

# Biogenic Isoprene Emissions in Australia

Jesse Greenslade<sup>1</sup> and Jenny Fisher<sup>1</sup>

<sup>1</sup>University of Wollongong, Australia

**Correspondence:** NAME (EMAIL)

**Abstract.** TEXT

## 1 Introduction

Biogenic volatile organic compounds (BVOC) affect the oxidative capacity of the atmosphere and their emissions are largely driven by what type of vegetation is in the area (Kefauver et al., 2014). In the troposphere, BVOC emissions affect hydroxyl radical ( $\text{OH}$ ) cycling, ozone ( $\text{O}_3$ ) production, secondary organic aerosol (SOA) production, and methane lifetime. Australian forests are strong emitters of isoprene, the primary BVOC emitted globally (Guenther et al., 2006; Messina et al., 2016). Isoprene is relatively difficult to measure due to its high reactivity and short lifetime. Poor measurement coverage of isoprene, isoprene products, and isoprene emissions within Australia means that emissions are poorly understood. The lack of knowledge about emissions makes it difficult to estimate the subsequent atmospheric processes.

Emission models used to derive estimates of isoprene fluxes are based on understanding the emissions from different plant species (phenotypes) in varying conditions (see Section ??). Guenther et al. (2012) estimated global biogenic isoprene emissions at roughly  $535 \text{ Tg yr}^{-1}$ , while Sindelarova et al. (2014) estimated around  $411 \text{ Tg yr}^{-1}$ . Reactions following emissions are complex, and are sensitive to other trace gases in the atmosphere. Uncertainties in several important products such as ozone and SOA are increased due to both isoprene measurement difficulties and its complicated subsequent chemical mechanisms.

Isoprene emissions in the frequently used MEGAN model may be overestimated in Australia (see Section ??) since they are based on measurements taken from a few young trees (Winters et al., 2009) that may emit more than older trees (Emmerson et al., 2016). The sample of trees used to construct the MEGAN emissions model included 4 types of Eucalyptus, which are not representative of the hundreds of species that make up Australian forests. Additionally, how these species react to biological and meteorological stresses is unclear (Winters et al., 2009; Fortems-Cheiney et al., 2012). Emissions estimates are necessary inputs for atmospheric chemistry models and improving these estimates for Australia is a primary goal of this thesis.

This chapter describes and implement a *top-down* technique using satellite measurements of HCHO to calculate surface isoprene emissions. HCHO is a primary product of most BVOC (including isoprene) oxidation, and is measured by satellites via remote sensing. In situ isoprene concentration measurements are costly and sparse within Australia, while satellite HCHO data are plentiful and freely available, making this technique very attractive. Top-down techniques have informed isoprene emission inventories in North America (Abbot, 2003; Palmer et al., 2003, 2006; Millet et al., 2006, 2008), South America (Barkley et al., 2013), Europe (Dufour et al., 2008; Curci et al., 2010), Africa (Marais et al., 2012), Asia (Fu et al., 2007;

Stavrakou et al., 2014), India (Surl et al., 2018), and even globally (Shim et al., 2005; Fortems-Cheiney et al., 2012; Bauwens et al., 2016). In this thesis the technique is applied focusing solely on Australia for the first time.

## 1.1 Aims

Recent work suggests that modelled emissions may be overestimated in southeast Australia (Emmerson et al., 2016). This chapter aims to improve the understanding of isoprene emissions over the whole of Australia, clarifying the spatial distribution of bias and how these biases impact modelled chemistry. Isoprene emissions are estimated in Australia using a top-down technique based on OMI HCHO measurements and GEOS-Chem modelled yields. This a posteriori top-down estimate is evaluated against bottom-up a priori estimates and also briefly compared against available ground-based measurements. The GEOS-Chem model is modified to run with the a posteriori isoprene emissions to determine potential impact on modelled chemistry. Goodness of fit between in situ, satellite, and modelled HCHO is determined before and after scaling emissions estimates.

First an outline is provided showing why current isoprene emissions estimates are inadequate and how they can be improved. This thesis discusses literature that shows how the estimates may be too high, and describes how emissions may be calculated using satellite datasets. Section 2 lays out how new isoprene emissions are estimated, with results examined in Section 3. Section 3 includes a comparison of updated satellite HCHO columns (Chapter ??) to available measurements, and an examination of how these changes in emissions would affect ozone concentrations in Australia. Uncertainties for each step along the way are quantified in Section 4.

## 1.2 Existing emissions estimates

MEGAN is one of the most widely used sources for estimating biogenic isoprene emissions. However, along with other models that rely on measured plant emission rates, it is poorly calibrated for Australian conditions. Emissions of isoprene ( $C_5H_8$ ) appear to be overestimated within Australia (Sindelarova et al., 2014; Stavrakou et al., 2014; Bauwens et al., 2016; Emmerson et al., 2016), although the lack of measurements of isoprene emission rates in Australia makes this overestimation difficult to characterise. The bias may be as high as a factor of 6, and is not easy to fix (see Section ??).

Recently Bauwens et al. (2016) estimated isoprene emissions with a top-down technique using the IMAGESv2 global CTM. They calculated emissions that create the closest match between model and OMI satellite vertical columns, and compared these a posteriori data with their a priori (model) data and independent datasets. For Australia they found MEGAN based estimates ranging from  $38 \text{ Tg C yr}^{-1}$ - $94 \text{ Tg C yr}^{-1}$ , and a posteriori emissions of  $36 \text{ Tg C yr}^{-1}$ . This wide range demonstrates the effect of unknowns and uncertainties involved in isoprene emission estimation. Their results are compared to results here in Section 3.1.1. In this thesis a top-down emission estimate is developed, analysed and then compared against MEGAN. How changed emissions affect modelled ozone levels over Australia is also looked into.

### 1.3 Top-down isoprene emissions estimates

In the remote troposphere HCHO production is dominated by methane oxidation, while in the continental boundary layer production is largely due to non-methane VOCs (NMVOCs) (Abbot, 2003; Kefauver et al., 2014). This leads to a causal relationship between enhanced HCHO concentrations and NMVOC emissions at low (< 1 km) altitudes. NMVOCs are generally short lived (< 1 hr), and the most prominent of these is isoprene. Isoprene is emitted and enters the atmosphere in the gas phase, where it begins a complex series of reactions. HCHO is produced with high yield in many reactions beginning with isoprene oxidation (discussed in more detail in Section ??) and it has a lifetime of a few hours (Kefauver et al., 2014).

Top-down estimates determine emissions of a particular species through careful analysis of the measurable products of that species. This generally takes advantage of longer-lived products that may reach an equilibrium in the atmosphere. Continental tropospheric HCHO enhancements can be directly linked to biogenic isoprene emissions. Since 1997, when the Global Ozone Monitoring Experiment (GOME) satellite measurements were first used to measure HCHO, satellites have been used to provide a total column measurement of HCHO, enabling isoprene emissions estimation by top-down methods (Thomas et al., 1998; Palmer et al., 2001; Bauwens et al., 2016). Using satellite information to improve estimates of biogenic emissions has been highlighted as a valuable use of satellite derived datasets (Streets et al., 2013). Here NASA's OMHCHO product based on measurements from the OMI instrument onboard the Aura satellite (see Section ??) is the basis for a top-down biogenic isoprene emission estimate over Australia.

There are two top-down isoprene emission estimation techniques, Bayesian and linear, which are discussed briefly here. Both the linear and Bayesian techniques assume that modelled chemistry is accurate and only try to correct precursor emissions. 75 This is an additional source of uncertainty given existing uncertainties in chemical mechanisms.

#### 1.3.1 Bayesian inversion

Bayesian inversion optimises model parameters in order to minimise the difference between model output and an (ideally) independent dataset such as satellite measurements. Emissions of isoprene (and other precursors to HCHO) will form part of the set of model parameters that are adjusted to make the model HCHO output most closely match satellite measurements. 80 These inversions can be set up to account for effects from transport and allow source attribution (e.g., Curci et al., 2010; Fortems-Cheiney et al., 2012).

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

$$y = \mathbf{K}x + b + \epsilon \quad (1)$$

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

This method was used by Shim et al. (2005) to optimise isoprene emissions in areas with high HCHO concentrations. They  
90 showed model underestimation of isoprene emissions by 14-46%, which was reduced to 3-25% after applying satellite based improvements. More recently Kaiser et al. (2018) showed a 40% bias in MEGAN isoprene emissions over the southeast US using a Bayesian inversion based on OMI HCHO.

An advantage (over the linear method described below) of the Bayesian method is that it can account for pyrogenic and anthropogenic emissions, as these form part of the state variable  $x$  as well as transport. However, biases in the underlying  
95 model will still propagate through to the a posteriori emission estimation (Curci et al., 2010). More limiting is the fact that the Bayesian method is computationally expensive. In this work the Bayesian method is not used due to the computational costs surpassing the resources available.

### 1.3.2 Linear inversion

The linear technique is performed in this thesis. Vertical columns of HCHO from satellite and modelled yield from isoprene  
100 allow the inference of local (grid space) isoprene emissions (Palmer et al., 2003; Millet et al., 2006; Marais et al., 2012; Surl et al., 2018). The primary assumption of the linear inversion technique is that HCHO and its precursors (primarily isoprene) are in a linear steady state relationship. This allows one to link isoprene emissions to HCHO measurements using production and loss rates. Essentially a linear relationship between total column HCHO ( $\Omega$ ) enhancement above a background level ( $\Omega_0$ ) and isoprene emissions ( $E_{isop}$ ) is determined:

$$105 \quad \Omega = S \times E_{isop} + \Omega_0$$

This uses modelled vertical columns and emissions to estimate the slope ( $S$ ). Then this modelled  $S$  is applied to satellite measurements of  $\Omega$  ( $\Omega_{sat}$  and  $\Omega_{sat,0}$ ) to determine  $\hat{E}_{isop}$ :

$$\hat{E}_{isop} = \frac{\Omega_{sat} - \Omega_{sat,0}}{S}$$

This is described further in Section 2, with an outline in Section 2.1.

110 The calculation requires reaction rates and yields from isoprene to HCHO, which can be determined most readily using chemical modelling. The method for calculating isoprene emissions from HCHO is laid out in Palmer et al. (2003), taking into account the expected lifetime and reaction rates of the precursor VOCs and HCHO. In their work, isoprene emissions fluxes over the US were derived using the GOME satellite instrument. The method has since been applied to many regions using GOME, SCIAMACHY, OMI, and GOME-2 satellite data (e.g., Abbot, 2003; Barkley et al., 2013; Stavrakou et al., 2014; Surl  
115 et al., 2018).

The linear inversion assumes fast HCHO yield from isoprene and no precursor transport, which is unrealistic in certain scenarios. For example, high wind speeds can transport precursors, or low  $\text{NO}_x$  concentrations can slow HCHO production (Palmer et al., 2006; Surl et al., 2018). Filtering out data that do not match assumptions is required but can limit the utility of this technique, and leads to some dependence on environmental factors. Uncertainties in the technique are discussed in more  
120 detail in Section 4.1. Nonetheless, a major benefit is that the simple nature of the inversion requires very little computational

power after acquiring satellite and model datasets, even over large amounts of gridded data. This allows an inversion using more than 8 years of satellite and model data, capturing inter-annual variability over all of Australia. With the computational resources available this would not have been possible using the Bayesian inversion.

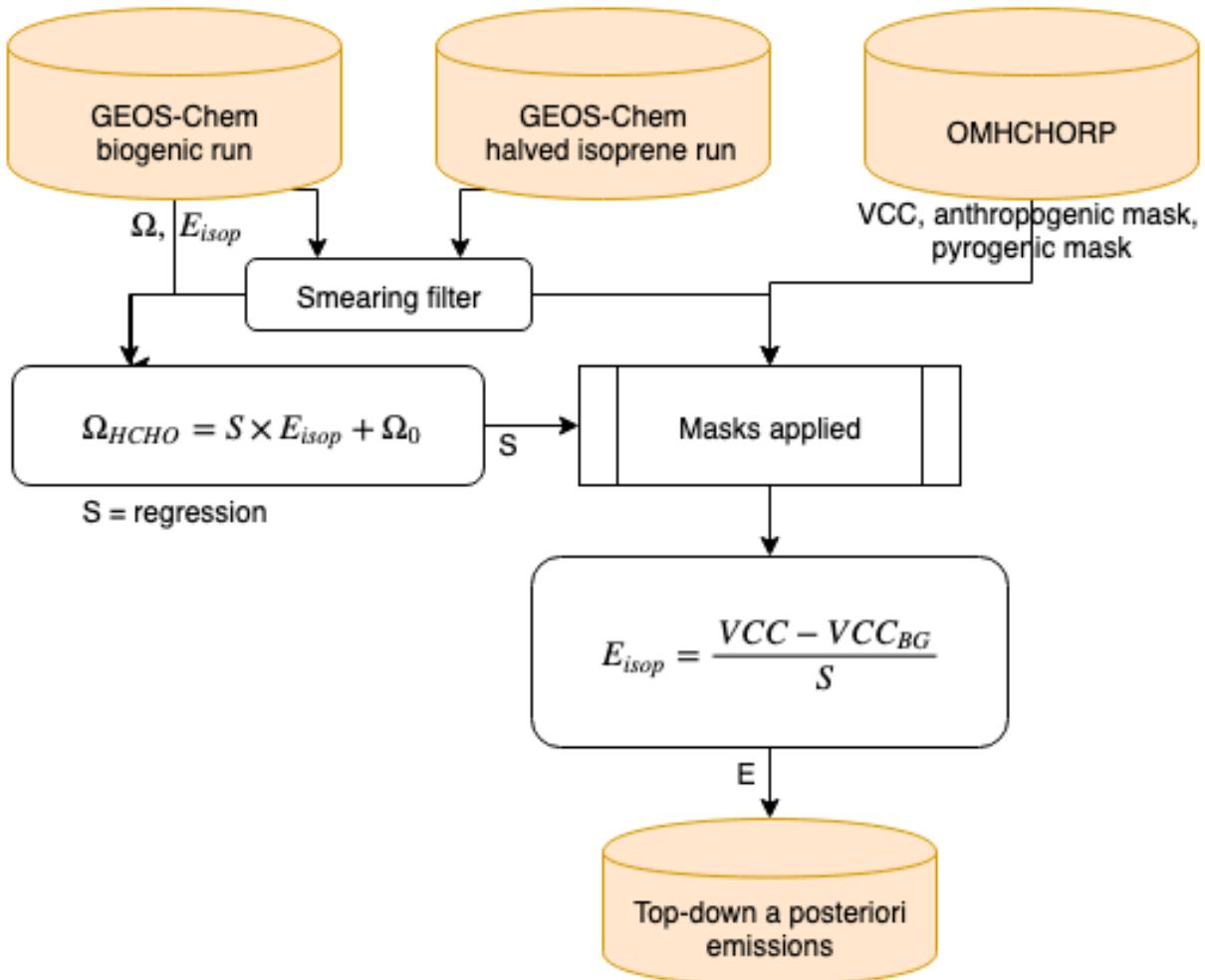
## 2 Methods

- 125 The method of Palmer et al. (2001) is broadly followed to create a biogenic isoprene emissions estimate over Australia. A relationship is modelled between biogenic-only midday tropospheric columns of HCHO and GEOS-Chem midday biogenic isoprene emission rates, and then this relationship is applied to satellite measured HCHO total columns to derive a new isoprene emissions estimate. Daily modelled values averaged between 13:00-14:00 LT are used to match the overpass time of the Aura satellite. Then the slope is calculated using reduced major axis regression between the a priori isoprene emissions (those from  
130 GEOS-Chem,  $E_{GC}$ ) and tropospheric HCHO columns in each model grid box each month. There is very little HCHO above the tropopause, so differences between total and tropospheric column are negligible. In this work total and tropospheric column HCHO are used interchangeably and referred to using  $\Omega$ .

### 2.1 Outline

This section provides an overview of the steps involved in creating a top-down emissions estimate. This process is summarised  
135 in Figure 1.

1. Corrected vertical columns ( $\Omega_{OMI}$ ; saved in the OMHCHORP dataset) are calculated (see Section ??) using level two OMI HCHO satellite data (see Section ??), along with GEOS-Chem model runs (see Section ??). Satellite columns are binned into both  $0.25^\circ \times 0.3125^\circ$  and  $2^\circ \times 2.5^\circ$  horizontal resolutions. In this step model background values (columns over the remote Pacific) are used to correct the vertical columns, which is explained in Section ??.
2. Level three satellite data are used to make anthropogenic, fire, and smoke influence masks (see Section 2.3). These are applied to remove  $\Omega_{OMI}$  that may be influenced by pyrogenic or anthropogenic sources.
3. A mask is created showing where the HCHO production is not dominated by local isoprene emissions. This is determined by calculating smearing over Australia using two model runs with differing isoprene emissions. The smearing value is determined as  $\hat{S} = \Delta\Omega_{GC}/\Delta E_{OMI}$ : the ratio of the differences between model runs of HCHO columns and isoprene emissions. A full description of the creation of this smearing filter is given in Section 2.3.3.
4. GEOS-Chem modelled biogenic emissions of isoprene ( $E_{GC}$ ) along with biogenic columns of HCHO ( $\Omega_{GC}$ ), both averaged over  $2^\circ \times 2.5^\circ$  horizontally and 13:00-14:00 LT temporally, are used to calculate a reduced major axis linear regression slope  $S$  ( $\Omega_{GC} = S \times E_{GC} + \Omega_{GC,0}$ ). Calculation of this modelled slope is explained in Section 2.5.
5. Satellite HCHO  $\Omega_{OMI}$  and  $S$  then form the basis for the top-down estimate of biogenic isoprene emissions ( $E_{OMI}$  atoms C cm $^{-2}$  s $^{-1}$ ). This is the a posteriori, and calculation details are given in Section 2.8.



**Figure 1.** Top-down isoprene emissions estimate formation using OMHCHORP and GEOS-Chem outputs.

6. A posteriori top-down emissions  $E_{OMI}$  are compared against a priori emissions, and analysed in conjunction with independent observations from in situ measurements (MUMBA and SPS). Results are examined in Section 3.
7. GEOS-Chem is run using the a posteriori emissions (see Section 2.9), and HCHO, O<sub>3</sub>, isoprene, and NO<sub>x</sub> outputs are compared to a priori values and to campaign and satellite measurements where possible (Sections 3.3).

## 155 2.2 Masks and reprocessed satellite HCHO

Satellite data pixels are read from OMHCHO, the level 2 OMI HCHO dataset, AMFs are recalculated, and then pixels are gridded daily into  $0.25^\circ \times 0.3125^\circ$  horizontal bins. This forms the intermediate product OMHCHORP, which is fully described in Section ???. This dataset includes gridded satellite HCHO columns ( $\Omega_{OMI}$ ), along with pixel counts (how many satellite data

points were used for each grid box) to allow averaging, re-binning, and uncertainty analysis. In this thesis the OMI product is  
160 used as the overpass time (13:30 local time) is closer to peak isoprene emissions than that of GOME or GOME-2.

In order to determine biogenic HCHO enhancements from  $\Omega_{OMI}$ , filters for non-biogenic sources are required. These  
165 masks are described in Section 2.3, and a brief overview is provided here. Methane oxidation is a major part of the formation  
of background HCHO; however, the linear regression used to estimate isoprene emissions effectively removes this source (by  
assuming a constant background), which means a methane contribution filter is not required. Anthropogenic, pyrogenic, and  
smoke influence masks are created from three satellite products: NO<sub>2</sub> from OMNO2d, fire counts from MOD14A1, and AAOD  
from OMAERUVd, respectively.

1. The fire mask is created daily using non-zero MODIS fire counts over the prior 2 days that occur in local or adjacent grid  
squares at  $0.25^\circ \times 0.3125^\circ$  horizontal resolution.
2. Influence from transported smoke plumes is removed by flagging grid squares where OMI aerosol absorption optical  
170 depth (AAOD, from OMAERUVd) is greater than 0.03.
3. A filter for anthropogenic influence is created daily using OMNO2d NO<sub>2</sub> tropospheric column amounts, masking any  
grid squares with greater than  $2.0 \times 10^{15}$  molec cm<sup>-2</sup> on any particular day, along with grid squares where the yearly  
average is above  $1.5 \times 10^{15}$  molec cm<sup>-2</sup>.

The recalculated corrected vertical columns are saved to OMHCHORP dataset both before and after applying the filters to  
175 allow filter analysis.

### 2.3 Filtering Data

A major goal in this work is to infer biogenic isoprene emissions from HCHO columns. Isoprene is the dominant source of  
HCHO over continental land masses; however, other precursor NMVOCs can contribute to observed HCHO. The main interference in the biogenic isoprene signal in HCHO comes from fire smoke plumes and major anthropogenic source regions.  
180 Biomass burning can be a large local or transported (via smoke plumes) source of HCHO, glyoxal (CHOCHO), and other  
compounds that influence HCHO levels. Anthropogenic emissions from power generation, transport, and agriculture can influence  
these levels as well. So to infer biogenic isoprene emissions, pyrogenic and anthropogenic influences should be removed  
(where possible) from modelled and measured data. In GEOS-Chem in this thesis, pyrogenic and anthropogenic emissions  
185 are turned off in the model run used to estimate biogenic yields. In the OMI HCHO satellite product, potentially pyrogenic or  
anthropogenic affected pixels are masked.

In this work anthropogenic and pyrogenic influences on the OMHCHO satellite HCHO columns are removed by masking active fires, high AAOD, and high NO<sub>2</sub> levels measured by satellite. Active fires and suspected smoke plumes are masked, and together termed the pyrogenic filter. NO<sub>2</sub> measurements are used to mask potential anthropogenic influence. Details of each filter  
190 are provided in the following sub-sections. A summary of yearly filtering over Australian land squares at  $0.25^\circ \times 0.3125^\circ$  resolution is provided in Table 1. Figure 2 shows an example year of anthropogenic and pyrogenic filtering, highlighting how

**Table 1.** Satellite pixels remaining after filtering by active fires, smoke, and anthropogenic masking. In parenthesis are the portion of pixels filtered.

Year	Pixels <sup>a</sup>	Active fires	Anthropogenic	Smoke	Total <sup>b</sup>
2005	3.9e+06	3.5e+06( 9.2%)	3.8e+06( 1.3%)	3.8e+06( 1.7%)	3.4e+06(11.5%)
2006	3.8e+06	3.4e+06(11.2%)	3.7e+06( 2.4%)	3.7e+06( 3.2%)	3.2e+06(14.8%)
2007	3.7e+06	3.4e+06( 9.9%)	3.7e+06( 2.0%)	3.7e+06( 2.6%)	3.3e+06(13.0%)
2008	3.5e+06	3.3e+06( 8.1%)	3.5e+06( 1.4%)	3.5e+06( 1.1%)	3.2e+06(10.0%)
2009	2.7e+06	2.4e+06( 9.1%)	2.6e+06( 1.8%)	2.6e+06( 1.7%)	2.4e+06(11.6%)
2010	2.0e+06	1.9e+06( 6.5%)	2.0e+06( 1.4%)	2.0e+06( 0.5%)	1.9e+06( 8.1%)
2011	1.9e+06	1.7e+06(12.5%)	1.9e+06( 1.9%)	1.9e+06( 2.7%)	1.6e+06(15.2%)
2012	2.5e+06	2.2e+06(14.3%)	2.5e+06( 1.8%)	2.5e+06( 3.3%)	2.1e+06(17.4%)

a: how many land pixels are read over Australia, after cloud fraction filtering.

b: Total removed pixels accounts for overlap between filters.

many days and pixels are removed across Australia. The anthropogenic filter completely removes grid squares over Brisbane, Melbourne, and Sydney. Other major cities in Australia either do not emit enough NO<sub>2</sub> to be clearly identifiable by satellite or are too dispersed and do not breach the threshold to be filtered as anthropogenic. The anthropogenic filter removes from ~0.25%–~5% of grid squares each day. Pyrogenic filtering removes from ~5%–~20% of the available grid square measurements per day. This filter tends to cover forested areas (as they are more prone to burning) as well as some hot spots that are likely due gas flaring or burning.

195

### 2.3.1 Pyrogenic filter

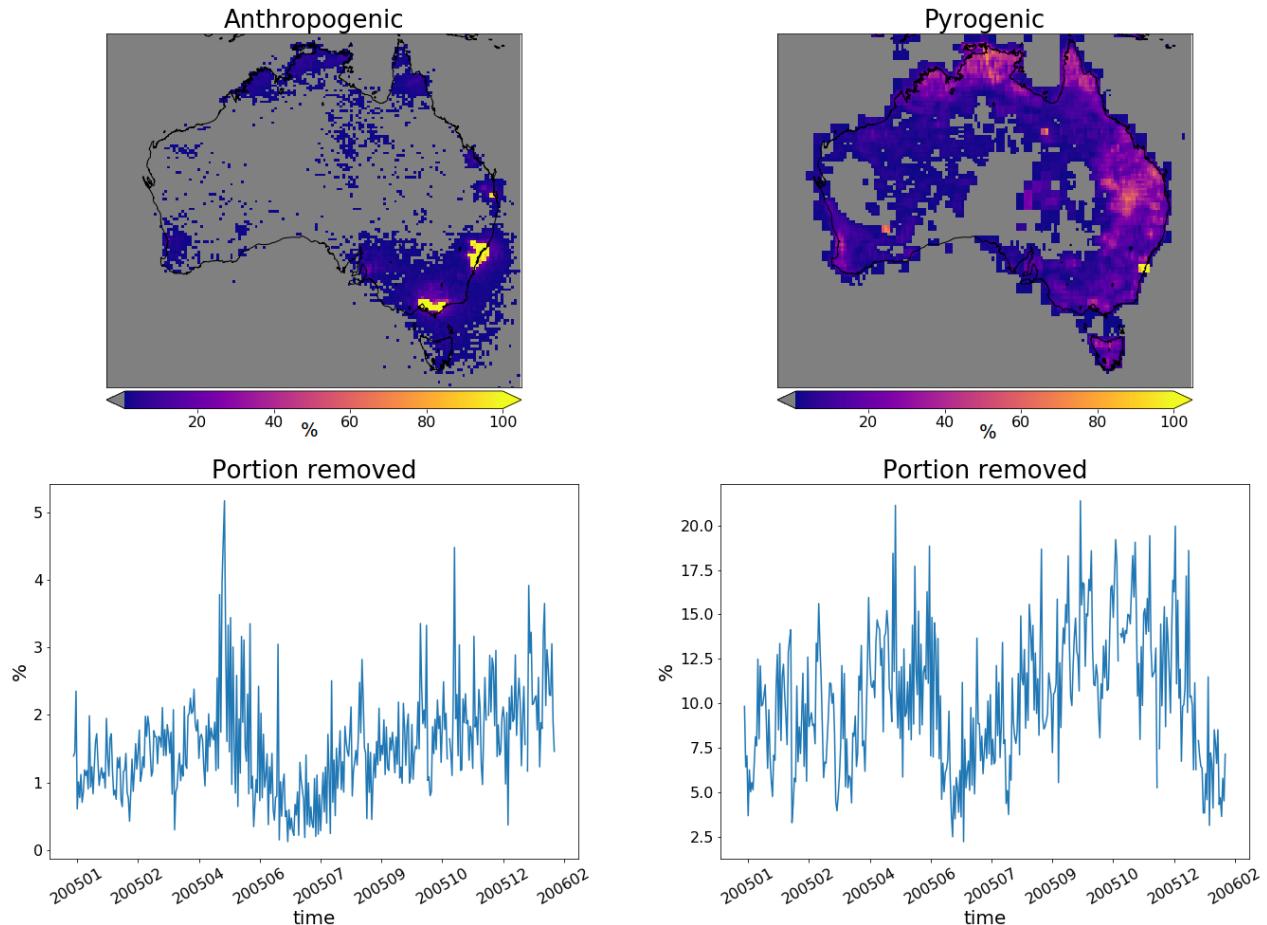
MODIS fire counts are used in conjunction with smoke AAOD enhancements from OMI to remove data points that may be affected by fires or fire smoke plumes. The MODIS fire counts come from a combination of measurements from the Terra and 200 Aqua satellites (Terra overpasses at 10:30, 22:30; Aqua at 13:30, 01:30 local time). Satellite AAOD from product OMAERUVd (described in Section ??) is analysed over Australia to determine a suitable filter threshold. AAOD is used instead of an alternative product AOD as it is less sensitive to the presence of clouds (Ahn et al., 2008).

205

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

1. MOD14A1 daily gridded Aqua/Terra combined fire counts ( $1 \times 1 \text{ km}^2$ ) are binned into  $0.25^\circ \times 0.3125^\circ$  bins (matching the resolution of binned  $\Omega$ ).
2. A rolling mask is formed that removes  $\Omega$  if one or more fires are detected in a grid square, or in the adjacent grid square, up to 2 days previously. This includes the current day, making 3 days of fires in total being filtered out on each day.

Pixels filtered: 2005



**Figure 2.** Top row shows grid squares filtered out by anthropogenic (left) and pyrogenic (right) influence masks over 2005. Bottom row shows portion of Australian grid squares filtered out each day.

- 210        3. AAOD at 500 nm is mapped from OMAERUVd ( $1 \times 1^\circ$  resolution) onto the  $0.25^\circ \times 0.3125^\circ$  resolution.
4. An AAOD threshold of 0.03 is determined through visual analysis of AAOD distributions over several days, including days with and without influence from active fires, dust, and transported smoke plumes. Grid squares with AAOD over this threshold are considered potentially affected by transported fire smoke.

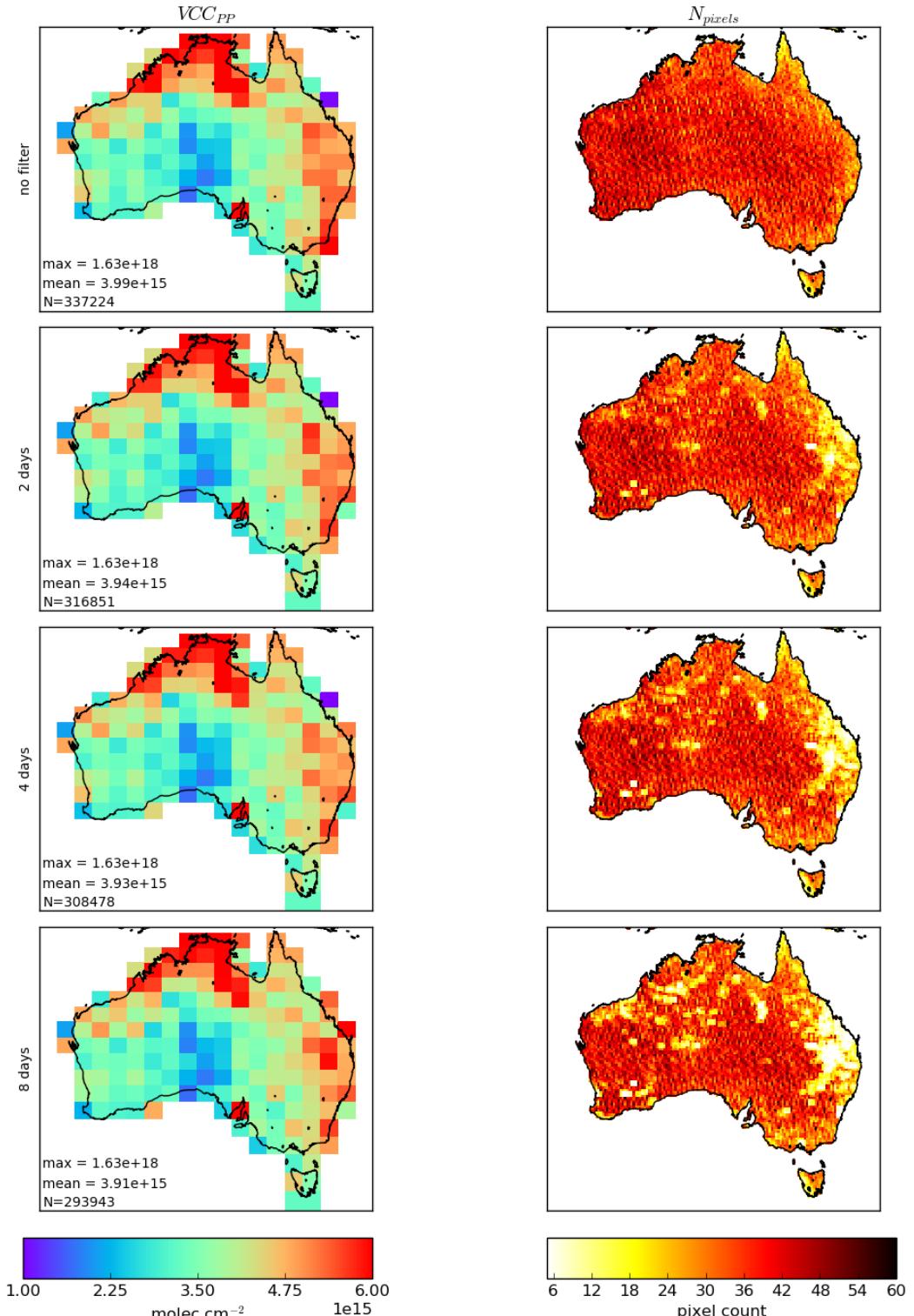
215        This method of masking fires can be compared to Marais et al. (2012) and Barkley et al. (2013): Marais et al. (2012) removed pixels colocated with non zero fire counts in any of the prior eight days, within grid squares with  $1^\circ \times 1^\circ$  resolution, and Barkley et al. (2013) used fires from the preceding and current day, within local or adjacent grid squares, at  $0.25^\circ \times 0.3125^\circ$  resolution.

220        Figure 3 shows the effects from filtering HCHO vertical columns with different fire filtering. Vertical column amounts averaged over January, 2005, and the pixel counts are shown side by side. The figure shows the effects of applying an increasingly strict fire filter. Each row has a stricter fire filter applied from top to bottom, with no fire filter on the first row up to filtering pixels from squares with fires up to 8 days prior. Increasing prior days used when creating the fire filter leads to a serious reduction in available pixels, with a wider range of areas dropping to near zero pixel counts. The overall decrease in vertical column HCHO is minor, dropping from  $\sim 4$  to  $\sim 3.9 \times 10^{15}$  molec  $\text{cm}^{-2}$  between applying no filter and the strictest filter. In this month the fires are most prominent across Queensland and northern NSW, and a large section of the north eastern coast becomes largely filtered out. Increasing the strength of the active fire filter increases the regional impacts on HCHO columns.

225        In this work two prior days of active fires are masked.

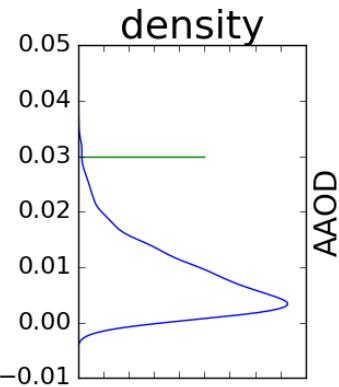
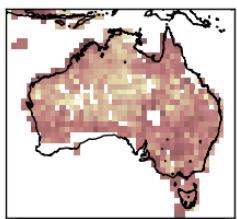
Determining the AAOD due to smoke can be difficult since both smoke and dust absorb UV radiation (Ahn et al., 2008; Marais et al., 2012). AAOD filtering is designed to remove pixels affected by smoke; however, it may occasionally remove pixels affected by dust. Dust in Australia is highly episodic and false positives in the smoke filter should not affect more than a few days per month, especially over regions with high tree coverage (Shao et al., 2007). The threshold is determined through 230 analysing AAOD over Australia in 4 scenarios: normal conditions, active local fires, transported fire smoke, and large scale dust storms. An example of these scenarios and the AAOD distributions is shown in Figure 4. This figure shows AAOD maps and distributions, along with satellite imagery on the same day in column 4 (from <https://worldview.earthdata.nasa.gov/>). The threshold of 0.03 is chosen as it removes very little of a normal day (row 1), and events with higher AAOD show peaks which appear to be distinct from the normal day distribution above this value.

235        Figure 5 shows the extent of pixels filtered out by the pyrogenic filter over 2005 and 2012. One clear hot spot is located over Port Kembla (south of Sydney), most likely due to the flame that burns over the blast furnace stack throughout the year (<https://www.bluescopeillawarra.com.au/community/skylineimages/>). Another hot spot can be seen in Western Australia over Kalgoorlie, where a large open cut gold mine is always open and blasting daily (<https://superpit.com.au/wp-content/uploads/2015/01/Blasting-Information-Sheet.pdf>). In Western Queensland over Mount Isa there is again a mining hotspot, 240 with a blast furnace and several smoke stacks (<http://www.mountisamines.com.au/en/about-mim/Pages/home.aspx>). A large area in southern Queensland/northern NSW is also heavily filtered, potentially due to gas flaring in the Surat Basin, which has thousands of petrol and gas wells (<http://www.ga.gov.au/scientific-topics/energy/resources/petroleum-resources/gas>). The highest concentrations of removed pixels lie along the northern and eastern coastlines, and correspond with forested areas (see

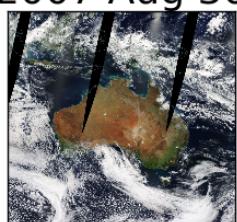


**Figure 3.** Recalculated OMI vertical HCHO columns and pixel counts for January 2005. Column 1:  $VCC_{PP}$  amounts after applying the active fire filter. The filter is applied at  $0.25^\circ \times 0.3125^\circ$  resolution, before output is averaged into  $2^\circ \times 2.5^\circ$  bins. Column 2: Pixel counts after applying the active fire filter (summed into  $0.25^\circ \times 0.3125^\circ$  bins). Row 1-4: increasing number of prior days that have active fires are

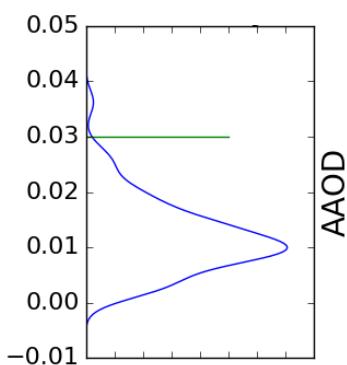
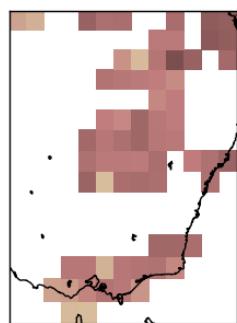
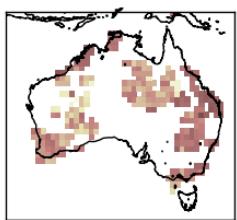
Normal day  
2007 Aug 30



2007 Aug 30



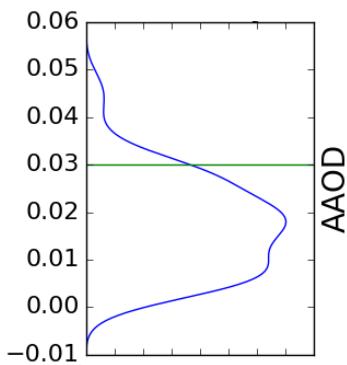
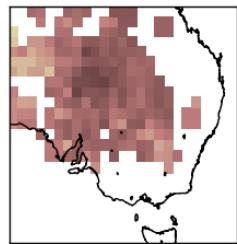
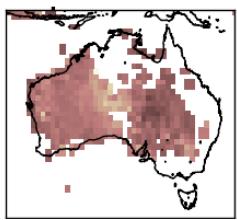
Black Saturday  
2009 Feb 19



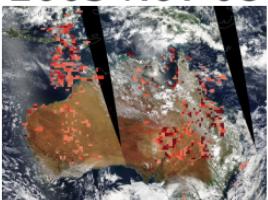
2009 Feb 19



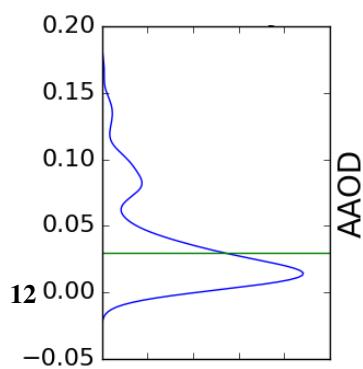
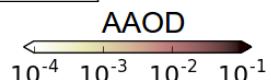
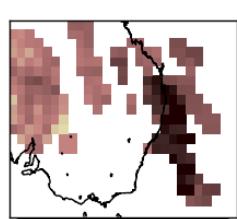
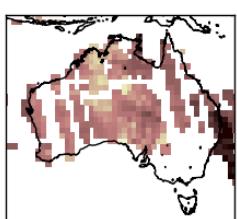
Transported plume  
2005 Nov 08



2005 Nov 08



Dust storm  
2009 Sep 23



2009 Sep 23



Figure 6), which suggests that forest fires are being masked properly. Central Australia is largely unmasked, which could be  
245 due to a lack of sufficient vegetation to create fires and smoke visible by satellite. The proliferation of petrol or gas wells (see Figure 7) may also lead to AAOD enhancement wherever activity stirs up dust, and could be mistaken as active fires wherever gas flaring occurs. In 2005, 388 gas wells existed in Queensland; however, more than 2000 wells (cumulative) were approved by 2013 (Carlisle, 2012). The comparison between fire filtering in 2005 and 2012 (Figure 2.31, bottom plot) does not indicate that the proliferation of gas wells in the intervening years had any strong effect on data filtering.

250 Additional work was performed to validate the pyrogenic filter through an examination of the relationship between temperature and HCHO (Appendix ??). HCHO should be exponentially related to temperature if it is driven by isoprene, with outliers possibly indicative of pyrogenic influence, especially when high emissions are shown on low temperature days. However, fires typically occur on hotter days, and HCHO destruction occurs faster on hotter days, confounding the clear occurrence of outliers (e.g. Zheng et al., 2015). At  $0.25^\circ \times 0.3125^\circ$  resolution, application of the pyrogenic filter slightly strengthens the correlation  
255 over the wider south eastern Australian region, with  $r$  increasing from  $\sim 0.31$  to  $\sim 0.38$ . In this thesis the pyrogenic filter is applied at the higher horizontal resolution ( $0.25^\circ \times 0.3125^\circ$ ), and this analysis suggests that the filter should strengthen the relationship between total column HCHO and its biogenic precursors over Australia.

### 2.3.2 Anthropogenic filter

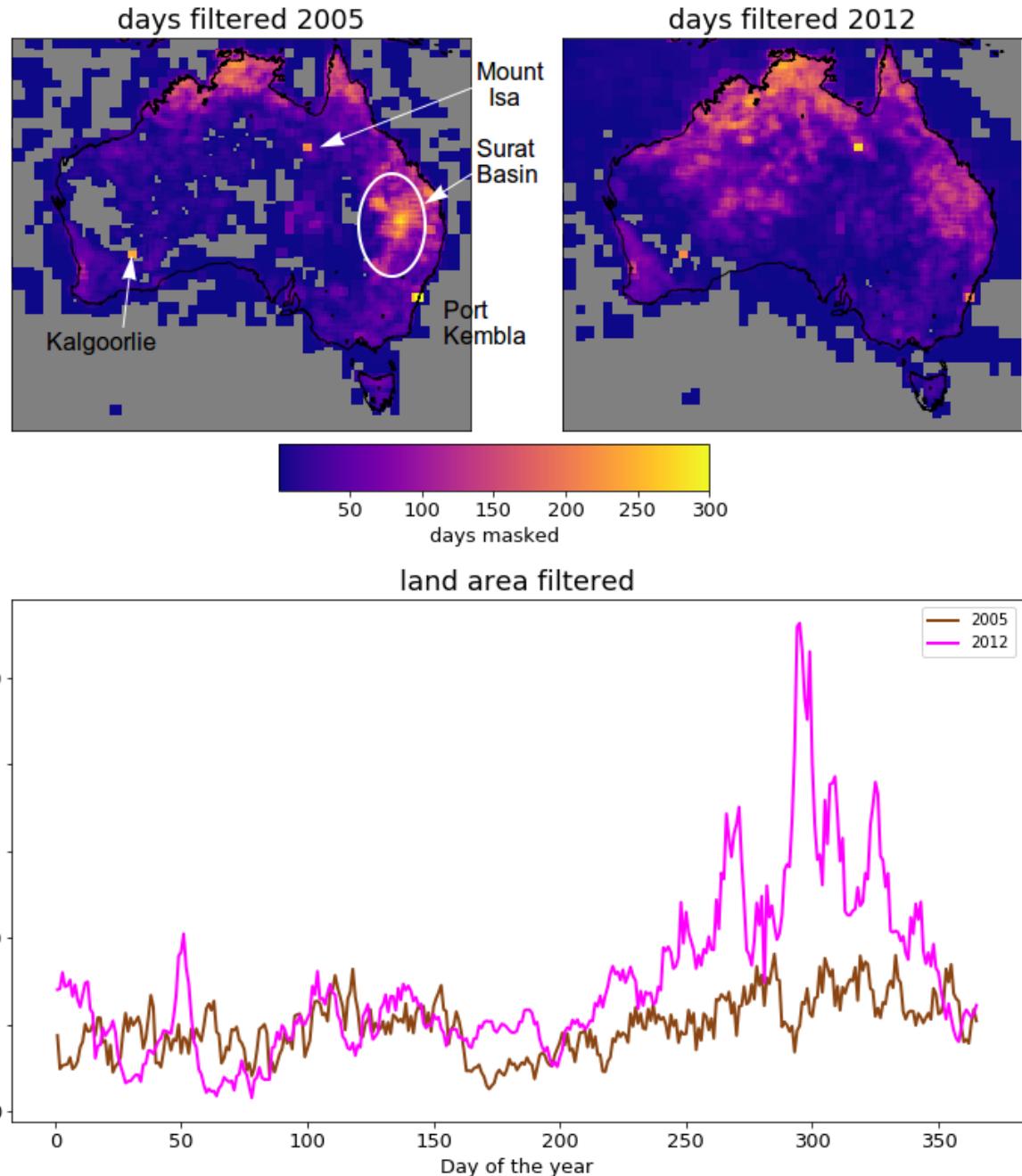
Enhanced  $\text{NO}_2$  concentrations indicate anthropogenic influence over Australia.  $\text{NO}_2$  is used as a proxy for potential anthropogenic NMVOC emissions, as these could bias the inversion performed in Chapter ???. A filter is designed using the tropospheric  $\text{NO}_2$  columns in the OMNO2d product.  
260

$\text{NO}_2$  columns near several major cities in south eastern Australia over 2005 are used to determine a suitable threshold for anthropogenic influence. The mean and standard deviation of tropospheric  $\text{NO}_2$  over Australia of measured by satellite is shown in Figure 8. Anthropogenic influences are clearly visible near major cities in Australia, and some influence can be seen  
265 along nearly every coastline.

The anthropogenic filter is created for each year in two steps:

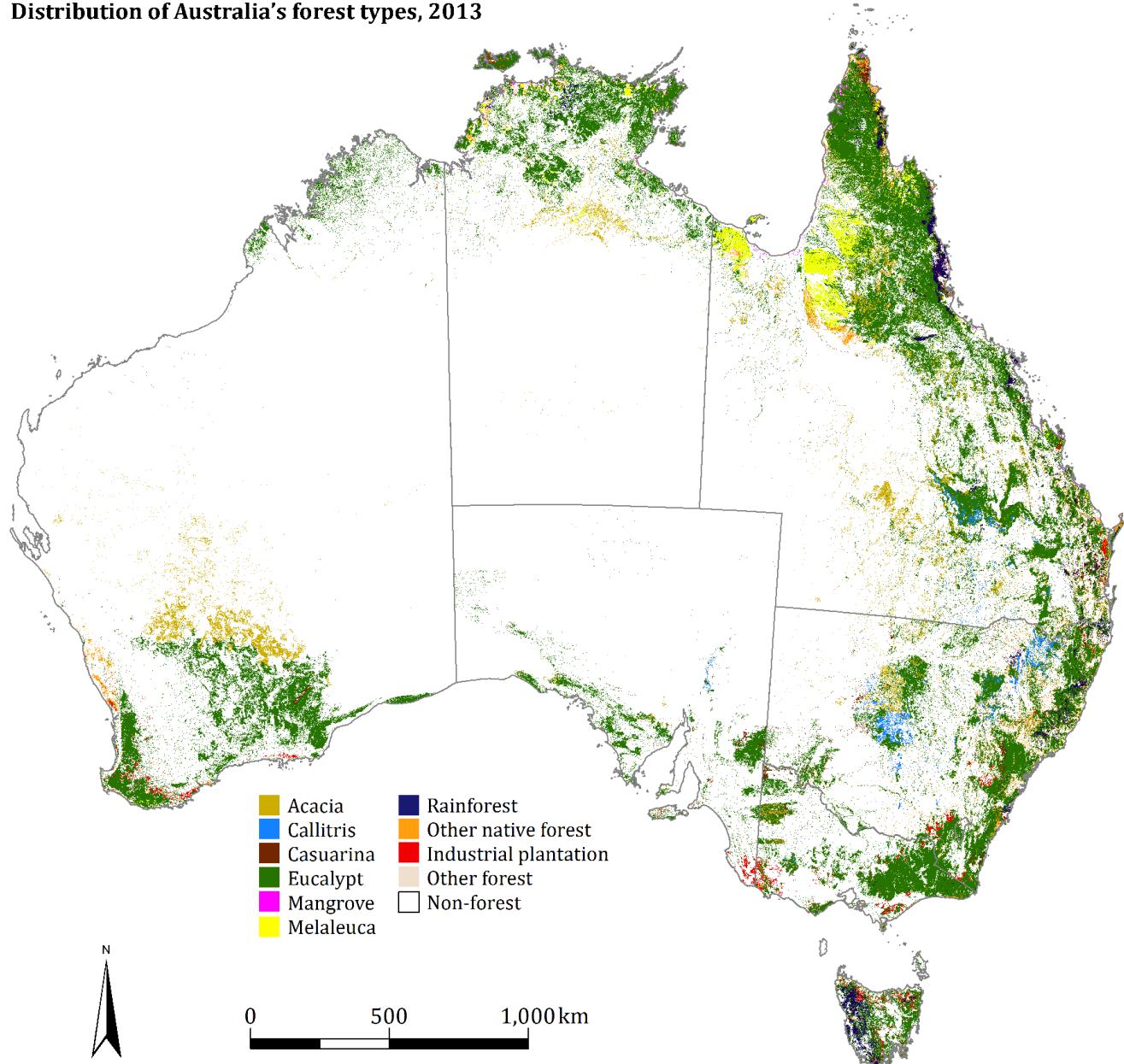
1. Daily grid squares with  $\text{NO}_2$  greater than  $2.0 \times 10^{15}$  molec  $\text{cm}^{-2}$  are flagged as anthropogenic.
2. After taking the yearly average for each grid square, any tropospheric  $\text{NO}_2$  columns greater than  $1.5 \times 10^{15}$  molec  $\text{cm}^{-2}$  are flagged for the whole year.

270 These thresholds are chosen subjectively to ensure the removal of definite anthropogenic influence while not removing too many data points over nearly uninhabited portions of Australia. These thresholds completely remove grid squares over major cities that are likely emitting NMVOCs year round, and also frequently remove grid squares downwind. The effects of applying this filter to the OMNO2d product itself can be seen in Figure 9. Areas over Sydney and Melbourne tend to oscillate around the upper quartile of Australian  $\text{NO}_2$  tropospheric columns once the filter is applied. Many coastal regions show some effect  
275 from the filtering, and Sydney, Melbourne, and the La Trobe Valley (east of Melbourne) are completely removed.



**Figure 5.** Top: percentage of pixels filtered out by fire and smoke masks in 2005 (left) and 2012 (right). Bottom: percentage filtered out each day from land squares in Australia for the two years shown.

## Distribution of Australia's forest types, 2013

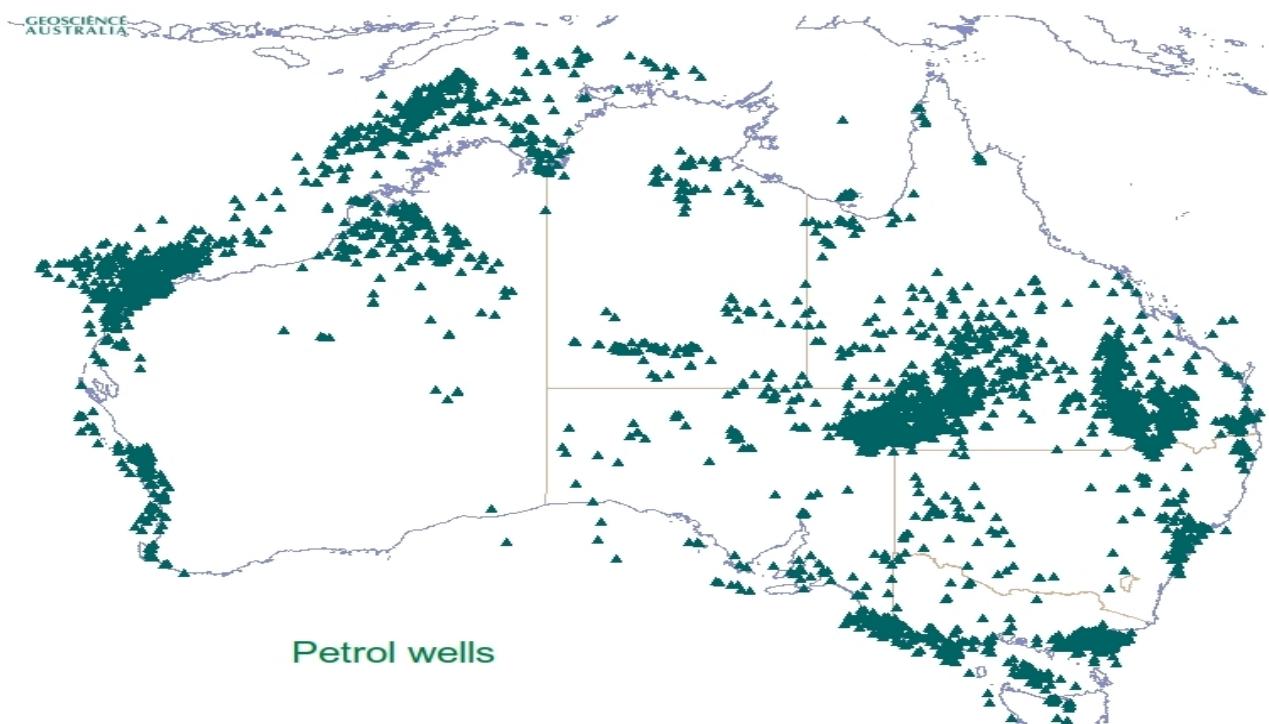
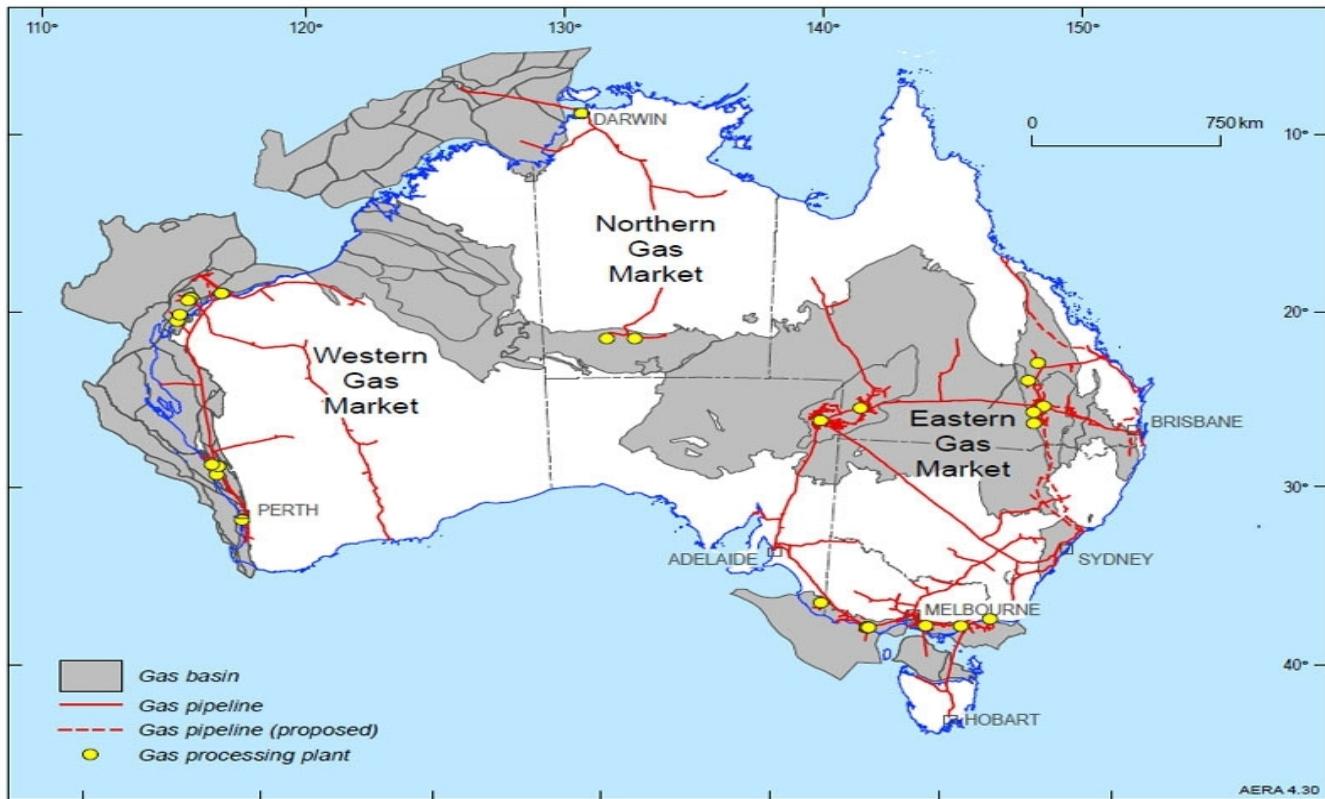


Projection: Albers equal-area with standard parallels 18°S and 36°S

Source: ABARES (2016) Map compiled by ABARES 2016

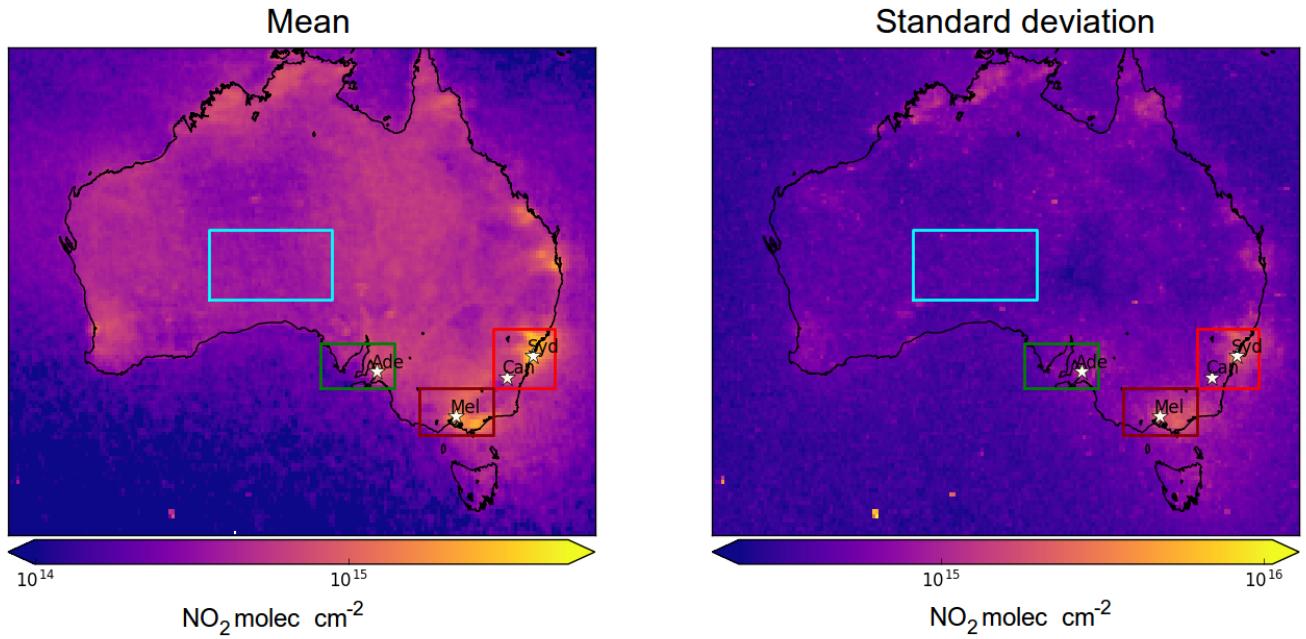
© Commonwealth of Australia 2016. This material is licensed under a Creative Commons Attribution 3.0 Australia Licence.

**Figure 6.** Forest coverage, coloured by predominant tree species.



**Figure 7.** Top: gas fields and pipelines (2018) for Australia (<http://www.ga.gov.au/scientific-topics/energy/resources/petroleum-resources/gas>). Bottom: petrol Well locations over Australia (as of 2018) (<http://dbforms.ga.gov.au/www/npm.well.search>).

## OMNO2d NO<sub>2</sub> total columns 2005



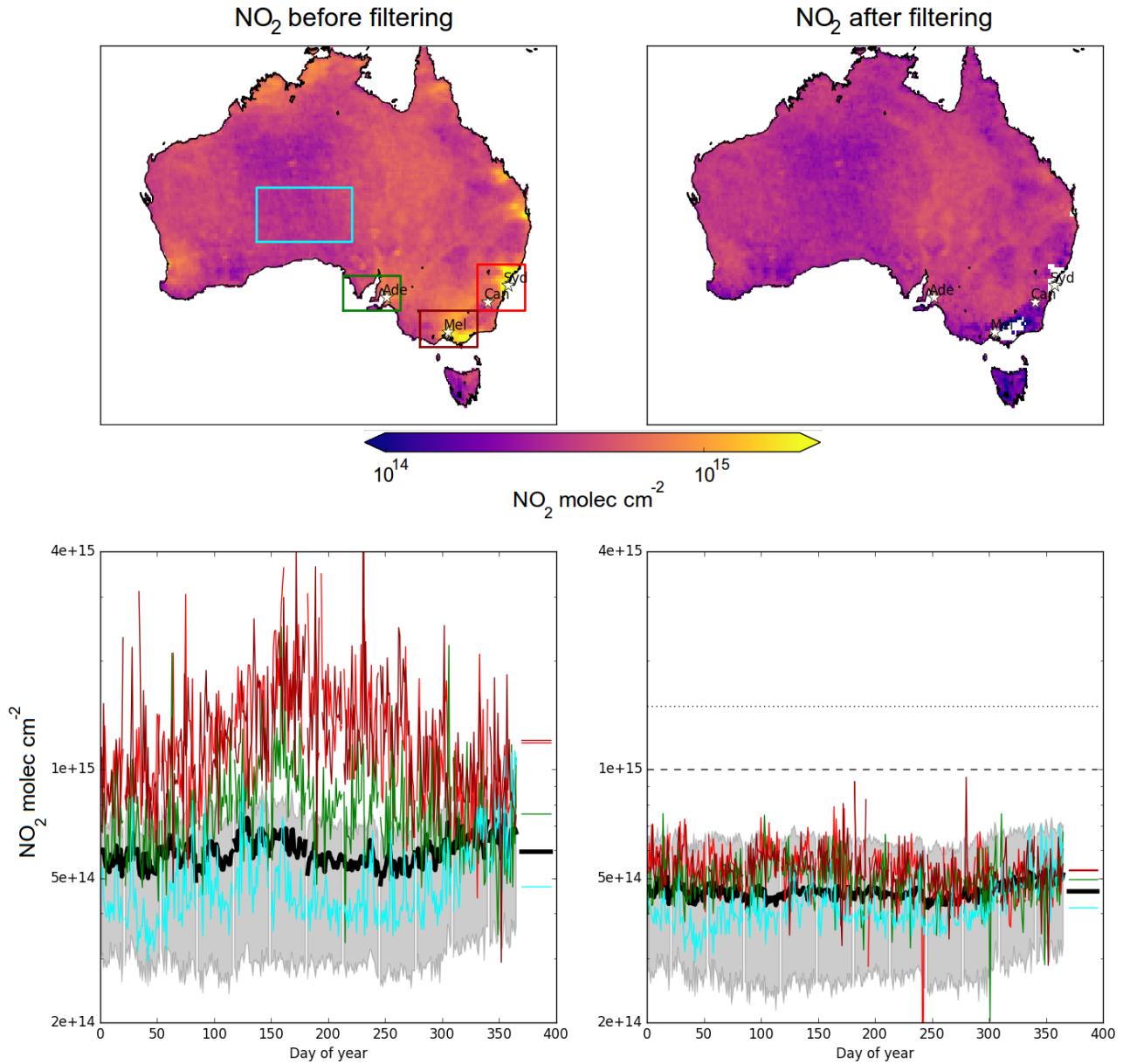
**Figure 8.** Mean (left) and standard deviation (right) of OMNO2d daily  $0.25^\circ \times 0.25^\circ$  tropospheric cloud filtered NO<sub>2</sub> columns.

**Table 2.** NO<sub>2</sub> averages ( $10^{14}$  molec cm<sup>-2</sup>) by region before and after filtering for anthropogenic emissions using 2005 data from the OMNO2d product.

Region	NO <sub>2</sub>	NO <sub>2</sub> after filtering	% Data lost
Aus	4.9	4.5	2.1%
Sydney	10.4	5.9	25.3%
Adelaide	6.2	5.4	4.0%
Central Aus	3.8	3.7	0.5%
Melbourne	9.6	5.3	25.4%

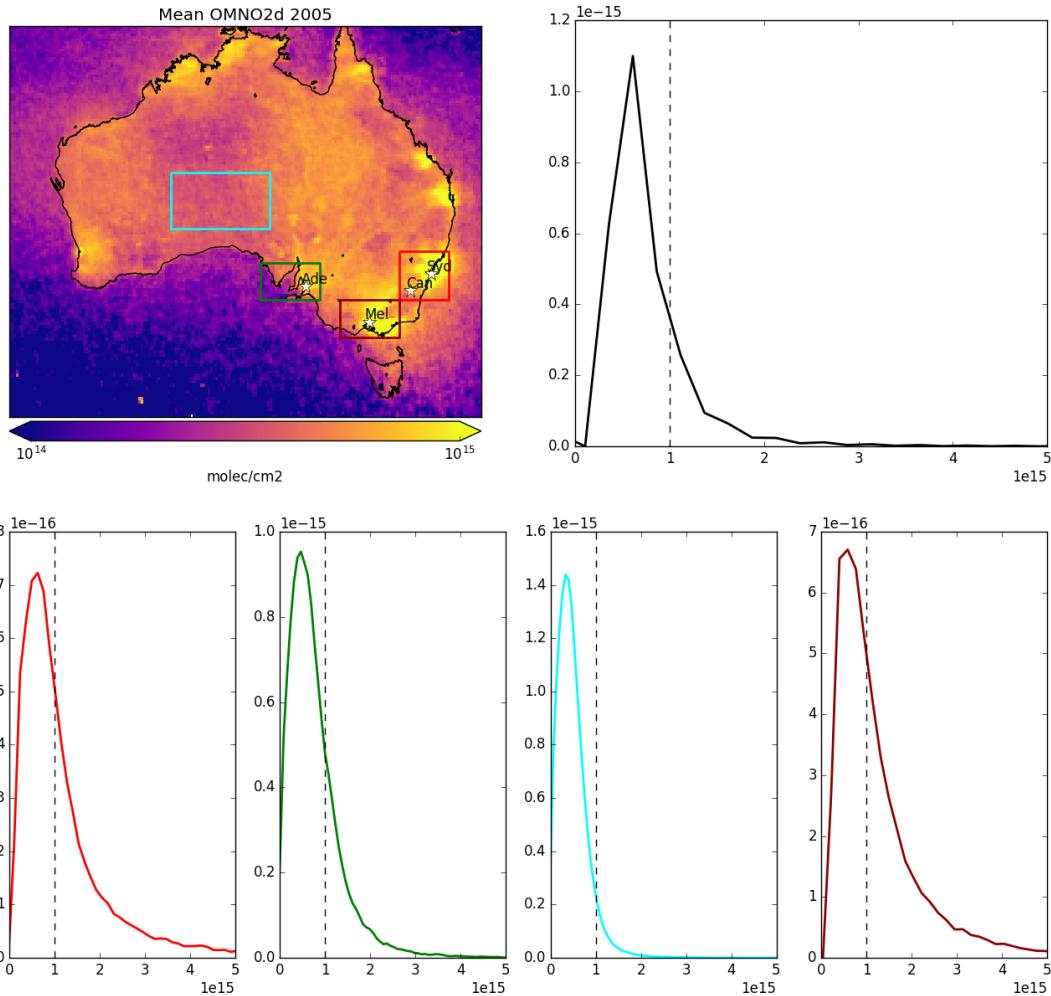
The same regions as in Figure 9 are shown again in Figure 10, with NO<sub>2</sub> pixel densities for each region shown, along with the thresholds of  $1.5$  and  $2 \times 10^{15}$  molec cm<sup>-2</sup>. The reduction in NO<sub>2</sub> columns along with the portion of available data removed over Australia and each subset region is listed in Table 2. Roughly a quarter of the available data points are removed over Sydney and Melbourne, decreasing the mean NO<sub>2</sub> amounts by  $\sim 50\%$ , while not too much information is lost overall.

## OMNO2d threshold effects 2005



**Figure 9.** Top row: 2005 OMNO2d  $\text{NO}_2$  column mean before (left) and after (right) applying the anthropogenic threshold filters as described in the text. Bottom row: time series for Australia, and each region (by colour) shown in the bottom panels, with mean for that region shown as horizontal bars within the bottom panels at the right (over 360–400 on the x-axis). The time series before and after applying the anthropogenic filter are shown on the left and right respectively. The thicker black line shows the mean value over all of Australia, and the grey shaded area shows the inter-quartile range. The dashed and dotted horizontal lines show daily and yearly mean thresholds respectively.

## OMNO2d NO<sub>2</sub> columns 2005



**Figure 10.** 2005 OMNO2d NO<sub>2</sub> column means (top left) and distributions (top right) for Australia, and each region shown in the area map (by colour). Vertical dashed lines show the threshold for anthropogenic influence; any columns above this value are filtered out. The vertical axis is normalised so that area under the curve adds to unity, and as such is not important except as a visual measure of the relative width between the distributions.

280 **2.3.3 Smearing filter**

Smearing is a measure of how much HCHO in a given grid box was produced from isoprene emitted in a different (upwind) grid box. Smearing affects emissions estimates as HCHO enhancements downwind of where precursor emissions occurred lead to misinterpretation of local emissions. In high  $\text{NO}_x$  ( $>\sim 1 \text{ ppb}$ ) environments, isoprene has a lifetime on the order of 30 minutes, and HCHO can be used to map isoprene emissions with spatial resolution from 10-100 km (Palmer et al., 2003).  
 285 In low  $\text{NO}_x$  conditions, isoprene has a longer lifetime (hours) and may form HCHO further from the source area (Fan and Zhang, 2004; Liu et al., 2016, 2017). Over Australia,  $\text{NO}_x$  levels are generally low and smearing is therefore expected to be important. Smearing limits the horizontal resolution of the linear top-down inversion process, as a finer resolution increases sensitivity to transport. Horizontal transport *smears* the HCHO signal so that its source location would need to be calculated using wind speeds and loss rates (Palmer et al., 2001, 2003). In this chapter smearing affected grid squares are filtered out prior  
 290 to application of Equation 10.

**2.3.4 Calculation of smearing**

Smearing has been analysed in several publications (e.g., Martin, 2003b; Palmer et al., 2003; Millet et al., 2006; Stavrakou et al., 2009; Marais et al., 2012; Barkley et al., 2013; Zhu et al., 2014; Wolfe et al., 2016; Surl et al., 2018) and is often calculated using the method used in this thesis, as first described by Palmer et al. (2003). This involves using two model runs,  
 295 one of which has isoprene emissions scaled globally by a constant (generally from 0.5 to 2). From Section 2.5, Equation 10 states that the modelled slope ( $S$ ) is the yield of HCHO per C of emitted isoprene divided by the HCHO loss rate ( $S = \frac{Y_{isop}}{k_{HCHO}}$ ). Using two runs of GEOS-Chem with differing isoprene emissions but otherwise identical provides:

$$\begin{aligned} Run_1 : \Omega_{HCHO} &= SE_{isop} + \Omega_0 \\ Run_2 : \Omega'_{HCHO} &= S'E'_{isop} + \Omega'_0 \end{aligned} \quad (2)$$

There are several assumptions that need to be understood, as these are what is tested by the smearing calculation. The initial  
 300 assumption is that the system is at steady state, with no transport of isoprene affecting HCHO columns. This is the basis for equation 2. It is also assumed that background values ( $\Omega_0$ ) are from oxidation of methane and other long-lived VOCs, so that  $\Omega_0 = \Omega'_0$ . Between these two runs only the  $E$  term is changed. Chemistry is unchanged so that the yield and loss rate should not change between the two runs:

$$S = S' = \frac{Y_{isop}}{k_{HCHO}} \quad (3)$$

305 Equations 2 may then be combined as follows:

$$\begin{aligned} Run_1 - Run_2 : \Omega_{HCHO} - \Omega'_{HCHO} &= SE_{isop} - S'E'_{isop} + \Omega_0 - \Omega'_0 \\ \Omega_{HCHO} - \Omega'_{HCHO} &= S(E_{isop} - E'_{isop}) \\ \Delta\Omega_{HCHO} &= S\Delta E_{isop} \\ \hat{S} \equiv & \frac{\Delta\Omega_{HCHO}}{\Delta E_{isop}} \approx \frac{Y_{isop}}{k_{HCHO}} \end{aligned} \quad (4)$$

**Table 3.** Smearing filters or typical slopes ( $S$ ) from literature.

Source	min. (s)	max. (s)	type <sup>a</sup>	Region
Palmer et al. (2003)	1270	2090	Range	North America <sup>b</sup>
Marais et al. (2012)		4000	Limits	Africa
Barkley et al. (2013) <sup>c</sup>	1300	1800	Limits	South America
Surl et al. (2018)	2200	4900	Range	India
This Thesis	800	4600	Limits	Australia

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

b: Slope range for summer only.

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

310 This allows the combination of outputs from the two runs to determine where  $\hat{S}$  diverges from expected values for  $S$ .

Potential smearing is masked by checking a daily modelled value for  $\hat{S} \approx Y_{isop}/k_{HCHO}$  against thresholds. By assuming that midday HCHO lifetime ( $\tau = 1/k_{HCHO}$ ) typically falls within 1.5 to 4 hrs (as seen in the USA; Palmer et al. (e.g., 2006); Wolfe et al. (e.g., 2016)) and isoprene-to-HCHO yield (HCHO per isoprene carbon emitted) lies within the range 0.2 to 0.4 (scenarios estimated in Palmer et al. (2003)), one can set a simple bound on  $\hat{S}$  of  $[0.2 \times 1.5, 0.4 \times 4]$  hrs or 1080 to 5760 seconds.

315 As  $\text{NO}_x$  levels across Australia are relatively low, and lower  $\text{NO}_x$  levels reduce the prompt yield (Palmer et al., 2003; Wolfe et al., 2016). In this work the threshold range is the simple bound shown here, reduced by 20% and rounded to the nearest hundred leading to a bounding range of 800 to 4600 for  $\hat{S}$ . This range strikes a balance between unlikely modelled yields and the amount of data lost to filtering. Table 3 compares the smearing filter for Australia used in this thesis to typical values used in previous work for other regions.

320 GEOS-Chem is run with normal isoprene emissions and with isoprene emissions halved, then Equation 4 ( $\hat{S} = \frac{\Delta \Omega_{HCHO}}{\Delta E_{isop}}$ ) provides  $\hat{S}$ . Here  $\Delta$  represents the difference (daily 1300-1400 local time) between default and scaled runs. If  $\hat{S}$  sits outside the 800-4600 range then that grid square day is removed from both  $S$  and subsequent a posteriori emission calculations. A relatively large change in  $\Omega_{HCHO}$  compared to local emissions ( $\hat{S} > 4600$ ) suggests HCHO production is from non-local isoprene emissions. Alternatively, a relatively low value of  $\hat{S}$  ( $\hat{S} < 800$ ) suggests emissions from the local grid square are 325 being exported before they form HCHO.

### 2.3.5 $\text{NO}_x$ dependence

$\text{NO}_x$  concentration directly affects the fate of VOCs in the atmosphere, influencing HCHO production by isoprene and therefore the amount of smearing. In low  $\text{NO}_x$  environments, reported HCHO yields from isoprene are around 0.2 - 0.3 C per C (or 100-

**Table 4.** Isoprene to HCHO yields and lifetime.

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

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

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

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

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

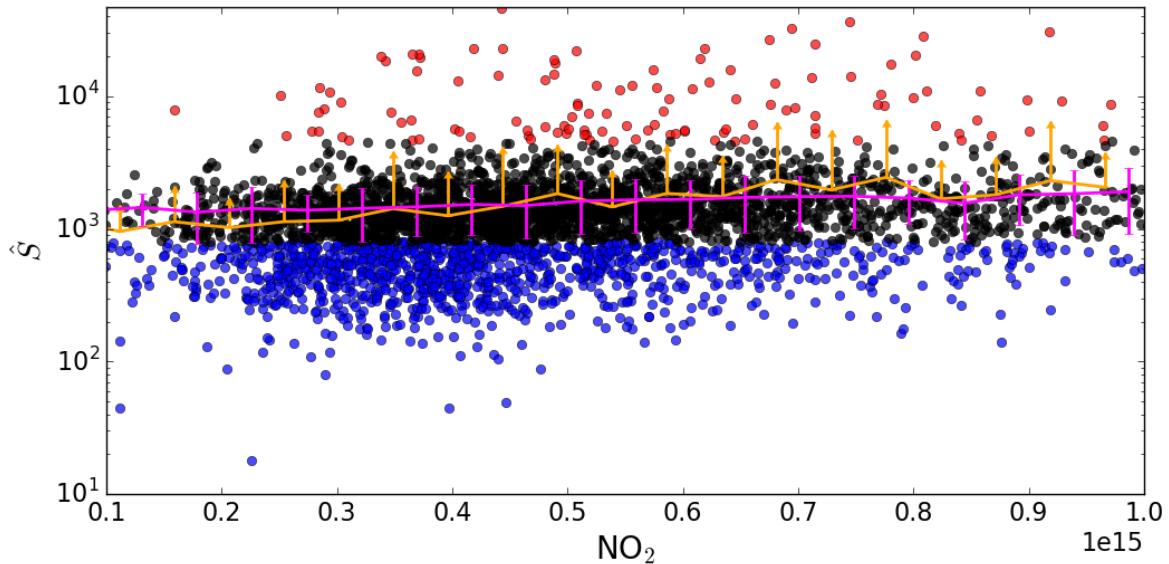
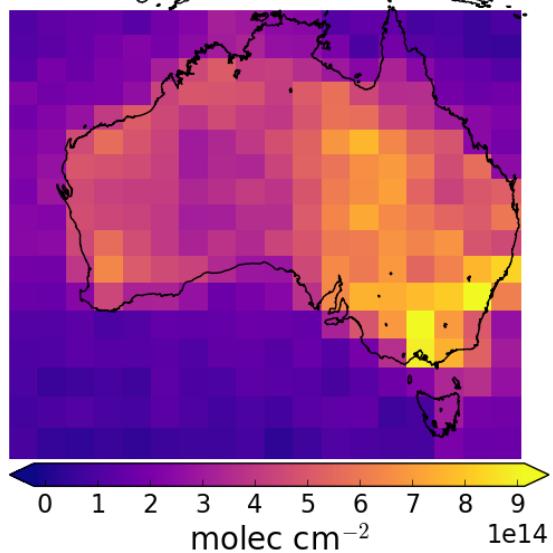
150 molar %), while in high NO<sub>x</sub> environments this value becomes two to three times higher (Palmer et al., 2003; Wolfe et al., 2016). Some values for HCHO yield from prior literature are shown in Table 4.

The effect of NO<sub>2</sub> on smearing can be seen in Figure 11. This plot shows how smearing over Australia compares to satellite NO<sub>2</sub>, with smearing distributions binned by NO<sub>2</sub> both with and without filtering for smearing. At lower NO<sub>2</sub> levels smearing is more frequently below the lower threshold. These low values decrease in frequency, and have less affect on the mean  $\hat{S}$  at around  $5 \times 10^{14}$  molec cm<sup>-2</sup> NO<sub>2</sub>.

## 335 2.4 GEOS-Chem emissions

In this work MEGAN (version 2.1) is run as a module within GEOS-Chem (version 10.01) (see Section ??). The chemical model is driven by GEOS-5 meteorological fields at 0.5°x 0.666° horizontal resolution. GEOS-Chem output is averaged onto 47 vertical levels at 2° × 2.5°, based on chemistry and transport calculated every 30 and 15 minutes respectively. Isoprene

OMNO2d on 200501



**Figure 11.** Top: OMNO2d tropospheric NO<sub>2</sub> columns averaged into  $2^\circ \times 2.5^\circ$  horizontal bins for Jan, 2005. Bottom: Scatter plot of NO<sub>2</sub> against smearing calculations from GEOS-Chem ( $\hat{S}$ ), with points above and below the smearing threshold range of 800–4600 s coloured red and blue respectively. Points are binned by NO<sub>2</sub> with and without the smearing filter applied (orange and magenta respectively). Overplotted is the mean and standard deviation (error bars) within each bin. Due to the logarithmic Y scale only the positive direction of standard deviations is shown for unfiltered data.

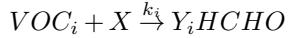
emissions from the default *tropchem* simulation are referred to as the a priori emissions. When shown as part of a formula the

340 a priori emissions are denoted as  $E_{GC}$ .

## 2.5 Relationship between isoprene emissions and formaldehyde

Tropospheric HCHO production is primarily due to the oxidation of VOC precursor species ( $VOC_i$ ). Background concentrations are driven by methane; a longer lived ( $\sim 8$  yr) VOC. Over continental land masses, the variability in HCHO is driven by shorter lived precursor emissions (Chance et al., 2000; Palmer et al., 2003). HCHO is then produced quickly from short-lived

345 intermediates:



where  $X$  is an oxidant,  $Y_i$  is HCHO yield (per C atom in  $VOC_i$ ), and  $k_i$  is the reaction rate constant. In specific conditions described below, HCHO total columns ( $\Omega$ ; molec  $\text{cm}^{-2}$ ) can be linearly related to isoprene emissions. In Australia the effective molar HCHO yield from isoprene has not been extensively studied, while in other continents this value varies from 1-3  
350 depending on local  $\text{NO}_x$  concentrations (e.g., Palmer et al., 2006; Millet et al., 2006; Bauwens et al., 2016; Surl et al., 2018).

The isoprene to HCHO relationship is derived using several assumptions that are outlined here. The first assumption is that HCHO is at steady state, which implies production ( $P_{HCHO}$ ) and loss ( $L_{HCHO}$ ) are equivalent:

$$\frac{d\Omega}{dt} = 0 = P_{HCHO} - L_{HCHO} \quad (5)$$

This is reasonable during midday when isoprene emissions are steady and  $\Omega$  has had time to stabilise. The second assumption

355 is that loss (with loss rate constant  $k_{HCHO}$ ) is only first order, such as from photolysis and oxidation:

$$L_{HCHO} = k_{HCHO}\Omega \quad (6)$$

This assumption means that loss due to transport must be negligible as it is not first order. This assumption is reasonable for large enough grid boxes as transport becomes negligible relative to the linear (first order) terms. Production and loss are on the order of minutes, and grid box sizes in this work are rectangular with  $\sim 200$  km edge lengths. Monthly averaged wind speeds

360 rarely exceed 20 km  $\text{hr}^{-1}$  over Australia, meaning HCHO and precursor transport remain minor. Transport can still be an issue, however, and is handled by applying a smearing filter described in Section 2.3.3.

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

$$P_{HCHO} = \sum_i Y_i E_i \quad (7)$$

365 By combining Equations 5, 6, and 7,  $\Omega$  can be related to precursor emissions:

$$k_{HCHO}\Omega = \sum_i Y_i E_i$$
$$\Omega = \frac{1}{k_{HCHO}} \sum_i Y_i E_i \quad (8)$$

Finally, it is assumed isoprene emissions are driving changes in  $\Omega$  (as assumed elsewhere, e.g., Palmer et al., 2003; Millet et al., 2008; Marais et al., 2014; Stavrakou et al., 2015) and lump other terms together:

$$\sum_i Y_i E_i = Y_{isop} E_{isop} + \sum_{i \neq isop} Y_i VOC_i \quad (9)$$

370 This assumption is reasonable only over continental land masses, and can be false when pyrogenic or anthropogenic emissions influence the HCHO column. These scenarios are filtered to remove pyrogenic and anthropogenic influence using independent satellite measurements (see Section 2.3). The linear relationship between isoprene emissions and  $\Omega$  is determined by equating  $P_{HCHO}$  and  $L_{HCHO}$  from Equations 7 and 6, substituting Equation 9, and assuming that the lumped terms make up the background:

$$\begin{aligned} k_{HCHO}\Omega &= Y_{isop} E_{isop} + \sum_{i \neq isop} Y_i VOC_i \\ \Omega &= \frac{Y_{isop}}{k_{HCHO}} E_{isop} + \Omega_0 \\ 375 &= S \times E_{isop} + \Omega_0 \end{aligned} \quad (10)$$

Here  $S$  is the slope of the regression between  $E_{isop}$  and  $\Omega$  :  $S \equiv \frac{Y_{isop}}{k_{HCHO}}$ .

## 2.6 Calculation of modelled slope

To determine  $S$ , the link between biogenic isoprene and midday column HCHO, GEOS-Chem is used. The term  $E_{GC}$  is used when discussing isoprene emissions estimated within GEOS-Chem and  $\Omega_{GC}$  is used to represent simulated HCHO total 380 column. The method to calculate  $S$  using GEOS-Chem follows roughly the following three stages:

1. Hourly gridded model outputs  $E_{GC}$  (atoms C cm $^{-2}$  s $^{-1}$ ), along with  $\Omega_{GC}$  ( molec cm $^{-2}$ ) at 13:30 LT daily are extracted from the biogenic-emissions-only run.
2. Filtering removes gridded output on days where grid squares are likely to be affected by smearing (see Section 2.6.1).
3. A reduced major axis regression slope is determined between  $\Omega_{GC}$  and  $E_{GC}$  for every month of modelled output (one 385 value per day) for each grid square (Figure 12).

Modelled background concentrations can be ignored here as they do not affect slope calculation. This effectively provides the monthly gridded slope ( $S$ ) between biogenic isoprene emissions and HCHO columns, in units of seconds.

Figure 12 (top left) shows how  $S$  varies spatially over Australia for an example mid-summer month. High  $S$  values suggest high sensitivity of HCHO to isoprene emissions. Some areas can be seen to be very sensitive to emissions, such as the west 390 coast and Eyre basin, which is likely due to the low isoprene and HCHO levels in those areas. The regression coefficients also vary spatially (top right), and some areas show little correlation. It is likely that this is due to weather, transport, and a lack of local emission sources. The slopes in the bottom panel show a small sample of scatter and regression plots. These can range widely due to differences in emission and yield parameters, which plays a role in the smearing filters described in Section

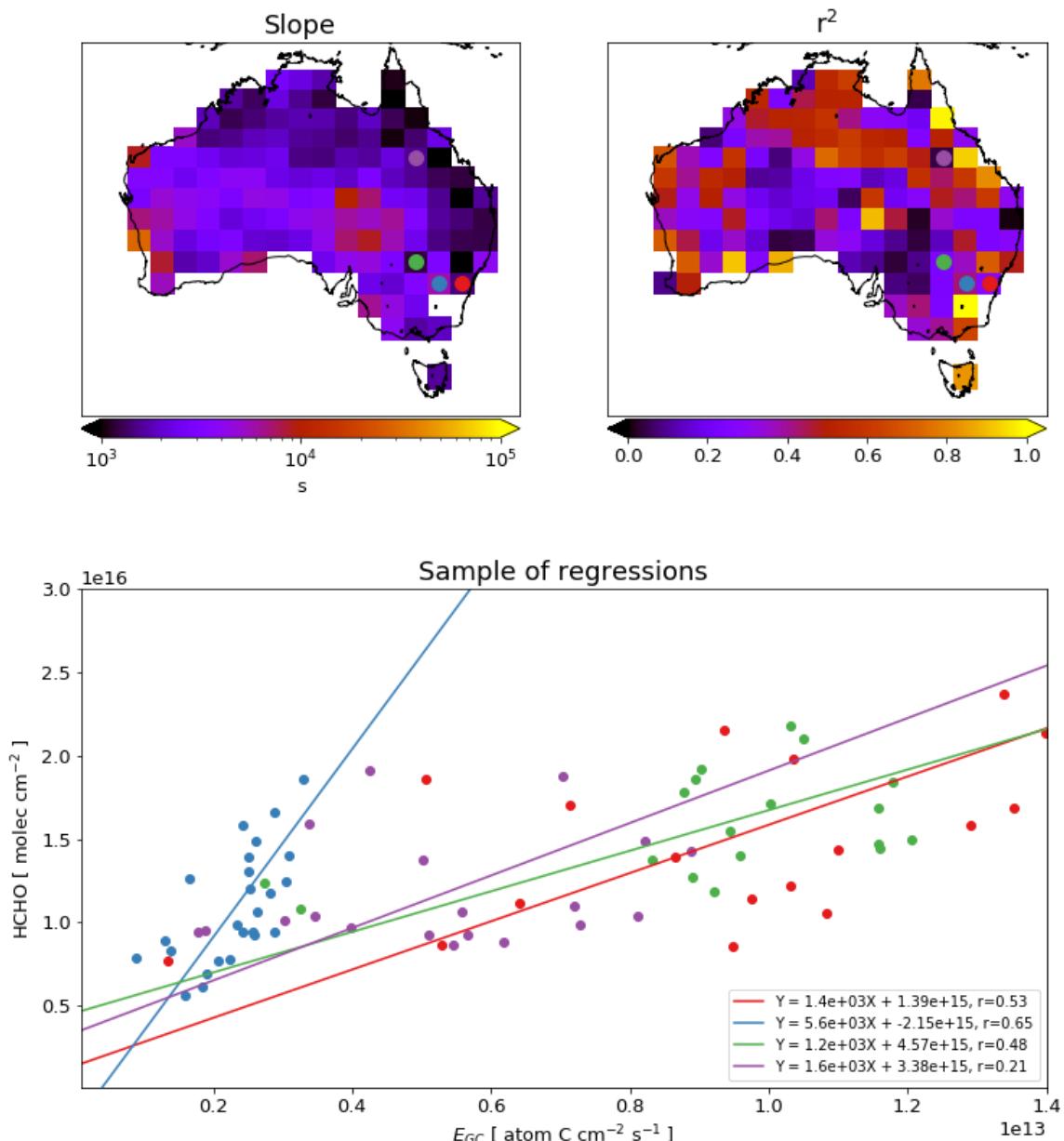
2.3.3. Due to the  $2^\circ \times 2.5^\circ$  horizontal resolution of GEOS-Chem, calculations over coastal grid boxes that are mostly oceanic  
395 are often discarded as the change in HCHO is not dominated by emissions of isoprene, as is assumed for equation 10.

### 2.6.1 Smearing

One issue with slope calculation is potential transport (also known as smearing), either of isoprene transported in from outside  
the local grid box (before any HCHO is formed), or of HCHO formed by local emissions but transported out of the local grid  
box. A *smearing* filter is created (see Section 2.3.3) which filters out grid squares on days when transport is likely. Days where  
400 smearing is expected to affect local levels of HCHO are removed before calculating  $S$ , and a simple analysis is performed on  
how the filter affects monthly slope, correlation, and uncertainty. Figure 13 shows the calculated slope for 2005-2012, along  
with its 95% confidence interval for the  $2^\circ \times 2.5^\circ$  grid box containing Sydney. The monthly and multi-year monthly averages  
are shown before and after the smearing filter is applied. The filter slightly reduces the amplitude of the seasonal cycle,  
raising the January minimum and lowering the June and July maximum. Filtering slightly improves the correlation coefficient  
405 throughout the year. This is likely due to smearing being prevalent in areas where the relationship between a priori emissions  
and HCHO columns is already be weak (due to low emissions or unsuitable meteorological conditions). More data are filtered  
in summer, presumably due to higher biogenic isoprene emissions over summer, making transport more noticeable on windy  
days. Anthropogenic precursor emissions are not halved between the two runs, and they make up relatively more of the total  
HCHO precursor concentration in winter (when biogenics are lower). This may lower the smearing filter signal over Sydney  
410 and other densely populated areas (especially in winter). This plot has been repeated for several grid squares over Australia (not  
shown) sometimes showing more filtering throughout winter months, and sometimes over summer months. When calculating  
top-down emissions the smearing-filtered slope ( $S$ ) is generally used for each grid square month. The multiple year monthly  
averaged slope is used instead when the regression coefficient ( $r$ ) is less than 0.4, or the number of data points used in the  
regression ( $n$ ) is less than 10. When  $r$  for the multiple year slope is also lower than 0.4 (does not happen in the example grid  
415 square), no estimation is performed. This happens only infrequently, and only in locations with very limited isoprene emissions  
such as the lake Eyre basin.

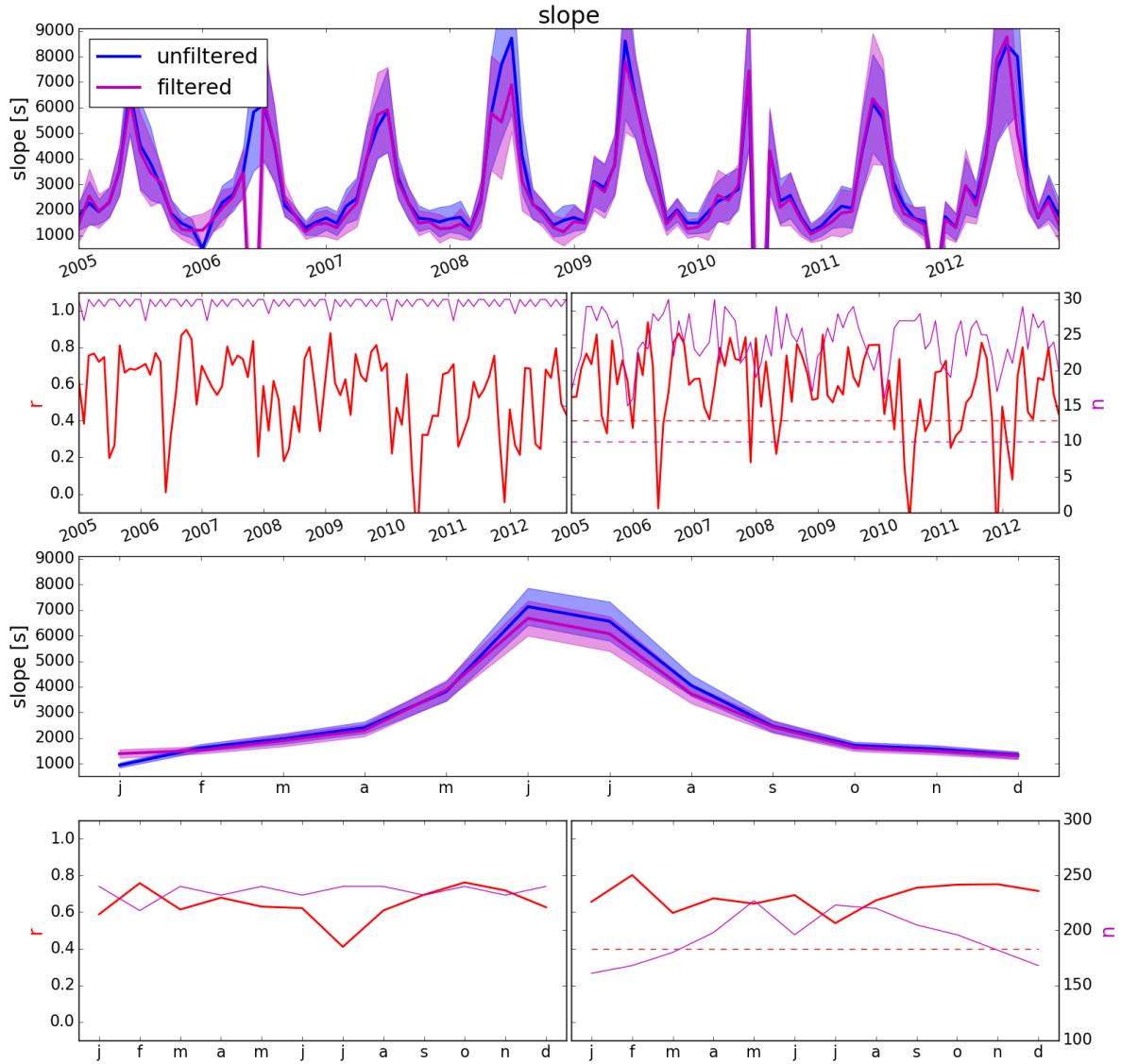
Smearing can be dependent on local or regional weather patterns, as greater wind speeds will reduce the time any emitted  
compound stays within the local grid box. As such smearing can vary both spatially and temporally. Smearing is also sensitive  
420 to time of day, season, and latitude, as lower insolation results in slower photolysis. Figure 14 shows smearing and how  
frequently grid squares are filtered using the smearing filter. The smearing filter is more active in winter and spring, especially  
at higher latitudes. Grid squares along the east coast are filtered frequently in all seasons, as well as grid squares in the high  
smearing lake Eyre region in northeastern South Australia. Data loss from smearing is approximately 30% over the entirety  
of Australia during summer, and 40% during winter. The data loss is generally higher towards the north east and southern  
coastlines, and at lake Eyre.

## HCHO trop column vs isoprene emissions Jan, 2005



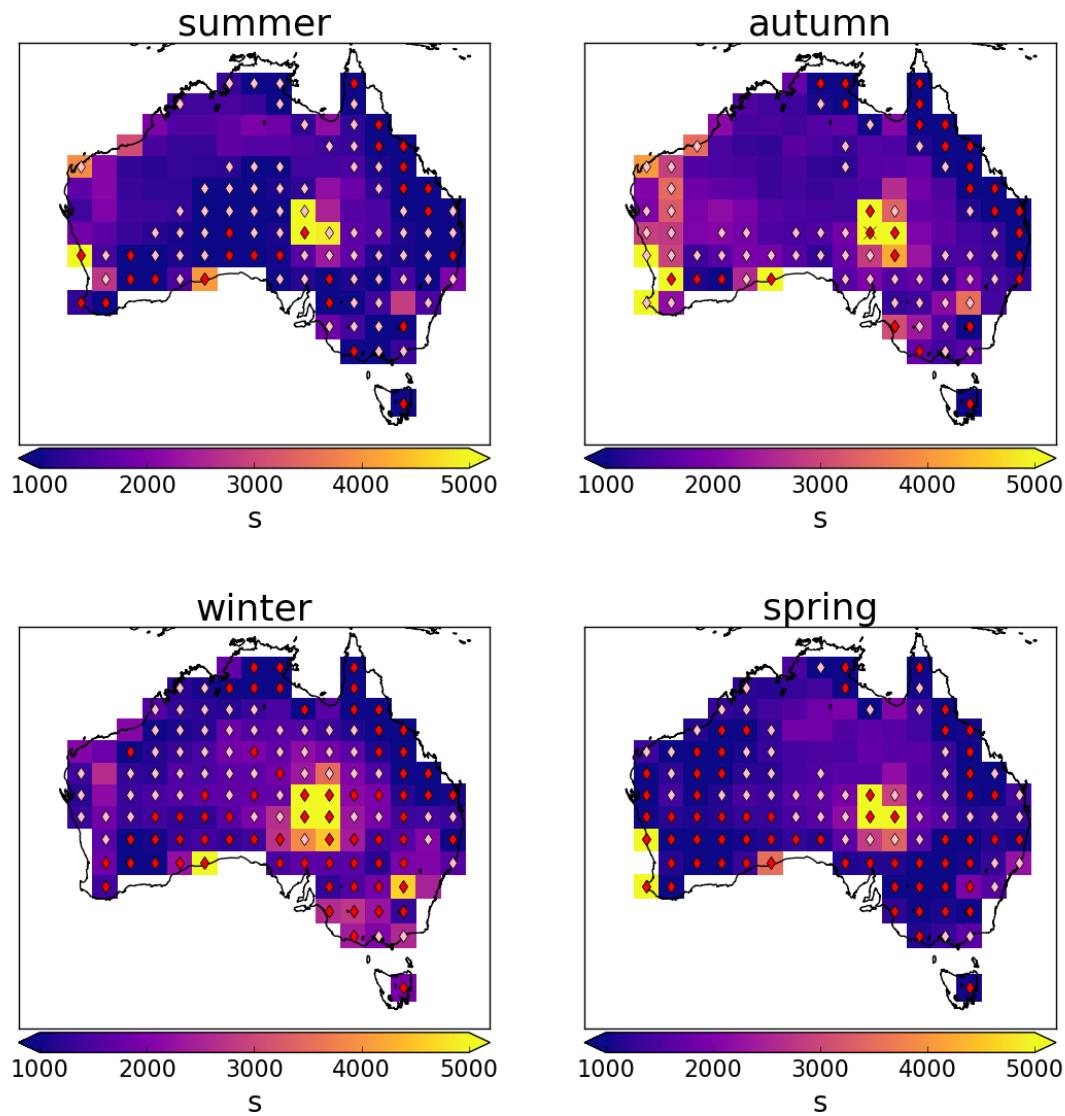
**Figure 12.** Top left: reduced major axis slope between modelled tropospheric column HCHO and isoprene emissions ( $E_{GC}$ ) using midday (13:00-14:00 LT) values over for January 2005, per grid square at  $2^\circ \times 2.5^\circ$  horizontal resolution. Top right: Squared reduced major axis correlation coefficient for regression in top left. Bottom: Sample of correlations from four grid squares. Coloured dots in top panels correspond

## Model Slope -34.000,150.000



**Figure 13.** Row 1: monthly slope along with 95% confidence interval both before (blue) and after (magenta) applying the smearing filter for the model grid square containing Sydney over 2005-2012. Row 2: regression coefficient and data point counts before (left) and after (right) performing filtering for slopes shown in row 1. Additionally, limits for  $r$  and  $n$  used in slope utilisation (see text) are shown with dashed lines. Row 3: slope and confidence interval using the multi-year dataset for each month. Row 4: regression coefficient and data point counts for row 3.

## smearing filter application



**Figure 14.** Seasonally averaged smearing ( $\hat{S}$ , see text) over 2005. Diamonds represent grid squares which have had at least 10 (pink) or 30 (red) days removed due to the smearing filter over the season. Red crosses show where the filter has removed all data.

425 **2.7 Modelled background HCHO**

There are two simple ways to determine the modelled background HCHO, one of which involves running the model with no isoprene emissions. Since variation in HCHO columns is assumed to only depend on isoprene emissions, the background term is theoretically identical to the simulated HCHO without isoprene emissions. The other method uses HCHO over the remote Pacific Ocean at matching latitudes and times, which emulates how the background is determined for the satellite measured HCHO. Figure 15 shows the background total column HCHO calculated in these two different manners, and how they compare to each other and normal levels for an example month (January, 2005). The difference between these definitions is approximately 15%, and is mostly caused by non-isoprene HCHO precursors (such as monoterpenes). Non-isoprene HCHO precursor emissions are neglected in the top-down estimation technique used in this work, which introduce some uncertainty (Section 4.2). Generally in summer months the continental HCHO levels are over 3 times higher than background levels, so the choice of background definition has only a very small impact on the final results. For consistency with the satellite data, backgrounds are determined using the remote Pacific. Background HCHO for any latitude in this thesis is calculated by averaging longitudinally (140°W to 160°W) the matching latitudes over the remote Pacific.

430 **2.8 Calculation of Emissions**

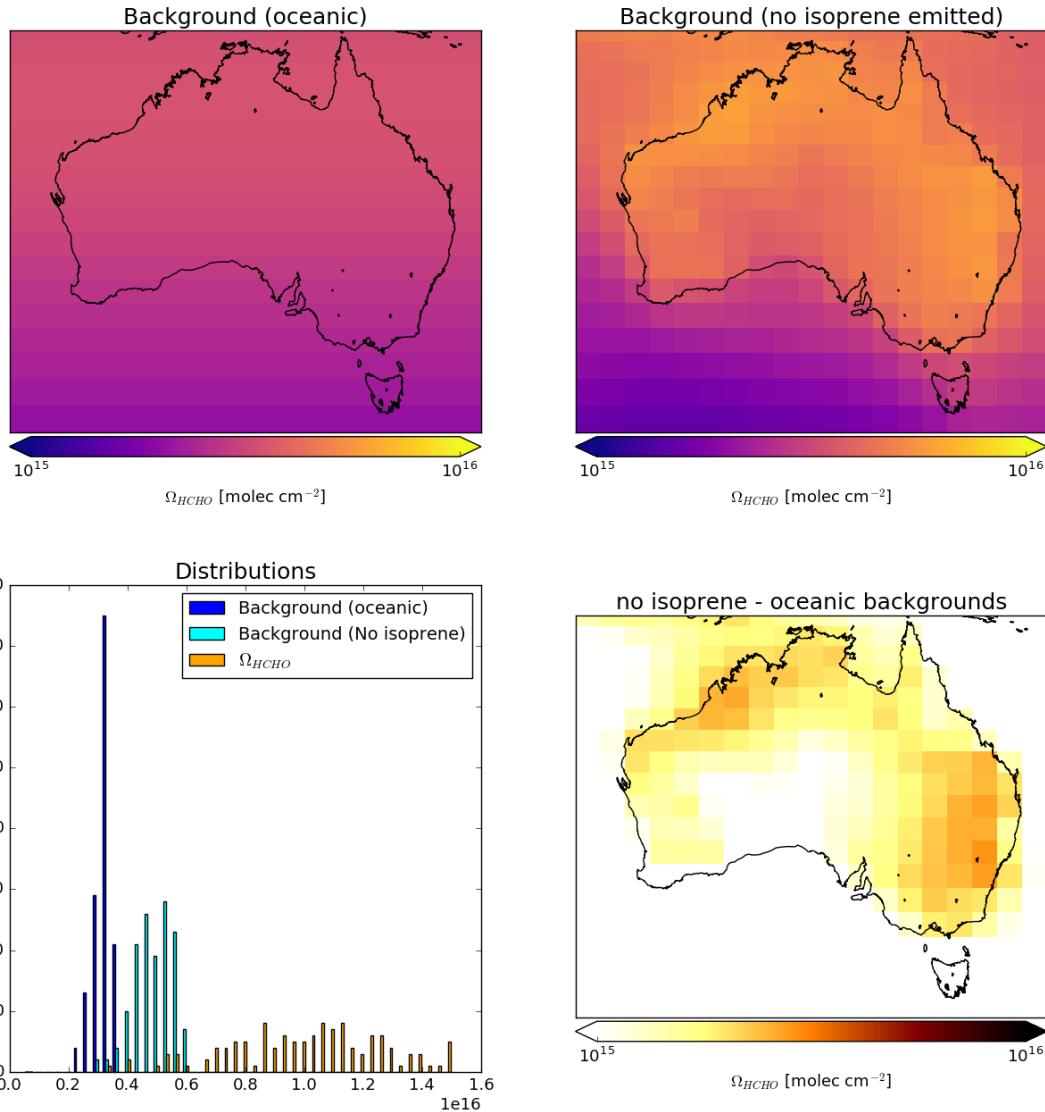
Top-down emissions estimates are calculated using OMHCHO (see Section ??) slant columns and an updated AMF calculated 440 using code written by Paul Palmer and Randal Martin, with modifications by Luke Surl (see Section ??). In Chapter ?? there are three calculations of the AMF referred to as  $AMF_{OMI}$  (the original OMI AMF),  $AMF_{GC}$  (AMF recalculated with GEOS-Chem shape factors), and  $AMF_{PP}$  (AMF recalculated using GEOS-Chem shape factors and scattering weights). The new emissions are calculated (using the  $AMF_{PP}$ ) are referred to as the a posteriori from here onward, or  $E_{OMI}$  in formulae.

A posteriori emissions are calculated using the linear relationship described in Section 2.5 using the modelled slope  $S$  445 calculated in the prior section and satellite HCHO columns recalculated in ??:

$$\begin{aligned}\Omega_{OMI} &= S \times E_{OMI} + \Omega_0 \\ E_{OMI} &= \frac{\Omega_{OMI} - \Omega_0}{S}\end{aligned}\quad (11)$$

This is the same as equation 10, except now the satellite HCHO ( $\Omega_{OMI}$  and  $\Omega_0$ ) is used.  $\Omega_0$  is calculated using  $\Omega_{OMI}$  in the remote Pacific averaged monthly and longitudinally, for each latitude. This leaves  $E_{OMI}$  as the only unknown once the satellite measurements are processed to match the temporal and horizontal resolution of  $S$ .

450 One potential issue in this top-down estimation technique is the low number of valid satellite measurements that may occur due to the higher solar zenith angles in winter and at higher latitudes. Another issue is that negative emissions are commonly calculated in areas wherever measured HCHO columns are lower than background amounts (as  $E_{OMI} = \frac{\Omega_{OMI} - \Omega_0}{S}$ ) due to uncertainty in the background calculation. These negative calculated emissions are set to zero, as negative emissions physically meaningless in this calculation. Relative uncertainty in these grid squares is set to 100% for later estimations of uncertainty 455 (Section 4.1).



**Figure 15.** Top left: Background total column HCHO based on longitudinally averaged values over the remote Pacific Ocean. Top right: Background total column HCHO based on GEOS-Chem output with isoprene emissions set to zero. Bottom left: Distributions of column amounts over land squares in Australia from the two background definitions and the monthly HCHO average from GEOS-Chem (standard tropchem run). Bottom right: Absolute difference between the two background defined value maps.

## 2.9 Running GEOS-Chem using a posteriori emissions

After creating the a posteriori isoprene emissions estimate (at monthly resolution), GEOS-Chem is re-run with biogenic emissions scaled to match the new estimate. This is performed by applying a seasonal scaling factor  $\alpha$ , based on the multi-year monthly average difference between midday a priori and a posteriori emissions at  $2^\circ \times 2.5^\circ$  horizontal resolution.  $\alpha$  is the ratio  
460 between the multi-year averaged monthly emissions from GEOS-Chem  $E_{GC}$  and the a posteriori  $E_{OMI}$ :

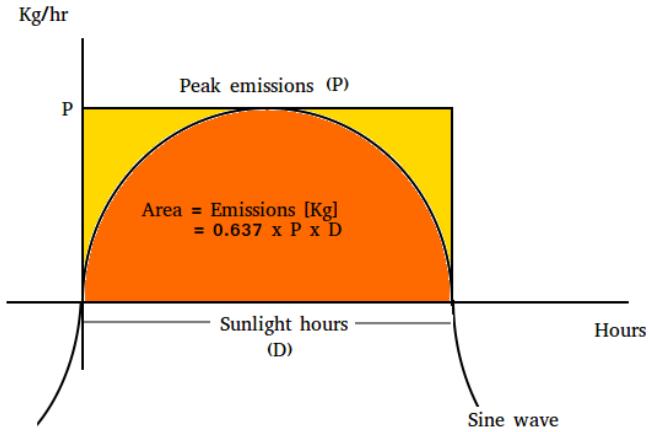
$$\alpha = \frac{E_{OMI}}{E_{GC}} \quad (12)$$

This seasonal scaling retains shorter time-scale variability and meteorological dependencies built into the parameterisations of the MEGAN model, while ensuring the multi-year monthly averaged emissions match the a posteriori totals. This method provides a preliminary estimate of the effects of scaling isoprene emissions towards the satellite based estimate, and running  
465 this analysis at a finer temporal resolution could be performed in future work. Initially  $\alpha$  is uniformly set to 1 globally. Where top-down emissions exist and  $E_{GC}$  is non-zero,  $\alpha$  is set using Equation 12.  $\alpha$  is applied through the emissions module in GEOS-Chem where isoprene emissions are calculated. First, the new midday (13:00-14:00 LT) emissions (per grid box) are combined forming a multi-year monthly mean, which can be compared to the a priori equivalent. Missing values for  $\alpha$  when  
470  $E_{GC}$  are zero are a negligible issue since the dominant discrepancies between estimates occur during summer when high emission rates are overestimated.

To create the  $\alpha$  scaling factor, modelled and top-down emissions are compared, which requires a change of units. Top-down emissions calculated in this work are in units of atom C  $\text{cm}^{-2} \text{ s}^{-1}$ , while modelled emissions are in  $\text{kg cm}^{-2} \text{ s}^{-1}$ . In order to calculate the top-down emissions in kg, each grid square is multiplied by its area, and then daily emissions are assumed to follow a sine wave peaking at the midday value. The sine wave approximates a daily insolation amplitude, allowing one to  
475 convert peak emission rates into daily emission amounts through integration. Figure 16 shows how the daily approximation of total emitted isoprene per grid square is calculated. Daytime hours are estimated per month, from 14 hrs (Jan) to 10 hr (Jul) (<https://en.wikipedia.org/wiki/Daytime>). This approximation is required since OMI observations occur at midday, when isoprene emissions are at their diurnal peak. GEOS-Chem emissions are similarly multiplied by area, but then integrated over time using hourly output to derive emissions in kg.

## 480 3 Results

Australia covers roughly  $7.7 \times 10^6 \text{ km}^2$ , with heterogeneous environmental conditions. The results presented in this section are therefore frequently split into five regions that are differentiated by colour, as shown in Figure 17. These regions are large enough to reduce the uncertainty with at least 10 grid squares in each area, and small enough to be somewhat homogeneous at the continental scale. Oceanic grid squares are filtered out of regional averages that follow.

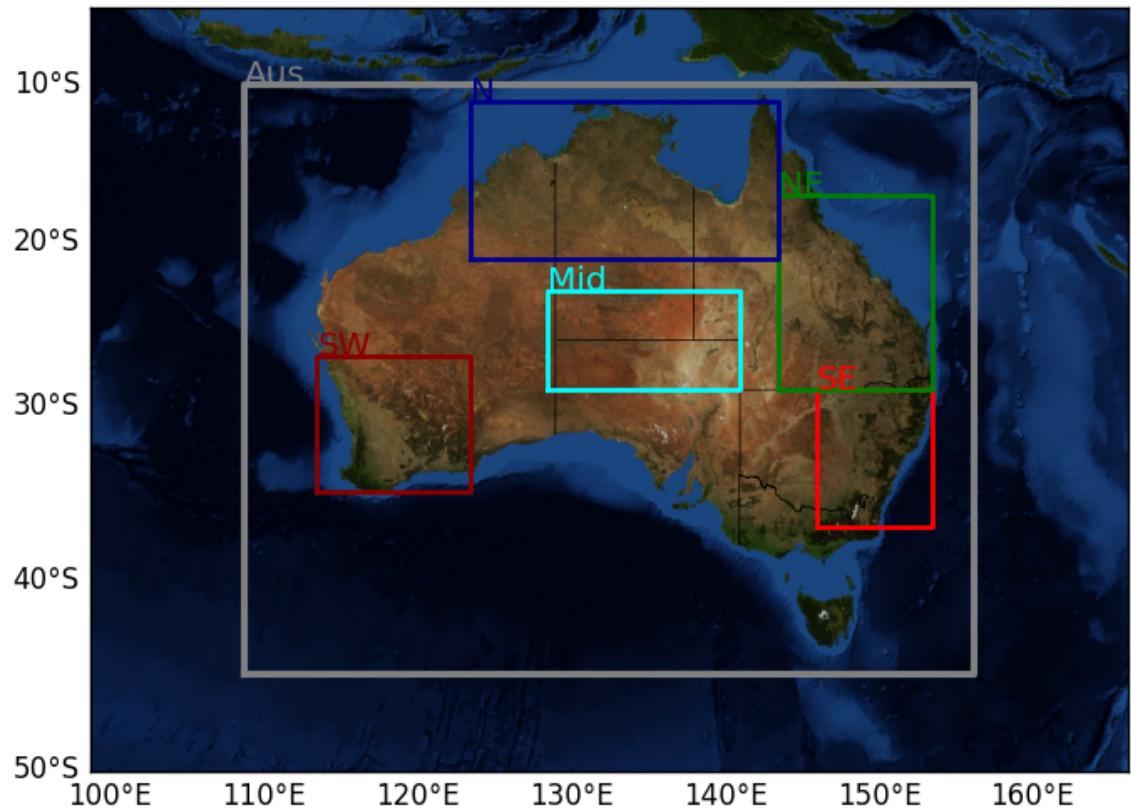


**Figure 16.** Total daily isoprene emissions (in kg) is represented by the area under the sine wave.

#### 485 3.1 A posteriori emissions

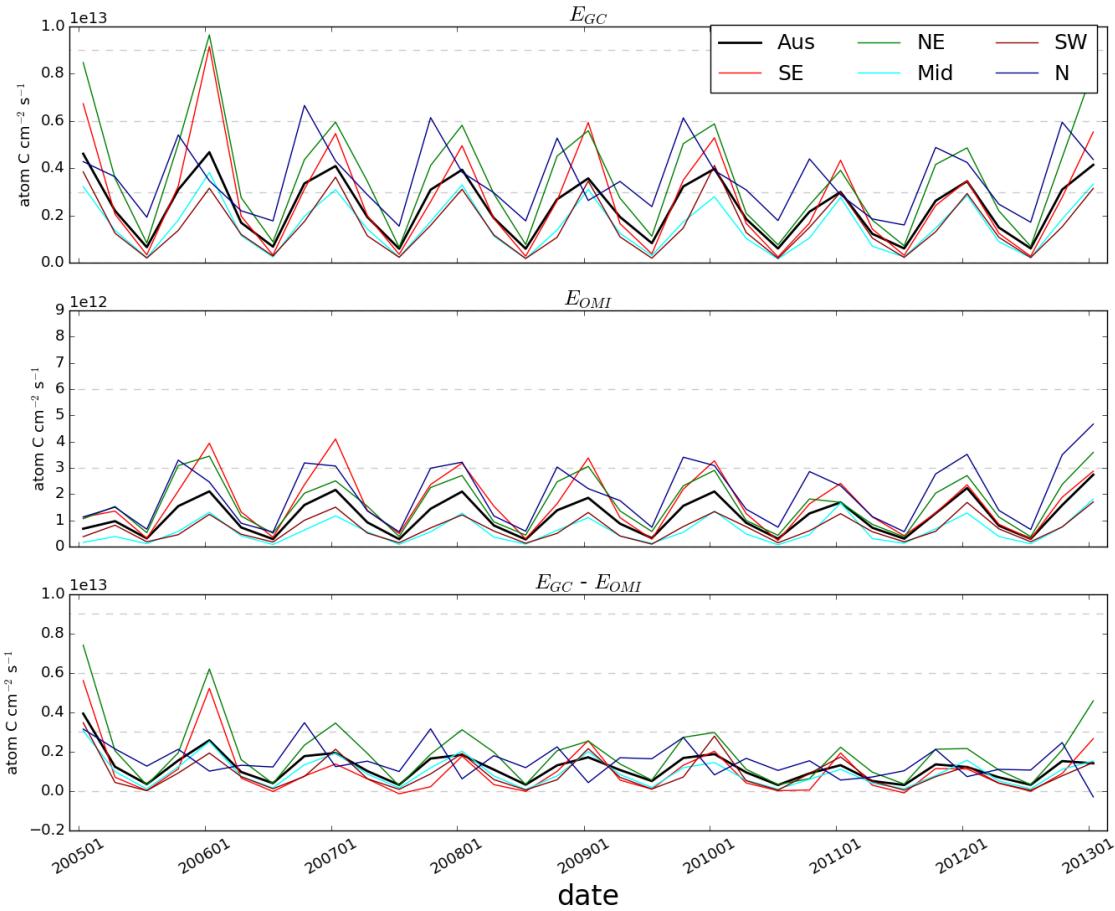
Figure 18 shows a priori emissions over Australia along with a posteriori emissions calculated as described in the prior sections. This figure shows the time series of seasonal area averaged midday emissions, and their absolute differences. Differences between a priori and a posteriori estimates are seasonally and spatially diverse. A seasonal overestimate peaking in summer can be seen in all regions except for the northern region, which is generally overestimated but less so in summer. Figure 490 19 shows how the yearly decrease is distributed spatially over Australia, with  $E_{GC}$  and  $E_{OMI}$  in  $\text{Tg yr}^{-1}$  calculated as a multi-year mean. Across all of Australia large reductions of total emissions are seen using the new top-down estimate.

Figure 20 shows the multi-year seasonal emissions for each region for the a priori and a posteriori emissions side by side. The a priori is approximately twice that of the a posteriori. Absolute differences are highest in spring and summer, when emissions are generally greatest. The exception is the northern region, where the largest overestimates occur in spring, and 495 large differences are seen in all seasons. Uncertainty is summarised for the a posteriori using the mean monthly uncertainty per grid square (calculated as described in 4.1) within each region, divided by the square root of 24 (8 years times 3 months per season). The uncertainty shown for each region is the average over the grid squares within that region, without reduction by the square root of the number of grid squares. Uncertainty is relatively small since 8 years of data are combined into the seasonal mean. The potential bias (see Section 4.1) is added to the uncertainty and displayed using horizontal bars. The highest 500 over-estimates (occurring in spring in the Northern region, and summer in all others) lie outside the potential bias caused by satellite underestimates of HCHO (Section 4.1). However, many overestimated seasons within each region are within this wide range of potential bias and uncertainty. Some monthly grid squares are very uncertain ( $> 200\%$ ) and these are removed for this plot, causing an increase in a posteriori emissions of  $< 1\%$ . There are very few of these grid squares, and they appear to be grid

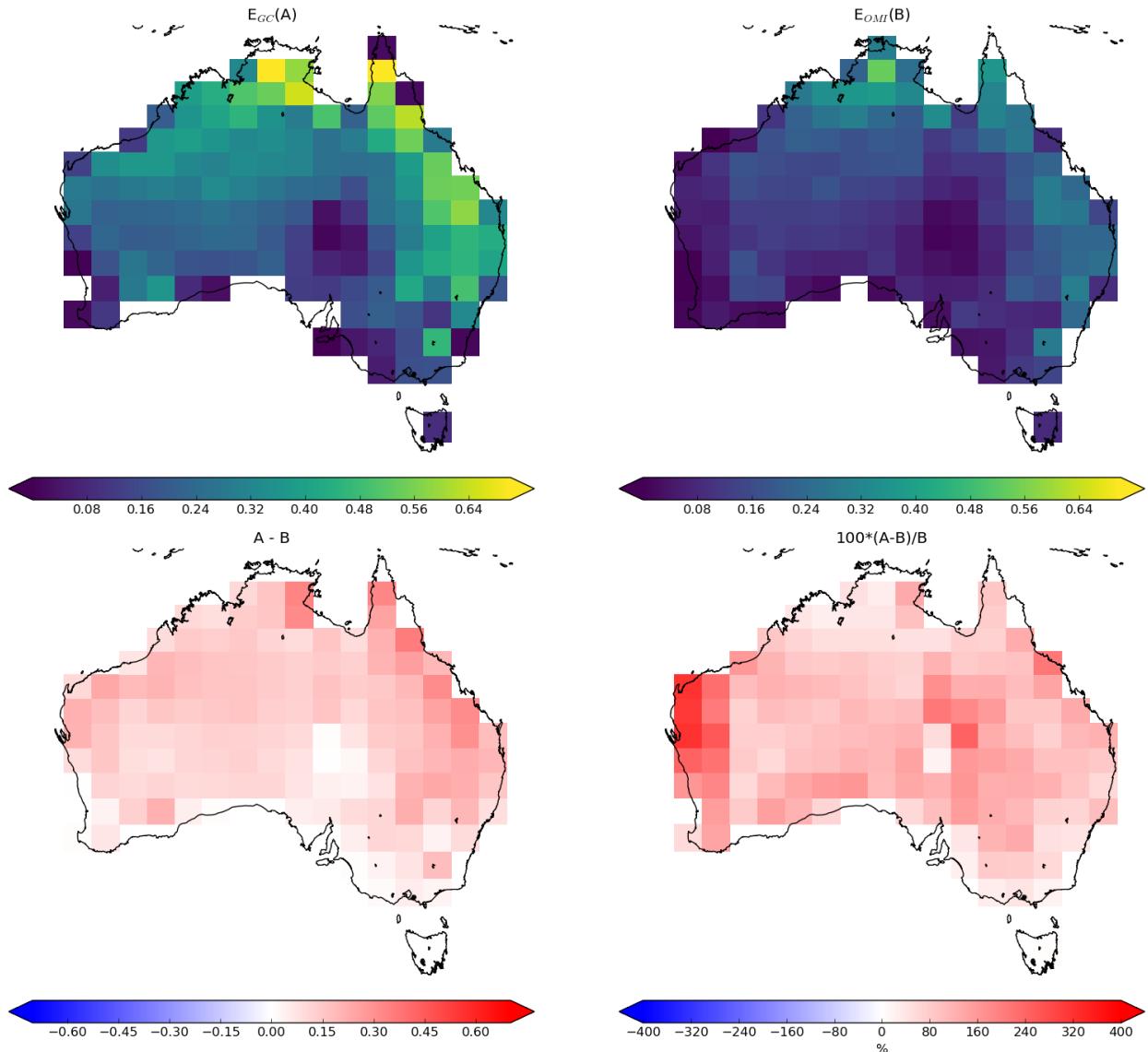


**Figure 17.** Sub-regions used in subsequent figures: Northern, North Eastern, South Eastern, South Western, and Middle. Australia-wide averages will be black or grey, while averages from within the coloured rectangles will match the colour shown here. Oceanic grid squares are filtered out from any area averages made in this chapter.

## Midday emissions

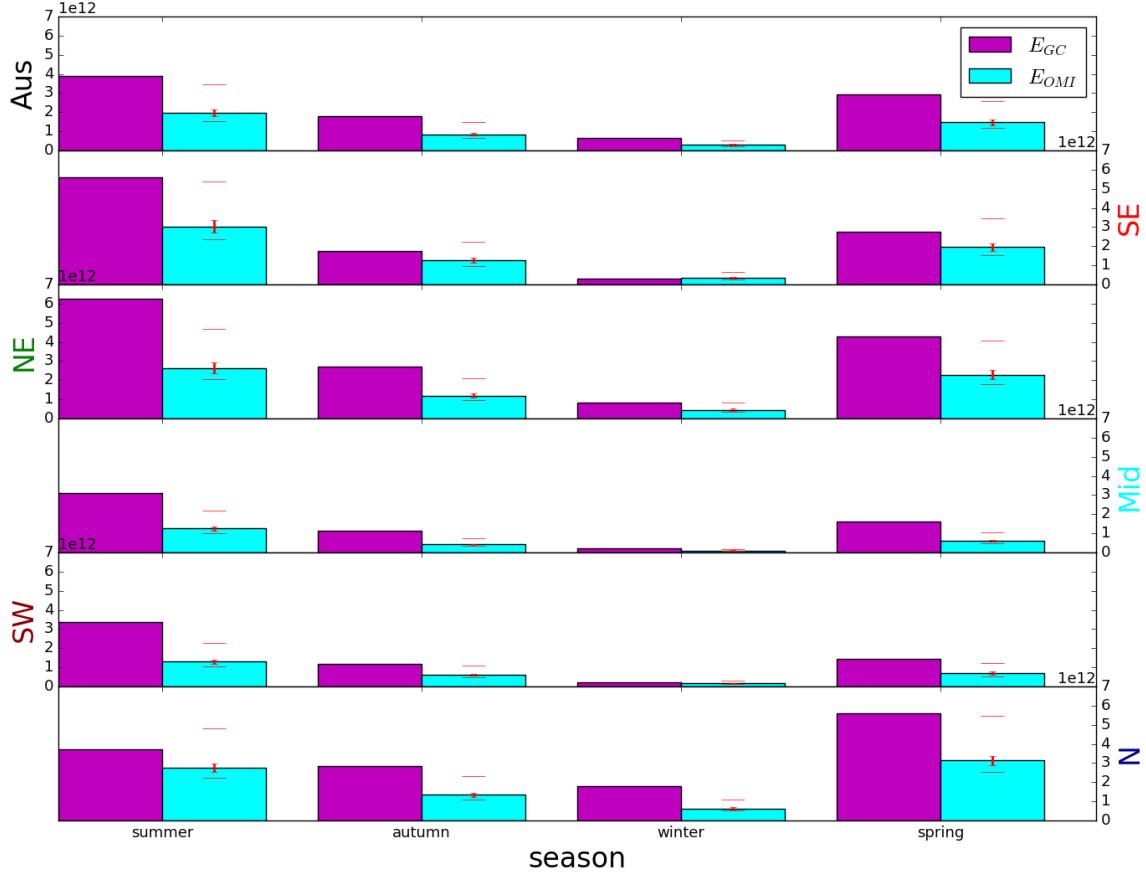


**Figure 18.** Row 1: Biogenic emissions of isoprene from GEOS-Chem (a priori,  $E_{GC}$ ). Row 2: Emissions calculated using the OMI top-down inversion (a posteriori,  $E_{OMI}$ ). Row 3: Absolute differences between the first two rows. Midday emissions are averaged for each season (DJF, MAM, JJA, SON), and colours represent averaged areas from subregions shown in Figure 17. Grey dashed horizontal bars are added to highlight the scale differences between rows.



**Figure 19.** Top row: multi-year mean a priori emissions in  $\text{Tg yr}^{-1}$  from  $E_{GC}$  (GEOS-Chem; running MEGAN) and from  $E_{OMI}$  (top-down emissions) respectively. The total flux from  $E_{OMI}$  is calculated assuming a sinusoidal daily cycle, with daylight hours prescribed for each month as described in Section 2.9). Bottom left and right show the absolute and relative differences, respectively.

## Midday emissions [ $\text{atom C cm}^{-2} \text{ s}^{-1}$ ]



**Figure 20.** Regional multi-year seasonal mean a priori emissions (magenta) compared to a posteriori emissions (cyan). Error bars show the regionally averaged uncertainty (calculated as described in 4.1). Additional horizontal dashes show the uncertainty plus effects from potential HCHO biases (discussed in Section 4.1) from satellite underestimation (40%) and monthly clear sky overestimation (13%).

squares with very low emission rates and low pixel counts. The cause for these highly uncertain grid squares is not analysed further, but may be worth examining in future work.

Figure 21 shows the multi-year monthly mean and inter-quartile range of midday isoprene emissions estimates in each region. Months from September to April show the most difference between a posteriori and a priori. The most overlap is seen in the south-eastern region, where high summer emissions along with high variance occur in both the a priori and a posteriori. The highest variance is seen in both eastern regions, potentially due to diversity within the regions which include high density cities, large forests, and rural areas. This overestimate by the a priori may be caused by some mixture of overly high emission factors and high emission sensitivity to temperatures and soil moisture (Emmerson et al., 2016, 2019). Over the entirety of

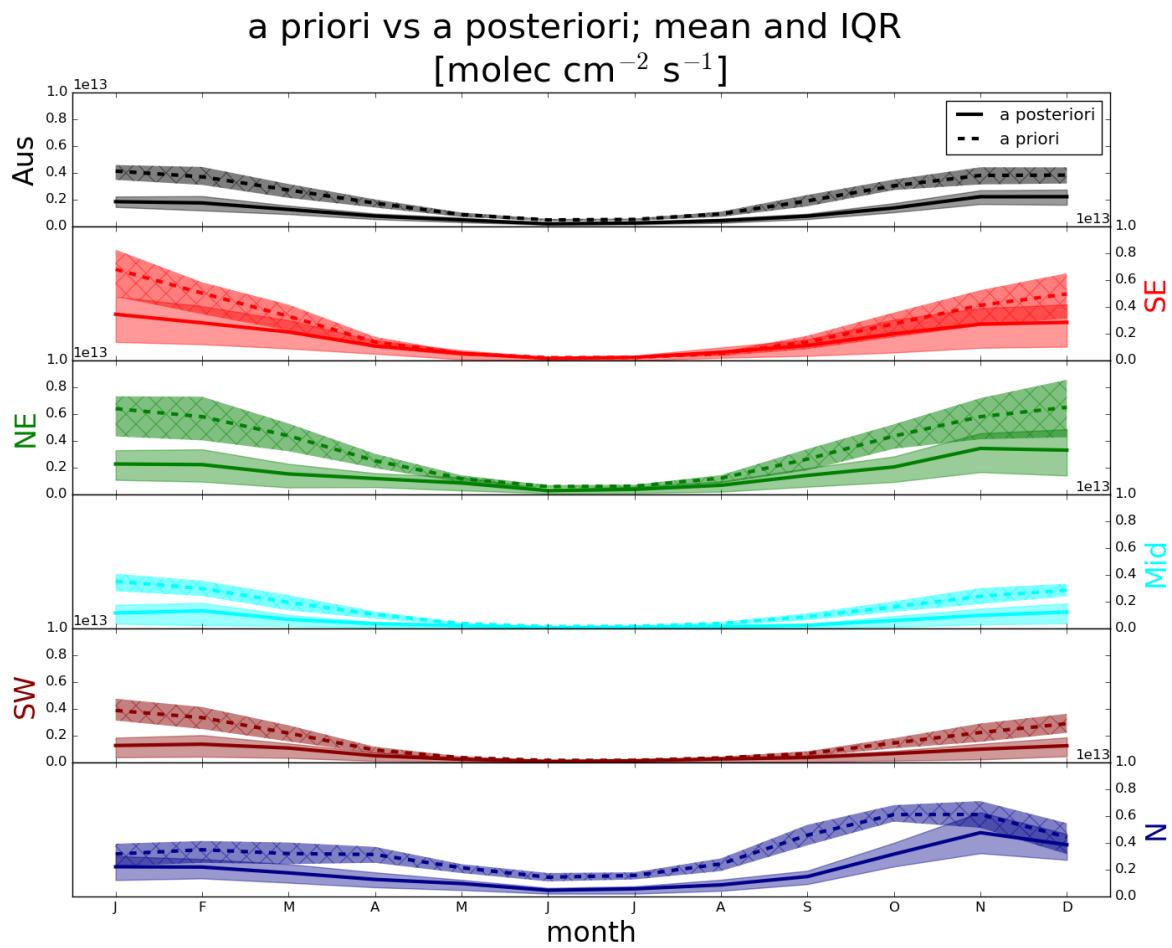
Australia the seasonal cycle of emissions is shown to be overestimated by the a priori, and further analysis and measurements are required to determine the cause.

While most regions show similar overestimates, the northern region of Australia follows a different cycle of bias. Northern  
515 Australian emissions appear to be overestimated throughout the year, with the smallest bias in early summer. One potential reason is that during the wet season (November-April) there are fewer satellite measurements due to increased cloud coverage, along with a different ecosystem response to sunlight and temperature (e.g., Surl et al., 2018). This is evidenced by the approximately 20% lower summer OMI pixel count (before filtering) in both northern regions (see Section 4.3). Low measurement counts in summer in the northern region could lead to a low bias in the a posteriori emissions estimate from the drier regions  
520 associated with lower biogenic emissions being over-represented. In situ measurements in both monsoon and non-monsoon seasons are required before robust conclusions can be drawn in this region.

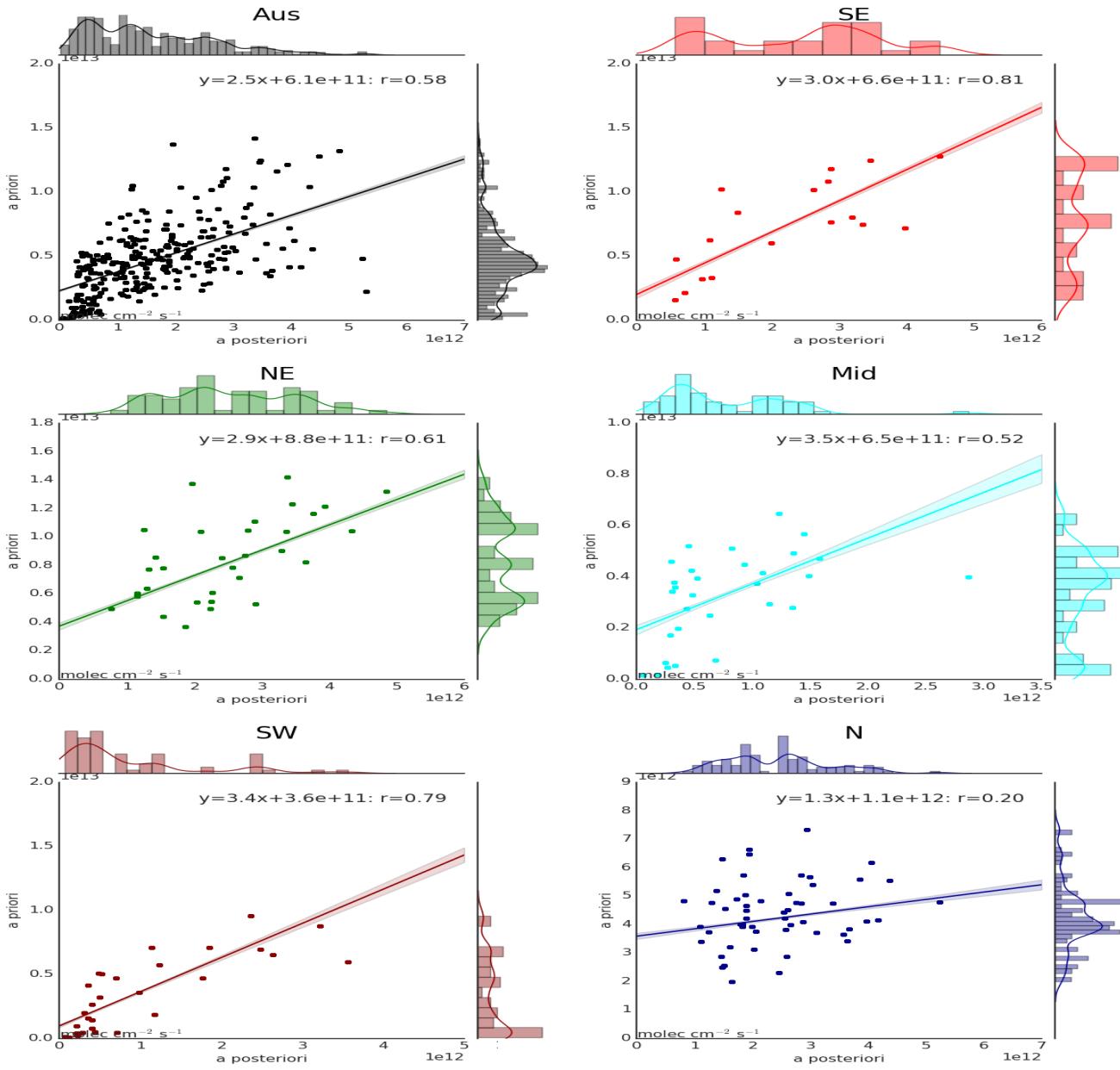
There is only weak correlation between daily estimates of the a priori and a posteriori ( $r < 0.2$ ), and a priori distributions show more variance (figure in Appendix ??). This is likely due to the filtering applied to satellite data (e.g., whenever cloud coverage exceeds 40%) which reduces the count and spread of a posteriori emission calculations. Figure 22 shows how the  
525 distributions of a posteriori emissions compare to a priori emissions in each region during summer months (December - February) with zeros removed from both distributions. This figure also shows the regressions between monthly averages of the same data. In the summer monthly averages, the linear regression coefficient  $r$  ranges from 0.2 to 0.81 depending on which region is being compared. The highest correlations between a priori and a posteriori emissions are in the southeast ( $r = 0.81$ ) and southwest ( $r = 0.79$ ) regions, followed by the northeast ( $r = 0.61$ ) region. This is likely due to isoprene emissions in these  
530 regions being dominated by the biogenic sources (large forests) that the top-down emission estimate is based upon. Although the northern region also contains large areas of forest, the correlation is the worst. In part this is due to misrepresented forest emissions in the model, exacerbated by unrepresented responses to moisture and poorly modelled environmental stresses in this region which lies within the tropics and undergoes monsoonal weather and intense heat and drought seasons (e.g., Emmerson et al., 2016; Surl et al., 2018; Emmerson et al., 2019). If the same analysis is performed over winter months, the regression  
535 in the northern region is greatly improved (to  $r = 0.66$ ), while other regions remain approximately the same. If the modelled emissions distributions are correct (even if the absolute magnitudes are not), this suggests that satellite measurements are not capturing representative samples of monthly grid square averages in the north in summer, likely because of cloud coverage, and the summer inversion in this region may be biased.

### 3.1.1 Comparison to other estimates

540 Table 5 provides a brief comparison of isoprene emissions estimated in this chapter of  $39.2 \text{ Tg yr}^{-1}$  (bottom-up: from GEOS-Chem running MEGAN) and  $21 \text{ Tg yr}^{-1}$  (top-down) with recent Australian isoprene emission estimates performed elsewhere. Bauwens et al. (2016) performed a global inversion of OMI HCHO to produce a top-down isoprene emission estimate, and also examined three separate bottom-up inventories. Their MEGAN-MOHYCAN model run is most similar to the bottom-up GEOS-Chem emissions seen in this work, and shows similar isoprene emissions of  $38 \text{ Tg yr}^{-1}$ . MEGAN-MOHYCAN  
545 formed the a priori for their inversion, with the two other independent bottom-up models used for comparison. The highest



**Figure 21.** The multi-year monthly mean (lines) and inter-quartile-range (shaded) of midday (13:00–14:00 LT) isoprene emissions estimates. A priori emissions are shown by the dashed lines and hatched shaded areas show the inter-quartile-range. A posteriori emissions are shown using the solid lines, with inter-quartile-range shown by unshaded shaded areas. Colours denote the region over which the multi-year monthly average was taken, as in Figure 17.



**Figure 22.** Scatter plot of a priori emissions against a posteriori emissions using monthly averaged grid squares as regression data points. Data points are created using monthly averages (of midday emissions) for each grid box for each month of summer (DJF) within each region shown. Multiple years of data are used, meaning if a region has 10 grid boxes, the 8 years of data will add up to  $10 \text{ boxes} \times 3 \text{ months} \times 8 \text{ years} = 240$  data points minus filtered and zero emission squares. Plots are coloured by regions shown in Figure 17. The linear best fit regression is inset into each plot along with the line equation and regression coefficient. The normalised distribution of each population is shown at the top and right spine of each subplot, with the right spine (facing the a priori axis) using the a priori axis and scale, and the top spine using the a posteriori axis and scale.

emissions they saw, at  $94 \text{ Tg yr}^{-1}$ , came from the MEGAN-MACC model. This high estimate is attributed to neglect of the moisture stress effects and a different crop distribution database compared to MEGAN-MOHYCAN. Sindelarova et al. (2014) also examine bottom-up BVOC emissions using MEGAN-MACC. They examined 30 years of MEGAN-MACC emissions globally to determine sensitivity of the model to various parameters. Their estimate over Australia was similar to the MEGAN-  
550 MACC estimate seen in Bauwens et al. (2016), at  $80 \text{ Tg yr}^{-1}$ . The other bottom-up model Bauwens et al. (2016) compare with (GUESS-ES) uses a different emissions model, based on Niinemets et al. (1999), which estimated isoprene emissions of  $26 \text{ Tg yr}^{-1}$ . This was lower than both their a priori (MEGAN-MOHYCAN) and their a posteriori.

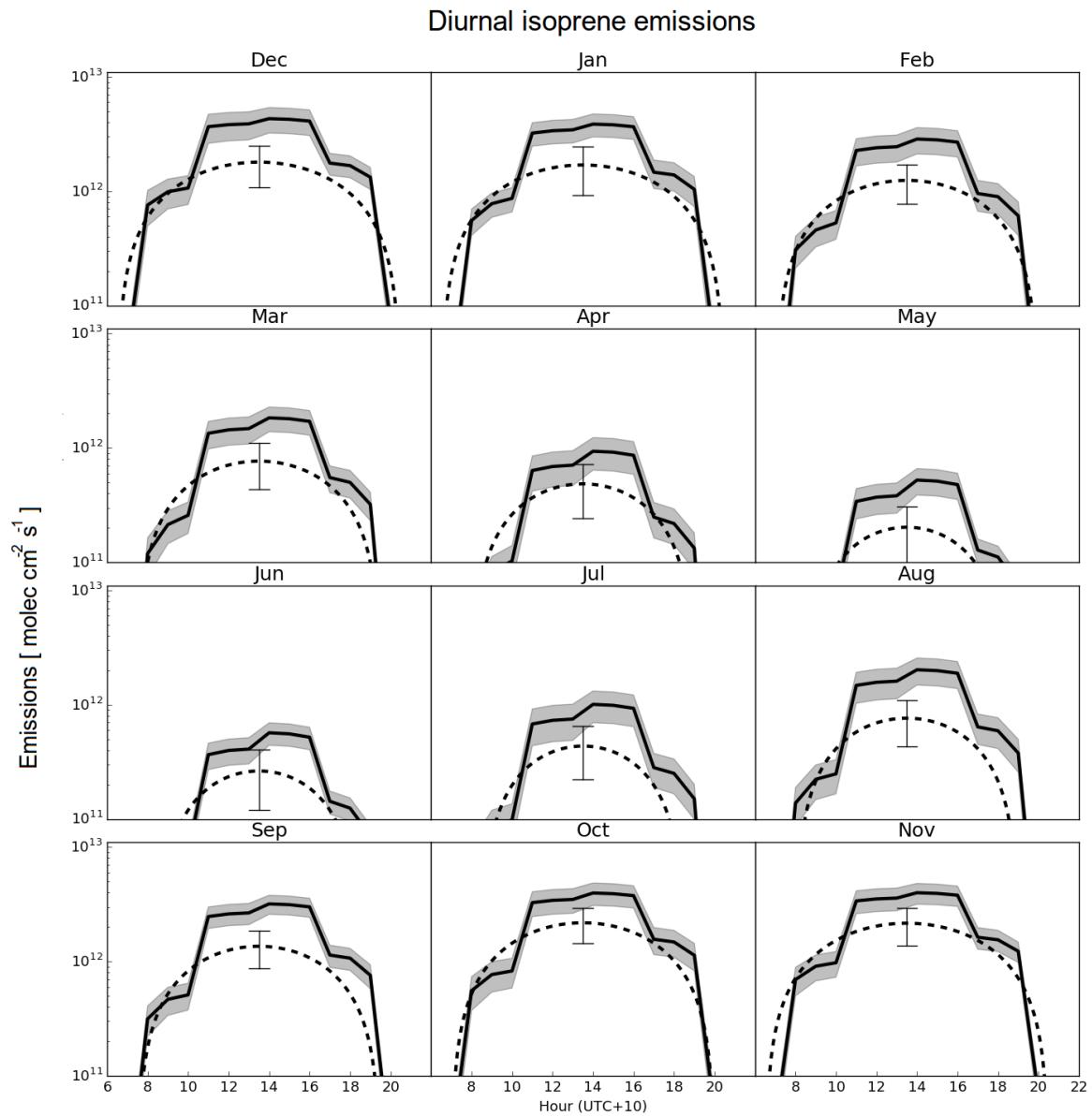
Table 5 shows total a posteriori emissions summed over Australia are lower than the a priori by  $\sim 46\%$ , reducing from  $39 \text{ Tg yr}^{-1}$  to  $21 \text{ Tg yr}^{-1}$ . The a posteriori estimated here to be  $21 \text{ Tg yr}^{-1}$  suggests isoprene emissions may be lower than  
555 all prior bottom up estimates. This is close to the lowest bottom-up estimate of Bauwens et al. (2016) of  $26 \text{ Tg yr}^{-1}$ . There are several major differences between work performed in this thesis and work done by Bauwens et al. (2016). They used a fully adjoint version of the IMAGESv2 CTM with emissions determined by MEGAN-MOHYCAN, and iteratively updated biogenic isoprene emissions and pyrogenic VOC emissions so that the CTM produced HCHO columns that matched (as near as possible) the column HCHO observed from OMI. Additionally they handle fire differently, attributing some of the HCHO to  
560 pyrogenic precursors rather than filtering out detectable fire influences as is performed here. The differences in results between Bauwens et al. (2016) and work performed here are most likely to be primarily driven by differences in modelled chemistry, rather than differences in top-down methodology. For instance, any model differences in  $\text{NO}_X$  emissions or chemistry will greatly affect local HCHO yield and change the link between satellite HCHO and emitted isoprene (e.g., Kaiser et al., 2018). Both techniques have the end result of reducing the differences between (respective) modelled and OMI total column HCHO,  
565 although the reduction seen here (46%) is much greater than that seen in Bauwens et al. (2016). The fact that both top-down techniques suggest a priori bottom-up emissions are too high is a good indicator that these inventories are indeed overestimating isoprene emissions.

### 3.1.2 Diurnal variability

Figure 23 shows the a priori daily emissions cycle for Australia compared to the estimated a posteriori emissions cycle over  
570 Australia. The conversion of midday a posteriori emissions ( $\text{ molec cm}^{-2} \text{ s}^{-1}$ ) into  $\text{Tg yr}^{-1}$  involves integration over an assumed sinusoidal diurnal emission cycle as described in Section 2.9. A priori emissions peak from approximately 11:00 LT to 16:00 LT, while outside these hours there is a non-sinusoidal drop in emissions to below the assumed a posteriori diurnal emission cycle. This means the conversion may be biased by this consistent difference between modelled a priori diurnal emission cycles and the assumed diurnal a posteriori cycle. This potential bias is not analysed further, and should be relatively  
575 small compared to other uncertainties as it only affects emissions towards the daily minima.

### 3.1.3 Trends

Figure 24 shows monthly deseasonalised a priori and a posteriori midday emission anomalies for each region. First the emissions are spatially averaged within each region to form a daily time series of midday emission rates. These are averaged into



**Figure 23.** The diurnal cycle of GEOS-Chem a priori emissions (solid line) averaged by month into hourly bins over from 2005 to 2013 along with top-down a posteriori (dashed line) emissions. Standard deviations for the monthly average are shaded for the a priori, and shown with error bars at 13:30 LT for the a posteriori. Top-down emissions shown here are based on monthly midday emissions being the peak of a sine wave which drops to zero after and before daylight hours (see Section 2.9).

**Table 5.** Isoprene emissions (Tg/yr) from Australia

Estimate <sup>a</sup>	Source	Year	Reference
39.2(4.0)	bottom-up	2005-2012	This thesis <sup>b</sup>
~ 80	bottom-up	1980-2010	Sindelarova et al. (2014) <sup>c</sup>
26-94	bottom-up	2005-2013	Bauwens et al. (2016) <sup>d</sup>
20.7(1.6)	top-down	2005-2012	This thesis <sup>e</sup>
36	top-down	2005-2013	Bauwens et al. (2016) <sup>f</sup>

a: Interannual standard deviation shown in parentheses.

b: GEOS-Chem with MEGAN diagnostics based on 3-hourly averages.

c: MEGAN run using MERRA meteorology.

d: Range shown here based on three models, two of which implement MEGAN.

e: Based on daily peak emissions integrated over a sinusoidal daily curve.

f: OMI based top-down inversion.

monthly data, and then the multi-year monthly mean is subtracted to form the anomaly time series. Any anomaly greater than  
 580 three standard deviations from the mean is removed (crosses in Figure 24). An ordinary least squares linear regression is then performed to look for any significant trend. A trend is considered significant if the p-value from a Wald test (equivalent to a t-test) is less than 0.05. Trend results for isoprene are summarised in Table 6. The a priori midday surface isoprene concentrations show a small decline outside of the southwest region over the 8 year period from 2005-2012. This decline reduces in scale by approximately a factor of 2-3 in the a posteriori emissions. There is no significant change in trend due to scaling isoprene  
 585 emissions. Since the scaling factor  $\alpha$  is applied to each grid square seasonally and not changing each year, changing trends are not expected.

### 3.2 Modelled impacts of reduced isoprene emissions

This section uses GEOS-Chem to determine how the improvements to biogenic isoprene emissions impact subsequent atmospheric chemistry and composition. A posteriori emissions are implemented in GEOS-Chem as described in Section 2.9.  
 590 Outputs from the scaled GEOS-Chem run using the a posteriori are denoted by superscript  $\alpha$ . For example, column HCHO from GEOS-Chem before and after scaling are denoted  $\Omega_{GC}$  and  $\Omega_{GC}^\alpha$  respectively. Figure 25 shows  $\alpha$  for January and June averaged over 2005-2012, along with the time series of  $E_{GC}$  and  $E_{OMI}$  and  $\alpha$  calculated for the  $2^\circ \times 2.5^\circ$  grid box containing Sydney and their multi-year seasonal average. GEOS-Chem emissions are scaled according to this multi-year averaged gridded  $\alpha$ , which is seasonal and spatially diverse.

**Table 6.** 2005-2012 trend<sup>a</sup> in surface isoprene mixing ratio (ppbvC yr<sup>-1</sup>) from simulations using a priori and a posteriori isoprene emissions.

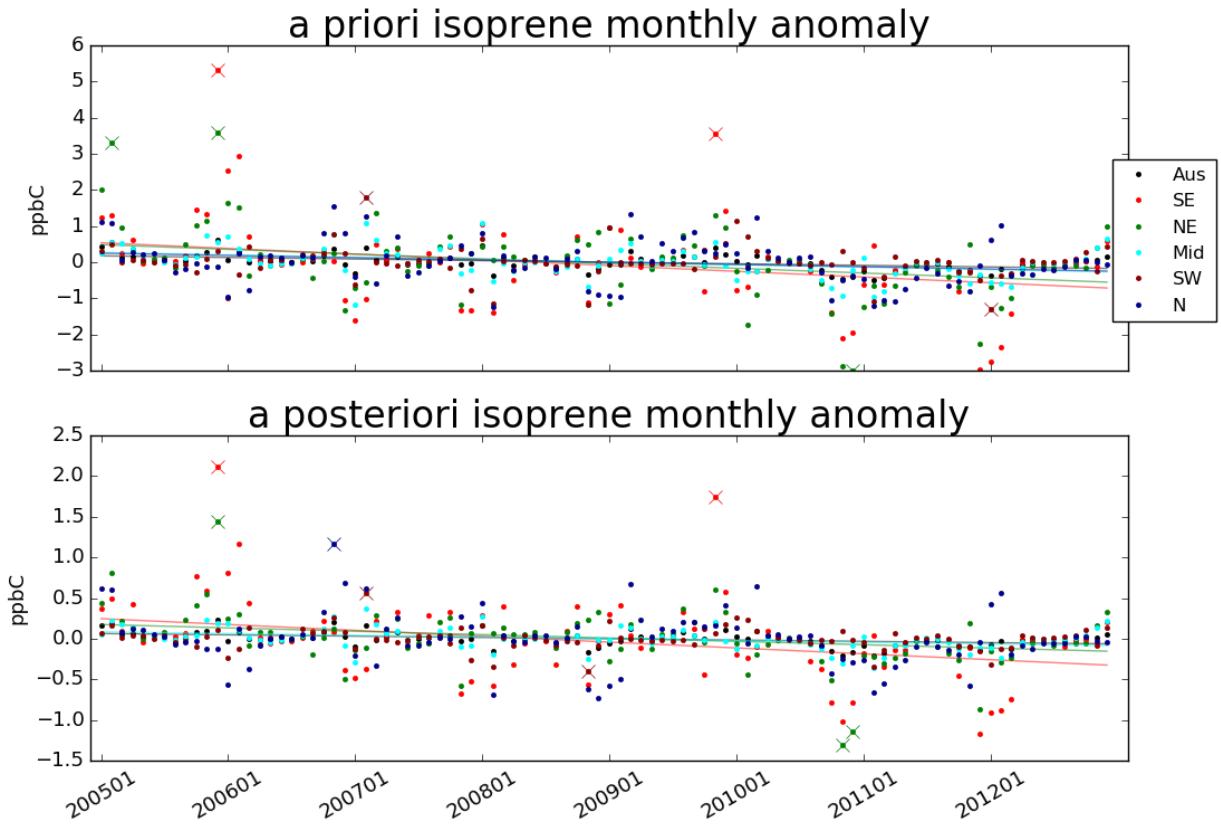
Region	a priori	a posteriori
Aus	<b><span style="font-weight:bold;">-.04</span></b>	<b><span style="font-weight:bold;">-.02</span></b>
SE	<b><span style="color:red; font-weight:bold;">-.15</span></b>	<b><span style="color:red; font-weight:bold;">-.07</span></b>
NE	<b><span style="color:green; font-weight:bold;">-.13</span></b>	<b><span style="color:green; font-weight:bold;">-.04</span></b>
Mid	<b><span style="color:blue; font-weight:bold;">-.07</span></b>	<b><span style="color:blue; font-weight:bold;">-.02</span></b>
SW	-.01	-.01
N	<b><span style="color:blue; font-weight:bold;">-.06</span></b>	-.02

a: Statistically significant (two sided test with  $\alpha = 0.1$ ) trends are bolded.

### 595 3.2.1 Implications for HCHO

As a preliminary check on the GEOS-Chem output, simulated  $\Omega_{GC}$  and  $\Omega_{GC}^\alpha$  are compared to  $\Omega_{OMI}$  over January and February, 2005 in Figure 26. In every region,  $\Omega_{GC}^\alpha$  is closer to  $\Omega_{OMI}$  with biases decreasing from  $\sim 100\%$  to  $\sim 50\%$  everywhere except the northern region, which has biases decreasing from  $\sim 50\%$  to  $\sim 25\%$ . Note that this is not an independent validation as  $\Omega_{OMI}$  drive the creation of  $\Omega_{GC}^\alpha$ , and as is expected the relationship is much improved. The remaining differences are 600 most likely driven by filtering and temporal averaging of the applied scaling factor  $\alpha$ . When looking over all of Australia from 2005-2012, the summer mean decreases from  $9.8 \times 10^{15}$  to  $7.4 \times 10^{15}$  molec cm<sup>-2</sup> after scaling isoprene emissions, while the satellite column is  $4.9 \times 10^{15}$ .

Figure 27 shows HCHO is most reduced in summer for the majority of Australia. In winter, a reduction is seen along the northern coast of Australia, most likely because emissions in this area are less affected by the seasonal decline. Seasonal means 605 and the standard deviations (between all grid boxes and days used in averaging) for each region are summarised in Figure 28. Both mean and variance are reduced after running GEOS-Chem with scaled isoprene, although the mean is lower still in OMI vertical columns. Model output standard deviations in summer range from  $\sim 20 - 30\%$ , while OMI standard deviations range from  $\sim 32 - 41\%$ . The highest OMI standard deviations occur in winter, ranging over  $\sim 46 - 73\%$ , which is the opposite of model output with standard deviations between  $\sim 12 - 27\%$ . This could be an effect of the increased winter uncertainty in 610 satellite output, which is amplified by the low column amounts in the season. Overall the standard deviation within model output appears to be too low in all regions, ranging from three quarters (south eastern region) to one fifth (middle region) of the standard deviations within the OMI HCHO column output. These standard deviations are formed from the monthly averaged

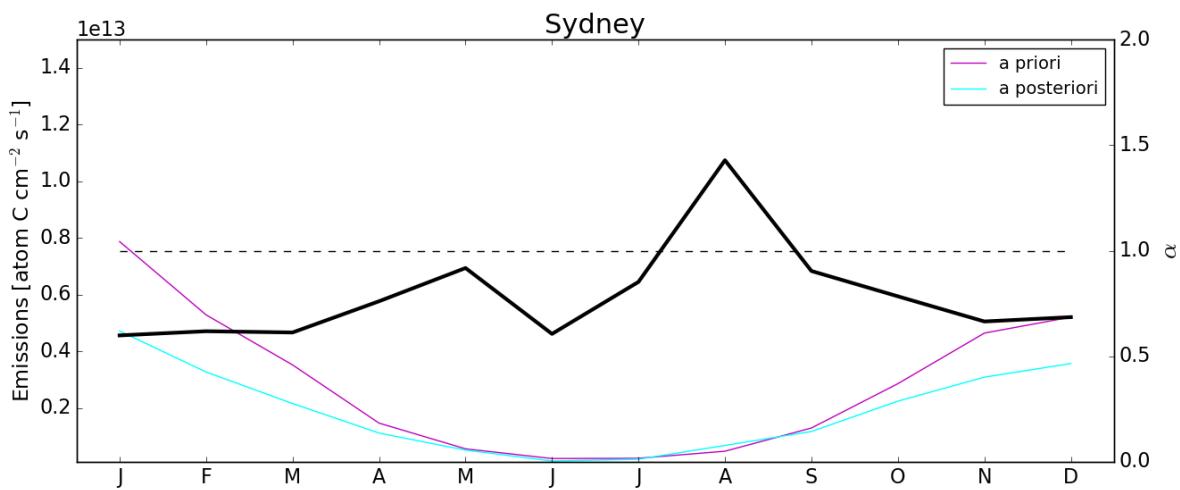
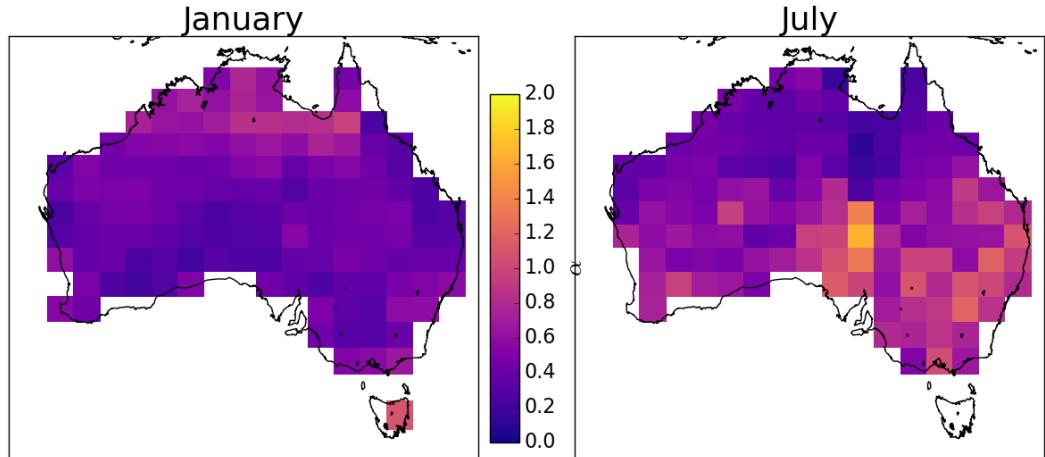


**Figure 24.** A priori (row 1) and a posteriori (row 2) monthly emissions anomalies from multi-year monthly mean, split by region (see Figure 17).

grid boxes used to create each multi-year seasonal mean. Their low value (relative to satellite measurements) suggests that the model does not capture the full range of variance over time and space of HCHO levels over Australia; however, variance 615 is heightened in satellite measurements due to the filtering that reduces the number of grid squares that form the multi-year seasonal mean.

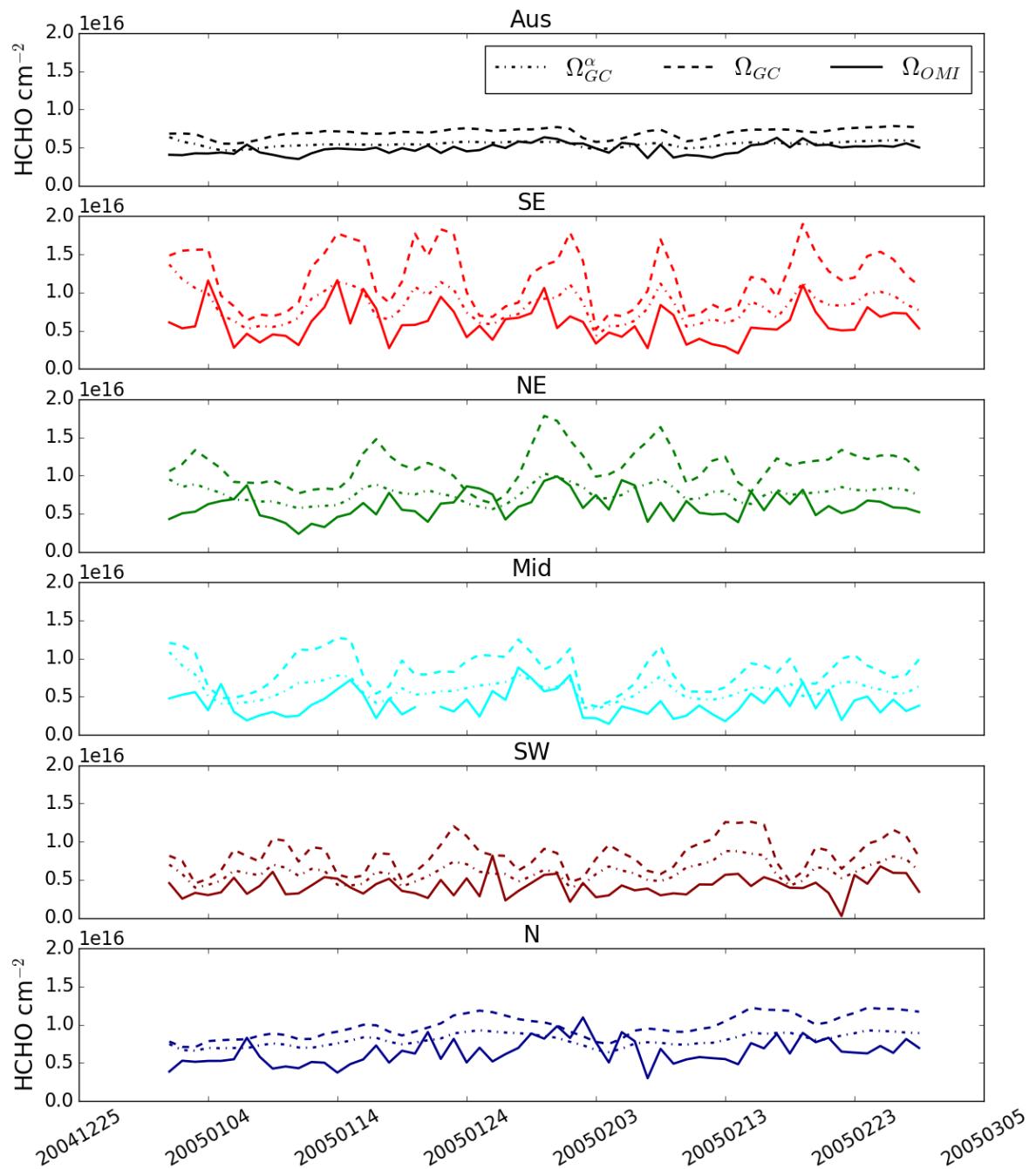
Decreasing isoprene emissions in the model lead to reduction in HCHO concentrations as one would expect. Figure 29 shows an example (over Wollongong) of the modelled HCHO profile before and after scaling isoprene emissions. This reduction of HCHO is greatest at the surface, and is highly correlated ( $r > .8$ ) in all regions to the reduction in isoprene emissions. The high 620 correlation is effectively a sanity check of the top-down method used in this thesis to estimate isoprene emissions.

$\alpha$  multi-year monthly average 2005-2012



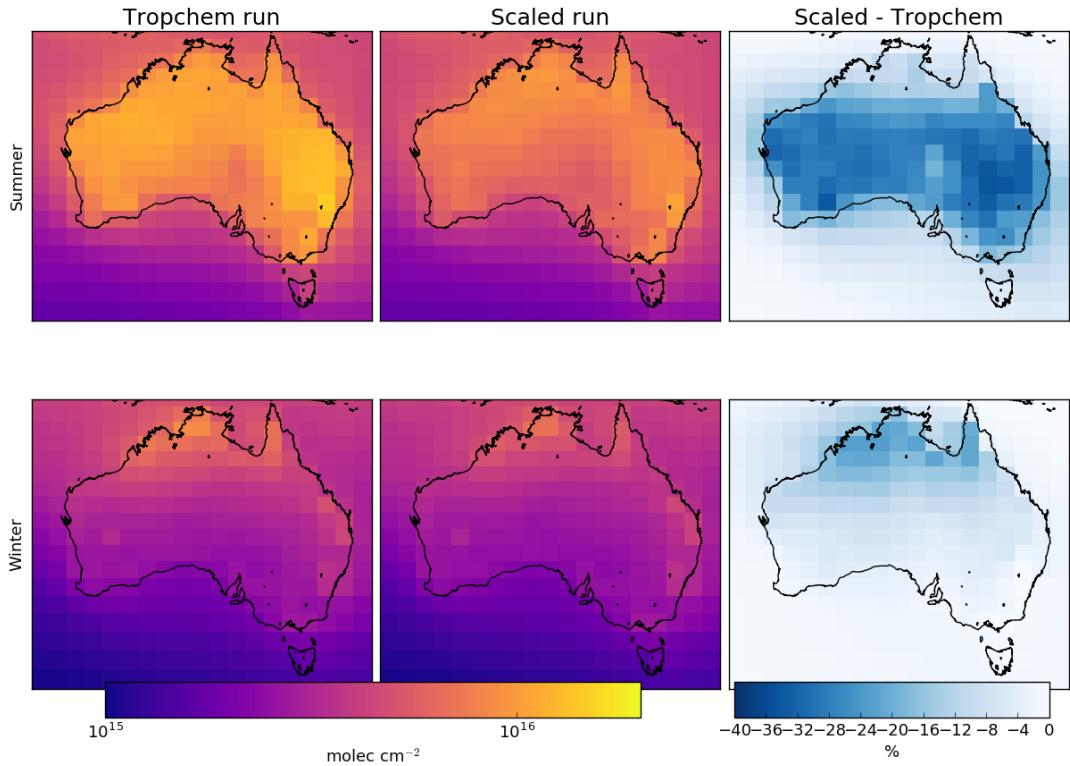
**Figure 25.** Row 1: Isoprene emissions scaling factor  $\alpha$  for the average January (left) and July (right) over 2005-2012. Row 2: a priori (magenta, left axis), a posteriori (cyan, left axis), and  $\alpha$  (black, right axis) multi-year monthly averages calculated for the  $2^\circ \times 2.5^\circ$  grid box containing Sydney.

### Daily mean $\Omega_{HCHO}$



**Figure 26.** Daily mean total column HCHO amounts from GEOS-Chem with a priori (new emissions run) and a posteriori (tropchem run) a posteriori scaled isoprene emissions, along with the recalculated OMI<sup>17</sup> HCHO columns. Each row shows the average over regions in Figure 17.

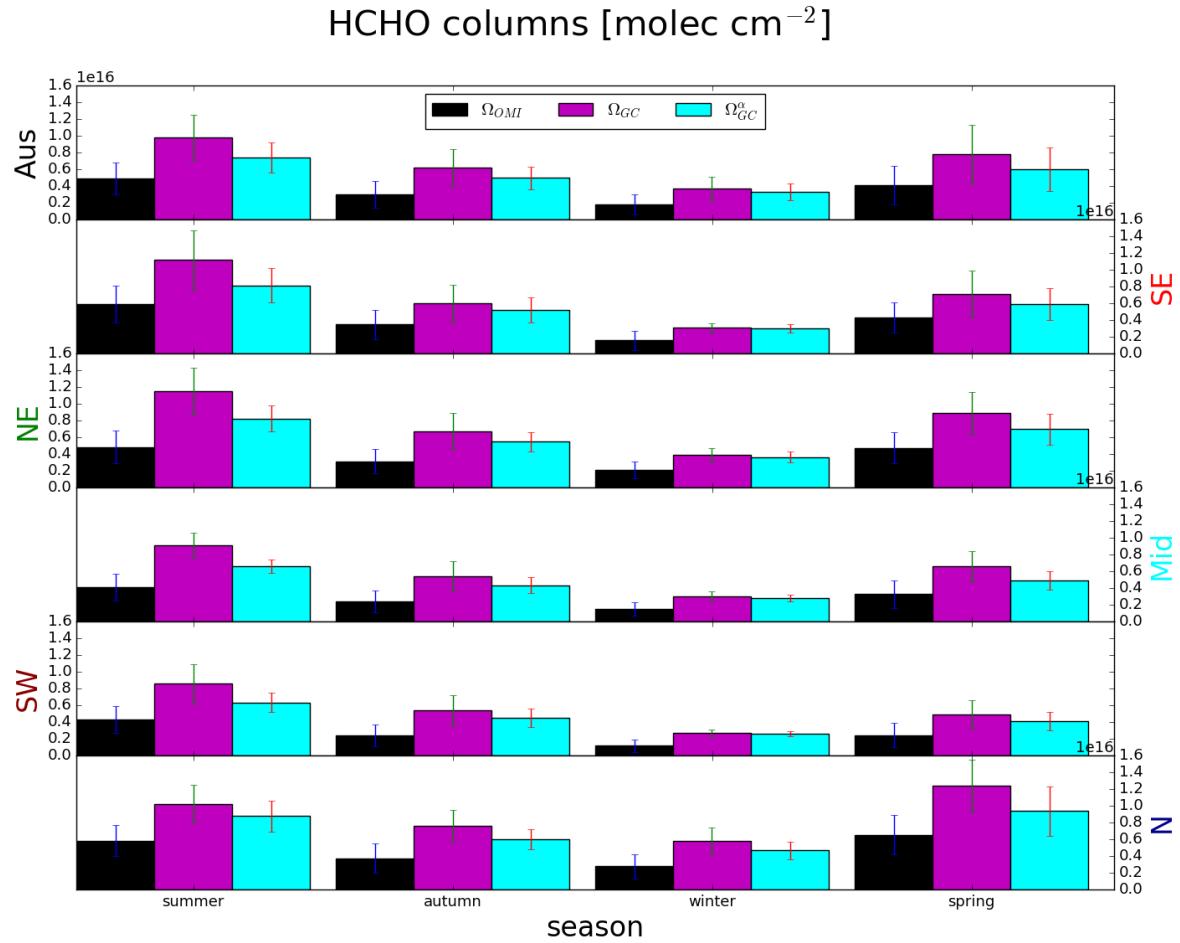
### Midday total column HCHO



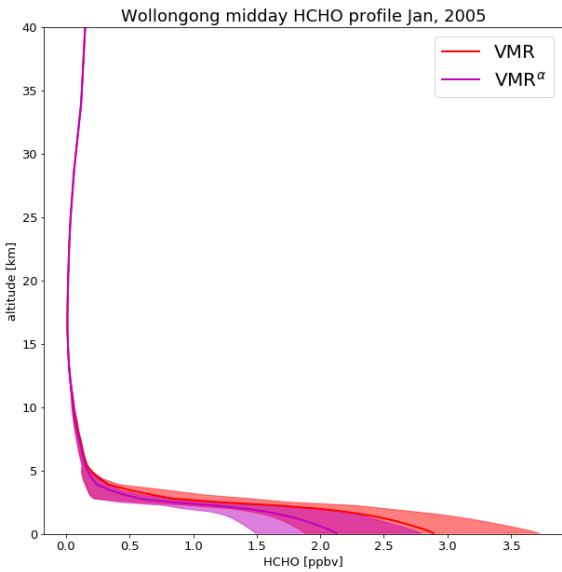
**Figure 27.** Total column HCHO a priori (left) and a posteriori (middle) isoprene emissions, and their relative differences (right). Top row shows summer (DJF) averaged total columns, while bottom row shows the winter (JJA).

#### 3.2.2 Implications for ozone

Isoprene oxidation can eventually lead to ozone formation, especially when isoprene-enriched air masses mix with polluted urban air masses that contain high NO<sub>x</sub>. Figure 30 shows surface level (up to ~ 150 m altitude) ozone concentrations over 2005 before and after scaling modelled isoprene emissions. Reducing isoprene emissions lowers surface ozone concentrations  
625 by ~ 1 ppbv in all regions in all seasons. The largest reductions (~ 3 ppbv) occur in the northern region in spring. A regression between the change in isoprene emissions, and the change in surface ozone was performed (as in the prior section with HCHO); however, only very weak correlations between reductions are apparent. While the overall decrease in surface ozone is clear, there is no direct correlation between monthly grid square averaged reductions in isoprene emission and surface ozone concentrations in the same grid square. This suggests that changes in isoprene emissions affect ozone in non-local grid squares due  
630 to transport, including grid squares where absolute emission reductions are not as strong. Figure 31 shows Australian summer



**Figure 28.** Regionally and seasonally averaged HCHO total columns from GEOS-Chem ( $\Omega_{GC}$ , and  $\Omega_{GC}^a$ ) and recalculated OMI measurements ( $\Omega_{OMI}$ ) side by side. Each row represents one region within Australia, while each column represents from left to right: summer, autumn, winter, spring. Standard deviations between grid boxes used to form the multi-year seasonal average are shown with error bars.



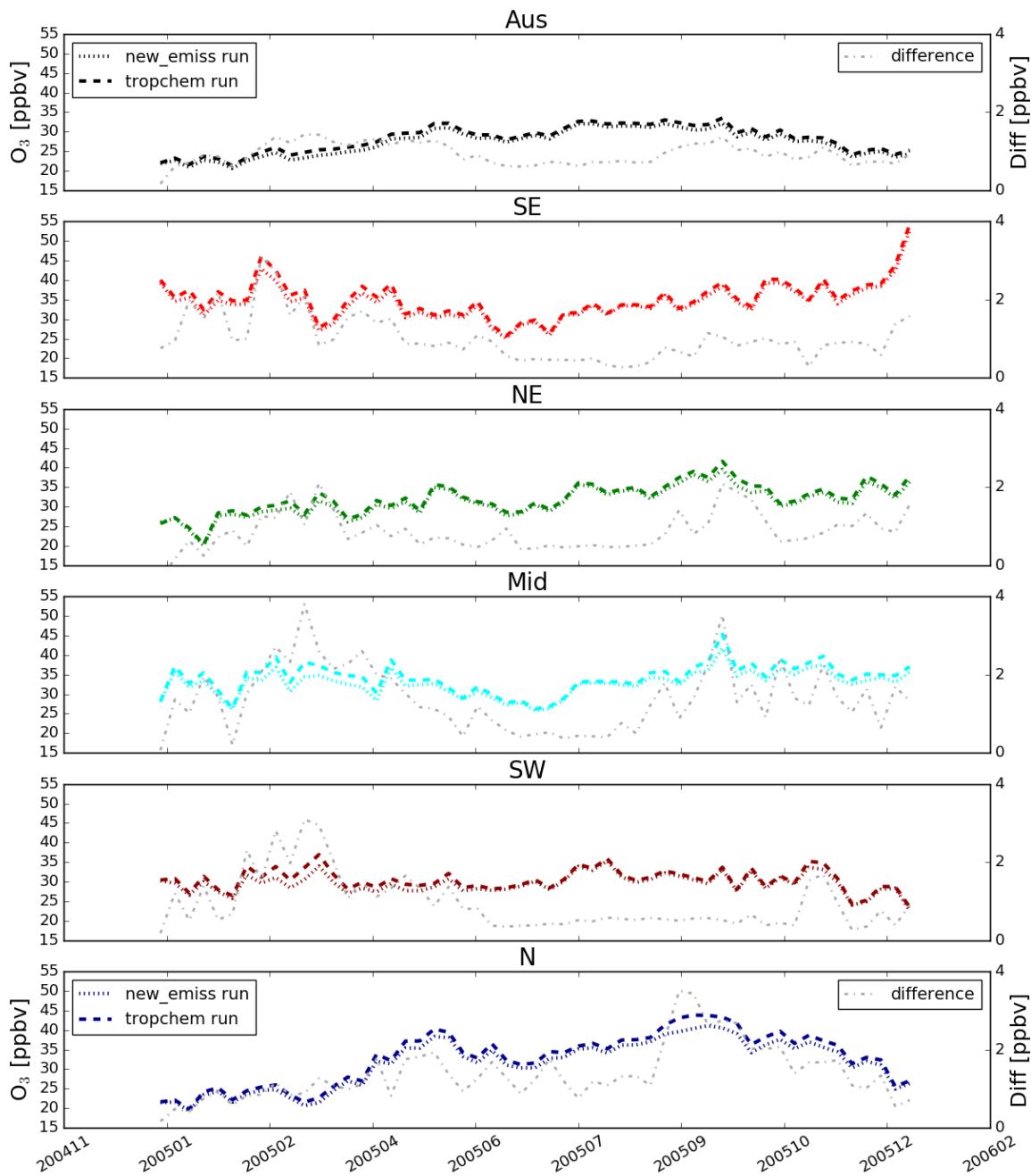
**Figure 29.** Monthly averaged HCHO profile over the  $2^\circ \times 2^\circ$  grid box containing Wollongong modelled by GEOS-Chem before (VMR) and after ( $\text{VMR}^\alpha$ ) scaling isoprene emissions. Shaded areas represent the inter-quartile range over the month.

surface level ozone with and without scaling isoprene emissions to match the multi-year top-down estimation. Surface ozone is reduced in summer by approximately 5%, with hotspots of reduction over southern New South Wales, Victoria, and Western Australia.

Ozone and fine particulate concentrations in Australian cities have not reduced over the last 10 years, unlike other atmospheric pollutants such as CO, NO<sub>2</sub>, and SO<sub>2</sub> (Keywood et al., 2016). This could be in part due to the downwind effects of isoprene emission, which are most likely to affect suburban fringes (e.g. western Sydney) of Australian cities which are surrounded by vegetation (Millet et al., 2016). Outside of densely populated regions, Australia is likely to be NO<sub>x</sub>-limited and changes in VOC emissions will have less of an impact on ozone production. Here the modelled ozone output is averaged over large areas that are mostly non-urban, which means detected ozone sensitivity to isoprene emissions is likely underestimated for cities. Figure 32 shows how ozone is reduced in summer and winter after scaling isoprene emissions. In summer, reductions are strongest in Sydney and Melbourne, but can also be seen in the west coast and central Australia. Winter reductions are more uncertain (relatively); however, they are strongest in the coastal and oceanic regions around Darwin.

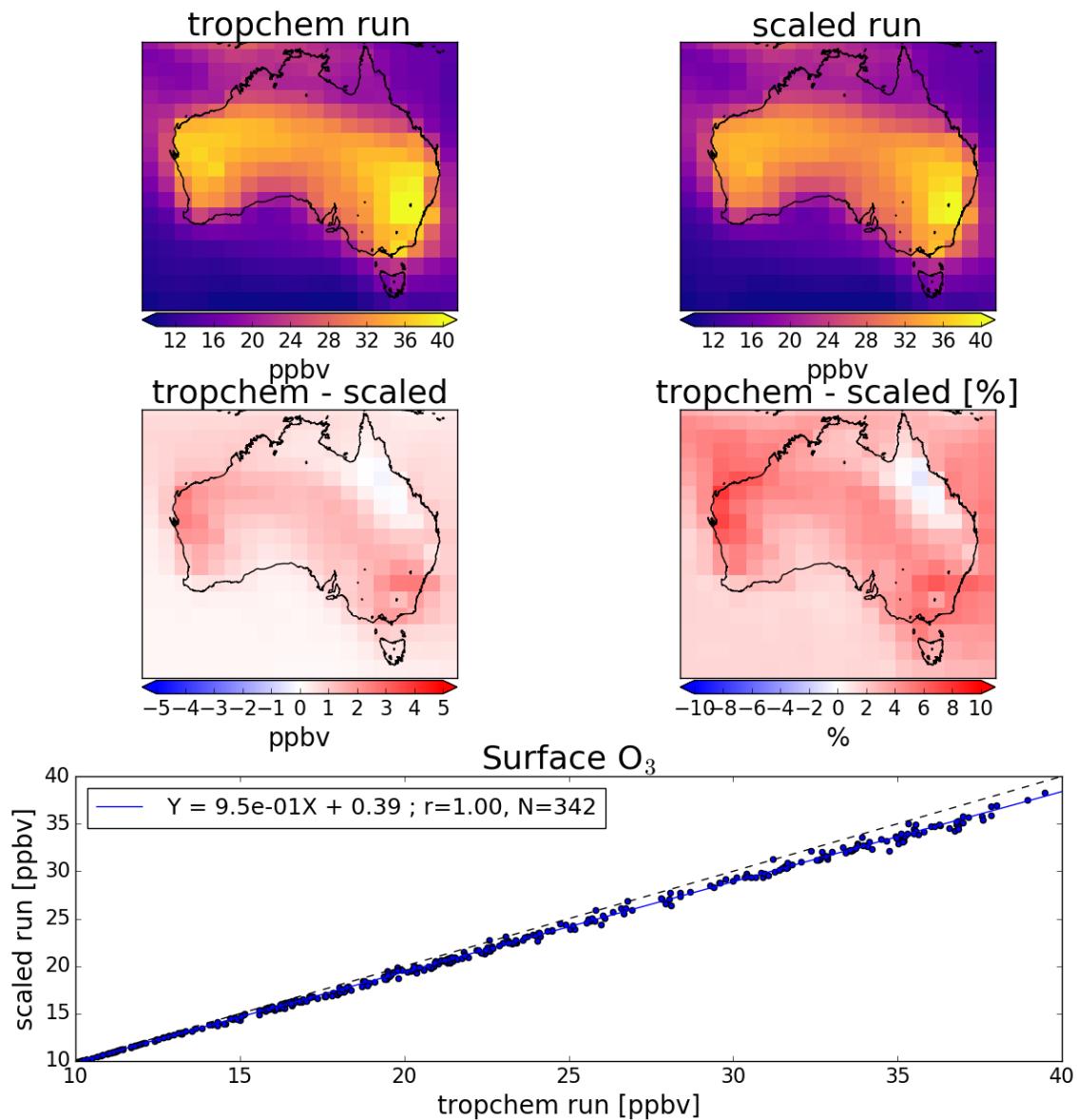
Table 7 shows regionally how the changed isoprene emissions affected ozone concentration yearly averages. The overall reduction is  $\sim 3\%$ , with the largest reduction occurring in the north region. The strongest reductions in isoprene emission do not coincide with those of ozone concentration. This is likely due to both transport (ozone formation occurs downwind

## Weekly mean surface O<sub>3</sub>

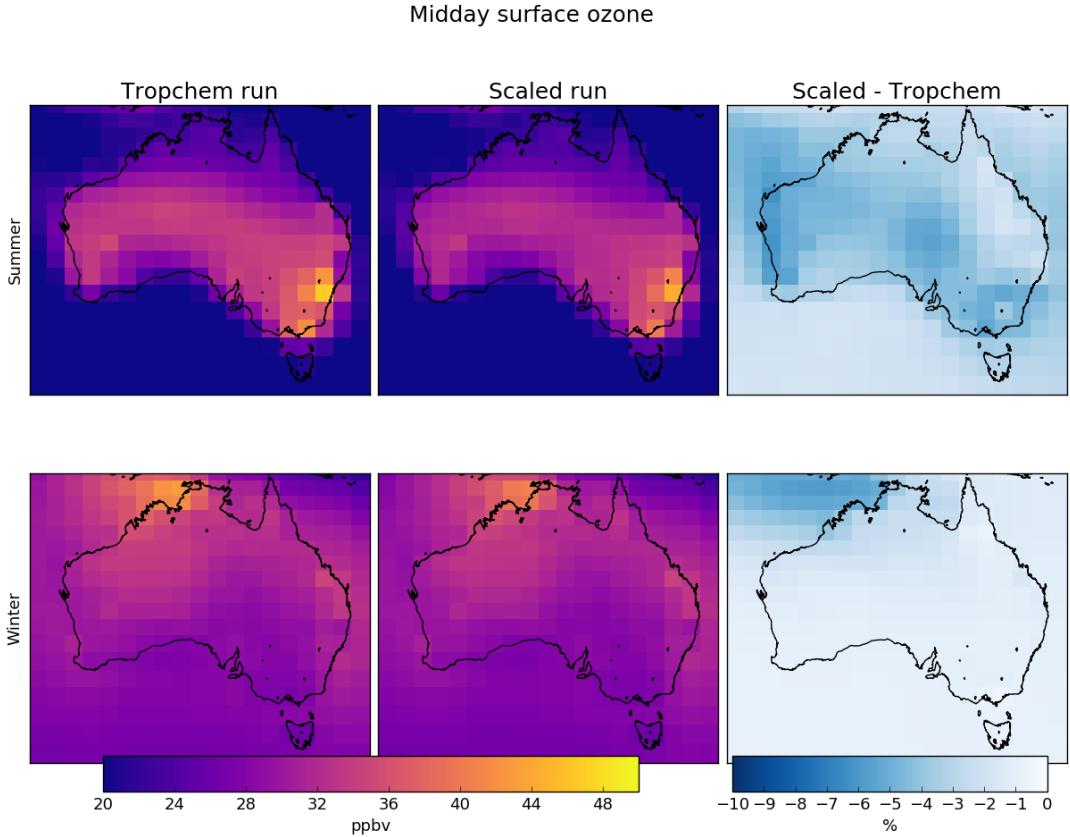


**Figure 30.** Surface ozone concentrations (ppb) per region over 2005. Concentrations are shown using the left axis, and absolute differences (tropchem run - scaled run) shown in grey using the right axis. **51**

GEOS-Chem surface ozone Summer\_05



**Figure 31.** Top row: Summer surface-level ozone before (left) and after (right) scaling isoprene emissions in GEOS-Chem to match top-down estimates. Middle row: absolute (left) and relative (right) differences. Bottom row: linear regression between the two model runs along with a black dashed line representing the 1-1 ratio. Each point represents one summer averaged land grid square (as shown in the top row).



**Figure 32.** Multi-year seasonal surface (up to  $\sim 150$  m) ozone before (left) and after (middle) scaling isoprene emissions, and their relative differences (right). Top row shows summer (DJF) averaged total columns, while bottom row shows the winter (JJA).

of isoprene emissions) and different  $\text{NO}_x$  concentrations within each region. In future work, the estimated emissions will be downscaled to better evaluate their impacts on ozone in different  $\text{NO}_x$  regimes (for instance over cities).

### 3.3 Comparison with in situ measurements

Comparison between ground-based measurements and large ( $2^\circ \times 2.5^\circ$ ) averaged model output suffers from representation error. Figure 33 shows the SPS and MUMBA measurement campaigns (see Section ??), along with the extent of the relevant  $2^\circ \times 2.5^\circ$  GEOS-Chem grid box, a rectangle with edge lengths of roughly 200 km $^2$ . The urban footprint of Sydney and Wollongong can be seen, along with ocean, forest, and rural regions, which will all affect the model output as it is based on the average of inputs within the grid box. Due to high uncertainty in components of the top-down emissions estimate, temporal resolution is also limited. MUMBA (Section ??), SPS1 and SPS2 (Section ??) provide approximately one month (two for

**Table 7.** Isoprene emissions from MEGAN and top-down estimate in Tg C a<sup>-1</sup>, along with ozone surface amounts in ppb.

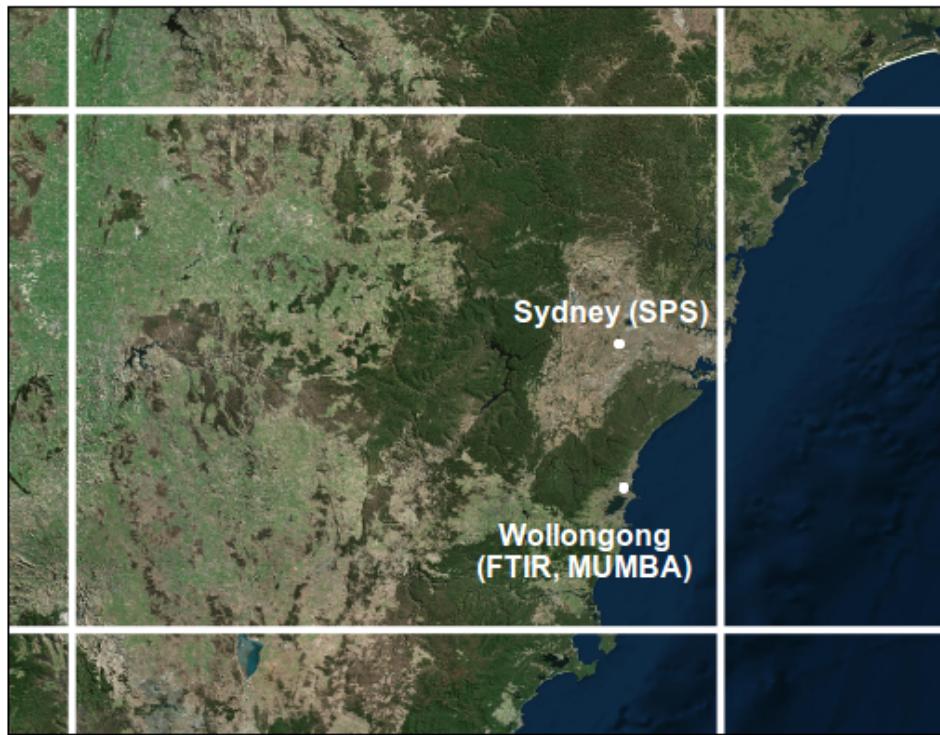
	AUS	SE	NE	MID	SW	N
Bottom-up						
Isoprene	34.67	2.80	6.33	2.31	2.04	9.95
Ozone	27.72	33.97	30.64	33.32	30.51	32.13
Top-Down						
Isoprene	18.26	1.61	3.02	1.03	0.99	6.31
Ozone <sup>a</sup>	26.93	33.10	29.99	32.19	29.69	30.92

a: Outputs from GEOS-Chem run with isoprene emissions scaled to match the top-down estimate.

655 MUMBA) of hourly or daily data, which are compared in this section against surface level concentrations from GEOS-Chem (midday output) before and after scaling the biogenic emissions.

Figure 34 shows GEOS-Chem output in the grid square containing Sydney and Wollongong campaign measurements. Measurements between 13:00 and 14:00 are averaged daily. In order to minimise introduced bias, any measurements below the instrument detection limit are set to half of the detection limit, as performed in Lawson et al. (2015). Midday measurements 660 from MUMBA (summer, Wollongong) compared against SPS (summer, autumn, Sydney) show lower isoprene and HCHO levels, and similar ozone levels, and also lower variance across all three species. Variation in SPS data are higher than modelled, which is likely due to a dependence on local meteorology, as plumes of HCHO or isoprene enriched air can be detected as they pass by while the modelled grid square averages these out. This higher variance is not seen in the MUMBA midday observations; however, this may be due to the small sample size of the measured data.

665 Coincident measured and modelled means and comparison statistics for each campaign for isoprene, HCHO, and ozone are summarised in Table 8. Isoprene correlations are greatest during the SPS1 campaign at  $r = 0.83, 0.84$ , nearly double those of SPS2 and MUMBA. The root mean square error (RMSE) in MUMBA is 1.7 ppb ( $\sim 300\%$ ) and approximately 0.9 ppb and 0.4 ppb in SPS1, SPS2 respectively ( $\sim 100\%$ ), before isoprene scaling is applied. In MUMBA the reduced isoprene emissions improves the model drastically, bringing the RMSE down to  $\sim 100\%$ , however this improvement is not seen in either of SPS1 670 or SPS2. Similarly for HCHO, the MUMBA RMSE is reduced from 2.0 ppb ( $\sim 130\%$ ) to 1.2 ppb ( $\sim 80\%$ ) by bringing down the mean modelled amounts. Modelled HCHO is low compared against measurements from Sydney (SPS1 and SPS2), and is further lowered through scaling. Without improved model resolution or further measurements, this is assumed to be due to local influences seen in the measurements but not captured by the model, such as HCHO enriched air from local sources. Modelled ozone concentrations are slightly reduced (3 – 8%) when scaling isoprene emissions, however the concentrations 675 remain 30 – 50% too high for Sydney and approximately 100% too high for Wollongong. This is likely due to the reduced

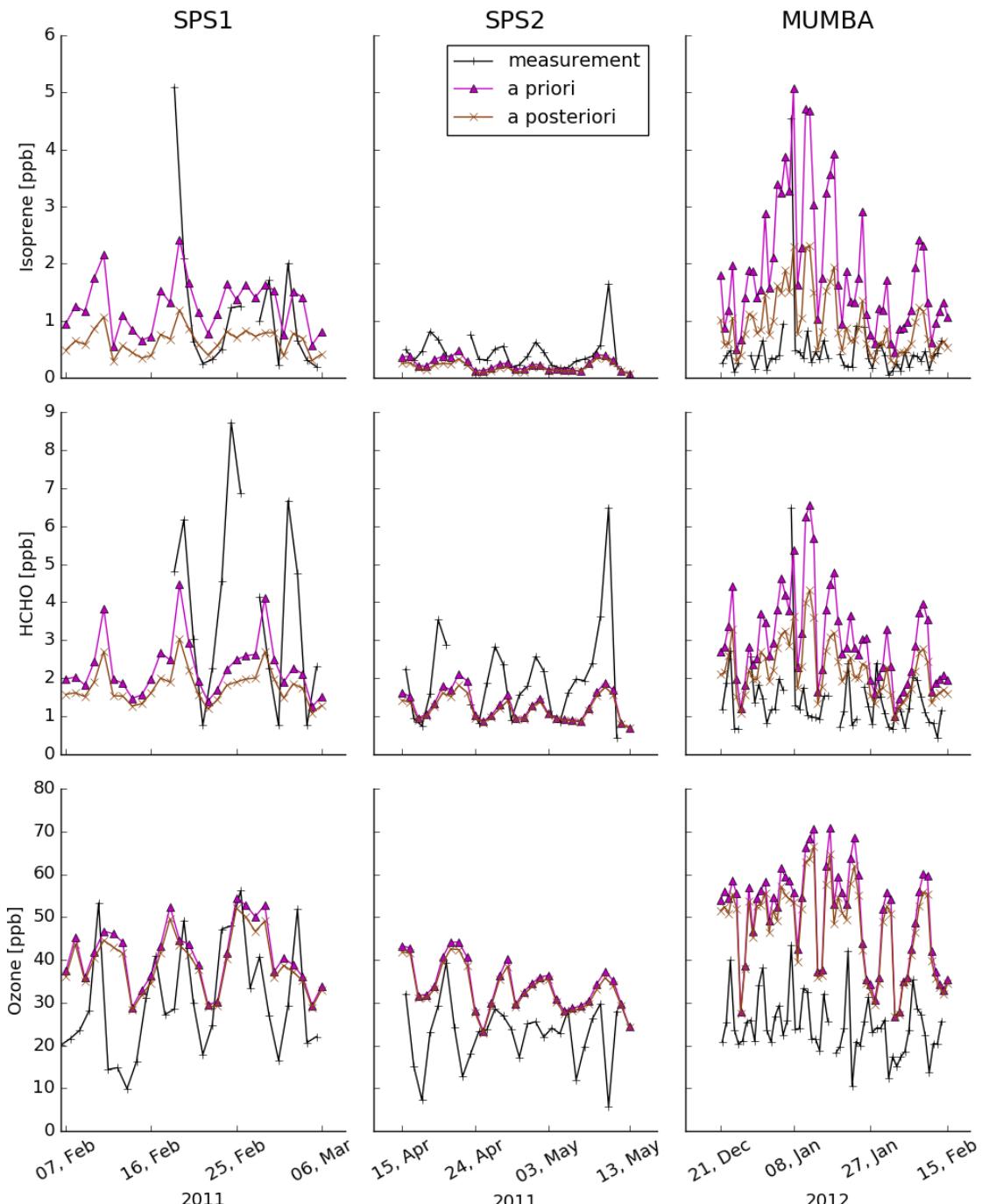


**Figure 33.** GEOS-Chem grid box ( $2^\circ \times 2.5^\circ$ ) containing Wollongong FTIR, SPS, and MUMBA campaign data.

ozone concentrations in cities compared to regional areas, as the modelled ozone concentrations are averages over the larger area which is mostly non-urban.

A spectrometer (FTIR) on the roof of the Chemistry building at the University of Wollongong measures total column HCHO in the atmosphere during clear sky conditions. This is the only non-satellite long-term measurement record of total column HCHO available in Australia (see Section ??). In order to compare modelled profiles against retrievals from the FTIR, modelled profiles are first convolved with the instrument averaging kernel and a priori. FTIR output is resampled to only include measurements taken at midday (13:00-14:00) and GEOS-Chem overpass outputs are interpolated onto matching vertical levels for days where FTIR output exists. Figure 35 shows total column HCHO from FTIR, and from GEOS-Chem before and after scaling isoprene emissions. While the comparison suffers from representational error between FTIR measurements and GEOS-Chem output (see Figure 33), one can see that the summer overestimate of HCHO from GEOS-Chem is removed by isoprene scaling. The mean summer underestimate of HCHO shown after scaling makes more sense due to the relatively dense

### GEOS-Chem vs campaign data



**Figure 34.** SPS1, SPS2, and MUMBA (left to right columns respectively) midday (13:00–14:00 local time) measurements of isoprene, HCHO, and ozone (top to bottom rows respectively). Shown in magenta and brown are the a priori and a posteriori GEOS-Chem surface outputs for the matching grid square at midday for days containing measurements.

**Table 8.** Campaign measurements compared to model output [ppb].

	SPS1			SPS2			MUMBA		
	mean	RMSE	r	mean	RMSE	r	mean	RMSE	r
Isoprene									
Meas	1.16			0.46			0.47		
GC	1.34	0.89	0.83	0.24	0.36	0.44	1.87	1.66	0.52
GC $^\alpha$	0.67	1.15	0.84	0.18	0.39	0.43	0.93	0.73	0.49
HCHO	mean	RMSE	r	mean	RMSE	r	mean	RMSE	r
Meas	3.92			2.02			1.41		
GC	2.38	2.63	0.47	1.28	1.28	0.77	3.00	1.97	0.44
GC $^\alpha$	1.83	3.01	0.54	1.19	1.35	0.78	2.21	1.17	0.47
Ozone	mean	RMSE	r	mean	RMSE	r	mean	RMSE	r
Meas	30.14			22.71			24.59		
GC	40.91	15.46	0.52	33.75	13.59	0.35	49.97	28.01	0.34
GC $^\alpha$	39.20	14.23	0.54	32.81	12.69	0.36	47.14	24.91	0.36

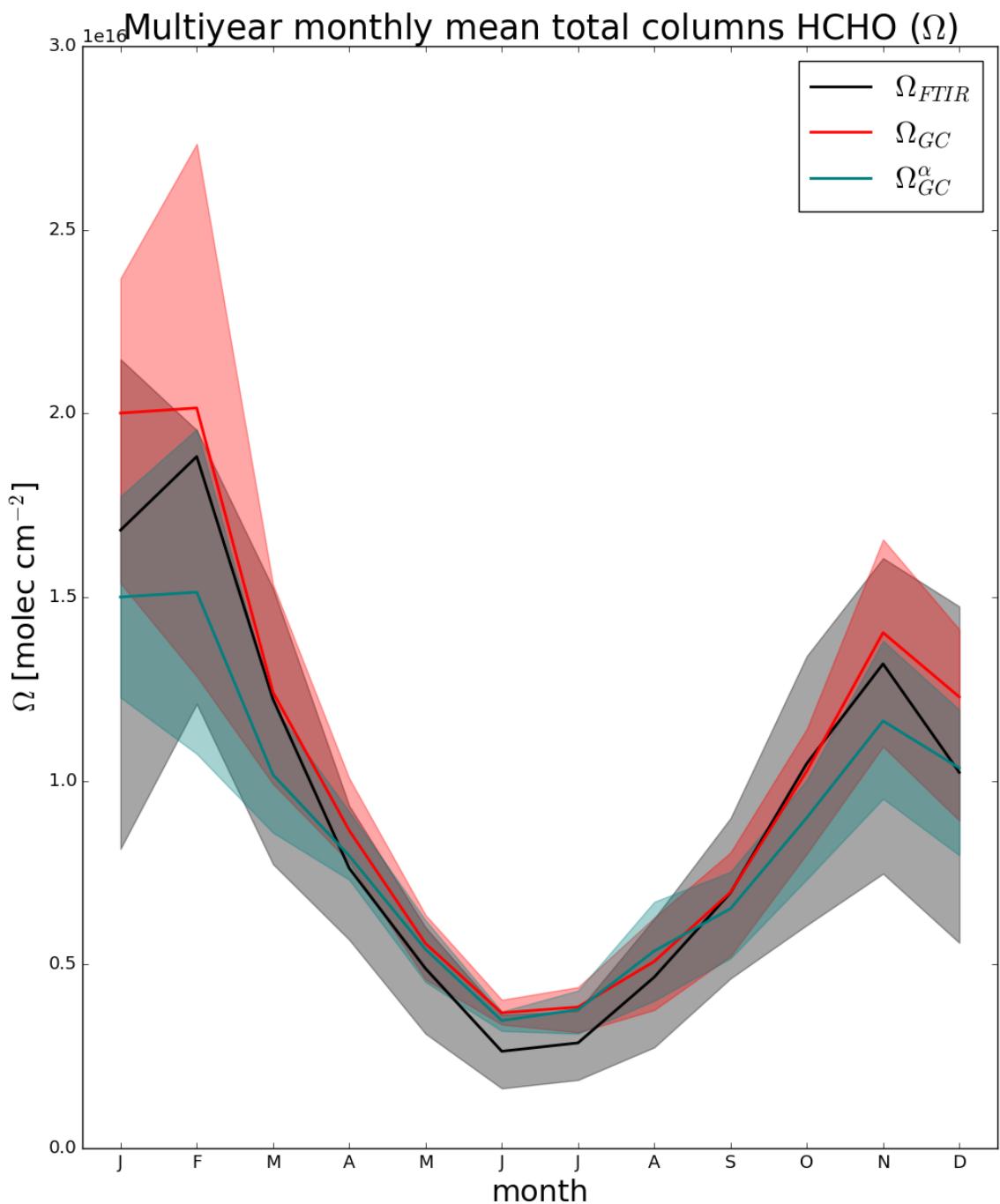
Meas: Measurements (midday means)

GC: GEOS-Chem (tropchem run)

GC $^\alpha$ : (GEOS-Chem scaled run)**Table 9.** Mean total column HCHO amounts in  $10^{15}$  molec cm $^{-2}$  at Wollongong.

Season	$\Omega_{FTIR}$	$\Omega_{GC}$	$\Omega_{GC}^\alpha$
summer	15.4	17.7	13.6
winter	3.68	4.41	4.49

forested areas around Wollongong, which would raise local HCHO concentrations above the average for the large grid square represented by GEOS-Chem. Another feature is the January and December dip in all three datasets. The FTIR has a lower data count and higher variance in these (holiday) months as measurements involve manual processes, increasing uncertainty for the 690 measurements. Analysis of this dip could be performed in future work; however, it would require an in depth examination of local and synoptic HCHO patterns, and higher resolution model data. Seasonal mean HCHO from the FTIR measurements and co-located, convolved, and resampled model output is summarised in Table 9. The original modelled summer vertical column is higher than the FTIR measurements by  $\sim 15\%$ , which decreases with isoprene scaling to  $\sim -12\%$ . In winter the column is modelled high by  $\sim 20\%$ , which increases to  $\sim 22\%$  after scaling.



**Figure 35.** Multi-year (from all days where both measurements and model data exist: 2007-2012) monthly mean total column ( $\Omega$ ) HCHO from the FTIR instrument at Wollongong (southeast Australia), and the colocated convolved GEOS-Chem equivalent before ( $\Omega_{GC}$ ) and after ( $\Omega_{GC}^\alpha$ ) scaling isoprene emissions. Shaded areas show inter quartile range.

**Table 10.** Relative uncertainty estimates.

Region	Summer			Winter		
	$\frac{\Delta E_{OMI}}{E_{OMI}}$	$\frac{\Delta \Omega_{OMI}}{\Omega_{OMI}}$	$\frac{\Delta S}{S}$	$\frac{\Delta E_{OMI}}{E_{OMI}}$	$\frac{\Delta \Omega_{OMI}}{\Omega_{OMI}}$	$\frac{\Delta S}{S}$
Aus	46%	21%	36%	51%	258%	35%
SE	54%	14%	37%	66%	141%	38%
NE	53%	18%	35%	51%	40%	37%
Mid	45%	38%	38%	47%	41%	35%
SW	43%	27%	32%	61%	96%	36%
N	39%	16%	33%	38%	22%	30%

a: Grid squares with monthly uncertainty over 200% are removed when calculating the mean uncertainty, which has small (< 1%) impacts on  $E_{OMI}$  overall, but removes many data points in winter.

## 695 4 Uncertainty

This section identifies and quantifies the overall uncertainties of calculating isoprene emissions using OMI HCHO observations and the GEOS-Chem model in the top-down method used in this chapter. However, these uncertainties lack verification against measurements. Even as the top-down inversion performed in this chapter attempts to work around the lack of measurements over Australia, it suffers from the lack of independent observations against which it can be verified.

700 The major source of uncertainty throughout the year comes from uncertainty in the modelled yield slope  $S$  (see Section 4.2); however, in winter, uncertainty from satellite column calculation becomes dominant (see Section 4.3). Monthly calculated slope uncertainty mostly lies within 30% to 50%, and for each grid box a month of model data is used, which means that slope uncertainty is not reduced by monthly averaging. Uncertainty from each OMI satellite measurement is relatively large (> 100%); however, averaging thousands of pixels in each grid square greatly reduces the monthly uncertainty. Uncertainty in 705 satellite HCHO is seasonally dependent, with better signal during the summer. Reliable OMI measurements are less frequent (leading to higher uncertainty) at high solar zenith angles, which are more frequent at higher latitudes and during winter. Table 10 shows the estimated uncertainty calculated in this work in summer and winter over each region described by Figure 17. The assumptions and calculations made to determine uncertainties in the top-down estimate ( $\Delta E_{OMI}$ ), the satellite column (710  $\Delta \Omega_{OMI}$ ) and the slope ( $\Delta S$ ) are described in the following subsections.

### 710 4.1 Top-down emissions

Important factors in the calculation of isoprene emissions using OMI HCHO include the modelled relationship between HCHO and isoprene and the satellite HCHO measurements. Uncertainty in each of these terms is quantified before being combined in quadrature to give the uncertainty estimate of the a posteriori. Additional biases may arise due to the filters applied to satellite data and model output, and where possible these are assessed.

715 The final determination of top-down emissions comes from Equation 11, repeated here:

$$E_{OMI} = \frac{\Omega_{OMI} - \Omega_{OMI,0}}{S}$$

Assuming each term is independent, the following quadrature rules are used to estimate random error in  $E_{OMI}$ :

$$z = x + y : \Delta z = \sqrt{(\Delta x)^2 + (\Delta y)^2} \quad (13)$$

$$z = x/y : \Delta z = z \sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2} \quad (14)$$

720 Which leads to the uncertainty estimation for the a posteriori emissions as follows

$$\Phi \equiv \Omega_{OMI} - \Omega_{OMI,0}$$

$$\Delta\Phi = \sqrt{(\Delta\Omega_{OMI})^2 + (\Delta\Omega_{OMI,0})^2} \quad (15)$$

$$\Delta E_{OMI} = E_{OMI} \times \sqrt{\left(\frac{\Delta\Phi}{\Phi}\right)^2 + \left(\frac{\Delta S}{S}\right)^2} \quad (16)$$

725  $\Delta E_{OMI}$  is calculated using the uncertainty in underlying terms:  $\Delta S$ ,  $\Delta\Omega_{OMI}$ , and  $\Delta\Omega_{OMI,0}$ . For  $\Delta S$  ( $\Omega_{GC} = S \times E_{GC} + \Omega_{OMI,0}$  from equation 10), the variance in the monthly linear regression of modelled isoprene emissions and column HCHO is used, shown in Section 4.2. For  $\Delta\Omega_{OMI}$  and  $\Delta\Omega_{OMI,0}$ , uncertainty comes from instrument fitting uncertainty, modelled AMF uncertainty, and uncertainty in the background correction terms, which are described and calculated in Section 4.3.

730 Figure 36 shows relative uncertainty over each region of Australia in monthly bins. Uncertainty in the southern regions increases between May and July due to increased error in the satellite measurements. Northern regions are impacted less by the seasonal satellite error, which generally is 10 – 20 % lower than the error from  $S$ . Figure 37 shows the spatial distribution of relative uncertainty in the a posteriori in summer and winter. Here the effects of satellite uncertainty at higher latitudes (especially in winter) can be seen through the increase in uncertainty with increasing latitude.

## 4.2 Model Uncertainty

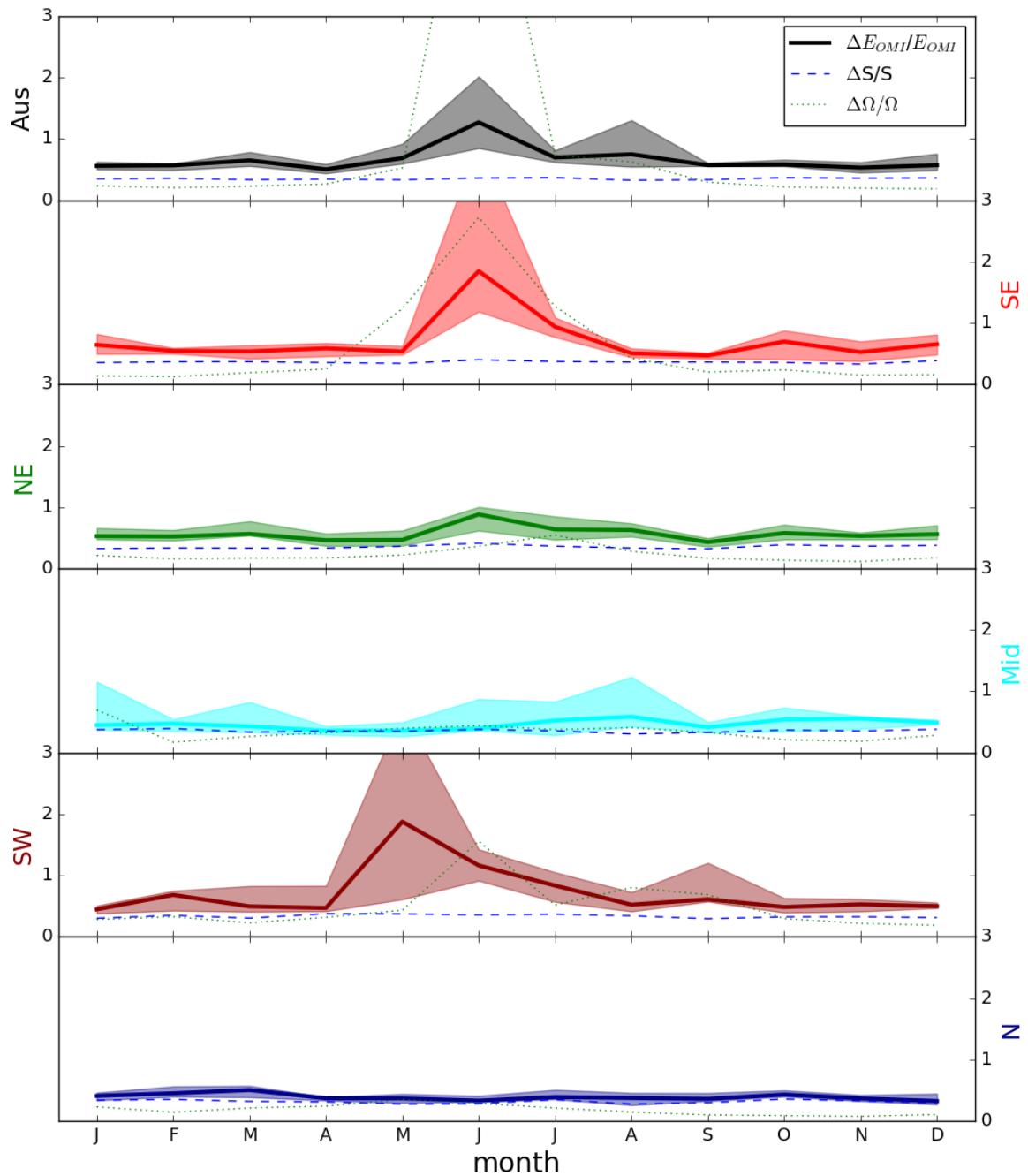
735  $E_{OMI}$  depends partly on the product it is trying to improve, as modelled yield is based on GEOS-Chem run with MEGAN emissions. The uncertainty in the reduced major axis regression slope between model HCHO ( $\Omega_{GC}$ ) and emissions ( $E_{GC}$ ) is used to estimate  $\Delta S$  in Equation 16. The upper bound of the 95% confidence interval of the regression slope ( $CI_{UB}$ ) is used to represent the relative uncertainty.

$$\frac{\Delta S}{S} = \frac{CI_{UB}}{S} - 1 \quad (17)$$

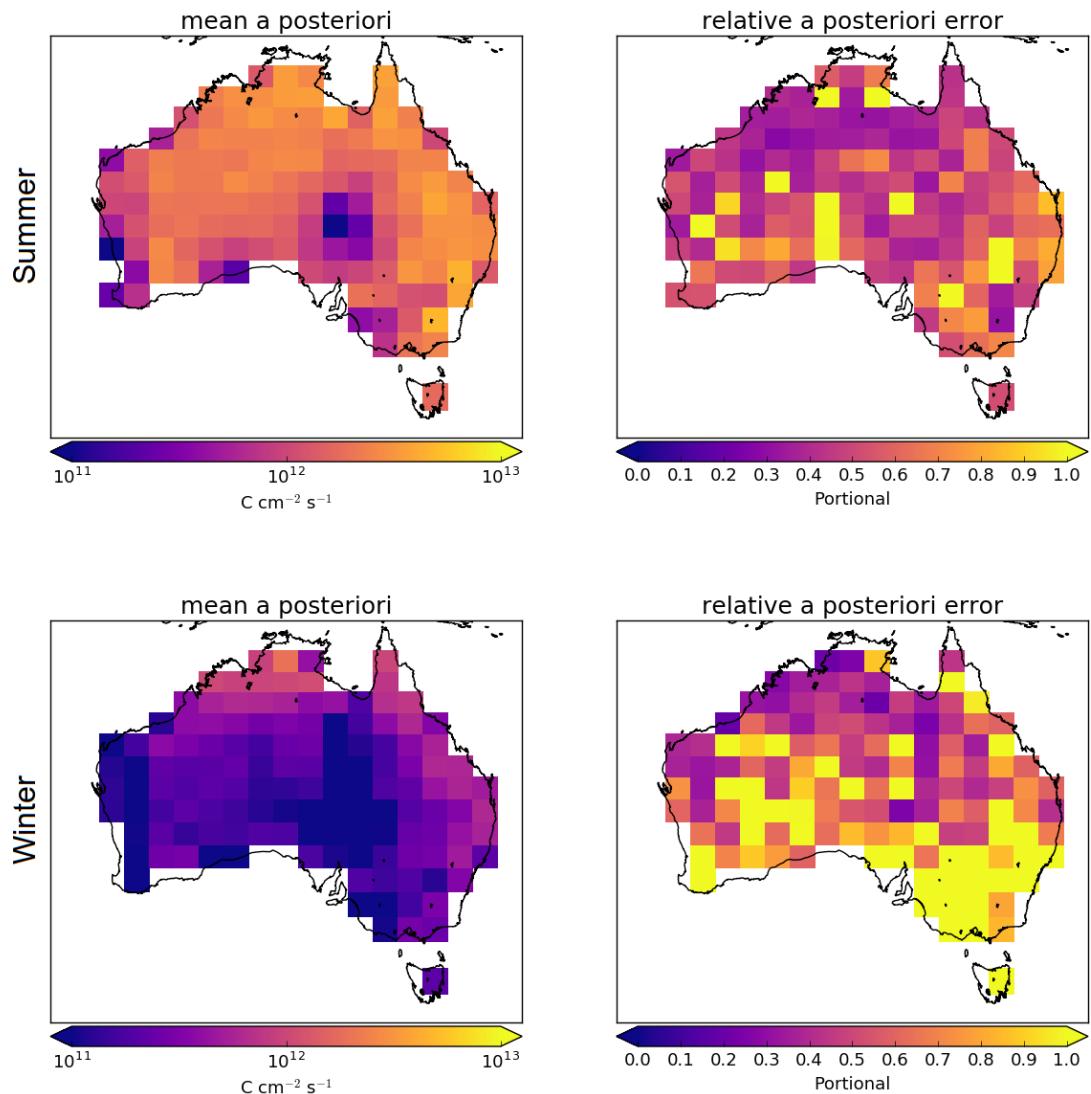
740 For example, if the interval upper bound is 30% higher than the slope, relative uncertainty is set to 0.3 (or 30%). The confidence interval for each month is based on the covariance matrix between  $\Omega_{GC}$  and  $E_{GC}$ , and the critical t-statistic considering  $n$  to be the number of days in the month (the significance value of the test is set to 0.025).

This is a simple method of approximating the uncertainty of this term, only accounting for monthly uncertainty of the slope calculation. It does not take into account uncertainty in the underlying model, nor representation uncertainties arising from

### relative error in monthly a posteriori



**Figure 36.** Median and inter-quartile range of multi-year monthly relative uncertainty in the a posteriori. Median relative uncertainty in  $S$  and  $\Omega$  are added as dashed and dotted lines respectively.



**Figure 37.** Summer (DJF, top row) and winter (JJA, bottom row) a posteriori emissions (left column) and relative error (right column).

temporal or spatial resolution, which are difficult to quantify. Figure 38 shows the relative uncertainty in  $S$  over Australia and  
745 for each region. There is little discernible seasonality to the relative error in  $S$ , which generally ranges from 0.3–0.4 (30–40%).  
For comparison, Palmer et al. (2006) found  $\frac{\Delta S}{S}$  to be 30% over the US after comparing with another chemical transport  
model and in situ measurements. To improve understanding of uncertainty in  $S$  would require further analysis of GEOS-Chem  
yield over Australia, including how it responds to environmental and meteorological parameters, and how representative this  
750 modelled quantity is when compared to measurements. Slope estimation may also be affected by non-isoprene biogenic VOC  
emissions such as monoterpenes, which are not analysed in this work. The effects of these biogenics will be in part seen by  
 $\Delta S$  (as they make up some of the variability); however this is an important source of HCHO that will need to be attributed in  
future work to improve confidence in the estimation of  $S$ .

Filtering for spatial smearing (see Section 2.6.1) reduces the number of data points making up the regression slope  $S$ . The  
process generally improves the linear relationship between isoprene emission and HCHO total column, which suggests it is  
755 working as intended to remove days when local biogenic emissions are not driving HCHO enhancement. Where  $S$  regressions  
have a correlation coefficient ( $r$ ) of less than 0.4, a multi-year average (or in the worst cases no value at all) is used in lieu of  
monthly  $S$  data. The thresholds for the smearing filter are based on literature values from other countries, which may prove  
to be unsuitable within Australia. To improve the understanding of smearing in Australia, a better approximation of HCHO  
lifetimes and yields, as well as  $\text{NO}_x$  seasonality and regional concentrations, is required. The filtering process is accounted for  
760 in the calculation of  $\Delta S$ , and the effect is a small reduction of the overall uncertainty from  $S$ .

Model biases are not analysed in this thesis, except to note that they would impact both preliminary OMI calculations (via  
use in calculating the air mass factor) and the modelled slope. Insufficient independent measurements in Australia make it  
impossible to quantify uncertainty at a national scale.

### 4.3 Satellite Uncertainty

765 Corrected vertical columns of HCHO from the OMI product are calculated using Equation ??:  $\Omega = \frac{SC - RSC}{AMF}$ . Error in satellite  
HCHO columns is determined by error in the three terms  $SC$ ,  $RSC$ , and  $AMF$ :

#### Fitting error from the OMI retrieval

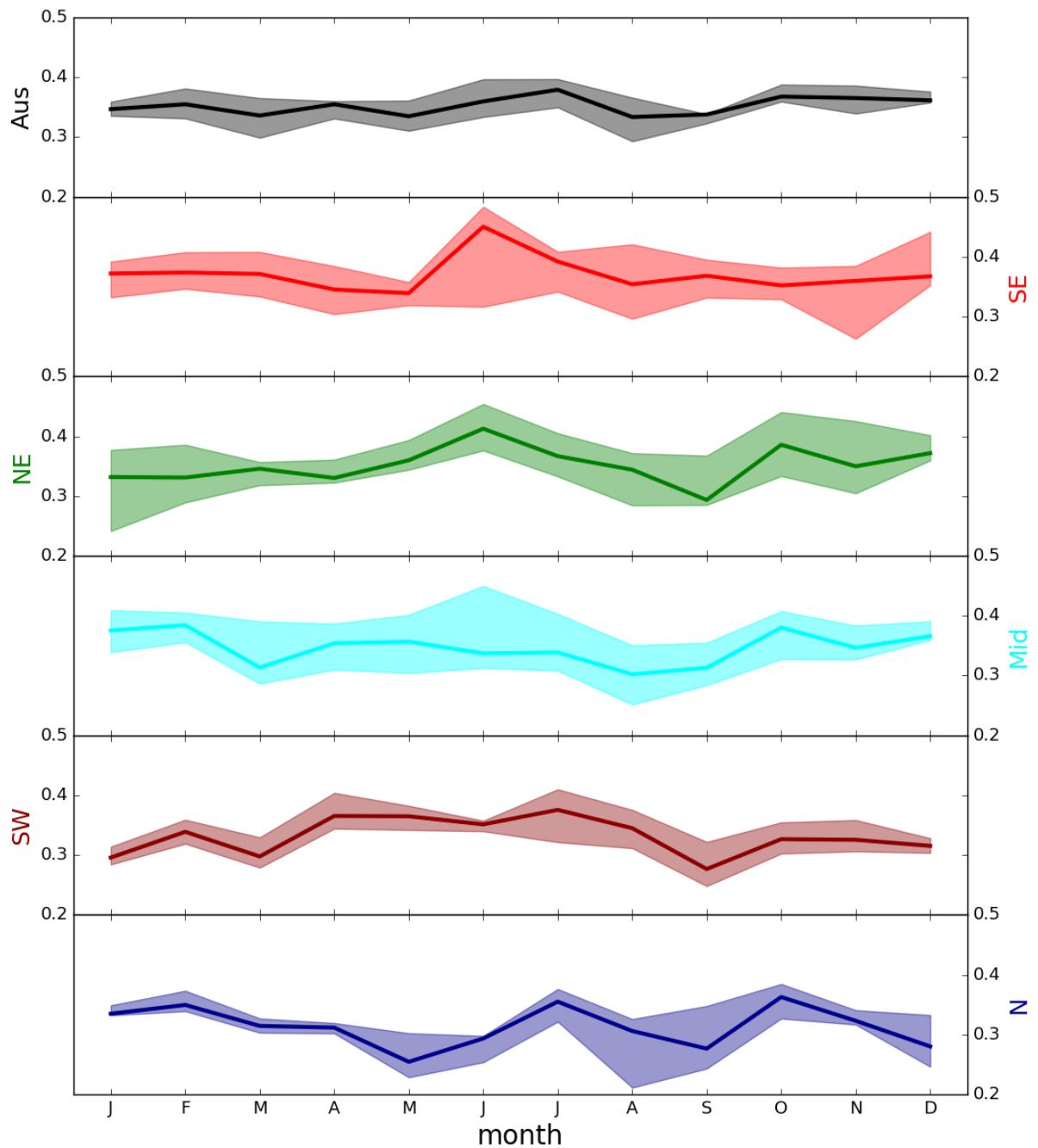
Fitting error represents the uncertainty in the DOAS technique used to estimate HCHO concentrations. Fitting error is  
provided in the OMHCHO product. This error is ascribed to the  $SC$  term.

770 Provided with the OMI product is the measurement of uncertainty in each pixel, calculated by the Smithsonian Astro-  
physical Observatory from the back scattered solar radiation fit (Gonzalez Abad et al., 2015; Abad et al., 2016). This is  
used as the  $\Delta SC$  in Equation 18. The relative fitting error per pixel ( $\frac{\Delta SC}{SC}$ ) ranges from around 20% to 150%, and is  
higher where low amounts of HCHO are detected, and at higher solar zenith angles (i.e. at high latitudes).

#### Uncertainty in AMF calculations

775 Air mass factors model vertical sensitivity to slant path measurements in the satellite instrument, and uncertainties arise  
predominantly from uncertain cloud parameters (Palmer et al., 2006). Vertical columns from OMI are recalculated using

### relative error in monthly slope



**Figure 38.** Median and inter-quartile range of monthly binned uncertainty in  $S$ . 64

AMFs derived from GEOS-Chem (Section ??). AMF uncertainty can be determined through comparison of GEOS-Chem output to independently measured HCHO columns. In this work the AMF error is set to 30%, which approximates that seen by Palmer et al. (2006), since measurements over Australia are lacking. For comparison, Wollongong FTIR measurements are over-predicted by GEOS-Chem modelled vertical columns by  $\sim 15 - 20\%$ , but these are not representative at the continental scale simulated here. The assumed error of 30% is ascribed to the *AMF* term.

Palmer et al. (2006) calculated the error in AMF through combining estimates of error in the UV albedo database ( $\sim 8\%$ ), model error based on in situ measurements, cloud error (20–30% (Martin, 2003a)), and aerosol errors (< 20%), totalling AMF error of around  $\sim 30\%$  (calculated in quadrature). This error estimate can be compared with that of Curci et al. (2010), where the error in AMF calculations and background columns were respectively found to be 30% and 15%. Millet et al. (2008) also examined this uncertainty and determine an overall uncertainty ( $1\sigma$ ) of 25–27% in HCHO vertical columns with calculated AMFs where cloud fraction  $< 0.2$ .

### Uncertainty of HCHO background

OMI vertical columns are corrected using background (or reference sector) measurements. This is to account for instrument degradation, and adds some uncertainty to the column. In this work vertical columns are corrected using reference sector measurements combined with modelled HCHO, which is described in Chapter ?? (Section ??). Error from background uncertainty is ascribed to the *RSC* term.

The *RSC*, or background correction, is based on differences in the remote Pacific between daily HCHO slant columns measured by OMI and monthly averages from GEOS-Chem. The correction for each pixel is determined per latitude and OMI track; however, a couple of conservative simplifications are used here to estimate the error in this term. For each SC, the *RSC* is set to the mean correction matching the *SC* latitude over all tracks. For each day, the  $\Delta RSC$  is set to the standard deviation of the *RSC* over Australian latitudes ( $45^\circ\text{S}$  to  $10^\circ\text{S}$ ) in all tracks. Background error calculated in this way is on the order of 5%-10% after monthly averaging, however this error increases in the higher southern latitudes during autumn and winter to  $\sim 15\%$ . For comparison, the background error is assumed to be 15% in Curci et al. (2010) following (Dufour et al., 2008).

Calculation of uncertainty in the OMI vertical column HCHO ( $\Omega = \frac{SC - RSC}{AMF}$ ) is performed using quadrature equations 13 and 14. Error in the slant column ( $\Delta SC$ ) is combined with assumed relative AMF error ( $\frac{\Delta AMF}{AMF}$ ) of 30%, and background error to calculate  $\Delta\Omega_{OMI}$  (and  $\Delta\Omega_{OMI,0}$ ) as follows:

$$\begin{aligned} \Delta(SC - RSC) &= \sqrt{(\Delta SC)^2 + (\Delta RSC)^2} \\ \Delta\Omega &= \Omega \sqrt{\left( \frac{\Delta(SC - RSC)}{(SC - RSC)} \right)^2 + \left( \frac{\Delta AMF}{AMF} \right)^2} \\ \frac{\Delta\Omega}{\Omega} &= \sqrt{\frac{(\Delta SC)^2 + (\Delta RSC)^2}{(SC - RSC)^2} + \left( \frac{\Delta AMF}{AMF} \right)^2} \end{aligned} \quad (18)$$

The *RSC* term is described in Chapter ?? Section ?? . Negative columns can occur where column amounts are lower than *RSC*, and these are not removed so as not to introduce a bias. When monthly averages are less than zero, relative error

**Table 11.** Uncertainties in satellite total column HCHO.

uncertainty	location	notes
40%	North America	GOME, mostly due to cloud interference <sup>a</sup>
26%	North America	GOME, OMI, with cloud fraction less than 20% <sup>b</sup>
30%-40%	global	GOME-2 <sup>c</sup>
> 60%	Mid-latitude	GOME-2 in winters <sup>c</sup>
1%-10%	Australia	OMI, monthly uncertainty at $2^\circ \times 2.5^\circ$ <sup>d</sup>
50%-100+%	Australia	OMI, in winter at higher latitudes <sup>d</sup>

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

b: Millet et al. (2008)

c: De Smedt et al. (2008, 2012)

d: This work

is set to 100%. This only impacts the uncertainty calculations in winter for the non-northern regions, where occasional  
810 highly negative absolute uncertainty was seen when  $\Omega$  approached 0.

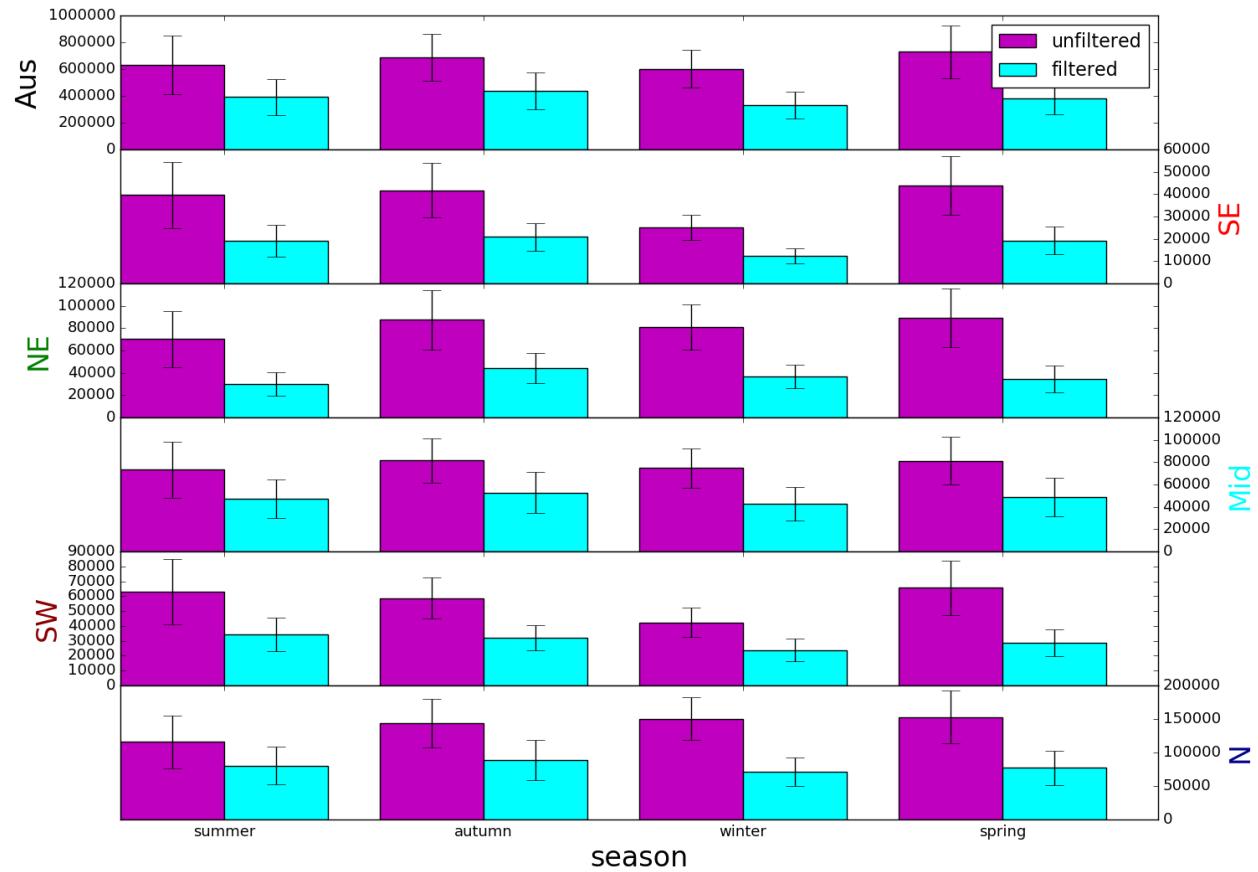
These sources of error can be reduced through spatial and temporal averaging, as they are assumed to be unbiased. Uncertainty  
815 is reduced by the square root of the number of pixels averaged over each  $2^\circ \times 2.5^\circ$  grid square for each day or month. For example, daily averaging reduces pixel uncertainty by a factor of 2-4. Figure 39 shows the pixel counts in each region before  
through May, June, and July. The lowest pixel counts occur in the southeast in winter, likely due to a mix of relatively high  
820 solar zenith angles and filtering of anthropogenic emissions. Northern regions have lower pixel counts in the summer, most  
likely due to increased cloud coverage which limits satellite measurement capabilities as discussed previously.

Figure 40 shows the relative uncertainty in monthly satellite columns for each sub-region and averaged over Australia.  
Uncertainty in winter at higher latitudes is greatly increased due to lower pixel counts, lower absolute column amounts, and  
825 higher fitting error.

Uncertainty in satellite HCHO ( $\Delta\Omega$ ) from literature and calculated here is listed in Table 11. De Smedt et al. (2012) found  
satellite HCHO uncertainty to be 30 – 40% for the GOME-2 instrument by combining slant column systematic and random  
errors. For mid latitude winters they found an excess of 60% uncertainty. OMI measurements will have similar uncertainty;  
however, the array of detectors provide more pixels which can be averaged to reduce this uncertainty.

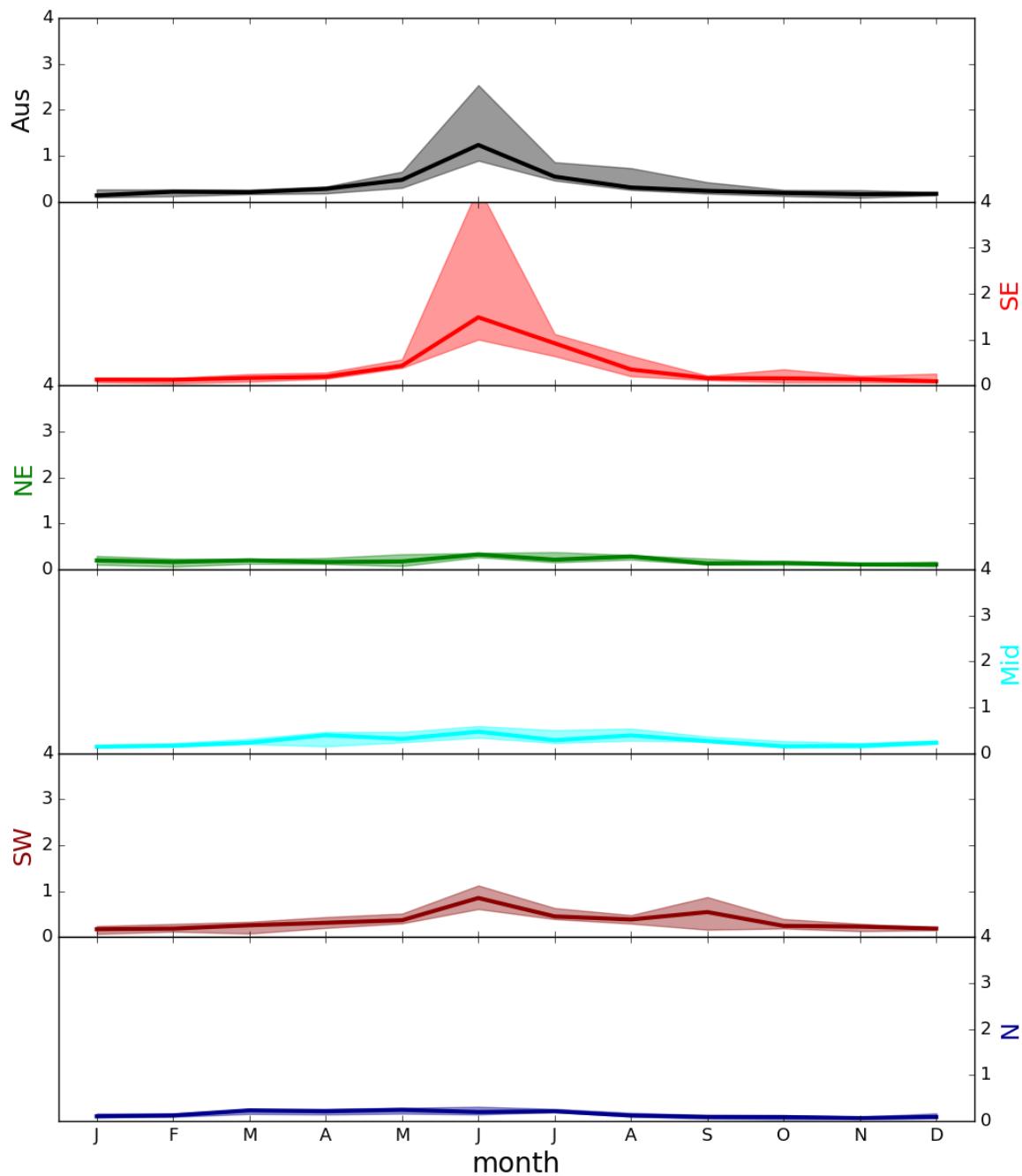
825 In order to calculate the bias or systematic error, an understanding of biases in the underlying terms is required, since  
there is little in the way of comparable measurements. OMI has been shown to underestimate observed HCHO by up to  $\sim 40\%$   
elsewhere in the world. For example OMI underestimated aircraft measurements by 37% in Guyana (Barkley et al., 2013). OMI  
underestimates range from 20-37% when compared against aircraft data over the southeast United States (Zhu et al., 2016).  
OMI validation against 7 sites (the most southern site being Reunion Island at  $20.9^\circ\text{S}$ ) using MAX-DOAS and FTIR retrievals

### Mean total-pixel-count per season per region



**Figure 39.** Mean and standard deviation (vertical error bars) of total pixel counts per region per season, before (magenta) and after (cyan) applying smearing, pyrogenic, and anthropogenic filters.

### relative error in monthly $\Omega$



**Figure 40.** Median and inter-quartile range for monthly binned relative uncertainty in satellite vertical columns.

830 showed up to 50% underestimates by satellite HCHO retrievals (De Smedt et al., 2015). The highest underestimate occurred during periods of high concentration (Vigouroux et al., 2009; De Smedt et al., 2015). Satellite HCHO may also suffer from ~13% overestimation when taking monthly averages due to only measuring on relatively cloud-free days (Surl et al., 2018). In this thesis, the a posteriori is linearly related to the satellite HCHO and any bias is directly transferred. The conclusion drawn here is that the isoprene emissions product in this work may be affected by satellite HCHO underestimation of up to 40%,  
835 and also by monthly HCHO overestimation of 13%, which gives a potential bias of 1/0.6 to 1/1.13. This may be complicated further if the satellite bias over Australia does not match the bias over the remote Pacific at corresponding latitudes. However, bias over Australia cannot be quantified due to insufficient measurements. GEOS-Chem biases would affect the recalculation of HCHO, but they cannot be quantified and so are not included in this thesis. Lacking suitable measurements to estimate satellite bias over Australia, the potential bias range compiled from the literature (−13% to +40%) is applied to the mean emissions in  
840 each month when calculating the uncertainty range shown in the results (see Figure 20).

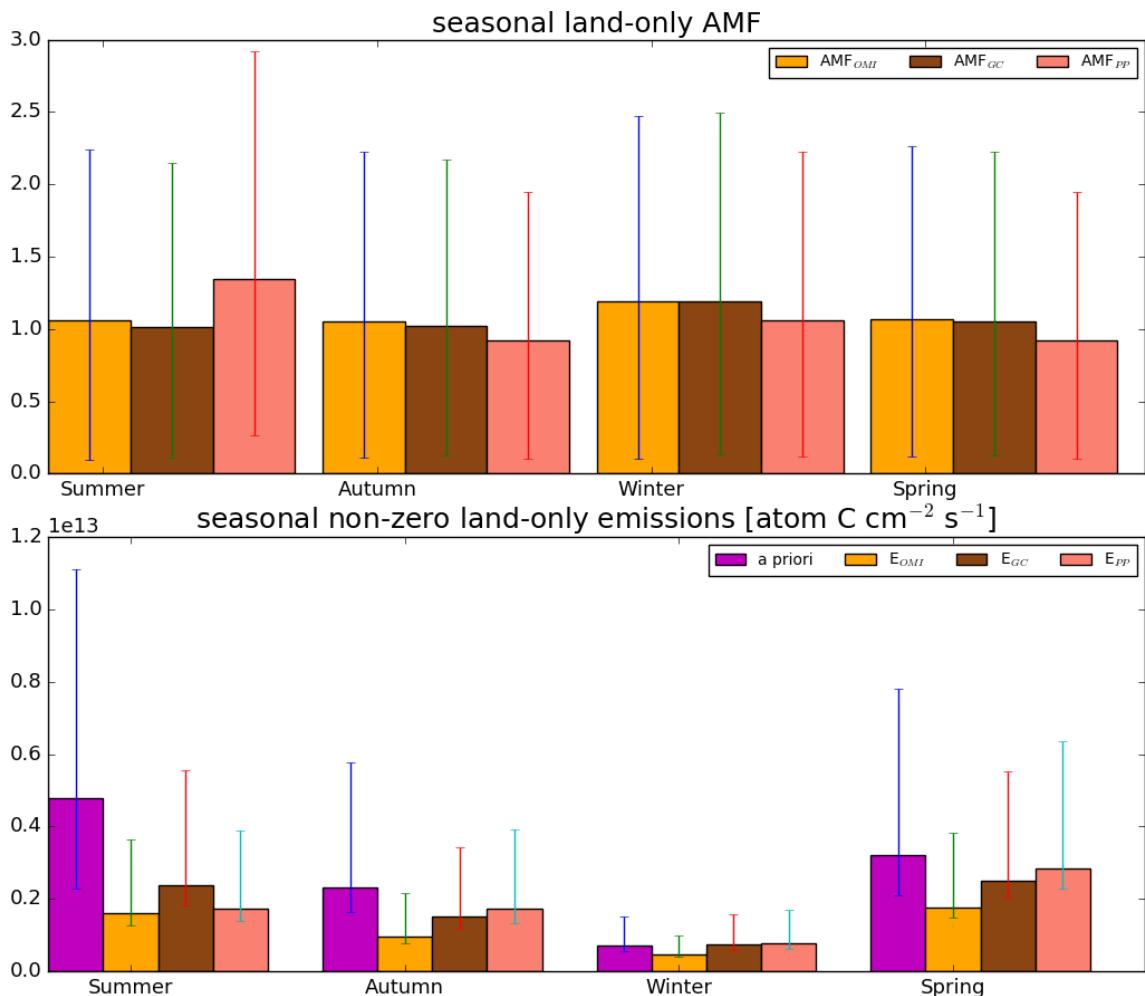
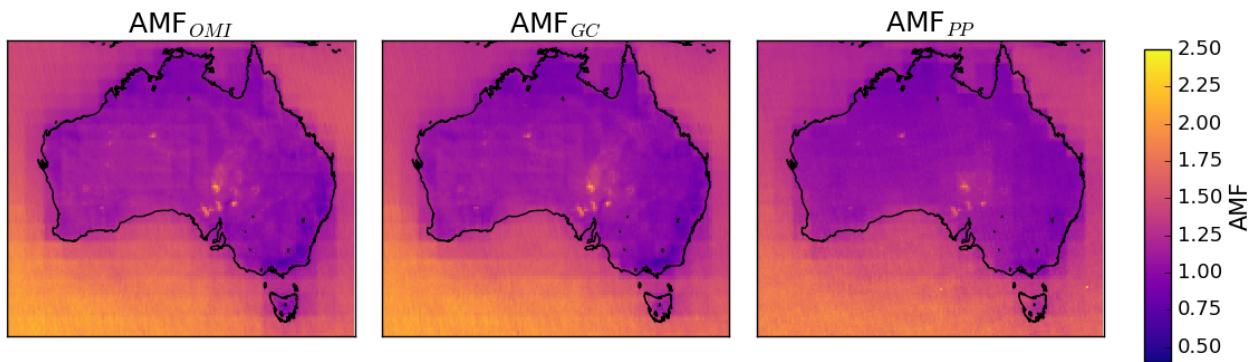
#### 4.4 Sensitivity to AMF recalculation

This section examines the sensitivity of the top-down isoprene emission estimates ( $E_{OMI}$ ) to the AMF recalculation method. The a posteriori emissions change linearly with recalculated vertical columns, which are calculated in three different ways: using the AMF provided in the OMHCHO product ( $AMF_{OMI}$ ), recalculating AMF shape factors but keeping the original  
845 scattering weights ( $AMF_{GC}$ ), or recalculating both shape factors and scattering weights ( $AMF_{PP}$ ), as described in Section ??.

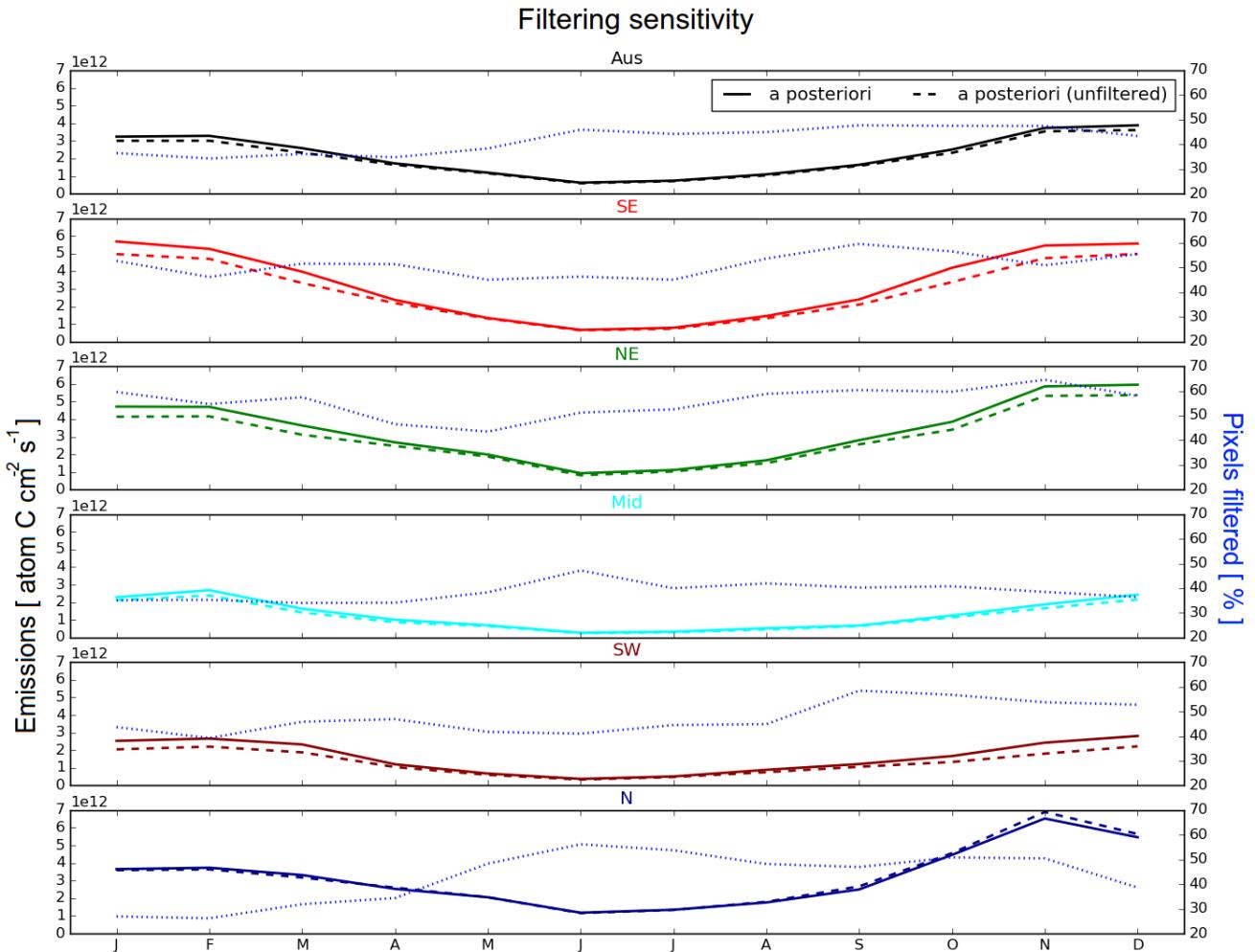
Figure 41 shows AMFs and emissions recalculated using each of these three methods over 2005. This figure only depicts grid squares with non-zero emissions, since in this work the regions of interest are those where emissions are relatively more substantial within Australia. The fully recalculated  $AMF_{PP}$  is higher in summer, but lower in other seasons compared against  
850 the other recalculations. The direct effect of a lower AMF is an increased vertical column, which should lead to higher emission estimates, and vice versa. This can be seen in all seasons except for summer, although even here the emissions estimates based on  $AMF_{PP}$  are higher than if they are calculated using the original  $AMF_{OMI}$ . Emission estimates vary widely over Australia, but the sensitivity to AMF recalculation technique is proportional and potentially non-linear. Changes in AMF of 5–30% cause changes in emission estimates of 5–50%. Further analysis is warranted and should be a focus of future work.

#### 855 4.5 Sensitivity to filtering

Figure 42 shows emissions estimates with and without filtering for smearing, anthropogenic, and pyrogenic influences. The overall effect of filtering is to slightly raise emissions in all non-summer months, with relatively little change to the mean in winter months. This is true in all regions except for northern Australia, which shows a slight decrease in spring. Eastern regions are most frequently filtered, with more than 50% of the available good pixels filtered throughout the year, while other regions  
860 vary between 20% and 60% with lower filtering rates in summer months. Seasonality in the fire filter may result matching seasonality of the Australian bushfire season, which spreads from the north to the south from July to March (TODO: Andrew Dowdy 2018 climatological variability of fire weather in australia).



**Figure 41.** Top row: averaged OMI Satellite AMF for 2005, from the MHCHO dataset (left,  $AMF_{OMI}$ ), recalculated using GEOS-Chem shape factors (middle,  $AMF_{GC}$ ), and recalculated using GEOS-Chem shape factors and scattering weights (right,  $AMF_{PP}$ ). Middle row: mean and inter-quartile range of each AMF over 2005 for each season. Bottom row: mean and inter-quartile range of non-zero a priori



**Figure 42.** Multi-year monthly mean values for a posteriori emission estimates calculated with (solid) and without (dashed) applying filters for anthropogenic, pyrogenic, and smearing influences. The portion of pixels within each region which are filtered is shown on the right axis with a blue dotted line.

## 5 Conclusions and implications

Very few ground-based measurements of BVOC concentrations are available in Australia. Emission models use largely un-  
865 verified extrapolations for emission factors and parameterisations in Australia, leading to overestimated isoprene emissions (Emmerson et al., 2016). This leads to uncertainty and error when modelling atmospheric ozone and other trace gases.

In this chapter an isoprene emissions estimate (a posteriori) is created and tested based on OMI satellite measurements of HCHO, a high-yield product of isoprene oxidation. The a posteriori emissions' effects on ozone and HCHO were tested by running GEOS-Chem with a seasonal (multi-year monthly averaged) gridded ( $2^\circ \times 2.5^\circ$ ) scaling factor calculated from the a  
870 posteriori applied to the a priori emissions. Uncertainty in the primary components of the top-down emissions calculation were calculated where possible and potential biases identified.

The a posteriori isoprene emission estimate showed that GEOS-Chem coupled with the MEGAN emissions model over-estimates emissions in summer by a factor of 2–5. Total yearly Australian emissions are reduced from  $39 \text{ Tg yr}^{-1}$  to  $21 \text{ Tg yr}^{-1}$  (decrease of  $\sim 46\%$ ) in the a posteriori. The overestimate is spatially and temporally diverse and leads to model biases in  
875 HCHO, and ozone. Running GEOS-Chem using scaled emissions based on OMI HCHO columns reduced the model HCHO overestimate (when compared to OMI HCHO) in summer by half, from  $\sim 100\%$  down to  $\sim 50\%$ , with most of the difference occurring outside the northern region. Model vertical column HCHO variance is somewhat ( $\sim 10\% - 50\%$ ) lower than that seen by the OMI satellite, and this difference increased after scaling isoprene emissions. Scaling GEOS-Chem emissions also lowered simulated surface ozone concentrations by  $\sim 5\%$ . A posteriori uncertainty was shown to be on the order of 50%  
880 (monthly, per grid square), with large satellite-based uncertainty in winter and a potential bias coming from satellite data of  $-13\%$  to  $+40\%$ . The primary uncertainty in the a posteriori emissions comes from the monthly modelled isoprene-to-HCHO yield ( $\sim 30\% - 50\%$  uncertainty), although at higher latitudes in winter the satellite uncertainty becomes restrictively high ( $> 100\%$ ).

The reason for the a priori overestimate remains unclear, and both global and Australian emissions estimates for isoprene  
885 range widely. The bias is in part due to the MEGAN emission model's parameterisation of how isoprene responds to parameters such as leaf area index, plant specific emission factors, and meteorological factors. In Australia, a mixture of poorly defined emission factors (e.g., Emmerson et al., 2016), unaccounted for soil moisture forcing (e.g., Sindelarova et al., 2014; Emmerson et al., 2019) and poorly understood forest responses to meteorological stresses likely drive uncertainty and model biases (Jiang et al., 2018; Emmerson et al., 2019). A potential improvement to the MEGAN soil parameterisation has been described (Jiang  
890 et al., 2018); however, a detailed map of soil properties is not currently available in Australia, and this would be required to apply the parameterisation here. An alternative parameterisation of soil moisture and drought effects on isoprene emissions was recently implemented in a different CTM by Emmerson et al. (2019) and applying this parameterisation in GEOS-Chem should be a priority for future work in Australia.

In the US, bias between OMI and in situ measurements is as high as 40%, but bias across Australia cannot be determined as  
895 there are not enough independent observations. This leads to a wide uncertainty range for estimated top-down emissions (see Figure 20) and limits potential top-down refinements to isoprene emissions. Ground-based and aircraft VOC,  $\text{NO}_x$ , HCHO,

and ozone measurements over large areas at relatively fine temporal resolution would help quantify the currently unknown satellite biases while additionally providing constraints for bottom-up models. In the northern region in particular, emissions are affected by monsoonal forcing, but increased cloud coverage during the monsoon limits satellite coverage. This makes  
900 characterisation of forest emissions and their response to sunlight, temperature, and moisture even more important in these areas. In addition to measurements, further analysis determining the sensitivity of modelled emissions to model resolution and changing soil moisture parameters (and parameterisations) would provide the foundation to improve GEOS-Chem and MEGAN modelled isoprene emissions along with oxidation products like HCHO and ozone.

In the future, other satellites (e.g., GOME-2, TROPOMI) could be used to improve emission estimates further, with differing  
905 overpass times potentially allowing a measure of diurnal emission patterns. Additionally, an adjoint version of GEOS-Chem over Australia could provide improved estimates of isoprene emission taking into account transport. An adjoint based inversion could also provide an evaluation of how resolution-limited the linear top-down emission estimates are over Australia. The emission estimate created in this chapter could also be refined to higher temporal resolution, with further analysis of uncertainty.  
910 Oversampling techniques could be applied near populated areas in order to improve the understanding of isoprene, HCHO, and ozone relationships over cities (e.g., Surl et al., 2018). Furthermore, linking the top-down isoprene emissions to the underlying emission factors for Australia and implementing the changes in GEOS-Chem would improve the understanding of the natural atmosphere over this relatively remote portion of the planet.

*Code and data availability.* TEXT

*Author contributions.* TEXT

915 *Competing interests.* TEXT

*Acknowledgements.* Model runs, data retrieval, analysis and error calculations and attribution in this Chapter are performed by the Author, with feedback and direction provided by Jenny Fisher and Clare Paton-Walsh. AMF and satellite corrections are performed by the Author, following discussions with Chris Miller, Paul Palmer, Luke Surl, and Gonzalez Abad.

## References

- 920 Abad, G. G., Vasilkov, A., Seftor, C., Liu, X., and Chance, K.: Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval, *Atmospheric Measurement Techniques*, 9, 2797–2812, <https://doi.org/10.5194/amt-9-2797-2016>, 2016.
- 925 Abbot, D. S.: Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space, *Geophysical Research Letters*, 30, 1999–2002, <https://doi.org/10.1029/2003GL017336>, <http://doi.wiley.com/10.1029/2003GL017336>, 2003.
- Ahn, C., Torres, O., and Bhartia, P. K.: Comparison of Ozone Monitoring Instrument UV Aerosol Products with Aqua/Moderate Resolution Imaging Spectroradiometer and Multiangle Imaging Spectroradiometer observations in 2006, *Journal of Geophysical Research Atmospheres*, 113, 1–13, <https://doi.org/10.1029/2007JD008832>, 2008.
- 930 Atkinson, R. and Arey, J.: Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review, *Atmospheric Environment*, 37, [https://doi.org/10.1016/S1352-2310\(03\)00391-1](https://doi.org/10.1016/S1352-2310(03)00391-1), 2003.
- Barkley, M. P., Smedt, I. D., Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D., Guenther, A., Paulot, F., Marais, E., and Mao, J.: Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns, *Journal of Geophysical Research Atmospheres*, 118, 6849–6868, <https://doi.org/10.1002/jgrd.50552>, <http://dx.doi.org/10.1002/jgrd.50552>, 2013.
- 935 Bauwens, M., Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Wiedinmyer, C., Kaiser, J. W., Sinelarova, K., and Guenther, A.: Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations, *Atmospheric Chemistry and Physics*, pp. 1–45, <https://doi.org/10.5194/acp-2016-221>, <http://www.atmos-chem-phys-discuss.net/acp-2016-221/>, 2016.
- Carlisle, W.: The coal seam gas rush, <http://www.abc.net.au/news/2011-11-24/coal-seam-gas-by-the-numbers-map/3664318>, 2012.
- 940 Chance, K., Palmer, P. I., Spurr, R. J. D., Martin, R. V., Kurosu, T. P., and Jacob, D. J.: Satellite observations of formaldehyde over North America from GOME, *Geophysical Research Letters*, 27, 3461–3464, <https://doi.org/10.1029/2000GL011857>, <http://dx.doi.org/10.1029/2000gl011857>, 2000.
- Curci, G., Palmer, P. I., Kurosu, T. P., Chance, K., and Visconti, G.: Estimating European volatile organic compound emissions using satellite observations of formaldehyde from the Ozone Monitoring Instrument, *Atmospheric Chemistry and Physics*, 10, 11 501–11 517, <https://doi.org/10.5194/acp-10-11501-2010>, <http://www.atmos-chem-phys.net/10/11501/2010/acp-10-11501-2010.pdf>, 2010.
- 945 De Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, *Atmospheric Chemistry and Physics*, 8, 4947–4963, <https://doi.org/10.5194/acp-8-4947-2008>, <http://www.atmos-chem-phys.net/8/4947/2008/>, 2008.
- De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J. F., Lerot, C., Theys, N., Valks, P., Hao, N., and Van Der A, R.: Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, *Atmospheric Measurement Techniques*, 5, 2933–2949, <https://doi.org/10.5194/amt-5-2933-2012>, <http://www.atmos-meas-tech.net/5/2933/2012/amt-5-2933-2012.pdf>, 2012.
- 950 De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J.-F., and Van Roozendael, M.: Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, *Atmospheric Chemistry and Physics*, 15, 12 519–12 545, <https://doi.org/10.5194/acp-15-12519-2015>, <http://www.atmos-chem-phys.net/15/12519/2015/acp-15-12519-2015.pdf>, 2015.

- Dufour, G., Wittrock, F., Camredon, M., Beekmann, M., Richter, A., and Burrows, J.: SCIAMACHY formaldehyde observations: constraint for isoprene emissions over Europe?, *Atmospheric Chemistry and Physics*, 8, 19 273–19 312, <https://doi.org/10.5194/acpd-8-19273-2008>, 2008.
- Emmerson, K. M., Galbally, I. E., Guenther, A. B., Paton-Walsh, C., Guerette, E.-A. A., Cope, M. E., Keywood, M. D., Lawson, S. J., Molloy, S. B., Dunne, E., Thatcher, M., Karl, T., and Maleknia, S. D.: Current estimates of biogenic emissions from eucalypts uncertain for southeast Australia, *Atmospheric Chemistry and Physics*, 16, 6997–7011, <https://doi.org/10.5194/acp-16-6997-2016>, <http://www.atmos-chem-phys.net/16/6997/2016/>, 2016.
- Emmerson, K. M., Palmer, P. I., Thatcher, M., Haverd, V., and Guenther, A. B.: Sensitivity of isoprene emissions to drought over south-eastern Australia : Integrating models and satellite observations of soil moisture, *Atmospheric Environment*, 209, 112–124, <https://doi.org/10.1016/j.atmosenv.2019.04.038>, <https://doi.org/10.1016/j.atmosenv.2019.04.038>, 2019.
- Fan, J. and Zhang, R.: Atmospheric oxidation mechanism of isoprene, *Environmental Chemistry*, 1, 140–149, <https://doi.org/10.1071/EN04045>, <http://dx.doi.org/10.1071/en04045>, 2004.
- Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K., and Fried, A.: The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system, *Atmospheric Chemistry and Physics*, 12, 6699–6721, <https://doi.org/10.5194/acp-12-6699-2012>, <http://www.atmos-chem-phys.net/12/6699/2012/acp-12-6699-2012.pdf>, 2012.
- Fu, T.-m., Jacob, D. J., Palmer, P. I., Chance, K., Wang, Y. X., Barletta, B., Blake, D. R., Stanton, J. C., and Pilling, M. J.: Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone, *Journal of Geophysical Research*, 112, D06 312, <https://doi.org/10.1029/2006JD007853>, <http://doi.wiley.com/10.1029/2006JD007853>, 2007.
- Gonzalez Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., Suleiman, R., Gonz??lez Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., and Suleiman, R.: Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval, *Atmospheric Measurement Techniques*, 8, 19–32, <https://doi.org/10.5194/amt-8-19-2015>, 2015.
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmospheric Chemistry and Physics*, 6, 3181–3210, <https://doi.org/10.5194/acp-6-3181-2006>, <http://dx.doi.org/10.5194/acp-6-3181-2006>, 2006.
- Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions, *Geoscientific Model Development*, 5, 1471–1492, <https://doi.org/10.5194/gmd-5-1471-2012>, 2012.
- Jiang, X., Guenther, A., Potosnak, M., Geron, C., Seco, R., Karl, T., Kim, S., Gu, L., and Pallardy, S.: Isoprene emission response to drought and the impact on global atmospheric chemistry, *Atmospheric Environment*, 183, 69–83, <https://doi.org/10.1016/J.ATMOSENV.2018.01.026>, <https://www.sciencedirect.com/science/article/pii/S1352231018300402>, 2018.
- Kaiser, J., Fried, A., St. Clair, J. M., Zhang, L., González Abad, G., Travis, K. R., Jacob, D. J., Wisthaler, A., Crounse, J. D., Fisher, J. A., Zhang, X., and Zhu, L.: High-resolution inversion of OMI formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: application to the southeast US, *Atmospheric Chemistry and Physics*, 18, 5483–5497, <https://doi.org/10.5194/acp-18-5483-2018>, <http://www.atmos-chem-phys.net/18/5483/2018/acp-18-5483-2018.pdf>, 2018.

- Kefauver, S. C., Filella, I., and Peñuelas, J.: Remote sensing of atmospheric biogenic volatile organic compounds (BVOCs) via satellite-based formaldehyde vertical column assessments, International Journal of Remote Sensing, <http://www.tandfonline.com/doi/abs/10.1080/01431161.2014.968690#.VkqEubNM6lM>, 2014.
- 995 Keywood, M., Emmerson, K., and Hibberd, M.: Australia state of the environment 2016, <https://doi.org/10.4226/94/58b65c70bc372>, <https://soe.environment.gov.au/theme/atmosphere>, 2016.
- Lawson, S. J., Selleck, P. W., Galbally, I. E., Keywood, M. D., Harvey, M. J., Lerot, C., Helmig, D., and Ristovski, Z.: Seasonal in situ observations of glyoxal and methylglyoxal over the temperate oceans of the Southern Hemisphere, Atmospheric Chemistry and Physics, 15, 223–240, <https://doi.org/10.5194/acp-15-223-2015>, <http://www.atmos-chem-phys.net/15/223/2015/>, 2015.
- 1000 Liu, Y., Brito, J., Dorris, M. R., Rivera-Rios, J. C., Seco, R., Bates, K. H., Artaxo, P., Duvoisin, S., Keutsch, F. N., Kim, S., Goldstein, A. H., Guenther, A. B., Manzi, A. O., Souza, R. A. F., Springston, S. R., Watson, T. B., McKinney, K. A., and Martin, S. T.: Isoprene photochemistry over the Amazon rainforest, Proceedings of the National Academy of Sciences, 113, 6125–6130, <https://doi.org/10.1073/pnas.1524136113>, <http://www.pnas.org/content/113/22/6125.abstract>, 2016.
- Liu, Z., Nguyen, V. S., Harvey, J., Müller, J. F., and Peeters, J.: Theoretically derived mechanisms of HPALD photolysis in isoprene oxidation, 1005 Physical Chemistry Chemical Physics, 19, 9096–9106, <https://doi.org/10.1039/c7cp00288b>, 2017.
- Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet, D. B., Barkley, M. P., Paulot, F., and Mao, J.: Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns, Atmospheric Chemistry and Physics, 12, 7475–7520, <https://doi.org/10.5194/acp-12-6219-2012>, <http://dx.doi.org/10.5194/acp-12-6219-2012>, 2012.
- 1010 Marais, E. A., Jacob, D. J., Guenther, A., Chance, K., Kurosu, T. P., Murphy, J. G., Reeves, C. E., and Pye, H. O. T.: Improved model of isoprene emissions in Africa using Ozone Monitoring Instrument (OMI) satellite observations of formaldehyde: implications for oxidants and particulate matter, Atmospheric Chemistry and Physics, 14, 7693–7703, <https://doi.org/10.5194/acp-14-7693-2014>, <http://dx.doi.org/10.5194/acp-14-7693-2014>, 2014.
- Martin, R. V.: Global and regional decreases in tropospheric oxidants from photochemical effects of aerosols, Journal of Geophysical Research, 108, 4097, <https://doi.org/10.1029/2002JD002622>, <http://doi.wiley.com/10.1029/2002JD002622>, 2003a.
- 1015 Martin, R. V.: Global inventory of nitrogen oxide emissions constrained by space-based observations of NO<sub>2</sub> columns, Journal of Geophysical Research, 108, 4537, <https://doi.org/10.1029/2003JD003453>, <http://doi.wiley.com/10.1029/2003JD003453>, 2003b.
- Messina, P., Lathiere??re, J., Sindelarova, K., Vuichard, N., Granier, C., Ghattas, J., Cozic, A., and Hauglustaine, D. A.: Global biogenic volatile organic compound emissions in the ORCHIDEE and MEGAN models and sensitivity to key parameters, Atmospheric Chemistry and Physics, 16, 14 169–14 202, <https://doi.org/10.5194/acp-16-14169-2016>, <http://www.atmos-chem-phys.net/16/14169/2016/acp-16-14169-2016.pdf>, 2016.
- Millet, D. B., Jacob, D. J., Turquety, S., Hudman, R. C., Wu, S., Fried, A., Walega, J., Heikes, B. G., Blake, D. R., Singh, H. B., Anderson, B. E., and Clarke, A. D.: Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission, J. Geophys. Res., 111, <https://doi.org/10.1029/2005jd006853>, TODO, 2006.
- 1020 Millet, D. B., Jacob, D. J., Boersma, K. F., Fu, T. M., Kurosu, T. P., Chance, K., Heald, C. L., and Guenther, A.: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, Journal of Geophysical Research Atmospheres, 113, 1–18, <https://doi.org/10.1029/2007JD008950>, 2008.
- Millet, D. B., Baasandorj, M., Hu, L., Mitroo, D., Turner, J., and Williams, B. J.: Nighttime Chemistry and Morning Isoprene Can Drive Urban Ozone Downwind of a Major Deciduous Forest, Environmental Science and Technology, 50, 4335–4342, <https://doi.org/10.1021/acs.est.5b06367>, <https://pubs.acs.org/doi/pdfplus/10.1021/acs.est.5b06367>, 2016.

- 1030 Niinemets, U., Tenhunen, J. D., Harley, P. C., and Steinbrecher, R.: A model of isoprene emission based on energetic requirements for isoprene synthesis and leaf photosynthetic properties for Liquidambar and Quercus, *Plant, Cell and Environment*, 22, 1319–1335, <https://doi.org/10.1046/j.1365-3040.1999.00505.x>, 1999.
- Palmer, P. I., Jacob, D. J., Chance, K., Martin, V., and Kurosu, P.: Air mass factor formulation for spectroscopic measurements from satellites' Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, *Journal of Geophysical Research*, 106, 2001.
- 1035 Palmer, P. I., Jacob, D. J., Fiore, A. M., Martin, R. V., Chance, K. V., and Kurosu, T. P.: Mapping isoprene emissions over North America using formaldehyde column observations from space, *J. Geophys. Res.*, 108, <https://doi.org/10.1029/2002jd002153>, <http://dx.doi.org/10.1029/2002jd002153>, 2003.
- Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., Guenther, A., Wiedinmyer, C., Stanton, J. C., Pilling, M. J., Pressley, S. N., Lamb, B., and Sumner, A. L.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, *J. Geophys. Res.*, 111, D12315, <https://doi.org/10.1029/2005JD006689>, <http://dx.doi.org/10.1029/2005JD006689>, 2006.
- 1040 Shao, Y., Leys, J. F., McTainsh, G. H., and Tews, K.: Numerical simulation of the October 2002 dust event in Australia, *J. Geophys. Res.*, 112, <https://doi.org/10.1029/2006jd007767>, <http://dx.doi.org/10.1029/2006jd007767>, 2007.
- Shim, C., Wang, Y., Choi, Y., Palmer, P. I., Abbot, D. S., and Chance, K.: Constraining global isoprene emissions with Global Ozone 1045 Monitoring Experiment (GOME) formaldehyde column measurements, *Journal of Geophysical Research Atmospheres*, 110, 1–14, <https://doi.org/10.1029/2004JD005629>, 2005.
- Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J. F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, *Atmospheric Chemistry and Physics*, 14, 9317–9341, <https://doi.org/10.5194/acp-14-9317-2014>, 2014.
- 1050 Stavrakou, T., Müller, J.-F., Smedt, I. D., Roozendael, M. V., van der Werf, G. R., Giglio, L., and Guenther, A.: Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns, *Atmospheric Chemistry and Physics*, 9, 1037–1060, <https://doi.org/10.5194/acp-9-1037-2009>, <http://dx.doi.org/10.5194/acp-9-1037-2009>, 2009.
- Stavrakou, T., Müller, J. F., Bauwens, M., De Smedt, I., Van Roozendael, M., Guenther, A., Wild, M., and Xia, X.: Isoprene emissions over 1055 Asia 1979–2012: Impact of climate and land-use changes, *Atmospheric Chemistry and Physics*, 14, <https://doi.org/10.5194/acp-14-4587-2014>, 2014, <https://www.atmos-chem-phys.net/14/4587/2014/acp-14-4587-2014.pdf>, 2014.
- Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Van Roozendael, M., De Mazière, M., Vigouroux, C., Hendrick, F., George, M., Clerbaux, C., Coheur, P.-F., and Guenther, A.: How consistent are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI?, *Atmospheric Chemistry and Physics*, 15, 11 861–11 884, <https://doi.org/10.5194/acp-15-11861-2015>, <http://www.atmos-chem-phys.net/15/11861/2015/acp-15-11861-2015.html>, 2015.
- 1060 Streets, D. G., Canty, T., Carmichael, G. R., de Foy, B., Dickerson, R. R., Duncan, B. N., Edwards, D. P., Haynes, J. A., Henze, D. K., Houyoux, M. R., Jacob, D. J., Krotkov, N. A., Lamsal, L. N., Liu, Y., Lu, Z., Martin, R. V., Pfister, G. G., Pinder, R. W., Salawitch, R. J., and Wecht, K. J.: Emissions estimation from satellite retrievals: A review of current capability, *Atmospheric Environment*, 77, 1011–1042, <https://doi.org/10.1016/j.atmosenv.2013.05.051>, <http://dx.doi.org/10.1016/j.atmosenv.2013.05.051>, 2013.
- Surl, L., Palmer, P. I., and Abad, G. G.: Which processes drive observed variations of HCHO columns over India?, *Atmospheric Chemistry 1065 and Physics*, 18, 4549–4566, <https://doi.org/10.5194/acp-18-4549-2018>, <https://doi.org/10.5194/acp-18-4549-2018>, 2018.

- Thomas, W., Hegels, E., Slijkhuis, S., Spurr, R., and Chance, K.: Detection of biomass burning combustion products in Southeast Asia from backscatter data taken by the GOME spectrometer, *Geophysical Research Letters*, 25, 1317–1320, <https://doi.org/10.1029/98GL01087>, <http://onlinelibrary.wiley.com/doi/10.1029/98GL01087/epdf>, 1998.
- Vigouroux, C., Hendrick, F., Stavrakou, T., Dils, B., De Smedt, I., Hermans, C., Merlaud, A., Scolas, F., Senten, C., Vanhaelewijn, G.,  
1070 Fally, S., Carleer, M., Metzger, J.-M., Müller, J.-F., Van Roozendael, M., and De Mazière, M.: Ground-based FTIR and MAX-DOAS observations of formaldehyde at Réunion Island and comparisons with satellite and model data, *Atmospheric Chemistry and Physics Discussions*, 9, 15 891–15 957, <https://doi.org/10.5194/acpd-9-15891-2009>, 2009.
- Winters, A. J., Adams, M. A., Bleby, T. M., Rennenberg, H., Steigner, D., Steinbrecher, R., and Kreuzwieser, J.: Emissions of isoprene, monoterpane and short-chained carbonyl compounds from Eucalyptus spp. in southern Australia, *Atmospheric Environment*, 43, 3035–  
1075 3043, <https://doi.org/10.1016/j.atmosenv.2009.03.026>, 2009.
- Wolfe, G. M., Kaiser, J., Hanisco, T. F., Keutsch, F. N., De Gouw, J. A., Gilman, J. B., Graus, M., Hatch, C. D., Holloway, J., Horowitz,  
1080 L. W., Lee, B. H., Lerner, B. M., Lopez-Hilfiker, F., Mao, J., Marvin, M. R., Peischl, J., Pollack, I. B., Roberts, J. M., Ryerson, T. B., Thornton, J. A., Veres, P. R., and Warneke, C.: Formaldehyde production from isoprene oxidation across NO<sub>x</sub> regimes, *Atmospheric Chemistry and Physics*, 16, 2597–2610, <https://doi.org/10.5194/acp-16-2597-2016>, <https://www.atmos-chem-phys.net/16/2597/2016/acp-16-2597-2016.pdf>, 2016.
- Zheng, Y., Unger, N., Barkley, M. P., and Yue, X.: Relationships between photosynthesis and formaldehyde as a probe of isoprene emission, *Atmospheric Chemistry and Physics*, 15, 8559–8576, <https://doi.org/10.5194/acp-15-8559-2015>, 2015.
- Zhu, L., Jacob, D. J., Mickley, L. J., Marais, E. A., Cohan, D. S., Yoshida, Y., Duncan, B. N., Gonzalez Abad, G., and Chance, K. V.: Anthropogenic emissions of highly reactive volatile organic compounds in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns, *Environmental Research Letters*, 9, 114 004, <https://doi.org/10.1088/1748-9326/9/11/114004>, <http://stacks.iop.org/1748-9326/9/i=11/a=114004?key=crossref.3d2869ee02fd4f0792f831ac8cbe117>, 2014.
- Zhu, L., Jacob, D. J., Kim, P. S., Fisher, J. A., Yu, K., Travis, K. R., Mickley, L. J., Yantosca, R. M., Sulprizio, M. P., De Smedt, I., Gonzalez  
1085 Abad, G., Chance, K., Li, C., Ferrare, R., Fried, A., Hair, J. W., Hanisco, T. F., Richter, D., Scarino, A. J., Walega, J., Weibring, P., and Wolfe, G. M.: 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, *Atmospheric Chemistry and Physics*, 0, 1–24, <https://doi.org/10.5194/acp-2016-162>, <http://www.atmos-chem-phys.net/16/13477/2016/acp-16-13477-2016.pdf>, 2016.