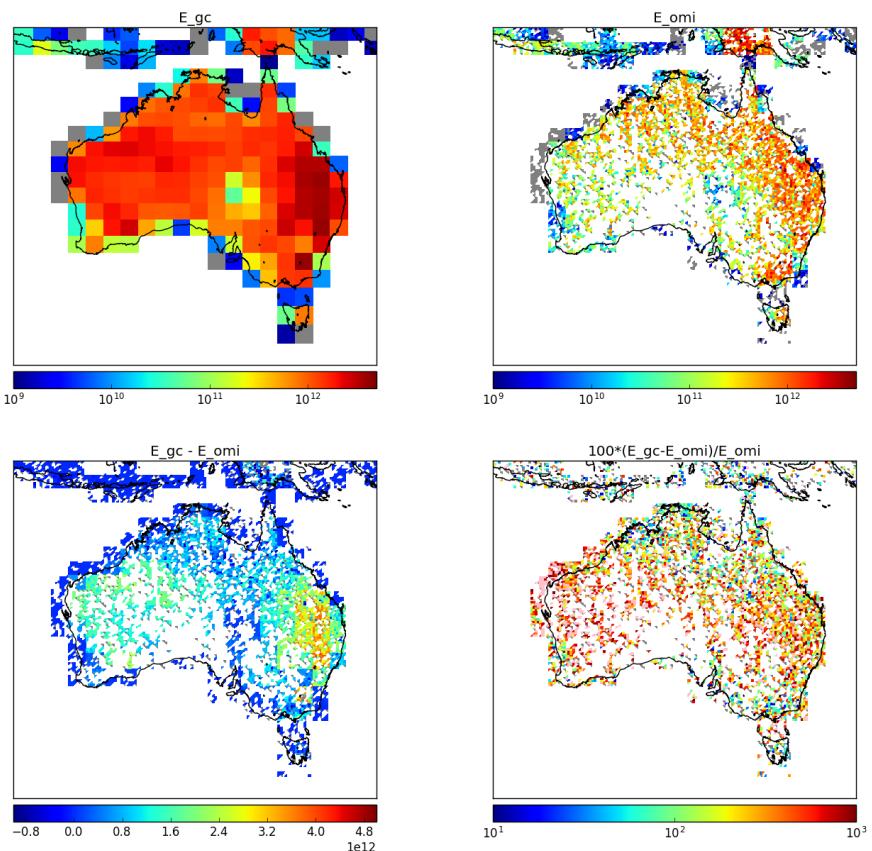
Calculated yields of HCHO can be classified using a box model which approximates specific environments, as described in Section ???. TODO: A table of different factors affecting emissions for three scenarios; urban, forest, shrublands is given in Table XX. The calculated yields for these scenarios is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

### 5.5.6 HCHO Products and yield

Australian forests are strong emitters of both isoprene and monoterpenes, which go on to form various products including secondary organic aerosols, oxygenated VOCs (OVOCs), ozone, OH, and  $\text{HO}_2$ . This production occurs over several steps, yields are often classed into at least two categories. First generation yield refers to the amount of HCHO produced per unit isoprene consumed by initial oxidation, total yield (sometimes molar yield) refers to time dependent yield of HCHO over multiple oxidation stages [121]. Wolfe et al. [121] define prompt yield as the change in formaldehyde measurement per unit change in initial isoprene emissions. Some argue that isoprene emissions are overestimated, due to the fact that they are based on relatively few measurements of isoprene emission factors [Winters2009, 32] TODO: read and cite paper mentioned in Fortems. Recently [Emmerson2017] showed



**Figure 5.1:** Top panel: isoprene emissions for January, 2005, shown in red, coplotted with tropospheric hcho columns, shown in magenta. Both series are daily averages over Australia. Bottom panel: (RMA) linear regressions from between emissions of isoprene and tropospheric hcho columns, sampled randomly from the  $2^\circ$  by  $2.5^\circ$  latitude longitude gridboxes over Australia for the month of January (2005).



**Figure 5.2:** 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.

that MEGAN estimates 3-6 times too much isoprene emissions, and 4 times too little monoterpenes when compared against 4 (relatively small scale) measurement campaigns in southeastern Australia.

Isoprene production of HCHO depends on several factors, importantly  $\text{NO}_X$  levels have a direct effect on the fate of VOCs in the atmosphere. At higher NO mixing ratios (at least a few hundred pptv), organic peroxy radicals ( $\text{RO}_2$ ) react mostly with NO. At low NO (less than 10's of pptv), reaction with  $\text{HO}_2$ , other  $\text{RO}_2$ , and isomerization dominate the fate of  $\text{RO}_2$ . In low  $\text{NO}_X$  environments, reported HCHO yields from isoprene are from XtoY%, while in high  $\text{NO}_X$  environments this value is XtoY% TODO: these values from table. For monoterpenes the yields are around X, Y% for low, high  $\text{NO}_X$  respectively. Emissions and yields for various species including some terpenes can be seen in table 5.2. Wolfe et al. [121] determine that going from  $\text{NO}_X = 0.1$  to 2.0 ppbv triples the prompt yield of HCHO, from 0.3 to 0.9 ppbv  $\text{ppbv}^{-1}$  due to isoprene, while the background HCHO doubles. They determine prompt yield as the change in HCHO per change in  $\text{ISOP}_0$ , using  $[\text{ISOP}]_0 = [\text{ISOP}] \exp(k_1[\text{OH}]t)$ ; where  $k_1$  is first order loss rate. This effectively relates HCHO abundance with isoprene emission strength. (TODO:and finish Wolfe2016 discussion paper for yields) TODO:go through atkinsonarey2003

Many of the HCHO yields from terpenoids are estimated through chamber studies which examine the products molecular mass and charge after mixing the compound of choice into a known volume of air. These conditions generally don't match those of the real world, where ambient air will have a cocktail of these compounds as well as various reactants.

A proton transfer reaction mass spectrometer (PTR-MS) can be used to determine gas phase evolution of terpene oxidation products. This is done through analysis of mass to charge ratios ( $m/z$ ) which can be used identify chemical compounds. Looking at Australian emissions from running GEOS-Chem and using yields provided by XYZ (TODO other table), we see that Australia may be more or less likely to do something TODO: this comparison sentence would be good to tie up tables and be copied to conclusions.

Conversions between HCHO per unit C yield and molar % yield from species X given by the equation  $Y_{molar\%} = 100 \times C_X \times Y_{HCHO \text{ per unit } C}$ , where  $C_X$  is how many Carbon are within species X (5 for isoprene, 10 for monoterpenes, etc...). For instance a 200% molar yield of HCHO from isoprene implies 1 Mole of  $\text{C}_5\text{H}_8$  becomes 2 Mole HCHO which is a 0.4 HCHO per unit C yield.

TODO: Fill out this table

Yields table looking at literature provided yields of HCHO.

**Table 5.2:** HCHO yields from various species averaged over Australia during Summer.

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

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

b:

c:

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

### 5.5.7 Accounting for smearing

Accounting for transport of the precursors is important, especially in low  $\text{NO}_X$  conditions in which isoprene has a longer lifetime (days). When estimating emissions of isoprene using one of its products, it is often assumed that isoprene has a short lifetime, however when low  $\text{NO}_X$  environments (which are prevalent in the Australian outback) this assumption can be wrong. Smearing (or spatial smearing) is a measure of how much formaldehyde (the product) was created from isoprene (the precursor) emissions in a different grid box. Smearing has been measured in order to account for this uncertainty in various works [Martin2003, 94, 80, 75, 9, 127, 121], often implementing the method designed in [94].

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

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

where U is wind-speed. Palmer [94] 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:

$$\frac{1}{k_{HCHO} - k_i} \left( k_{HCHO} \exp \left[ \frac{-k_i L_{s,i}}{U} \right] - k_i \exp \left[ \frac{-k_{HCHO} L_{s,i}}{U} \right] \right) = \frac{1}{e} \quad (5.7)$$

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  $k_{HCHO} \ll k_i$ .

**Table 5.3:** HCHO yields from various species, and lifetime against oxidation by OH.

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

a Atkinson and Arey [6]: Table 2, Yield from Isoprene reaction with OH, two values are from two referenced papers therein.

b Palmer [94]: lifetimes assume [OH] is 1e15 mol cm<sup>-3</sup>.

c [Lee2006]: Calculated through change in concentration of parent and product linear least squares regression. Estimates assume 20° C conditions.

d Wolfe et al. [121]: “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. [25]: One-day yields from oxidation modelled by CHIMERE, using MCM reference scheme.

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

TODO: outline of smearing reasoning (find where I wrote this, put it here)

TODO: calculation of smearing Similarly to smearing sensitivity calculations in [75], we run GEOS-Chem with isoprene emissions halved, then calculate  $\hat{S} = \frac{\Delta\Omega_{HCHO}}{\Delta E_{Isop}}$ , where  $\Delta$  represents the monthly mean departure over 1300-1400LT from default run values.

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

## 5.6 Results

### 5.6.1 Emissions affect on GEOS-Chem

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

### 5.6.2 Emissions comparisons

Some global numbers (TODO: where do I throw these?) Guenther et al. [41] Estimate global biogenic isoprene emissions at roughly 535 Tg yr<sup>-1</sup>, using MEGAN. Sindelarova et al. [109] Estimate around 594 Tg yr<sup>-1</sup>using MEGAN with MACC, showing isoprene as 69.2% of the total BVOC emissions, with monoterpenes at 10.9 Tg yr<sup>-1</sup>(10.9%). They show 41 Tg yr<sup>-1</sup>decrease in Australia when introducing soil moisture parameterisation.

When comparing the GEOS-Chem (which runs MEGAN) emissions to those calculated using our top-down inversion, we see a decrease over TODO: locations and seasons. TODO: table or figure showing summary of isoprene emissions changes over the whole of our time domain.

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

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

We examine the affect of decreased isoprene emissions on the correlation between modelled and satellite based HCHO columns. Figure TODO: shows the regressions

between GEOS-Chem tropospheric column amounts of HCHO and satellite columns for two runs of GEOS-Chem: a) using standard MEGAN emissions, b) using our updated emissions.

# Chapter 6

## Formaldehyde product over Australia

### 6.1 Australian Biogenic Volatile Organic Compounds (BVOCs)

#### 6.1.1 Isoprene, Monoterpenes

Isoprene has a large impact on the oxidative properties of the atmosphere, as it reacts quickly with the OH radical to form RO<sub>2</sub> and then in the presence of NO<sub>X</sub> various OVOCs(largely HCHO) and other SOAs(notably ozone) are formed.

Estimates of isoprene emissions require more work in order to generate confidence at a global scale. Due to isoprene being optically thin (see section 6.2.3) and having a short life time (around an hour) there are relatively few accurate measurements against which a comparison and verification can be made.

The emissions of isoprene have been modelled at around 500 Tg C yr<sup>-1</sup> in Guenther et al. [42] using MEGAN, and more recently around 465 Tg C yr<sup>-1</sup> in [78] using ORCHIDEE. The global emission factors model used to derive these estimates is based on modelling emissions from different plant species (phenotypes), and very few are used to set the emission factors of Australian forests.

Globally around 710 - 1150 Tg C yr<sup>-1</sup> of BVOCs are emitted [Lathiere2006, 42, 41]. 90% of these emissions come from plants and trees, with the most dominant species being isoprene (C<sub>5</sub>H<sub>8</sub>) (~ 50%), monoterpenes (C<sub>10</sub>H<sub>16</sub>), methanol (CH<sub>3</sub>OH), ethanol (C<sub>2</sub>H<sub>6</sub>O), acetaldehyde (CH<sub>3</sub>CHO), acetone ((CH<sub>3</sub>)<sub>2</sub>CO), ethene (C<sub>2</sub>H<sub>4</sub>) and propene (C<sub>3</sub>H<sub>6</sub>) (together making up ~ 30%) [41]. The larger 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. Another model (ORCHIDEE, with inputs similar to MEGAN) estimates 752 ± 16

Tg C yr<sup>-1</sup>, sensitive to terrestrial vegetation variations [Lathiere2006].

One problem with current estimates of biogenic VOC emissions in Australia is that the emission rates from various species of eucalypt and other flora are highly complex, depending on current and recent weather, temperature, tree age, health, etc. [41]. 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. Isoprenoid emissions remain to be verified in Australia and the few monoterpenes emission rates we have may be underestimated by a factor of 2-4 [Winters2009].

### 6.1.2 Biomass Burning

As biomass burning can be a large local or transported source of HCHO, CHOCHO, glyoxal, and other compounds we would like to use to determine BVOC emissions, it is advantageous to filter out this source. One complication when computing HCHO yield from VOC emissions is biomass burning interference, as smoke plumes can contribute to column HCHO. Influence from biomass burning can be removed through measurements of acetonitrile and CO (eg: [121, 79], or else removal of scenes coincident with satellite detected fire counts and aerosol absorption optical depth as done in [74]. Wolfe et al. [121] disregard HCHO measurements when acetonitrile  $\geq$  210 pptv and CO  $\geq$  300 ppbv, while acetonitrile  $\geq$  200 pptv is used to determine fire influence in [79].

### 6.1.3 MEGAN

One method used to estimate global isoprene (among other species) emissions is the Model of Emissions of Gases and Aerosols from Nature (MEGAN). MEGAN is a global model with resolution of around 1 km, and is used to generate the BVOC emissions used in various global chemistry models such as GEOS-Chem. MEGAN uses leaf area index, global meteorological data, and plant functional types (PFTs) to simulate terrestrial isoprene emissions. The various PFTs are used to generate emission factors which represent quantities of a compound released to the atmosphere through an associated activity. For example, an emission factor for isoprene within a forest would include the requirement of sunshine and suitable temperature. The schematic for MEGAN, taken from [Megan Website](#), is shown in figure 6.1

MEGAN output in Australia is adversely affected by poor emission factor estimation, which is based on plant type classifications (PFTs) and local conditions like soil moisture and weather. An example can be seen in Müller et al. [85] where MEGAN overestimates isoprene in northern Australia. Underestimates of monoterpenes emissions are also seen from MEGAN (TODO: ask Jenny what other paper

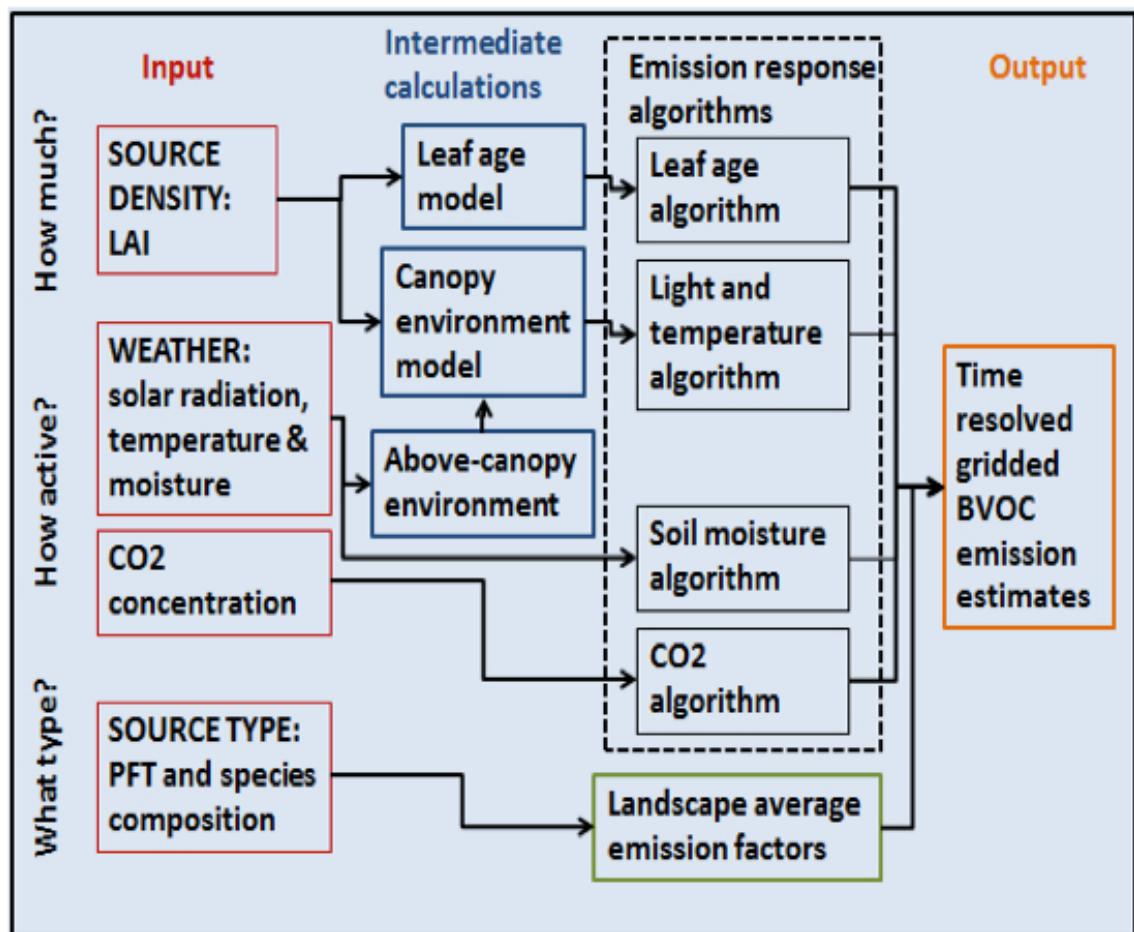


Figure 6.1: MEGAN schematic, copied from Megan'Website

showed this?), which may be due simply to underestimated emission rates for many Eucalypt species [Winters2009].

## 6.2 Satellite HCHO measurements

### 6.2.1 Satellite Retrievals

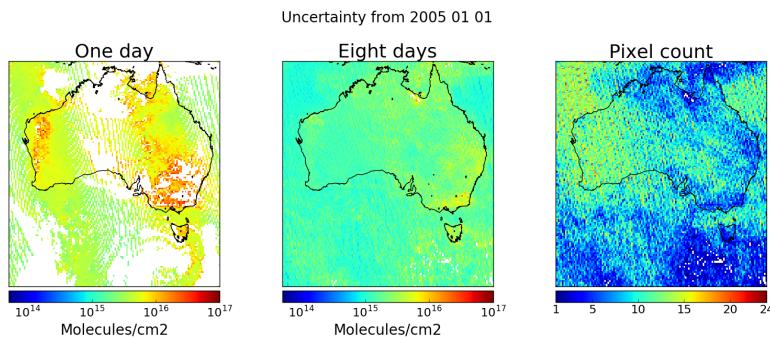
Satellites remotely sense atmospheric trace gases through irradiance measurements of solar light which has reflected off the earth's surface. These irradiances are affected by gases which exist along the reflected path of light between the detector, earth, and sun. The irradiance is then used to estimate 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).

One satellite is NASA's Earth Observing System's (EOS) Aura, which houses the Ozone Monitoring Instrument (OMI), a near-UV/Visible Charged Coupled Device (CCD) spectrometer. Aura orbits the earth in a polar sun-synchronous pattern, circling the earth on a plane coincident with the sun and the poles. OMI measurements are used to map several atmospheric trace gases, including NO<sub>2</sub>, SO<sub>2</sub>, BrO, HCHO, O<sub>3</sub>, and aerosols. OMI measurements occur from right to left on a band covering 115°, resulting in swaths of around 2600 km, with pixel sizes from 13x24 km<sup>2</sup> at nadir to 26x135 km<sup>2</sup> at the swath edges [39]. The swaths cover Earth daily, although half of these are at night time and contain no useful near-UV/Visible information. 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.

Atmospheric HCHO can be measured using Differential Optical Absorption Spectroscopy (DOAS), as long as trace gases with similar features near the same wavelength are accounted for. OMI fits HCHO slant column densities (SCD) along a solar backscatter path within the spectral window of 327.5-356.5 nm [15]. TODO: Go through Lee2005 or Volkamer2005 and detail the DOAS Retrieval of HCHO. 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. The full method details for slant column retrieval by OMI are outlined in section 6.2.2.

Uncertainty in a single pixel for OMI is quite high, roughly the same magnitude as HCHO background levels. Figure 6.2 shows uncertainty over Australia after one and eight days of averaging at 0.25° longitude by 0.3125° latitude. There are

several methods of calculating this, one of which is used here and compared against the provided uncertainty (TODO) as shown in Section 6.2.9. If we assume the uncertainty is random error, and not bias introduced through calculation techniques, then we are able to reduce the uncertainty through averaging. Random error can be reduced by either temporal or spatial averaging, decreasing uncertainty by a factor of  $1/\sqrt{N}$  where  $N$  is the number of observations being averaged. High resolution low detection limit estimates can be built up using “oversampling”, which averages satellite measurements over time. A good example can be seen in [127] where  $0.2^\circ$  by  $0.2^\circ$  resolution with high enough sensitivity to see anthropogenic HCHO is achieved with three summers worth of satellite data.



**Figure 6.2:** OMI uncertainty before and after gridding and averaging 8 days from Jan 1 2005 to Jan 8 2005. The third panel shows the number of pixels in each grid box after 8 days of averaging, before accounting for fire.

### 6.2.2 OMI Algorithm BOAS

The following information comes from the OMHCHO dataset documentation at **Kurosu2014** and Instrument [50]. 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  $c_0$ - $c_3$ ).

$$I(\lambda) = AI_0 \exp(-\sum_i S_i \sigma_i) + c_R \sigma_R + c_0 + c_1(\sigma - \bar{\sigma}) + c_2(\sigma - \bar{\sigma})^2 + c_3(\sigma - \bar{\sigma})^3 \quad (6.1)$$

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 [50].

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 6.2.6) 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.

### 6.2.3 Optical Depth ( $\tau$ )

Optical Depth, also called optical thickness, is the natural logarithm of the ratio of incident radiant power to transmitted radiant power through a material. In

the atmosphere we are interested in the optical depth of various chemical species, and we use incoming solar radiation to determine this. The difference between solar radiation at the top of the atmosphere and the Earth's surface defines the atmospheric optical depth along the path of observation.

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

where  $\phi_e^i$  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.

### 6.2.4 Scattering

Rayleigh and Mie scattering describe two kinds of particle effects on radiation passing through a medium. Rayleigh scattering is heavily wavelength dependent, and is the dominant form of scattering from particles up to roughly one tenth of the wavelength of the light. Mie scattering is more dominant from larger particles, and has less wavelength dependence. These scattering functions are described in detail at (TODO:section? reference?).

The effects of scattering are what gives us the information about substances in the atmosphere. The different particals and gases in the air have various properties which affect remote sensing devices such as a satellite, making them more or less sensitive at certain altitudes for detecting various species [Martin2002].

### 6.2.5 Absorption cross section and number density

TODO: Fill in this section, describe cross sections.

$\tau$  can be described using the attenuation cross section (the attenuation coefficient divided by its number density), with the following relation:

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

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

### 6.2.6 Air Mass Factors

DOAS retrieval columns are an integration of a trace gas over the instrument's viewing path, in order to convert this total to a vertically distributed column a

few assumptions and estimates are required. The vertical profile of a trace gas is assumed or estimated via a CTM, while its' scattering and radiative properties are calculated at all altitudes using an RTM. These properties are combined into a single array called the AMF which can be combined with a SC to produce a VC. AMFs are unique to each trace gas and due to their complexity and the influence of cloud cover they remain one of the larger error sources in remote sensing of BVOCs [95, 80]).

The latest OMI algorithm uses a shape factor determined from GEOS-Chem using 47 vertical levels at monthly temporal resolution and  $2^{\circ}$  latitude by  $2.5^{\circ}$  longitude horizontal resolution [39]. The GEOS-Chem model has been substantially updated since then, and using the more recent version V10.01 to recalculate the AMF is performed within this thesis, details are shown in section 6.2.8.

The vertical distribution of a trace gas determined by CTM is independent of the vertically dependent observation sensitivity provided by RTM, which prevents model contamination of remotely sensed data. Two examples of this are GOME-2 products on the MetOp-A satellite ([http://atmos.caf.dlr.de/gome/product\\_hcho.html](http://atmos.caf.dlr.de/gome/product_hcho.html)) and OMI products which use IMAGESv2 combined with LIDORT and GEOS-Chem with LIDORT for product processing respectively [50, 39]. The recalculation of OMI AMFs is explained in section 6.2.8).

Calculations of the AMF performed by different groups tend to agree fairly well, as long as all the apriori and ancillary data is similar. Large differences can occur depending on the apriori vertical profile, trace gas concentrations, and cloud properties [**Lorent2017**]. Choice of RTM and interpolation operations have a relatively small affect compared to the assumed state of the atmosphere, with high structural uncertainty introduced at this stage of AMF calculation - as shown in [**Lorent2017**].

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 [95], with modifications for clouds [**Martin2003**] and use of the LIDORT RTM [**Spurr2002**] is described by [**Martin2003**]. This code does not work as is when using OMI satellite data, and requires modifications as first performed by Luke Surl at Edinburgh University. Additionally dr. Surl acquired the tropopause heights averaged within satellite overpass times output by GEOS-Chem with modifications to the model diagnostic (ND51).

### 6.2.7 OMI HCHO data products

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

In order to filter fire or anthropogenic influence on HCHO data one generally requires the level 2 data along with extra data such as MODIS and AATSR fire counts, and national or regional industrial emissions estimates.

### 6.2.8 HCHO Vertical Column Calculation

The method used here largely follows that of Palmer et al. [95]. The Beer-Lambert law of extinction allows spectroscopic measurement of absorbing chemical species (absorbers) in the atmosphere:

$$I_B = I_{B_0} e^{-\tau_s} \quad (6.2)$$

where  $I_B$ ,  $I_{B_0}$  is backscattered intensity with and without the absorber respectively, and  $\tau_s$  is the optical thickness of the slant column absorber (see section 6.2.3).

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

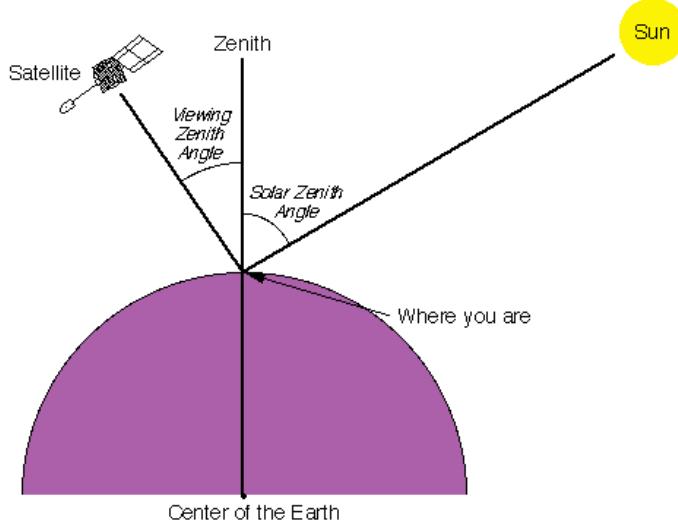
$$AMF = \frac{\Omega_s}{\Omega_v} = \frac{\tau_s}{\tau_v} \quad (6.3)$$

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

In the absence of atmospheric scattering a simple geometric AMF can be defined as a function of the solar zenith angle. The solar zenith angle ( $\theta_s$ ) and the satellite viewing angle ( $\theta_v$ ) are shown in image 6.3. However, in the UV-VIS region of the spectrum, Rayleigh and Mie scattering (see section 6.2.4) must be accounted for. Using equations 6.2 and 6.3 gives us an expression for the AMF which includes scattering:

$$AMF = \frac{\ln I_{B_0} - \ln I_B}{\tau_v} \quad (6.4)$$

We use  $\nabla I = I_B - I_{B_0}$  to represent the change in intensity due to the absorber.



**Figure 6.3:** Solar and viewing zenith angles, image copied from **SZA Image**, originally from a NASA website.

Note that for optically thin absorption,  $\nabla I / I_B \ll 1$ , and we can use:

$$AMF = \frac{\ln \left( 1 - \frac{\nabla I}{I_B} \right)}{\tau_v} \approx \frac{-\frac{\nabla I}{I_B}}{\tau_v} \quad (6.5)$$

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

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

which can be placed into equation 6.5:

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

and rewritten as:

$$AMF = \frac{-1}{\tau_v} \int_0^{\infty} \frac{\partial \ln I_B}{\partial \tau} \alpha(z) \eta(z) dz \quad (6.6)$$

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 6.2.5).

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

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

This is the same as  $\tau_v = \Omega_v \alpha_e$ , which we can place into equation 6.6 to obtain:

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

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

$$\omega(z) = - \frac{1}{AMF_G} \frac{\alpha(z)}{\alpha_e} \frac{\partial \ln I_B}{\partial \tau} \quad (6.7)$$

and vertical shape factor  $S_z(z)$  representing a normalized vertical number density profile:

$$S_z(z) = \frac{\eta(z)}{\Omega_v} \quad (6.8)$$

Now the AMF can be expressed as

$$AMF = AMF_G \int_0^\infty w(z) S_z(z) dz \quad (6.9)$$

It's worth noting that in the OMI satellite product, the provided  $\omega(z)$  term does not include the  $\frac{1}{AMF_G}$  term and a the calculation in equation 6.9 does not use the  $AMF_G$  term. 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.

For this equation  $\omega$  is a function of atmospheric scattering which can be calculated using an RTM to determine the absorption cross section and optical thickness along the path.  $S_z(z)$  can be calculated using an apriori vertical profile, which may be sourced from any atmospheric chemistry model. Note that for level two non-gridded OMI satellite HCHO total column data, the  $w(z)$ ,  $S_z(z)$ , and  $AMF_G$  are all provided.

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.

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

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

Substitution into 6.9 gives AMF using the sigma coordinates:

$$AMF = AMF_G \int_0^1 w(\sigma) S_\sigma(\sigma) d\sigma \quad (6.10)$$

Where  $S_\sigma$  is defined as a vertical shape factor representing a normalised mixing ratio:

$$S_\sigma(\sigma) = \frac{\Omega_a}{\Omega_v} C(\sigma) \quad (6.11)$$

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

When comparing satellite observations to a chemical model, one needs to recalculate the satellite AMF using their own modelled vertical gas profiles as the a-prior shape factor in order to remove any total column bias which may be due to the satellite's apriori. Another way of removing this bias is through deconvolution (TODO: EQNS) of the averaging kernal (AK) of the satellite instrument. The AK represents sensitivities to each species at multiple altitudes through the atmosphere and in the case of OMI, can be approximated from the scattering weights ( $\omega(z)$ ) function as follows:

$$AK(z) = \frac{\omega(z)}{AMF} \quad (6.12)$$

Note that this is an approximation for the OMI product, which does not include the AK but does include the  $\omega$  and AMF, as explained in Gonzalez Abad et al. [39].

### 6.2.9 Uncertainty in OMI total columns

Provided with the OMI product is the measurement of uncertainty in each pixel, calculated by SAO from the backscattered solar radiation fit [39, 1]. BIRA use another method, and calculate the standard deviation of HCHO over the remote pacific ocean as the uncertainty [24, 23]. 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. For an analysis of the uncertainty in the recalculation of the OMI HCHO vertical columns see section 6.3.10.

### 6.2.10 Reference sector correction for comparison of products to various models

HCHO products from OMI and SCIAMACHY both use a median daily remote pacific ocean radiance reference spectrum, over  $15^\circ\text{S}$ - $15^\circ\text{N}$ ,  $140^\circ$ - $160^\circ\text{W}$  where it is as-

sumed that the only significant source of HCHO is methane oxidation [DeSmedt2008, Kurosu2014, 9]. Since this oceanic background is used instead of a solar irradiance spectrum, in order to compare the output against different models, the vertical columns need to be corrected by an absolute amount. The corrected vertical column ( $\Omega_V$ ) is calculated as the slant column ( $\Omega_S$ ) minus the reference slant column ( $\Omega_{S_0}$ ) multiplied by the AMF, plus the modelled reference sector column ( $\Omega_{V_B}$ ):

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

This method is used in various papers, including DeSmedt2008, De Smedt et al. [24, 23], Barkley et al. [9], and Bauwens et al. [11].

Recently this correction was expanded (for OMI data) to include latitudinal and instrument track influence by Gonzalez Abad et al. [39]. The updated correction is explained in detail in section 6.3.9.

## 6.3 Recalculating HCHO from satellite(OMI) data over Australia

### 6.3.1 Process Outline

First satellite slant columns of formaldehyde for the years January 1st, 2005 - April 1st, 2013 are downloaded from NASA. 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. This data is first quality assured and undergoes basic analysis and filtering criteria as is done in several other studies [75, 9, 11, 128, eg.]. This filtering removes cloudy and uncertain data points, along with instrument problems such as the OMI row anomaly (see section 6.3.2). A full account of the filters used when reading OMI satellite data is given in section 6.3.2.

In order to reduce uncertainty and increase the utility of the satellite data we regrid it from pointwise single time data points to 8-day averages on a latitude longitude grid which matches our model input and output grid spacing. Using the 8-day average reduces the uncertainty in each datapoint significantly, TODO: cite marais and barkley uncertainty improvements, details of the uncertainty estimations is shown in section 6.3.10.

Once the slant columns are quality filtered and gridded, additional data sources need to be used to account for anthropogenic and pyrogenic sources of HCHO. MODIS fire counts are used in conjunction with NO<sub>2</sub> enhancements (also measured by satellite) to remove data points which may be affected by fires. TODO: If it is

easier to use OMI smoke smoke aaod I'll do it that way instead of using the NO<sub>2</sub>, write here if that is the case. One possible solution to anthropogenic filtering is the national pollution index (TODO:cite:<http://www.npi.gov.au>) which contains industrial HCHO and NO<sub>X</sub> emissions from 2003 to 2014. This has a negative affect on uncertainty, as fewer measurements are averaged over the 8-days. The affect of the fire filtering on uncertainty, and how many points are removed is shown in section 6.3.5.

Each satellite slant column measurement is corrected by some amount, based on the divergence from a modeled reference sector. The reference sector correction method corrects for several problems, however it introduces some apriori 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 [23]. This method also corrects for the errors introduced through correlations between BrO and HCHO absorption cross sections, which are especially significant at high latitudes [39].

The reference sector we use is defined over the pacific ocean at 140 to 160°W and 90°N to 90°S, as in Gonzalez Abad et al. [39]. HCHO concentrations are assumed to be at background levels over the pacific ocean, with their only source being CH<sub>4</sub> oxidation. A correction for each instrument pixel is created based on the difference between the background HCHO measurements from OMI and the GEOS-Chem modelled HCHO columns within the reference sector. This correction is calculated daily and applied to all good pixels based on their latitude. The full process for this is shown in section 6.3.9.

In order to visualise and analyse satellite column data it is generally transformed into vertical columns. This is done using AMF calculations as shown in section 6.2.8. Taking the biogenic slant columns, scattering weights, and apriori estimates of HCHO vertical profiles we determine vertical HCHO column amounts. This is an in depth process involving radiative transfer modelling in order to work out satellite sensitivities at various altitudes, as well as the effect from the local HCHO profile on those sensitivities. Several of these required data are available from the satellite data products, including the scattering weights and the zenith angles required to determine an AMF at any particular measured point. In this work the shape factor is recalculated from GOES-Chem, with the associated OMI per-pixel scattering weights unchanged. The satellite shape factor is replaced by GEOS-Chem's overpass time simulated HCHO profile, normalised and saved daily along with air density.

When comparing satellite measurements against models it is important to recognise the impact of the apriori shape factor on the total column values. This is due to the sensitivity of instruments varying vertically through the atmosphere, and how

the simulated distribution of HCHO is accounted for. In order to remove a possible bias caused by systematic differences between the old model and the current model, the shape factor used by the satellite is replaced using the profile from the current model before satellite total columns are recalculated (generally using equation 6.9). Both the shape factor and scattering weights of the satellite are recalculated using a combination of GEOS-Chem apriori profile information and satellite measurement data using code initially written by dr. Paul Palmer, which calculates the AMF after running the LIDORT radiative transfer calculations to determine apriori scattering, this is explained fully in section TODO: and how it's used is shown in section 6.3.8. Without performing this step a bias between modeled and measured total column values may be due to an apriori rather than actual chemistry or measurements.

### 6.3.2 Quality filtering OMI HCHO slant columns

TODO: Quality flags and cloud cover metric uses, and discussion, along with statistics like how many datapoints are removed.

The OMI 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 6.1 which is taken from **Kurosu2014**. Filtering bad or missing measurement pixels is preformed prior to any other filtering, this includes the datapoints affected by the row anomaly. This anomaly (<http://projects.knmi.nl/omi/research/product/rowanomaly-background.php>) affects radiance data at particular viewing angles, corresponding to a row on the CCD detectors, and is dynamic over time. The slant columns affected are flagged and easy to remove before further processing.

Clouds have various detrimental effects on slant column uncertainty and AMF calculation, so the cloudy data needs to be filtered. Any pixel with a cloud fraction of greater than 40% is also removed, after the pixel is used in determining the reference sector correction, as is done in Gonzalez Abad et al. [39] and De Smedt et al. [23]. Another way this has been performed is to remove pixels where the AMF lies outside a certain range: **Martin2002** filter AMFs below 0.5 in order to remove the effects of heavy cloud and optical thickness.

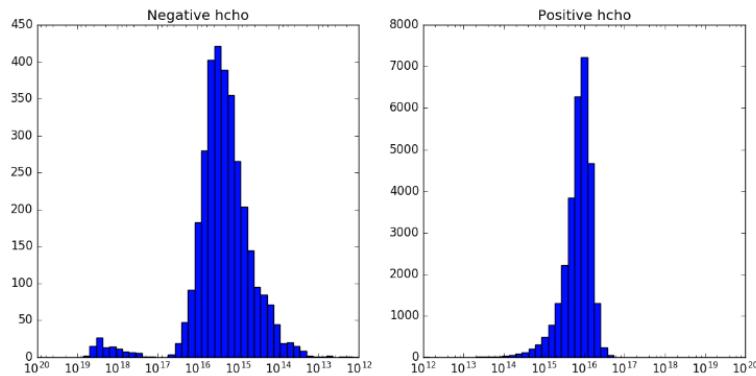
The cloud fraction with each pixel is provided with the OMHCHO dataset, however its source is the OMI cloud product, OMCLDO2. To give an idea of how much data is filtered out, around 30% of the pixels which remain after filtering out the bad or missing data are subsequently removed due to cloudiness.

Further filtering is performed to remove the measurements which are most likely to be unrealistic: pixels with a solar zenith angle greater than 60°, and those with column density outside the range  $-0.5 \times 10^{16}$  to  $10^{17}$  molecules cm<sup>-2</sup>. These are similar filters to those applied in (TODO: read Zhu et al. [128], add similar justi-

**Table 6.1:** OMI quality flag values table from **Kurosu2014**

<b>Value</b>	<b>Classification</b>	<b>Rational</b>
0	Good	Column value present and passes all quality checks; data may be used with confidence.
1	Suspect	Caution advised because one or more of the following conditions are present: <ul style="list-style-type: none"> <li>• Fit convergence flag is <math>&lt; 300</math> but <math>&gt; 0</math>: Convergence at noise level</li> <li>• Column <math>+2\sigma</math> uncertainty <math>&lt; 0 &lt;</math> Column <math>+3\sigma</math> uncertainty</li> <li>• Absolute column value <math>&gt;</math> Maximum column amount (<math>1\text{e}19</math> molec <math>\text{cm}^{-2}</math>)</li> </ul>
2	Bad	Avoid using as one of the following conditions are present: <ul style="list-style-type: none"> <li>• Fit convergence flag is <math>&lt; 0</math> : No convergence, abnormal termination</li> <li>• Column <math>+3\sigma</math> uncertainty <math>&lt; 0</math></li> </ul>
$< 0$	Missing	No column values have been computed; entries are missing

fication if succinct). This final filter is required due to currently unexplained large negative values which occur in the OMI HCHO product increasingly over time. Figure 6.4 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.



**Figure 6.4:** 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.

One final selection of the data is removed prior to calculation and analysis due to instrument sensitivity issues. This removed data is based on the latitude of the satellite measurement, high northern and southern latitudes are removed as the effect of both the high solar zenith angle and (over arctic regions) the high albedo cause anomalous readings which are hard to correct. The latitudes removed from analysis for OMI satellite data are those outside of  $0^\circ \pm 60^\circ$ . This is where the satellite measurements are less robust, and often publications will remove this same region (TODO: cites where they did this).

### 6.3.3 Reading OMHCHO daily slant columns

Each  $\sim 90$  minutes the AURA satellite sweeps over the sunny side of the planet, with OMI recording roughly 90 k pixels, of which around 50 k – 80 k are classified as good. Each pixel contains several important pieces of data which are needed for recalculation of the HCHO vertical column: the total column of HCHO (molecules  $\text{cm}^{-2}$ ), 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 apriori shape factor rather than NASA's included apriori shape factor. Each pixel and it's relevant data are saved in a long list, around 1.1 million pixels per day. As well as the data directly read from the OMI swath files, further information is added to each pixel. This is the new AMF calculated through replacing the apriori vertical profile with the newer GEOS-Chem simulated profile, which is described in section 6.3.7. The shape factors and scattering weights for each pixel lie along a z-axis which is vertically resolved to 47 layers and is shown in figure TODO: make figure showing this stuff and an example profile.

TODO: Show an example of OMI swaths.

### 6.3.4 Regridding to 0.25 by 0.3125 8-day averaged vertical columns

Regridded OMI HCHO columns from the are based on 14-15 daily swaths of measurements provided by NASA. Each swath contains roughly  $9 \times 10^4$  pixels, each of which contains various data including latitude, longitude vertical column HCHO, etc. In order to regrid these columns each pixel is mapped into a global grid of  $0.25^\circ \times 0.3125^\circ$  latitude by longitude (matching GEOS-5 native resolution) which may contain up to 15 entries from a particular day's orbits. Total vertical column amounts (both the satellite original and the columns reprocessed as follows) are averaged over each 8 days starting on January 1st 2005.

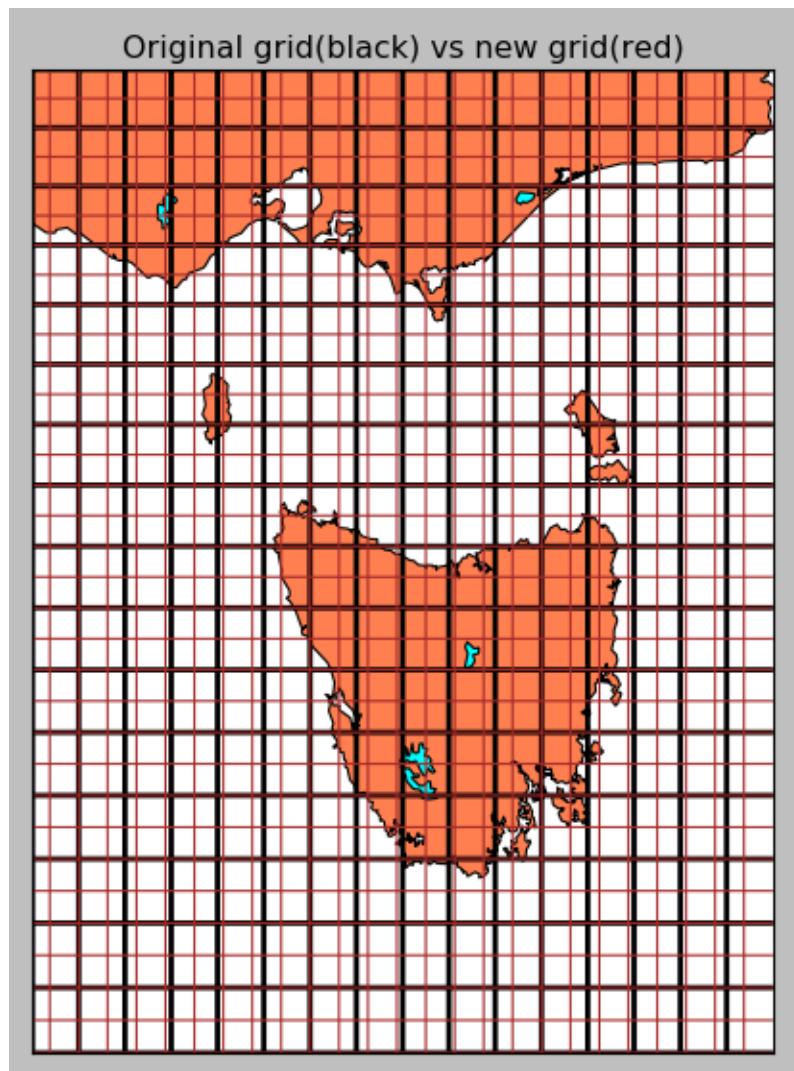
The process of regridding is performed by first reading all relevant information at each pixel for a single day into a list in python, which is then processed, with the shape factor read from GEOS-Chem output and AMF recalculated, before being saved as a gridded array of total columns and pixel counts. The total columns are averaged into each grid box for each day, and eventually averaged over eight day time periods.

TODO: time per regridding and reprocessing: This mapping requires some processing time as well as RAM and computer storage space, and has been performed on the National Computing Infrastructure (NCI) supercomputer cluster. In order to reprocess one year of swath files, X GB of daily data was downloaded and then transformed into Y GB of 8-day averaged gridded data. This takes around 90 minutes per day, and is very parallelisable as each day is completely independent. Using  $N \times 8$  concurrent processors on NCI's computer cluster running Python allows for very fast reprocessing of our entire timeline. As much as possible, processing is done using the HDF-EOS5 format, with GEOS-Chem output being read and processed from bitpunch to HDF-EOS5 prior to reprocessing. The scripts to regrid and reprocess the swath data set are available in the supplementary (TODO).

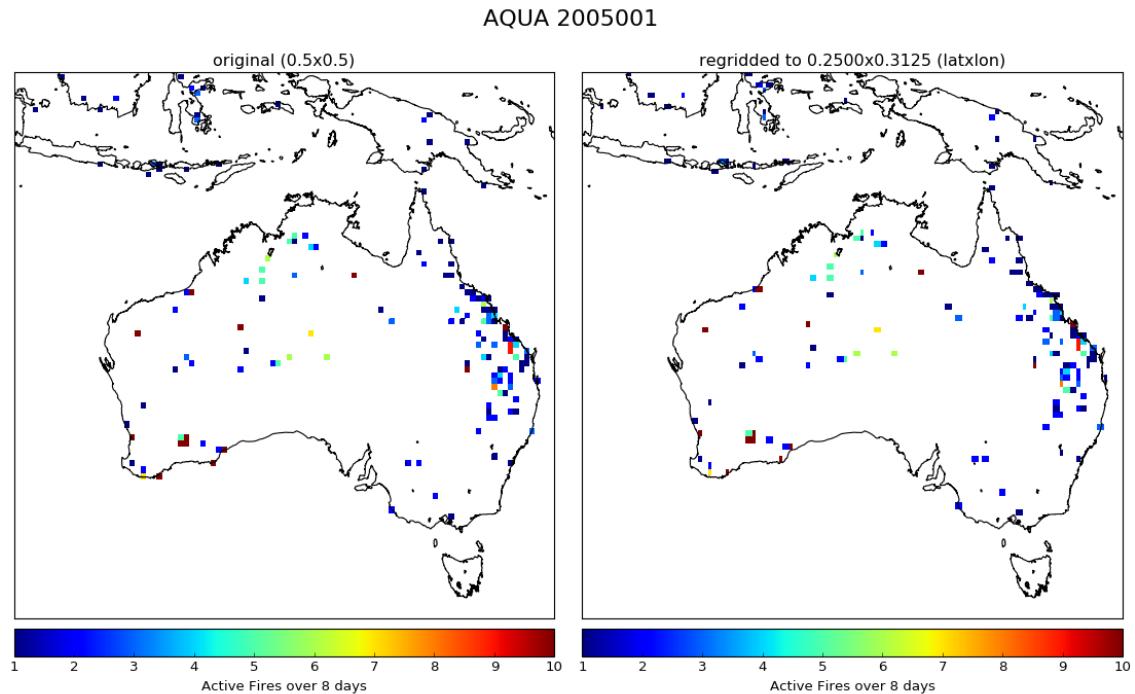
### 6.3.5 Filtering pyrogenic HCHO

TODO: How modis fire counts are used as well as statistics on removed data points.

On board NASA's AQUA satellite, the MODIS instrument is used to detect fire activity. The product used here is called MYD14C8H ([Giglio2006]), 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. An example of the change in resolution is provided in figure 6.5, where the grids are shown over a basic map of Tasmania. The direct affect of this interpolation is shown as an example in figure 6.6, which is showing the regridded MODIS fire count over Australia from January 2005 (avg of first 8 days) in two subplots.



**Figure 6.5:** Example of grid space change using  $0.5 \times 0.5$  and  $0.25 \times 0.3125$  latitude by longitude resolution.



**Figure 6.6:** 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.

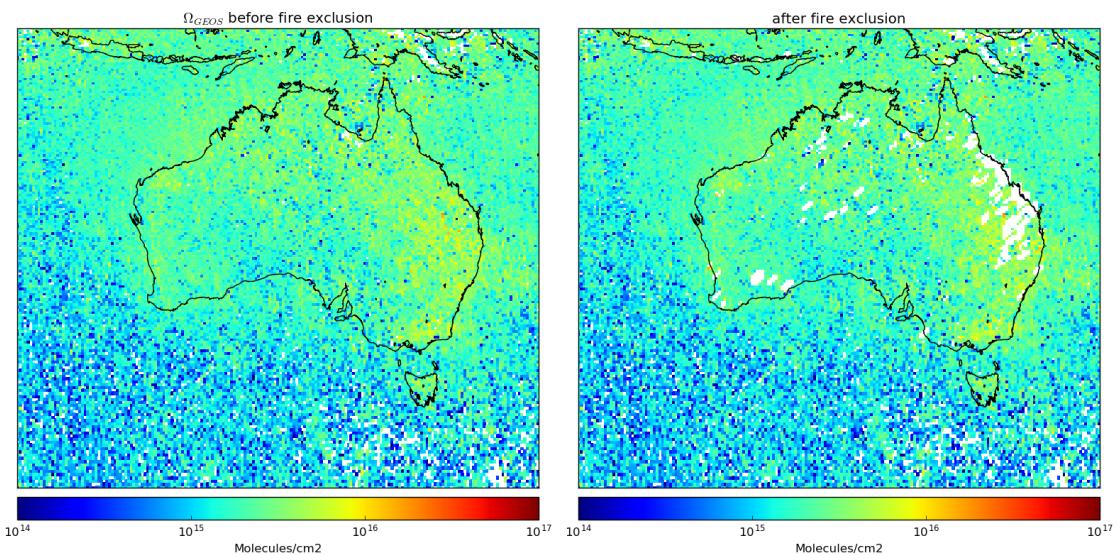
Figure 6.7 shows an example of the total column HCHO calculated using GEOS-Chem aprioris ( $\Omega_{GEOS}$ ) before and after using the MYD14C8H product to exclude fire influenced pixels. (TODO: show time series of how many pixels are removed and discuss if this causes any issues down the line)

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.

### 6.3.6 Filtering anthropogenic HCHO

TODO

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



**Figure 6.7:** 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.

### 6.3.7 Recalculating the AMF to create our own vertical HCHO columns

OMI's apriori shape factor is based on the GEOS-Chem (v9) model, which uses 47 layers between the earth's surface and the top of the atmosphere using a pressure-eta hybrid (the actual values are shown in table 8.1). Taking a more recent GEOS-Chem apriori shape factor and integrating along the vertical axis using equation 6.9 gives us a new AMF ( $AMF_n$ ). Since we are using the  $\omega$  provided by OMI, we remove the  $AMF_G$  term from this calculation. The integration is done in Python using a simple rectangular method, which multiplies the integrand midpoints by the change in height, and then takes the sum. This is identical to calculating the integral if we assume the integrand is linear between each measured point, and introduces no new uncertainty. All that remains for recalculating the total vertical column using our new apriori shape factor is to apply the new AMF and remove the old:

$$\Omega_{new} = \Omega \frac{AMF}{AMF_n}$$

The vertical column scattering weights and apriori shape factors provided in the OMHCHO 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! a)The new apriori profiles are monthly averages, which is the same temporal resolution used by the OMI apriori shape factors. b)The new apriori profiles are simulated daily and averaged over 8 days along with the recalculated total vertical columns.

A new AMF is determined using equation 6.9) with the apriori shape factor set by our GEOS-Chem model run. In order to reformulate the AMF, GEOS-Chem's 72 level vertical profile is transformed from ppb to a normalized number density profile in order to match equation 6.8. This conversion uses the following equation:

$$\eta_{HCHO} = ppb_{HCHO} \times \eta_a \times 10^{-9} \quad (6.13)$$

where  $\eta_{HCHO}$  is the number density of a HCHO, and  $ppb_{HCHO}$  is the molecules of that species per billion molecules of air. In order to normalize these vertical density profiles over the globe, we divide by the modelled total vertical column  $\Omega_{HCHO}$  which is determined by:

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

where  $P(z)$  is the pressure (hPa) at the bottom of altitude level  $z$ , the constant 2.12e13 is determined from equation (TODO: run through this number in another section?). In effect this equation sums over the molecules per  $\text{cm}^2$  in each altitude level.

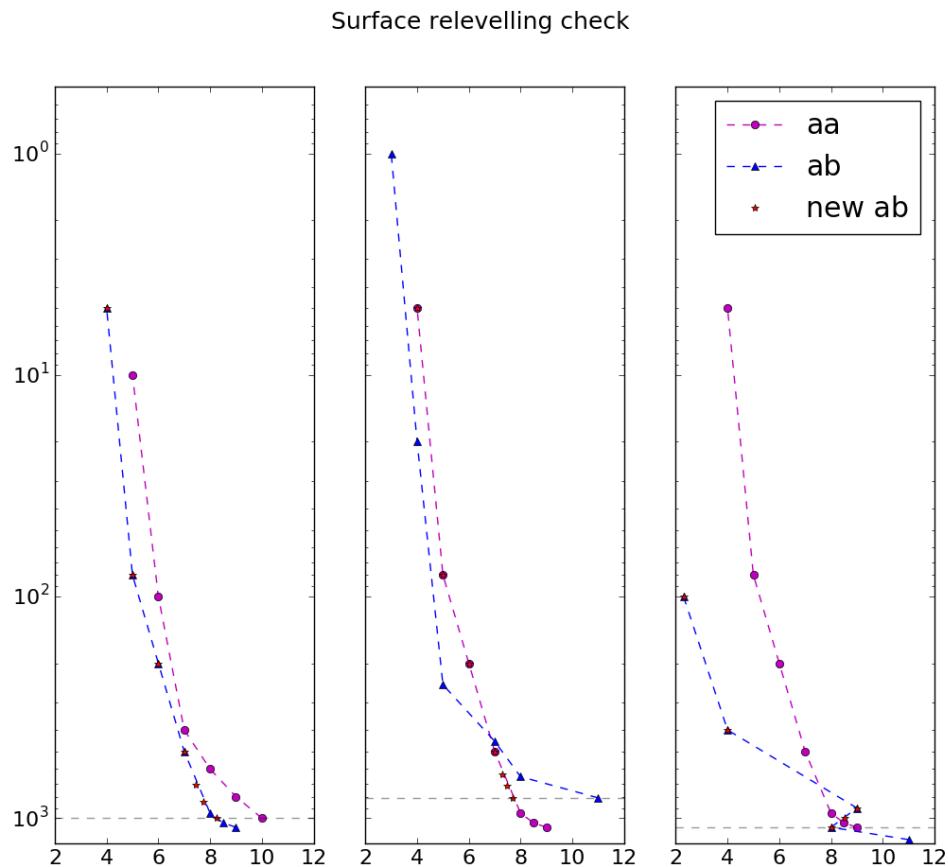
We have  $S_z(z)$  and  $\omega(z)$  over the vertical pressure coordinate  $z$  at all latitude and longitude points on whatever grid we wish. A conversion to the sigma ( $\sigma$ ) vertical coordinate is performed using  $P = \sigma(P_S - P_T) + P_T$ , where  $P_T$  is pressure at the top of the atmosphere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

$$S_\sigma(\sigma) = \frac{\Omega_a}{\Omega_v} C_{HCHO}(\sigma) \quad (6.14)$$

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. This equation comes from Palmer et al. [95], and is unitless since  $\Omega_a/\Omega_v$  is molecules of air per molecule of HCHO; the opposite of  $C_{HCHO}$ .

Pressure dimension from OMI are the surface pressures from each gridbox (of-

fine 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 6.8 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.



**Figure 6.8:** Constructed example of the initial interpolation of OMI's  $\omega$  onto a pressure dimension with mismatched surface pressure.

$S_\sigma(\sigma)$  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 6.9. 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 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 6.9. 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.

### 6.3.8 AMF code from Paul Palmer

TODO: describe how I use this here I use code originally written by Dr. Paul Palmer with various updates and modifications described in section (TODO:) as another way to recalculate the AMF using information from the satellite swaths and the GEOS-Chem overpass simulation output. These are used to recalculate the instrument sensitivity or scattering weights for each pixel, as well as the shape factor which together are integrated to give the pixel AMF.

GEOS-Chem outputs quantities averaged between 1200 and 1400 LT, including optical depths at several wavelengths (TODO: list), dust, and HCHO. I run a script on the satellite swaths which pulls out a subset of the pixel information into a daily csv file, which can be read by the AMF code as modified by Dr. Luke Surl, in conjunction with the GEOS-Chem outputs for each day. The AMF code is then run and produces a csv of recalculated AMFs which get read by my python code and

associated with the corresponding pixel.

### 6.3.9 Determination and application of the pacific ocean reference sector normalisation

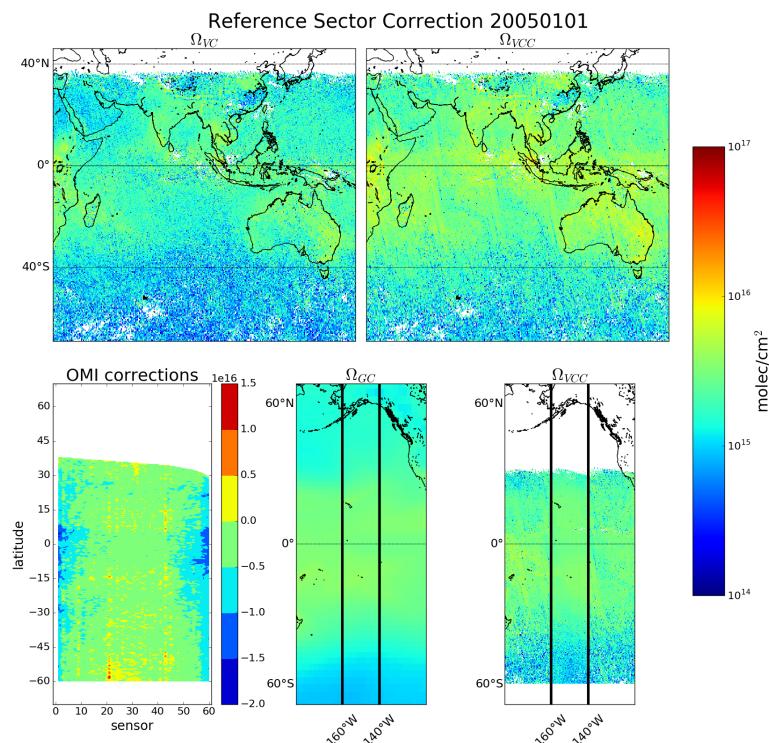
As is done in Gonzalez Abad et al. [39], a reference sector defined over the Pacific ocean is used to correct OMI instrument degradation. This correction is calculated based on satellite measurements over the pacific ocean reference sector; between 140°W and 160°W, covering every latitude. Corrections are made using apriori HCHO columns in the same reference sector modelled using GEOS-Chem. The apriori reference sector of HCHO vertical columns (VCs) is created by GEOS-Chem using 15 minute time resolution, with 2 by 2.5° latitude by longitude resolution. These simulated values use the GEOS-Chem output averaged between 1300 and 1400 local time at each grid box, in order to match the overpass time of OMI. The longitudinal average is taken within the apriori reference sector, as corrections are assumed to be longitudinally invariant. The modeled reference sector is interpolated latitudinally in for use in the OMI measurement correction array creation. Figure 6.9 the simulated reference sector VCs as an example, calculated on January 1st 2005. In this figure the vertical resolution is increased from 2° to 0.36°, through linear interpolation, in order to form 500 vertical bins which are used in correcting the satellite data. Each day, good satellite measurements taken over the reference sector are used to determine a correction array. The correction is based on the difference between measured slant column and the modeled slant column within the reference sector. The model does not produce slant columns associated with each measurement, however one is created by multiplying the VC with the associated slant column's AMF.

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

$$\text{Correction}(i, j) = SC_{HCHO}(i, j) - VC_{GEOS-Chem}(\text{lat}(j)) \times AMF_{OMI}(i, j) \quad (6.15)$$

where j represents a latitude index and  $VC_{GEOS-Chem}(\text{lat})$  represents the apriori reference sector vertical column HCHO at the latitude corresponding to j. Note that the correction is in molecules per cm<sup>2</sup>.

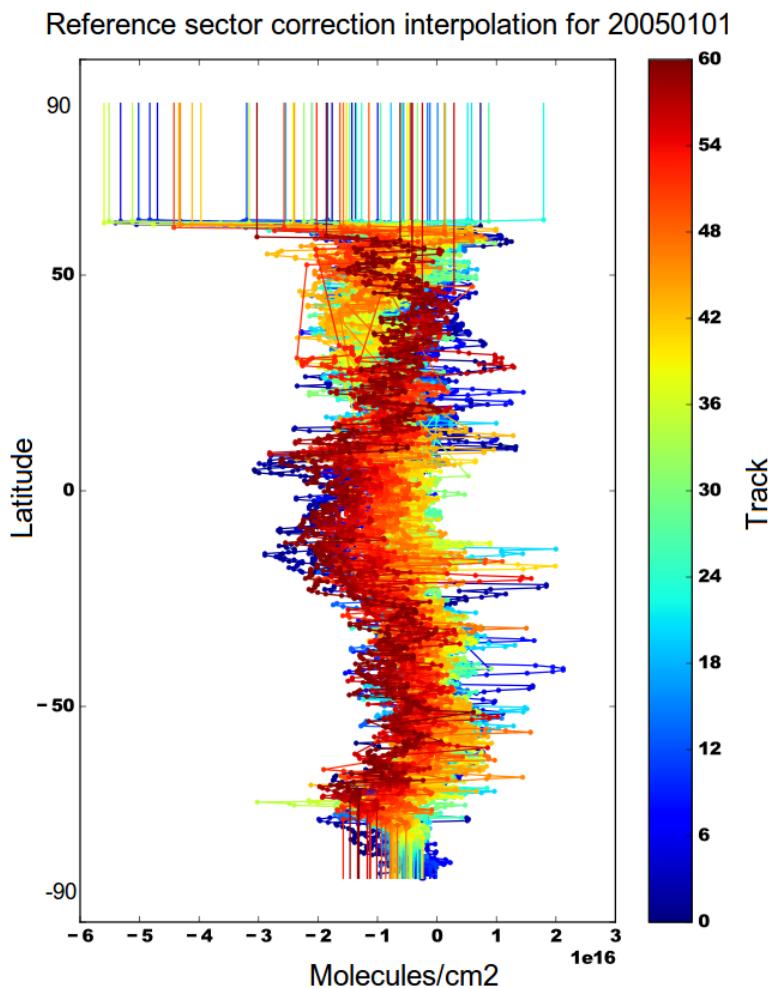
The reference sector correction is independently calculated for each of the 60 tracks, at each latitude where a good satellite measurement exists which used that



**Figure 6.9:** Example of remote pacific reference sector correction (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.

track. The  $\text{Correction}(i, \text{lat}(j))$  function is determined by binning corrections for each track into 500 equidistant latitude bands.

Due incomplete latitudinal coverage, the correction for each track is interpolated linearly between measurements, with corrections outside of the highest measured latitudes being equal to the corrections at the highest measured latitudes. Figure 6.10 shows an example of the 60 track corrections for January 1st 2005, the points are satellite measurements and the lines are the interpolations for each track.



**Figure 6.10:** Example of track correction interpolations for January 1st 2005, points represent satellite slant column measurements, with lines interpolating and extrapolating along the latitudinal direction.

Another way to look at this correction is given in the OMI corrections panel of figure 6.9, 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:

$$VCC(i, j) = \frac{SC_{HCHO}(i, j) - Correction(i, lat(j))}{AMF(i, j)} \quad (6.16)$$

Finally, for each day, the good satellite measurements are averaged into our own latitude longitude resolution bins along with the associated corrected SC, VC, VCC, AMF, and bin entry count. The bin entry count is used to create an 8-day average out of the one day averages, as it is the daily mean multiplied by the daily count summed over 8 days divided by the total count for each bin.

### 6.3.10 Estimation of error or uncertainty

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

- a** Fitting error from the OMI retrieval.
- b** Uncertainty in AMF calculations.
- c** Uncertainty of HCHO background.

a) is available in the OMI product and reduced through spatial and temporal averaging. Taking the eight day grided average with horizontal resolution of 0.25 by 0.3125 degrees (latitude by longitude) typically reduces uncertainty by a factor of 1.5 to 4. Another method for examining uncertainty of OMI is to analyse the standard deviation of the HCHO columns over the remote pacific. If we assume there is no HCHO variation from background levels over any 8-day period, then this method infers variations in the measuring instrument, and can be used as a metric for uncertainty as done in De Smedt et al. [24]. TODO: uncertainty calculation on remote pacific OMI. [80, 96] both examine OMI HCHO columns over North America and determine overall uncertainty to be 40%, with most of this coming from cloud interference.

b) is determined through an analysis of GEOS-Chem output, validated against the total column of HCHO at Wollongong using FTIR measurements from the (TODO: Nicholas Jones roof HCHO citation here). [96] 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%) [Martin2003], and aerosol errors (< 20%), totalling AMF error of around  $\sim 30\%$ . It is worth noting here that independent error estimates are added in quadrature, which means total error equals the root of the sum of the independent errors each squared ( $e_{Total} = \sqrt{\sum_i e_i^2}$ ). TODO: Paul palmer calculation and combination for overall Satellite VC uncertainty per pixel and gridded. TODO: Millet2008?

c) is also determined through a study of GEOS-Chem output, in relation to in-situ measurements. TODO: calculate this uncertainty. Compare this error estimate with that of Curci et al. [21], where the error in b) and c) are respectively found to be 30% and 15% based on their analysis of CHIMERE. [81] 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.

## 6.4 Validation and comparisons

### 6.4.1 Comparison with standard OMI product

Figure TODO: shows global and Australian HCHO eight day averaged total column maps for 1-8 January 2005, along with the reduced major axis (RMA) regression corellation and percentage difference. This comparison shows how reprocessing with an updated model can have a systematic influence on the total column.

### 6.4.2 Comparison with in-situ measurements

TODO: Describe Wollongong FTIR and junk Analyse comparison of gridbox with instrument!

### 6.4.3 Summary

First the OMI HCHO level 2 data was downloaded, and read using python creating a list of good pixels for each day. Next the associated AMF and reference sector correction for each good pixel was calculated using GEOS-Chem for the apriori shape factor, and using the provided scattering weights from OMI. Each 8 days the pixel list is averaged onto a  $0.25^\circ$  latitude by  $0.3125^\circ$  longitude grid. The new HCHO product along with counts and average uncertainty of pixels used in the grid square is also kept. The product includes the 8-day gridded averages of the old and new AMFs, the average correction sector from GEOS-Chem over the pacific ocean, and the old and new HCHO with and without the reference sector correction from [39] applied.

### 6.4.4 Conclusions

# Chapter 7

## Isoprene Emissions in Australia

### 7.1 GEOS-Chem isoprene mechanisms(moved to Modelling chapter)

#### 7.1.1 Outline(moved)

#### 7.1.2 Emissions from MEGAN(moved)

### 7.2 Isoprene emissions estimation

#### 7.2.1 Outline

Isoprene quickly forms HCHO in the atmosphere when in the presence of high levels of  $\text{NO}_X$ . However, over Australia  $\text{NO}_X$  levels are generally not high enough and we must take extra care that we can account for the transport or 'smearing' caused by slower HCHO formation. Smearing sensitive grid boxes within the model can be detected by running the model with two uniformly differing isoprene emission levels, then finding the grid boxes where the changed HCHO column is greater than can be attributed to local emission difference. Using equation ?? with two different isoprene emission levels:

$$\hat{S} = \frac{\Delta \Omega_{\text{HCHO}}}{\Delta E_{\text{ISOP}}}$$

Consider halving the isoprene emitted globally and rerunning the model, if the local grid HCHO is reduced by much more than half (factoring yield) then you can infer sensitivity to non-local isoprene emissions. This can be dependent on local or regional weather patterns, as greater wind speeds will reduce the time any emitted compound stays within the local grid box. As such smearing sensitivity is both spatially and temporally diverse, shown in figure TODO: is a picture of the smearing sensitivity over Australia.

Once the smearing sensitive grid squares are filtered out, application of equation ?? gives us an estimate of isoprene emissions across the nation.

Most recently a Bauwens et al. [11] undertook a similar process to what I am doing, although with slightly different focus, using the IMAGESv2 global CTM instead of GEOS-Chem. They calculate emissions which create the closest match between model and satellite vertical columns, and compare these postiori data with the apriori (satellite data) and independent data sets. (TODO: simple outline of what they did and how my focus is different, this paper will also need to be summarised in the LitReview)

### 7.2.2 HCHO Products and yield

### 7.2.3 CAABA/MECCA yield

### 7.2.4 Calculation of Emissions

### 7.2.5 Calculation of smearing effect

TODO: Smearing scale length,  $\hat{S}$  formula, and results of calculations in here. As shown in [94], smearing sensitivity can be calculated through multiple runs of the same model with the only difference being the isoprene emissions. I have run GEOS-Chem with and without  $E_{ISOP}$  multiplied uniformly by 0.5, and the grid boxes with the most affected  $\Omega_{HCHO}$  are those affected most by smearing. The smearing parameter ( $\hat{S}$ ) is defined as follows:

$$\hat{S} = \frac{\Delta\Omega_{HCHO}}{\Delta E_{ISOP}} \quad (7.1)$$

TODO: Plot shows smearing parameter over Australia.

### 7.2.6 Calculations of uncertainty

There are several factors which need to be considered when looking at the uncertainty in emissions estimates. Things with their own inherent uncertainty include the modelled apriori, 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.

Uncertainty in satellite measurements is generally provided along with the data, although uncertainty introduced through AMF calculation needs to be determined to give a representation of the confidence in vertical column amounts. The measurement uncertainty is shown in section 6.3.10, and amounts to  $\sim X\%$ . (TODO this

number when calculated)

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. TODO: find out how this is estimated in other papers, or else point to HCHO uncertainty and used some function of that.

The uncertainty for HCHO to isoprene mechanisms TODO: how to do this?

### 7.2.7 Extrapolating the circadian cycle

Isoprene emissions occur with regular daily cycles caused by things like local temperature, sunlight, drought, and other environmental factors (TODO: find/cite eucalypt isoprene paper, daily cycle plot if can find).

(TODO: following stuff, add some basic plots and error analysis eventually also) Using a model of the daily isoprene emissions fit to the offset determined by satellite HCHO based estimates, we produce a high temporal resolution isoprene emissions inventory. During days with more than one HCHO column measurement we can more confidently fit the cycle. For example EOS AURA's OMI measurements from 2004 can be combined with MetOp-A's GOME2 after October 2006, with daily overpasses by OMI and GOME2 at 1345 and 0930 respectively. This allows a better retrieval of the daily amplitude of isoprene emissions.

### 7.2.8 Comparison with MEGAN

TODO: Direct comparison here, maps of differences for some metrics(monthly average,?). comparison of model run results using different inventory shown in section (reference here)

## 7.3 New estimates affects on the Australian atmosphere

# Bibliography

- [1] Gonzalo González Abad et al. “Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval”. In: *Atmospheric Measurement Techniques* 9.7 (2016), pp. 2797–2812. ISSN: 18678548. DOI: 10.5194/amt-9-2797-2016.
- [2] Dorian S. Abbot. “Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space”. In: *Geophysical Research Letters* 30.17 (2003), pp. 1999–2002. ISSN: 0094-8276. DOI: 10.1029/2003GL017336. URL: <http://doi.wiley.com/10.1029/2003GL017336>.
- [3] a. Arneth et al. “Why are estimates of global isoprene emissions so similar (and why is this not so for monoterpenes)?” In: *Atmospheric Chemistry and Physics Discussions* 8.2 (2008), pp. 7017–7050. ISSN: 1680-7375. DOI: 10.5194/acpd-8-7017-2008.
- [4] M R Ashmore, Lisa. Emberson, and Murray Frank. *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, 2003, xiii, 372 p. : ISBN: 186094292.
- [5] Roger Atkinson. “Atmospheric chemistry of VOCs and NO(x)”. In: *Atmospheric Environment* 34.12-14 (2000), pp. 2063–2101. ISSN: 13522310. DOI: 10.1016/S1352-2310(99)00460-4.
- [6] Roger Atkinson and Janet Arey. “Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review”. In: *Atmospheric Environment* 37.SUPPL. 2 (2003). ISSN: 13522310. DOI: 10.1016/S1352-2310(03)00391-1.
- [7] James D Ayers and William R Simpson. “Measurements of N<sub>2</sub>O<sub>5</sub> near Fairbanks, Alaska”. In: *Journal of Geophysical Research: Atmospheres* 111.D14 (2006), n/a–n/a. ISSN: 2156-2202. DOI: 10.1029/2006JD007070. URL: <http://dx.doi.org/10.1029/2006JD007070>.

- [8] Jean-Luc Baray et al. “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 (2012). ISSN: 2156-2202. DOI: 10.1029/2011JD016717. URL: <http://dx.doi.org/10.1029/2011JD016717>.
- [9] Michael P. Barkley et al. “Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns”. In: *Journal of Geophysical Research Atmospheres* 118.12 (2013), pp. 6849–6868. ISSN: 21698996. DOI: 10.1002/jgrd.50552. URL: <http://dx.doi.org/10.1002/jgrd.50552>.
- [10] M Bauwens et al. “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*. 2013, pp. 17–20.
- [11] Maite Bauwens et al. “Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations”. In: *Atmospheric Chemistry and Physics Discussions* March (2016), pp. 1–45. ISSN: 1680-7375. DOI: 10.5194/acp-2016-221. URL: <http://www.atmos-chem-phys-discuss.net/acp-2016-221/>.
- [12] N Bei, G Li, and L T Molina. “Uncertainties in SOA simulations due to meteorological uncertainties in Mexico City during MILAGRO-2006 field campaign”. In: *Atmospheric Chemistry and Physics* 12.23 (2012), pp. 11295–11308. DOI: 10.5194/acp-12-11295-2012. URL: <http://www.atmos-chem-phys.net/12/11295/2012/>.
- [13] E. J. Brinksma et al. “Five years of observations of ozone profiles over Lauder, New Zealand”. In: *Journal of Geophysical Research* 107.D14 (2002), pp. 1–11. ISSN: 0148-0227. DOI: 10.1029/2001JD000737. URL: <http://doi.wiley.com/10.1029/2001JD000737>.
- [14] S. S. Brown et al. “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 (2009), pp. 3027–3042. ISSN: 16807316. DOI: 10.5194/acp-9-3027-2009.
- [15] K. Chance et al. “Satellite observations of formaldehyde over North America from GOME”. In: *Geophysical Research Letters* 27.21 (2000), pp. 3461–3464. ISSN: 00948276. DOI: 10.1029/2000GL011857. URL: <http://dx.doi.org/10.1029/2000GL011857>.

- [16] Kenneth E Christian, William H Brune, and Jingqiu Mao. “Global sensitivity analysis of the GEOS-Chem chemical transport model: ozone and hydrogen oxides during ARCTAS (2008)”. In: *Atmos. Chem. Phys* 17 (2017), pp. 3769–3784. DOI: 10.5194/acp-17-3769-2017. URL: [www.atmos-chem-phys.net/17/3769/2017/](http://www.atmos-chem-phys.net/17/3769/2017/).
- [17] John D Crounse et al. “Atmospheric Fate of Methacrolein. 1. Peroxy Radical Isomerization Following Addition of OH and O<sub>2</sub>”. In: *Physical Chemistry m* (2012).
- [18] John D. Crounse et al. “Autoxidation of organic compounds in the atmosphere”. In: *Journal of Physical Chemistry Letters* 4.20 (2013), pp. 3513–3520. ISSN: 19487185. DOI: 10.1021/jz4019207. URL: <http://pubs.acs.org/doi/abs/10.1021/jz4019207>.
- [19] John D Crounse et al. “Peroxy radical isomerization in the oxidation of isoprene-new”. In: *Phys. Chem. Chem. Phys. Phys. Chem. Phys.* 13.13 (2011), pp. 13607–13613. ISSN: 1463-9076. DOI: doi:10.1039/c1cp21330j. URL: <http://dx.doi.org/10.1039/C1CP21330J>.
- [20] PAUL J CRUTZEN, MARK G LAWRENCE, and ULRICH PÖSCHL. “On the background photochemistry of tropospheric ozone”. In: *Tellus A* 51.1 (1999), pp. 123–146. ISSN: 1600-0870. DOI: 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>.
- [21] G. Curci et al. “Estimating European volatile organic compound emissions using satellite observations of formaldehyde from the Ozone Monitoring Instrument”. In: *Atmospheric Chemistry and Physics* 10.23 (2010), pp. 11501–11517. ISSN: 16807316. DOI: 10.5194/acp-10-11501-2010.
- [22] J J Davenport et al. “A measurement strategy for non-dispersive ultra-violet detection of formaldehyde in indoor air : spectral analysis and interferent gases”. In: *Measurement Science and Technology* (2015), p. 15802. ISSN: 0957-0233. DOI: 10.1088/0957-0233/27/1/015802. URL: <http://dx.doi.org/10.1088/0957-0233/27/1/015802>.
- [23] I. De Smedt et al. “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 (2015), pp. 12519–12545. ISSN: 16807324. DOI: 10.5194/acp-15-12519-2015. URL: <http://www.atmos-chem-phys-discuss.net/15/12241/2015/%5Cnhttp://www.atmos-chem-phys-discuss.net/15/12241/2015/acpd-15-12241-2015.pdf>.

- [24] I. De Smedt et al. “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 (2012), pp. 2933–2949. ISSN: 18671381. DOI: 10.5194/amt-5-2933-2012.
- [25] G. Dufour et al. “SCIAMACHY formaldehyde observations: constraint for isoprene emissions over Europe?” In: *Atmospheric Chemistry and Physics* 8.6 (2009), pp. 19273–19312. ISSN: 1680-7324. DOI: 10.5194/acpd-8-19273-2008.
- [26] D. P. Edwards et al. “Satellite-observed pollution from Southern Hemisphere biomass burning”. In: *Journal of Geophysical Research* 111.14 (2006), pp. 1–17. ISSN: 01480227. DOI: 10.1029/2005JD006655.
- [27] Kathryn M. Emmerson et al. “Current estimates of biogenic emissions from eucalypts uncertain for southeast Australia”. In: *Atmospheric Chemistry and Physics* 16.11 (2016), pp. 6997–7011. ISSN: 1680-7324. DOI: 10.5194/acp-16-6997-2016. URL: <http://www.atmos-chem-phys.net/16/6997/2016/>.
- [28] H J Eskes and K F Boersma. “Averaging kernels for DOAS total-column satellite retrievals”. In: *Atmospheric Chemistry and Physics* 3.1 (2003), pp. 1285–1291. ISSN: 1680-7324. DOI: 10.5194/acp-3-1285-2003. URL: <http://dx.doi.org/10.5194/acpd-3-895-2003>.
- [29] EUMETSAT. *GOME2*. 2015. URL: <http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/GOME2/index.html>.
- [30] Jiwen Fan and Renyi Zhang. “Atmospheric oxidation mechanism of isoprene”. In: *Environmental Chemistry* 1.3 (2004), pp. 140–149. ISSN: 14482517. DOI: 10.1071/EN04045. URL: <http://dx.doi.org/10.1071/en04045>.
- [31] P. Forster et al. *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. 2007. 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)].
- [32] A. Fortems-Cheiney et al. “The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system”. In: *Atmospheric Chemistry and Physics* 12.15 (2012), pp. 6699–6721. ISSN: 16807316. DOI: 10.5194/acp-12-6699-2012. URL: <http://www.atmos-chem-phys.net/12/6699/2012/acp-12-6699-2012.pdf>.

- [33] B. Franco et al. “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 (2015), pp. 1733–1756. ISSN: 18678548. DOI: 10.5194/amt-8-1733-2015.
- [34] Tzung-may Fu et al. “Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone”. In: 112 (2007), pp. 1–15. DOI: 10.1029/2006JD007853.
- [35] J. D. Fuentes et al. “Biogenic Hydrocarbons in the Atmospheric Boundary Layer: A Review”. In: *Bulletin of the American Meteorological Society* 81.7 (2000), pp. 1537–1575. ISSN: 00030007. DOI: 10.1175/1520-0477(2000)081<1537:BHITAB>2.3.CO;2. arXiv: arXiv: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).
- [36] Louis Giglio, James T. Randerson, and Guido R. Van Der Werf. “Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4)”. In: *Journal of Geophysical Research* 118.1 (2013), pp. 317–328. ISSN: 21698961. DOI: 10.1002/jgrg.20042.
- [37] Marianne Glasius and Allen H. Goldstein. “Recent Discoveries and Future Challenges in Atmospheric Organic Chemistry”. In: *Environmental Science and Technology* 50.6 (2016), pp. 2754–2764. ISSN: 15205851. DOI: 10.1021/acs.est.5b05105.
- [38] Annemieke Gloudemans et al. “Evidence for long-range transport of carbon monoxide in the Southern Hemisphere from SCIAMACHY observations”. In: *European Space Agency, (Special Publication)* 33.SP-636 (2007), pp. 1–5. ISSN: 03796566. DOI: 10.1029/2006GL026804.
- [39] G. Gonzalez Abad et al. “Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval”. In: *Atmospheric Measurement Techniques* 8.1 (2015), pp. 19–32. ISSN: 18678548. DOI: 10.5194/amt-8-19-2015.
- [40] A Guenther et al. “Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature)”. In: *Atmospheric Chemistry and Physics* 6.11 (2006), pp. 3181–3210. DOI: 10.5194/acp-6-3181-2006. URL: <http://dx.doi.org/10.5194/acp-6-3181-2006>.

- [41] A. B. Guenther et al. “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 (2012), pp. 1471–1492. ISSN: 1991959X. DOI: 10.5194/gmd-5-1471-2012.
- [42] Alex Guenther et al. “A global model of natural volatile organic compound emissions”. In: *Journal of Geophysical Research* 100.D5 (1995), pp. 8873–8892. ISSN: 0148-0227. DOI: 10.1029/94JD02950. URL: <http://onlinelibrary.wiley.com/doi/10.1029/94JD02950/full>.
- [43] Alex Guenther et al. “Natural emissions of non-methane volatile organic compounds, carbon monoxide, and oxides of nitrogen from North America”. In: *Atmospheric Environment* 34.12-14 (2000), pp. 2205–2230. ISSN: 13522310. DOI: 10.1016/S1352-2310(99)00465-3.
- [44] C. Hak et al. “Intercomparison of four different in-situ techniques for ambient formaldehyde measurements in urban air”. In: *Atmospheric Chemistry and Physics Discussions* 5.3 (2005), pp. 2897–2945. ISSN: 1680-7316. DOI: 10.5194/acpd-5-2897-2005.
- [45] Michaela I Hegglin and Theodore G Shepherd. “Large climate-induced changes in ultraviolet index and stratosphere-to-troposphere ozone flux”. In: *Nature Geoscience* 2.10 (2009), pp. 687–691. DOI: 10.1038/ngeo604. URL: <http://dx.doi.org/10.1038/ngeo604>.
- [46] C N Hewitt et al. “Ground-level ozone influenced by circadian control of isoprene emissions”. In: *Nature Geoscience* 4.10 (2011), pp. 671–674. DOI: 10.1038/ngeo1271. URL: <http://dx.doi.org/10.1038/ngeo1271>.
- [47] Gerard Hoek et al. “Long-term air pollution exposure and cardio-respiratory mortality: a review”. In: *Environmental Health* 12.1 (2013), p. 43. DOI: 10.1186/1476-069x-12-43. URL: <http://dx.doi.org/10.1186/1476-069x-12-43>.
- [48] Nan-Hung Hsieh and Chung-Min Liao. “Fluctuations in air pollution give risk warning signals of asthma hospitalization”. In: *Atmospheric Environment* 75 (2013), pp. 206–216. DOI: 10.1016/j.atmosenv.2013.04.043. URL: <http://dx.doi.org/10.1016/j.atmosenv.2013.04.043>.
- [49] Guanyu Huang et al. “Validation of 10-year SAO OMI Ozone Profile (PROFOZ) Product Using Aura MLS Measurements”. In: *Atmospheric Measurement Techniques Discussions* (2017), pp. 1–25. ISSN: 1867-8610. DOI: 10.5194/amt-2017-92. URL: <https://www.atmos-meas-tech-discuss.net/amt-2017-92/>.

- [50] O M I Instrument. “OMI Algorithm Theoretical Basis Document Volume I”. In: I.August (2002), pp. 1–50.
- [51] *Intergovernmental Panel on Climate Change (IPCC): Climate Change: The Scientific Basis*. Tech. rep. Cambridge University Press, 2001. URL: <http://www.ipcc.ch/ipccreports/tar/>.
- [52] Daniel J Jacob. *Introduction to Atmospheric Chemistry*. Ed. by Daniel J Jacob. Princeton University Press, 1999. URL: <http://acmg.seas.harvard.edu/people/faculty/djj/book/index.html>.
- [53] P Jöckel et al. “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 (2006), pp. 5067–5104. DOI: 10.5194/acp-6-5067-2006. URL: <http://www.atmos-chem-phys.net/6/5067/2006/>.
- [54] Patrick Jöckel, Rolf Sander, and Jos Lelieveld. “Technical Note: The Modular Earth Submodel System (MESSy) a new approach towards Earth System Modeling”. In: *Atmospheric Chemistry and Physics Discussions* 4.6 (2005), pp. 7139–7166. ISSN: 1680-7324. DOI: 10.5194/acpd-4-7139-2004.
- [55] Peeters Jozef et al. “Hydroxyl Radical Recycling in Isoprene Oxidation Driven by Hydrogen Bonding and Hydrogen Tunneling: The Upgraded LIM1 Mechanism”. In: *Journal of Physical Chemistry* (2014).
- [56] M Kanakidou et al. “Physics Organic aerosol and global climate modelling : a review”. In: *Atmospheric Chemistry and Physics* 5 (2005), pp. 1053–1123.
- [57] Shawn C. Kefauver, Iolanda Filella, and Josep Peñuelas. “Remote sensing of atmospheric biogenic volatile organic compounds (BVOCs) via satellite-based formaldehyde vertical column assessments”. en. In: *International Journal of Remote Sensing* (2014). URL: <http://www.tandfonline.com/doi/abs/10.1080/01431161.2014.968690#.VkjEubNM61M>.
- [58] D Krewski et al. “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 (2009), pp. 5–36. ISSN: 1041-5505 (Print) 1041-5505 (Linking). URL: <http://www.ncbi.nlm.nih.gov/pubmed/19627030>.
- [59] Jesse H. Kroll and John H. Seinfeld. “Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere”. In: *Atmospheric Environment* 42.16 (2008), pp. 3593–3624. ISSN: 13522310. DOI: 10.1016/j.atmosenv.2008.01.003. URL: <http://www.sciencedirect.com.ezproxy.uow.edu.au/science/article/pii/S1352231008000253>.

- [60] Shi Kuang et al. “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* (2017), pp. 1293–1311. ISSN: 2169897X. DOI: 10.1002/2016JD026078. URL: <http://doi.wiley.com/10.1002/2016JD026078>.
- [61] Hyeong-Ahn Kwon et al. “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 (2017), pp. 4673–4686. ISSN: 1680-7324. DOI: 10.5194/acp-17-4673-2017. URL: <http://www.atmos-chem-phys.net/17/4673/2017/>.
- [62] Hanlim Lee et al. “Investigations of the Diurnal Variation of Vertical HCHO Profiles Based on MAX-DOAS Measurements in Beijing: Comparisons with OMI Vertical Column Data”. In: *Atmosphere* (2015). URL: 10.3390/atmos6111816.
- [63] J. Lelieveld et al. “Severe ozone air pollution in the Persian Gulf region”. In: *Atmospheric Chemistry and Physics* 9 (2009), pp. 1393–1406. ISSN: 1680-7324. DOI: 10.5194/acp-9-1393-2009.
- [64] J Lelieveld et al. “The contribution of outdoor air pollution sources to premature mortality on a global scale”. In: *Nature* 525.7569 (2015), pp. 367–371. DOI: 10.1038/nature15371. URL: <http://dx.doi.org/10.1038/nature15371>.
- [65] C Leue et al. “Quantitative analysis of NO x emissions from Global Ozone Monitoring Experiment satellite image sequences”. In: *J. Geophys. Res.* 106.D6 (2001), p. 5493. DOI: 10.1029/2000jd900572. URL: <http://dx.doi.org/10.1029/2000jd900572>.
- [66] Hiram Levy. “Photochemistry of the lower troposphere”. In: *Planetary and Space Science* 20.6 (1972), pp. 919–935. ISSN: 00320633. DOI: 10.1016/0032-0633(72)90177-8.
- [67] Meiyun Lin et al. “Climate variability modulates western US ozone air quality in spring via deep stratospheric intrusions.” In: *Nature communications* 6.May (2015), p. 7105. ISSN: 2041-1723. DOI: 10.1038/ncomms8105. URL: <http://www.nature.com/ncomms/2015/150512/ncomms8105/full/ncomms8105.html>.
- [68] Junhua Liu et al. “Causes of interannual variability of tropospheric ozone over the Southern Ocean”. In: *Atmospheric Chemistry and Physics Discussions* October (2016), pp. 1–46. ISSN: 1680-7316. DOI: 10.5194/ACP-2016-692.

- [69] Junhua Liu et al. “Causes of interannual variability over the southern hemispheric tropospheric ozone maximum”. In: *Atmospheric Chemistry and Physics* 17.5 (2017), pp. 3279–3299. ISSN: 1680-7324. DOI: 10.5194/acp-17-3279-2017. URL: <http://www.atmos-chem-phys.net/17/3279/2017/>.
- [70] Yingjun Liu et al. “Isoprene photochemistry over the Amazon rainforest”. In: *Proceedings of the National Academy of Sciences* 113.22 (2016), pp. 6125–6130. ISSN: 0027-8424. DOI: 10.1073/pnas.1524136113. URL: <http://www.pnas.org/content/113/22/6125.abstract>.
- [71] Alba Lorente et al. “Structural uncertainty in air mass factor calculation for NO<sub>2</sub> and HCHO satellite retrievals”. In: *Atmospheric Measurement Techniques* 2 (2017), pp. 1–35. ISSN: 1867-8610. DOI: 10.5194/amt-10-759-2017. URL: <https://www.atmos-meas-tech.net/10/759/2017/amt-10-759-2017.html>.
- [72] J. Mao et al. “Insights into hydroxyl measurements and atmospheric oxidation in a California forest”. In: *Atmospheric Chemistry and Physics* 12.17 (2012), pp. 8009–8020. ISSN: 16807316. DOI: 10.5194/acp-12-8009-2012.
- [73] Jingqiu Mao et al. “Ozone and organic nitrates over the eastern United States: Sensitivity to isoprene chemistry”. In: *Journal of Geophysical Research Atmospheres* 118.19 (2013), pp. 11256–11268. ISSN: 21698996. DOI: 10.1002/jgrd.50817.
- [74] E A Marais et al. “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 (2014), pp. 7693–7703. DOI: 10.5194/acp-14-7693-2014. URL: <http://dx.doi.org/10.5194/acp-14-7693-2014>.
- [75] E A Marais et al. “Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns”. In: *Atmospheric Chemistry and Physics* 12.3 (2012), pp. 7475–7520. DOI: 10.5194/acpd-12-7475-2012. URL: <http://dx.doi.org/10.5194/acpd-12-7475-2012>.
- [76] Margaret R. Marvin et al. “Impact of evolving isoprene mechanisms on simulated formaldehyde: An inter-comparison supported by in situ observations from SENEX”. In: *Atmospheric Environment* 164 (2017), pp. 325–336. ISSN: 13522310. DOI: 10.1016/j.atmosenv.2017.05.049. URL: <http://www.sciencedirect.com/science/article/pii/S1352231017303618>.

- [77] Gina M. Mazzuca et al. “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 (2016), pp. 14463–14474. ISSN: 16807324. DOI: 10.5194/acp-16-14463-2016.
- [78] Palmira Messina et al. “Global biogenic volatile organic compound emissions in the ORCHIDEE and MEGAN models and sensitivity to key parameters”. In: *Atmospheric Chemistry and Physics* 16.22 (2016), pp. 14169–14202. ISSN: 16807324. DOI: 10.5194/acp-16-14169-2016. URL: <http://www.atmos-chem-phys.net/16/14169/2016/acp-16-14169-2016.pdf>.
- [79] Christopher Chan Miller et al. “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 (2016), pp. 1–25. ISSN: 1680-7375. DOI: 10.5194/acp-2016-1042. URL: <http://www.atmos-chem-phys-discuss.net/acp-2016-1042/>.
- [80] Dylan B Millet et al. “Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission”. In: *J. Geophys. Res.* 111.D24 (2006). DOI: 10.1029/2005jd006853. URL: TODO.
- [81] Dylan B. Millet et al. “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 (2008), pp. 1–18. ISSN: 01480227. DOI: 10.1029/2007JD008950.
- [82] Y Mintz. “The sensitivity of numerically simulated climates to land surface conditions”. In: *Land Surface Processes in Atmospheric General Circulation Models* (1982), pp. 109–111.
- [83] P. S. Monks et al. “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 (2015), pp. 8889–8973. ISSN: 1680-7324. DOI: 10.5194/acp-15-8889-2015. URL: <http://www.atmos-chem-phys.net/15/8889/2015/>.
- [84] J. F. Müller, J. Peeters, and T. Stavrakou. “Fast photolysis of carbonyl nitrates from isoprene”. In: *Atmospheric Chemistry and Physics* 14.5 (2014), pp. 2497–2508. ISSN: 16807316. DOI: 10.5194/acp-14-2497-2014.

- [85] J.-F. Müller et al. “Global isoprene emissions estimated using MEGAN ECMWF analyses and a detailed canopy environment model”. In: *Atmospheric Chemistry and Physics Discussions* 7.6 (2008), pp. 15373–15407. DOI: 10.5194/acpd-7-15373-2007. URL: <http://dx.doi.org/10.5194/acpd-7-15373-2007>.
- [86] G Myhre and D Shindell. *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., 2013.
- [87] Amin Nawahda. “Comments on ”Global crop yield reductions due to surface ozone exposure: 1. Year 2000 crop production losses and economic damage” and ” Global crop yield reductions due to surface ozone exposure: 2. Year 2030 potential crop production losses and economic”. In: *Atmospheric Environment* 71.13 (2013), pp. 408–409. ISSN: 13522310. DOI: 10.1016/j.atmosenv.2012.12.045. URL: <http://dx.doi.org/10.1016/j.atmosenv.2011.01.002>.
- [88] Tran B. Nguyen et al. “Atmospheric fates of Criegee intermediates in the ozonolysis of isoprene”. In: *Phys. Chem. Chem. Phys.* 18.15 (2016), pp. 10241–10254. ISSN: 1463-9076. DOI: 10.1039/C6CP00053C. URL: <http://xlink.rsc.org/?DOI=C6CP00053C>.
- [89] U. Niinemets et al. “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 (1999), pp. 1319–1335. ISSN: 01407791. DOI: 10.1046/j.1365-3040.1999.00505.x.
- [90] U. Niinemets et al. “The emission factor of volatile isoprenoids: Stress, acclimation, and developmental responses”. In: *Biogeosciences* 7.7 (2010), pp. 2203–2223. ISSN: 17264170. DOI: 10.5194/bg-7-2203-2010.
- [91] Narendra Ojha et al. “Secondary ozone peaks in the troposphere over the Himalayas”. In: *Atmospheric Chemistry and Physics Discussions* November (2016), pp. 1–25. ISSN: 1680-7375. DOI: 10.5194/acp-2016-908. URL: <http://www.atmos-chem-phys-discuss.net/acp-2016-908/>.
- [92] J Oltmans et al. “Ozone in the Pacific tropical troposphere from ozonesonde observations”. In: *Journal of Geophysical Research* 106.D23 (2001), pp. 32503–32525.

- [93] B.C.a Pak et al. “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 (2003), pp. 1–10. ISSN: 0148-0227. DOI: 10.1029/2002JD002343. URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0742322536&partnerID=40&md5=cafafef03b948fb456696583ed3ab9a5>.
- [94] Paul I Palmer. “Mapping isoprene emissions over North America using formaldehyde column observations from space”. In: *J. Geophys. Res.* 108.D6 (2003). DOI: 10.1029/2002jd002153. URL: <http://dx.doi.org/10.1029/2002jd002153>.
- [95] Paul I Palmer et al. “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 (2001).
- [96] Paul I Palmer et al. “Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column”. In: *J. Geophys. Res.* 111 (2006), p. D12315. ISSN: 0148-0227. DOI: 10.1029/2005JD006689. URL: <http://dx.doi.org/10.1029/2005JD006689>.
- [97] AMIE K. PATCHEN et al. “Direct Kinetics Study of the Product-Forming Channels of the Reaction of Isoprene-Derived Hydroxyperoxy Radicals with NO”. In: *International Journal of Chemical Kinetics* 31.5 (2007), pp. 493–499. ISSN: 13000527. DOI: 10.1002/kin.
- [98] F. Paulot et al. “Isoprene photooxidation: new insights into the production of acids and organic nitrates”. In: *Atmospheric Chemistry and Physics* 9.4 (2009), pp. 1479–1501. ISSN: 1680-7324. DOI: 10.5194/acp-9-1479-2009.
- [99] Fabien Paulot et al. “Unexpected Epoxide Formation in the”. In: *Science* 325 (2009), pp. 730–733. ISSN: 0036-8075. DOI: 10.1126/science.1172910.
- [100] Jozef Peeters and Jean-François Müller. “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 (2010), p. 14227. ISSN: 1463-9076. DOI: 10.1039/c0cp00811g. URL: <http://xlink.rsc.org/?DOI=c0cp00811g>.
- [101] P. R. Rowntree and J. A. Bolton. “Simulation of the atmospheric response to soil moisture anomalies over Europe”. In: *Quarterly Journal of the Royal Meteorological Society* 109.461 (1983), pp. 501–526. ISSN: 00359009. DOI: 10.

- 1002/qj.49710946105. URL: <http://doi.wiley.com/10.1002/qj.49710946105>.
- [102] R Sander et al. “Technical note: The new comprehensive atmospheric chemistry module MECCA”. In: *Atmospheric Chemistry and Physics* 5.2 (2005), pp. 445–450. ISSN: 1680-7324. DOI: 10.5194/acp-5-445-2005. URL: <http://www.atmos-chem-phys.net/5/445/2005/>.
- [103] A Sandu and R Sander. “Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1”. In: *Atmospheric Chemistry and Physics* 6.1 (2006), pp. 187–195. DOI: 10.5194/acp-6-187-2006. URL: <http://www.atmos-chem-phys.net/6/187/2006/>.
- [104] V. M. Erik Schenkeveld et al. “In-flight performance of the Ozone Monitoring Instrument”. In: *Atmospheric Measurement Techniques* 10.5 (2017), pp. 1957–1986. ISSN: 1867-8548. DOI: 10.5194/amt-10-1957-2017. URL: <https://www.atmos-meas-tech.net/10/1957/2017/>.
- [105] Stefan F. Schreier et al. “Estimates of free-Tropospheric NO<sub>2</sub> and HCHO mixing ratios derived from high-Altitude mountain MAX-DOAS observations at midlatitudes and in the tropics”. In: *Atmospheric Chemistry and Physics* 16.5 (2016). ISSN: 16807324. DOI: 10.5194/acp-16-2803-2016.
- [106] N E Selin et al. “Global health and economic impacts of future ozone pollution”. In: *Environmental Research Letters* 4.4 (2009), p. 044014. ISSN: 1748-9326. DOI: 10.1088/1748-9326/4/4/044014.
- [107] Changsub Shim et al. “Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements”. In: *Journal of Geophysical Research Atmospheres* 110.24 (2005), pp. 1–14. ISSN: 01480227. DOI: 10.1029/2004JD005629.
- [108] Raquel A Silva et al. “Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change”. In: *Environ. Res. Lett.* 8.3 (2013), p. 34005. DOI: 10.1088/1748-9326/8/3/034005. URL: <http://dx.doi.org/10.1088/1748-9326/8/3/034005>.
- [109] K. Sindelarova et al. “Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years”. In: *Atmospheric Chemistry and Physics* 14.17 (2014), pp. 9317–9341. ISSN: 16807324. DOI: 10.5194/acp-14-9317-2014. arXiv: arXiv:1011.1669v3.
- [110] T Stavrakou et al. “Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns”. In: *Atmospheric Chemistry and Physics* 9.3 (2009), pp. 1037–1060. DOI: 10.

- 5194/acp-9-1037-2009. URL: <http://dx.doi.org/10.5194/acp-9-1037-2009>.
- [111] T. Stavrakou et al. “How consistent are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI?” English. In: *Atmospheric Chemistry and Physics* 15.20 (2015), pp. 11861–11884. ISSN: 1680-7324. DOI: 10.5194/acp-15-11861-2015. URL: <http://www.atmos-chem-phys.net/15/11861/2015/acp-15-11861-2015.html>.
- [112] T. Stavrakou et al. “Isoprene emissions over Asia 1979-2012: Impact of climate and land-use changes”. In: *Atmospheric Chemistry and Physics* 14.9 (2014). ISSN: 16807324. DOI: 10.5194/acp-14-4587-2014.
- [113] D S Stevenson et al. “Multimodel ensemble simulations of present-day and near-future tropospheric ozone”. In: *Journal of Geophysical Research* 111.D8 (2006). DOI: 10.1029/2005jd006338. URL: <http://dx.doi.org/10.1029/2005JD006338>.
- [114] D. S. Stevenson et al. “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 (2013), pp. 3063–3085. ISSN: 16807316. DOI: 10.5194/acp-13-3063-2013.
- [115] T.F. 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*. Tech. rep. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. DOI: 10.1017/CBO9781107415324.
- [116] Andreas Stohl et al. “A new perspective of stratosphere-troposphere exchange”. In: *Bulletin of the American Meteorological Society* 84.11 (2003), pp. 1565–1573+1473. ISSN: 00030007. DOI: 10.1175/BAMS-84-11-1565.
- [117] R J VanDerA et al. “Trends seasonal variability and dominant NO x source derived from a ten year record of NO 2 measured from space”. In: *J. Geophys. Res.* 113.D4 (2008). DOI: 10.1029/2007jd009021. URL: <http://dx.doi.org/10.1029/2007jd009021>.
- [118] A Vasilkov et al. “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 (2017), pp. 333–349. DOI: 10.5194/amt-10-333-2017. URL: <http://www.atmos-meas-tech.net/10/333/2017/>.

- [119] C. Vigouroux et al. “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 (2009), pp. 15891–15957. ISSN: 1680-7316. DOI: 10.5194/acpd-9-15891-2009.
- [120] V Wagner. “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 (2002), p. 4029. ISSN: 0148-0227. DOI: 10.1029/2001JD000722. URL: <http://doi.wiley.com/10.1029/2001JD000722>.
- [121] G M Wolfe et al. “Formaldehyde production from isoprene oxidation across NO x regimes”. In: *Atmospheric Chemistry and Physics* 16.x (2016), pp. 2597–2610. DOI: 10.5194/acp-16-2597-2016. URL: [www.atmos-chem-phys.net/16/2597/2016/](http://www.atmos-chem-phys.net/16/2597/2016/).
- [122] P. J. Young et al. “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 (2013), pp. 5277–5298. ISSN: 16807316. DOI: 10.5194/acp-13-5277-2013.
- [123] X. Yue, N. Unger, and Y. Zheng. “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 (2015), pp. 11931–11948. ISSN: 16807324. DOI: 10.5194/acp-15-11931-2015.
- [124] Xu Yue et al. “Ozone and haze pollution weakens net primary productivity in China”. In: *Atmospheric Chemistry and Physics* 17.9 (2017), pp. 6073–6089. ISSN: 1680-7324. DOI: 10.5194/acp-17-6073-2017. URL: <https://www.atmos-chem-phys.net/17/6073/2017/>.
- [125] P. Zanis et al. “Summertime free-tropospheric ozone pool over the eastern Mediterranean/middle east”. In: *Atmospheric Chemistry and Physics* 14.1 (2014), pp. 115–132. ISSN: 16807316. DOI: 10.5194/acp-14-115-2014.
- [126] G. Zeng et al. “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 (2015), pp. 7217–7245. ISSN: 1680-7324. DOI: 10.5194/acp-15-7217-2015. URL: <http://www.atmos-chem-phys.net/15/7217/2015/>.

- [127] Lei Zhu et al. “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 (2014), p. 114004. ISSN: 1748-9326. DOI: 10.1088/1748-9326/9/11/114004. URL: <http://stacks.iop.org/1748-9326/9/i=11/a=114004?key=crossref.3d2869ee02fd4f0792f831ac8cbecc117>.
- [128] Lei Zhu et al. “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 (2016), pp. 1–24. ISSN: 1680-7375. DOI: 10.5194/acp-2016-162. URL: <http://www.atmos-chem-phys.net/16/13477/2016/acp-16-13477-2016.pdf>.

# **Chapter 8**

# **Appendix A**

## **8.1 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 8.1:** The 47 level vertical grid used by GEOS-Chem