

each method. Figure 2.28 shows vertical columns of HCHO for: column 1) the original satellite swaths, column 2) recalculated without changing the provided scattering weights, and column 3) fully recalculated vertical columns. Each grid square (at 0.25 by 0.3125°lat lon resolution) has been created by binning the recalculated satellite pixels within the month. The average pixels per land square is inset as text, changing due to how the fire filter is applied. Each row has a stricter fire filter applied from top to bottom, with no fire filter on the first row up to filtering pixels from squares with fires up to 8 days prior. This figure looks at March 2005 with biomass burning filtered differently in each row. Active fires over the last 0, 1, 2, 4, and 8 days are filtered as the row number increases.

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

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

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

2.7 Filtering Data

In order to examine only biogenic processes, pyrogenic and anthropogenic influences need to be removed from modelled and measured data. Biomass burning can be a large local or transported (via smoke plumes) source of HCHO, CHOCHO, glyoxal, and other compounds which influence levels of both HCHO and isoprene. Anthropogenic emissions from power generation, transport, and agriculture can influence these levels as well. Where possible these influences need to be removed so that calculations of purely biogenic emissions are not biased. In GEOS-Chem we can simply turn off pyrogenic and anthropogenic emissions, however in satellite datasets we need to mask potentially affected pixels.

Influence from biomass burning can be removed through measurements of acetonitrile and CO (eg. Wolfe et al. 2016; Miller et al. 2016), or else removal of scenes coincident with satellite detected fire counts and aerosol absorption optical depth as done in Marais et al. (2014). Marais et al. (2012) remove pixels colocated with non zero fire counts in any of the prior eight days, within grid squares with $1 \times 1^\circ$ resolution. Barkley et al. (2013) use fires from the preceding and concurrent day, within local or adjacent grid squares, with grid resolution of $0.25 \times 0.3125^\circ$. Wolfe et al. (2016) disregard HCHO measurements when acetonitrile > 210 pptv and CO > 300 ppbv, while acetonitrile > 200 pptv is used to determine fire influence in Miller et al. (2016). TODO: look at yearly correlation, compare to exponential curve and look for fire outliers As seen in TODO: citation, HCHO concentrations scale exponentially with temperature. This allows another method for detecting the influence of non-biogenic HCHO emission/creation by looking for outliers above the curve at low temperature. Zhu et al.

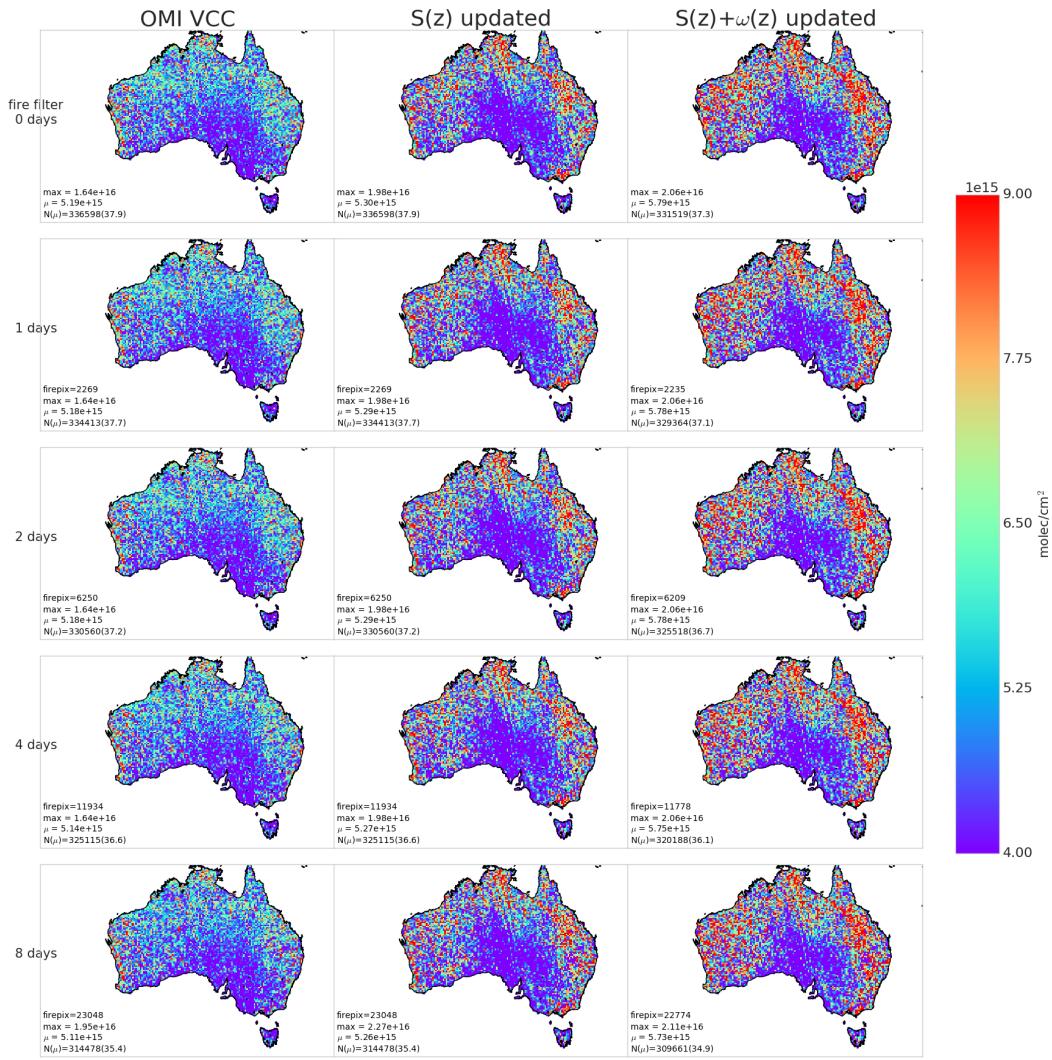


FIGURE 2.28: Column 1: Reference sector corrected HCHO vertical columns Ω from OMHCCHOv003. Column 2: Ω with recalculated apriori shape factors using GEOS-Chem v10.01. Column 3: Ω with recalculated apriori shape factors and scattering weights using GEOS-Chem v10.01 and LIDORT. Row 1-5: increasing number of prior days which have active fires included when masking fire influence.

TABLE 2.3: How many satellite pixels are filtered by pyrogenic and anthropogenic masking. Left to right the columns display year, how many land pixels are read over Australia, how many of these pixels are removed by the pyrogenic filter, how many are removed by the anthropogenic filter, and how many are removed in total. In parenthesis are the portion of pixels filtered.

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

(2013) has a similar analysis over south-eastern USA showing an exponential correlation of $HCHO = \exp(0.15 \times T - 9.07)$.

I use satellite data to account for anthropogenic and pyrogenic influences on the OMHCHO satellite HCHO columns. MODIS fire counts are used in conjunction with smoke AAOD enhancements (from OMI) to remove data points which may be affected by fires or fire smoke plumes. OMI NO₂ measurements are used to mask potential anthropogenic influence. These masks negatively affect uncertainty, as fewer measurements are available to be averaged. This section describes the creation and effects of filters used on satellite data.

A quick summary of how much data is filtered over Australian land squares is provided in table 2.3, and an quick check of how many pixels are filtered in January 2006 can be seen in Figure 2.29. The anthropogenic filter completely removes grid squares over Sydney and Melbourne, and high removal rates over Brisbane. Other major cities in Australia either do not emit enough NO₂ or are too spread out and do not breach the threshold to be filtered as anthropogenic.

2.7.1 Fire and smoke

The method used in this thesis follows that of Marais et al. (2012), and Barkley et al. (2013), with active fires filtered using fire counts, and smoke filtered out using smoke aerosol absorption optical depth (AAOD). We use the MODIS fire counts, detected from space using the combined product from Terra and Aqua (Terra at 10:30, 22:30 LT; Aqua at 13:30, 01:30 LT). Smoke plumes are filtered using smoke AAOD from product OMAERUVd, with the threshold determined through analysis of Australian AAOD distributions.

OMHCHO total column HCHO Ω is processed into a 0.25x0.3125°horizontal daily grid. Pyrogenic filters are created as follows. They are created or interpolated to the same the same horizontal resolution as Ω to simplify application. The following steps are performed in order to create the pyrogenic influence mask:

1. MOD14A1 daily gridded Aqua/Terra combined fire counts ($1 \times 1 \text{ km}^2$) are read, and binned into 0.25x0.3125°bins (matching the resolution of binned Ω).

Anthro and Fire filters applied on 20060101-20060131

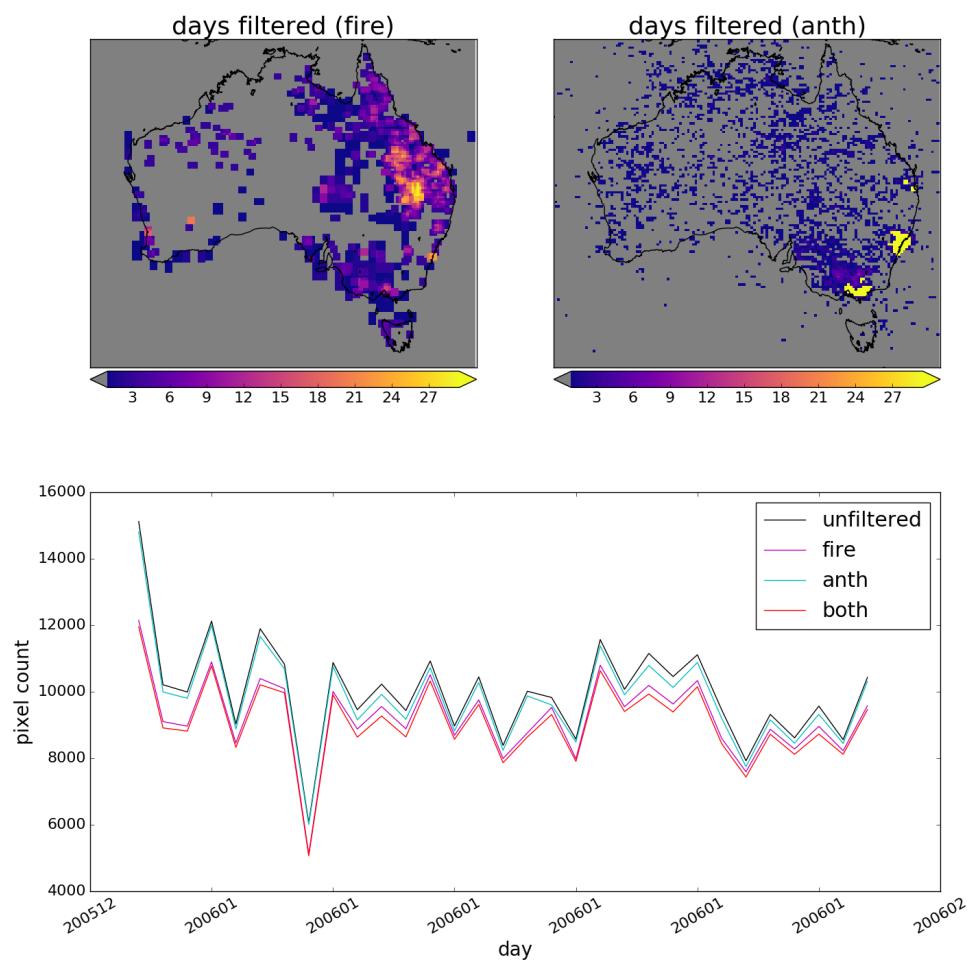


FIGURE 2.29: Top row shows grid squares filtered out by pyrogenic(left) and anthropogenic(right) influence masks during January 2006. Along the bottom is the time series of total pixels over Australian land squares with and without filtering the data.

2. A rolling mask is formed which removes Ω if one or more fires are detected in a grid square, or in the adjacent grid square, up to 2 days previously. This includes the 'current' day, making 3 days of fires in total being filtered out on each day.
3. AAOD at 500 nm is mapped from OMAERUVd ($1 \times 1^\circ$ resolution) onto the $0.25 \times 0.3125^\circ$ resolution grid.
4. An AAOD threshold of 0.03 is determined through visual analysis of AAOD distributions over several days, including days with and without influence from active fires, dust, and transported smoke plumes (see figure 2.30).
5. Grid squares with AAOD over this threshold are considered potentially affected by transported fire smoke.

Determining the AAOD due to smoke can be difficult since both smoke and dust absorb UV radiation (Ahn2008; Marais et al. 2012). AAOD is less sensitive to cloud contamination than AOD, and I use AAOD from the daily gridded level 3 satellite product OMAERUVd (Ahn2008) described in section 2.3.1.3 to provide a filter for smoke plumes. Although removing gridsquares with dust reduces how much data is available to analyse, it's considered a minor problem as dust in Australia is highly episodic and should not affect more than a few days per year, especially over regions with high tree coverage (Shao et al. 2007).

Filtering fire smoke using AAOD is done by removing OMHCHO gridsquares where the AAOD is above a 0.03, after the AAOD is mapped from $1 \times 1^\circ$ to the same $0.25 \times 0.3125^\circ$ resolution as our OMHCHO gridded product. The threshold is determined through analysing AAOD over Australia in 4 scenarios: normal conditions, active local fires, during influence from transported fire smoke, and large scale dust storms. Figure 2.30 shows AAOD (columns 1 and 2), with AAOD distribution in column 3, along with satellite imagery on the same day in column 4 (from <https://worldview.earthdata.nasa.gov/>). The scenarios listed are shown from row 1 to 4, and AAOD = 0.03 is demarcated by a horizontal line in the density plots in column 3.

Figure ?? shows what portion of pixels are filtered out by the pyrogenic filter. The top panel shows the spatial distribution of fire masks, with most pixels removed along the northern and eastern coastlines. A large portion of the filtered areas appear to correspond with forested areas (see figure ??), which suggests that forest fires are being masked properly. Central Australia is largely unmasked, which could be due to a lack of sufficient vegetation to create a large enough fire to be seen by satellite. The other potential cause of pyrogenic filtering is the proliferation of petrol or gas wells (see figure 2.32 and figure 2.33). The filtering shown here is for 2005, when 388 gas wells existed in Queensland, however more than 2000 wells (cumulative) were approved by 2013, so this may cause more filtering over the course of this thesis' timeline (Carlisle2012). To check this the filtering portion for 2012 is also plotted in Figure TODO: plot map of filtered squares for 2012 when it's been run by NCI. One clear hotspot is located over port Kembla (south of Sydney), most likely due to the flame which burns over the blast furnace stack throughout the year. Another hotspot can be seen in Western Australia over Kalgoorlie, where a large open cut gold mine "super pit" is always open and blasting daily. In Western Queensland over Mount Isa there is again a mining related hotspot. A large area in southern Queensland/northern NSW is also heavily filtered, potentially due to gas flaring in the Surat Basin, which has thousands of petrol and gas wells.

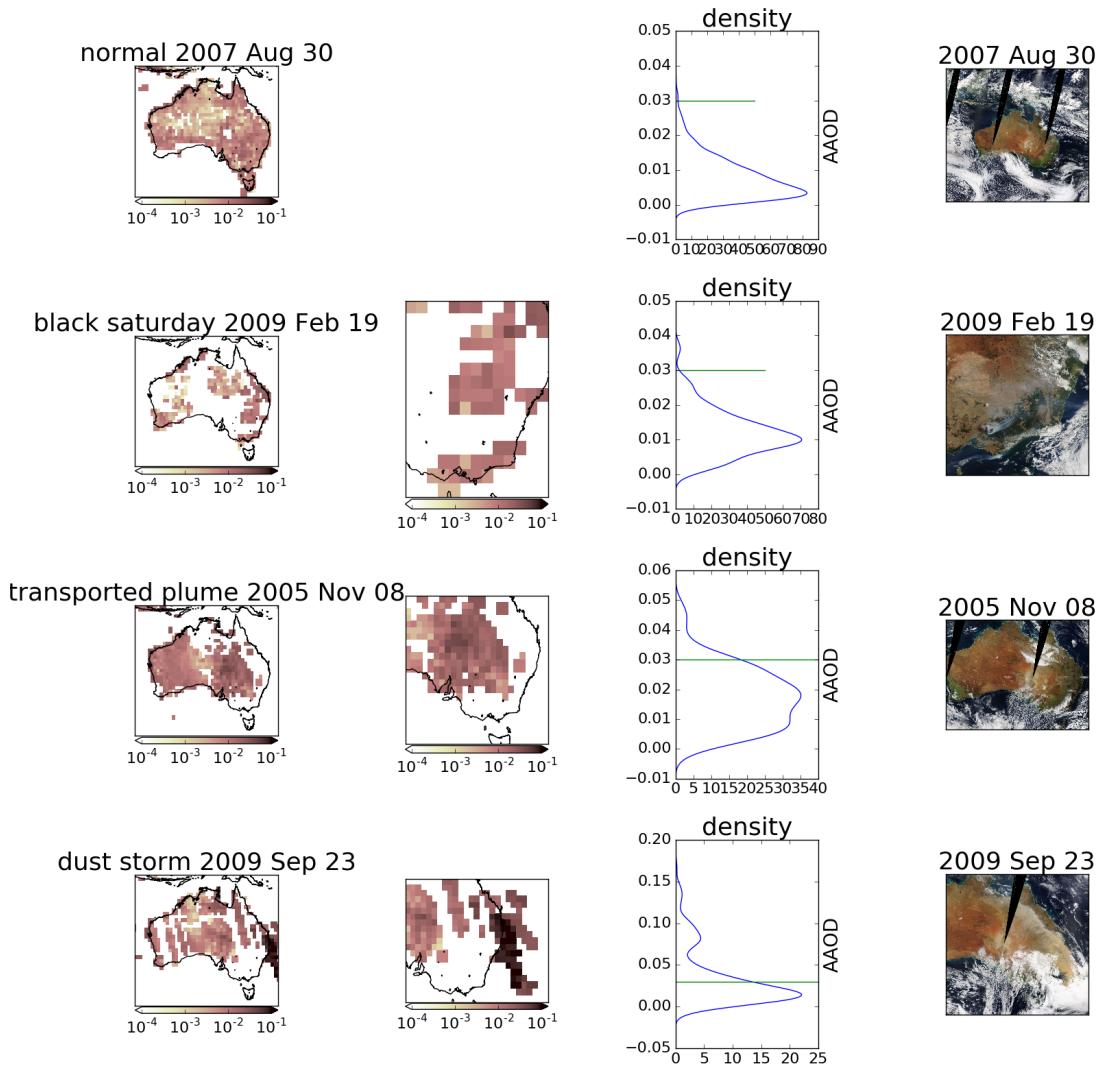


FIGURE 2.30: AAOD from OMAERUVd (columns 1, 2, 3) over Australia for four different scenarios (rows 1-4). Scenes from the same day are taken from the EOS Worldview website <https://worldview.earthdata.nasa.gov/>.

Pyrogenic filter: 20050101-20060101

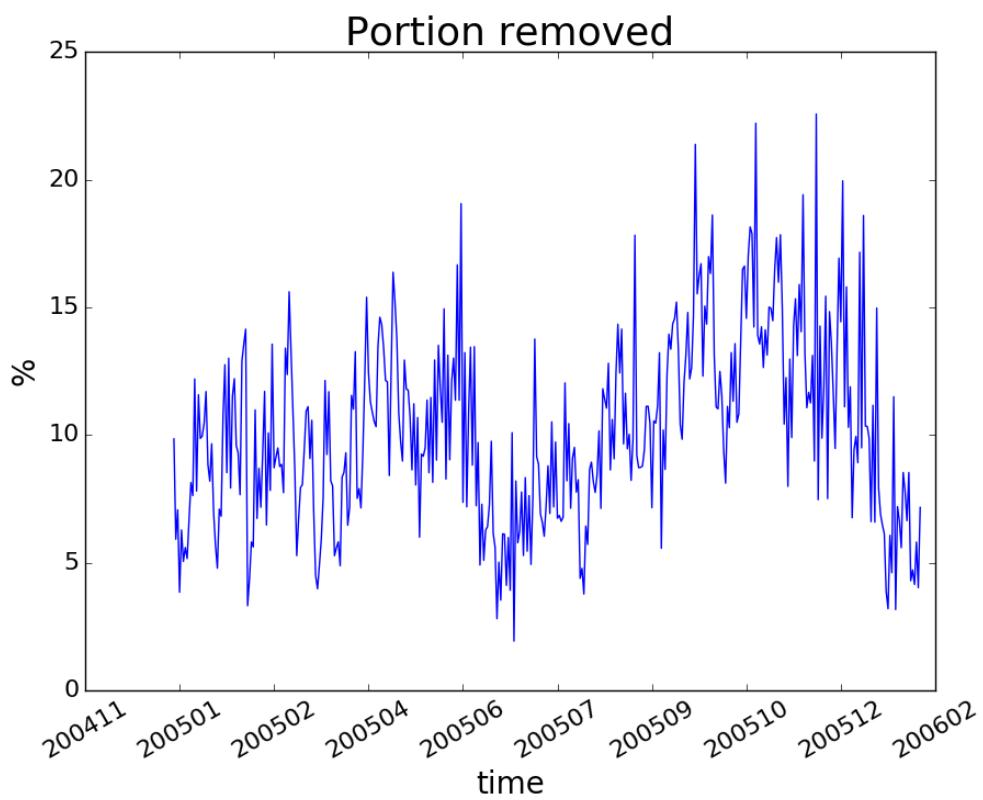
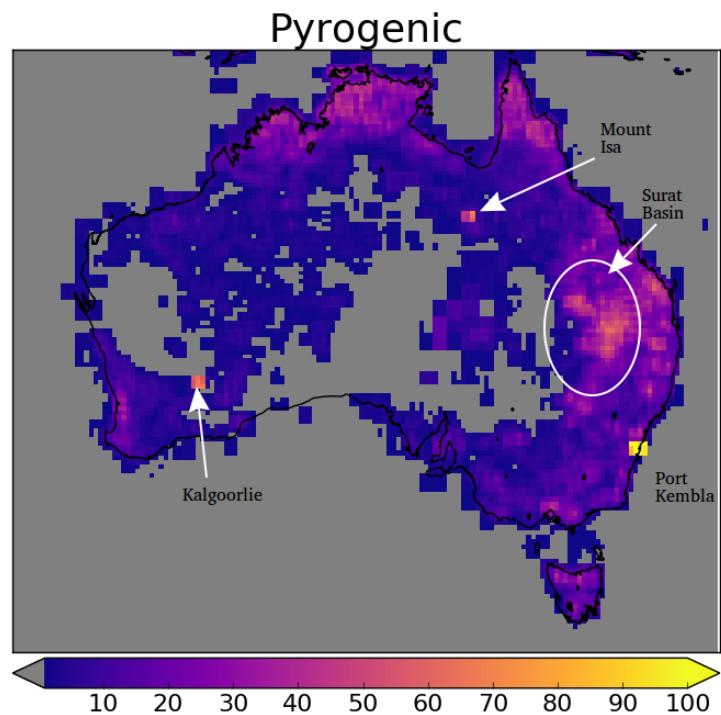


FIGURE 2.31: Top: Portion of 2005 filtered out by fire and smoke masks.
Bottom: portion filtered out each day from land squares in Australia.

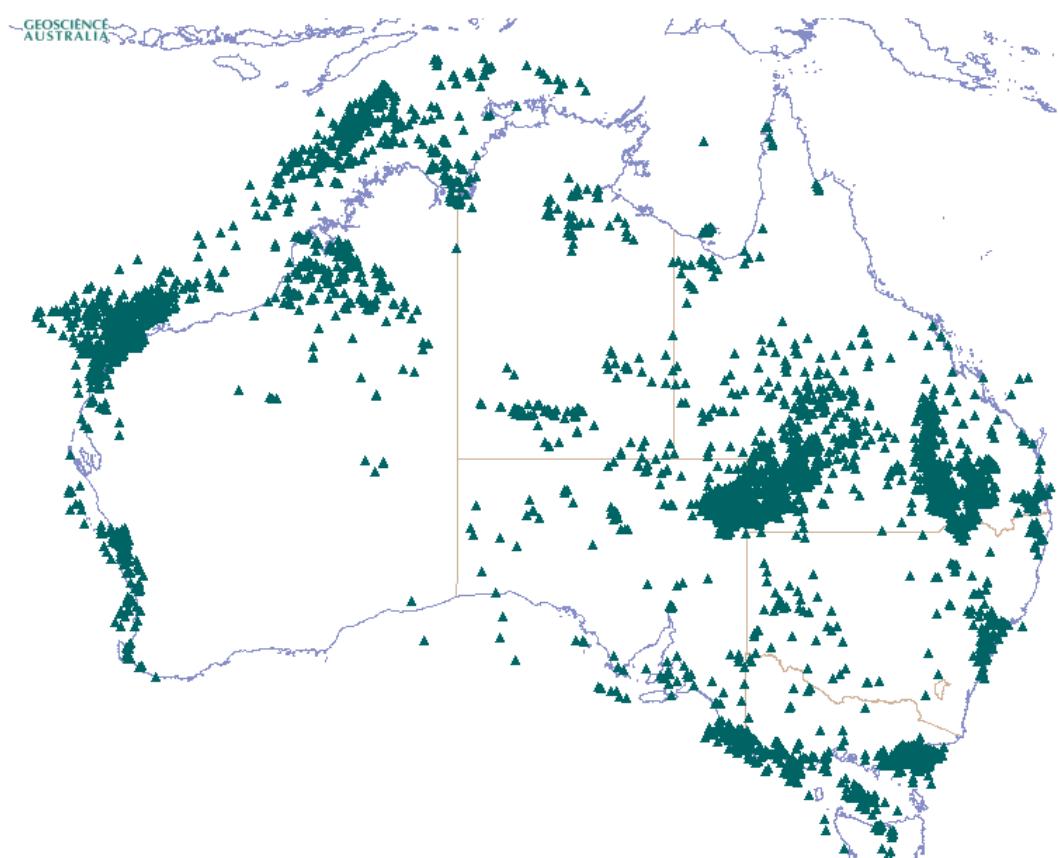


FIGURE 2.32: Petrol Well locations over Australia (current-2018) (<http://dbforms.ga.gov.au/www/npm.well.search>)

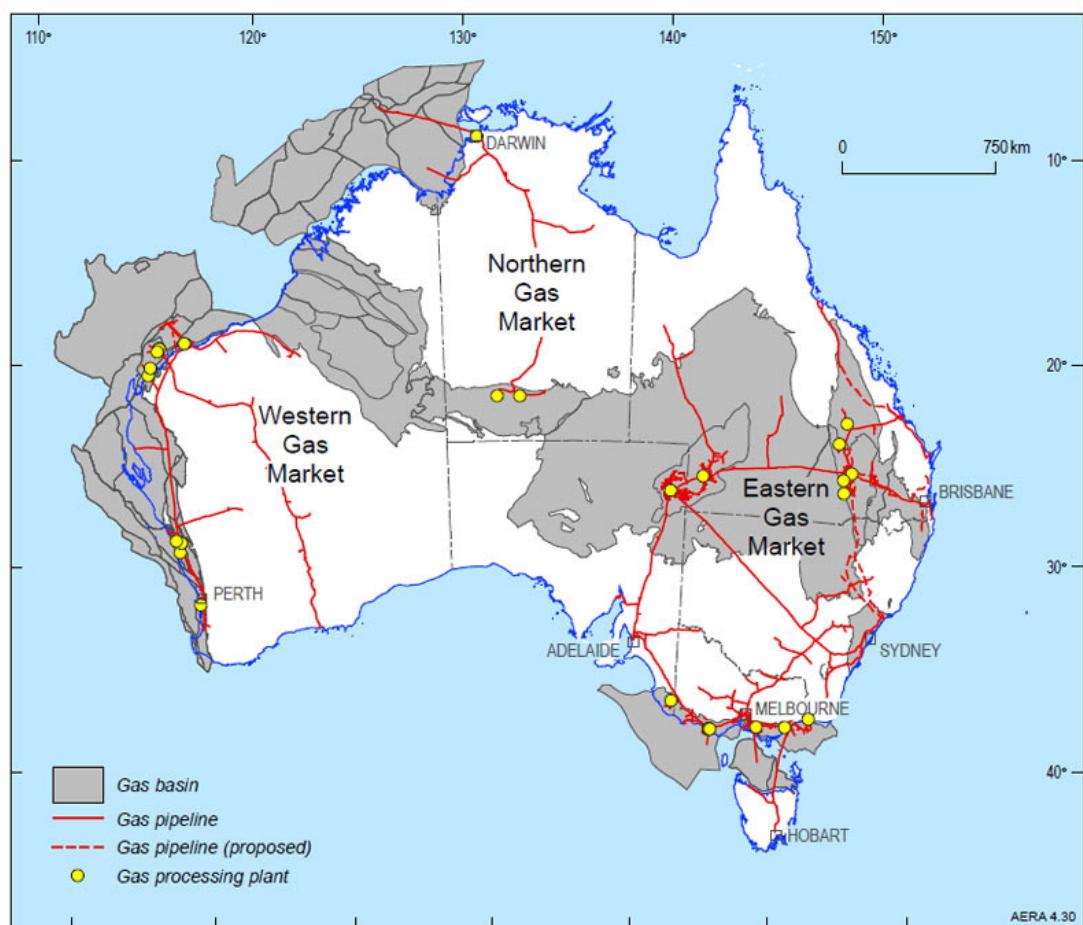


FIGURE 2.33: Gas fields and pipelines (2018) for Australia
(<http://www.ga.gov.au/scientific-topics/energy/resources/petroleum-resources/gas>)

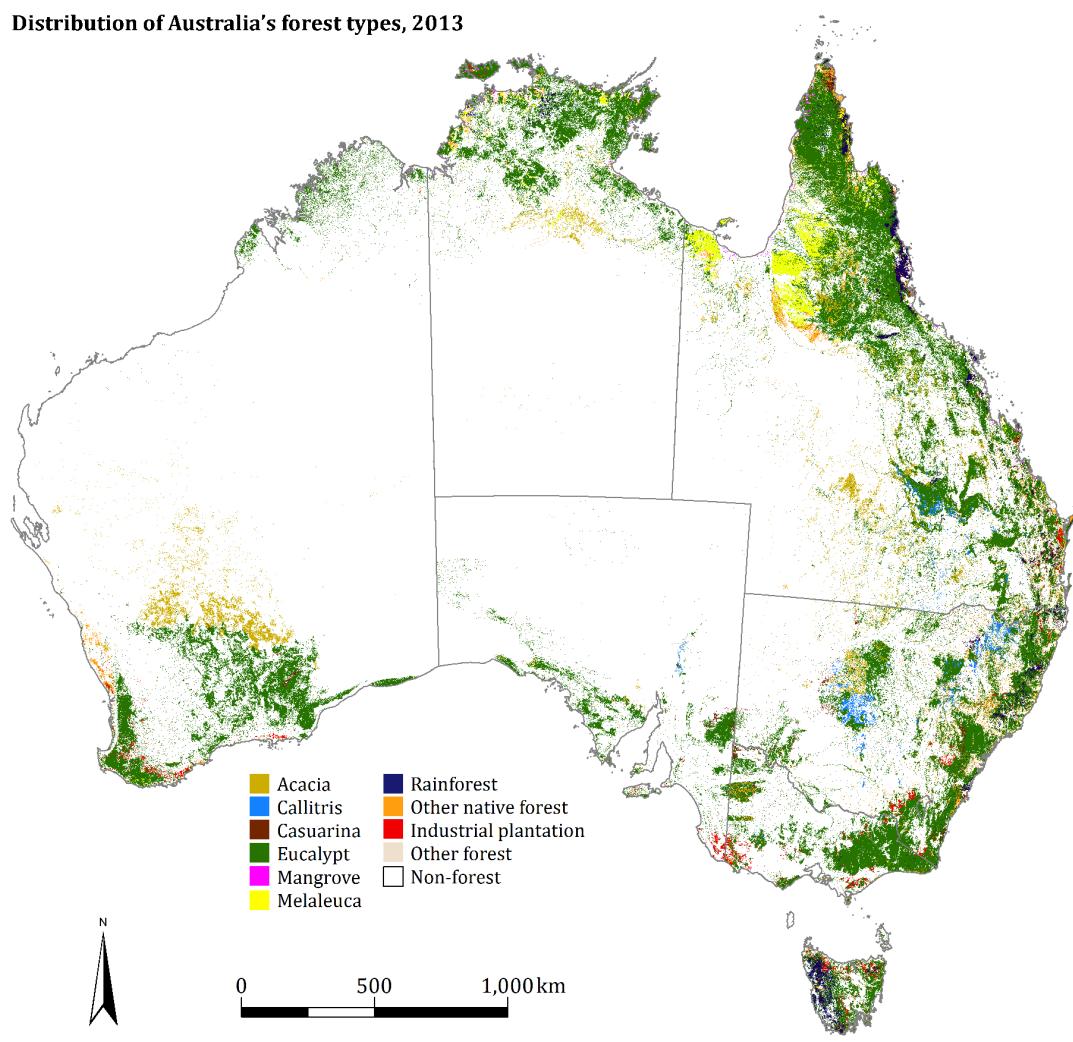


FIGURE 2.34: Forest coverage, coloured by predominant tree species.

??

2.7.1.1 Checking that fire masks are influencing pyrogenic HCHO

Looking at temperature can provide evidence of pyrogenic HCHO. HCHO precursors are heavily tied to temperature (TODO:cite), and model output shows how higher temperature leads to an increase in HCHO levels. Figures 2.37 - 2.39 show the relationship between modelled temperature, and satellite HCHO for January 2005 within subsets of Australia. A reduced major axis regression is used to determine the correlation between surface temperature (X axis) and HCHO (Y axis). Using the natural log of HCHO we can take the linear regression and then exponentiate each side in the equation $\ln Y = mX + b$ to get $Y = \exp mX + b$. This gives us the exponential fit as shown, with the correlation coefficient between $\ln HCHO$ and temperature. The distributions of exponential correlation coefficients and 'm' terms is shown in the embedded plot, with one datapoint available for each grid square where the regression is performed.

Figures 2.35 and 2.36 show the regressions between OMI HCHO total columns and temperature from GEOS-Chem output and CPC daily maximum temperatures respectively. Comparing against GEOS-Chem modelled surface temperatures first requires deresolution from 0.25×0.3125 to $2 \times 2.5^\circ$ latitude by longitude resolution. The left column in figure 2.35 shows scatter and RMA correlation within a single gridbox over 2 months from Jan 1 to Feb 28, 2005, without having applied either the fire nor anthropogenic masks to OMI HCHO columns. The right column shows the same correlation after applying the fire filter, affected datapoints are marked in teal (matching the red marked points in the left column). The analysis is repeated for Sydney, Canberra, and three gridsquares to the north-west, west, and southwest of Sydney (w1, w2, w3 respectively). Figure 2.36 shows the same analysis at higher resolution using CPC daily maximum temperatures (see 2.3).

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

2.7.2 NOx

Enhanced NO_2 concentrations can indicate anthropogenic influence over Australia. In order to filter out these influences on satellite HCHO measurements, a filter is designed using the OMNO2d product which includes tropospheric NO_2 columns.

OMNO2d from 2005 is used to determine a suitable threshhold for anthropogenic influence by looking at NO_2 columns near several major cities in the south eastern sector of Australia. The mean, standard deviation, and time series over Australia of

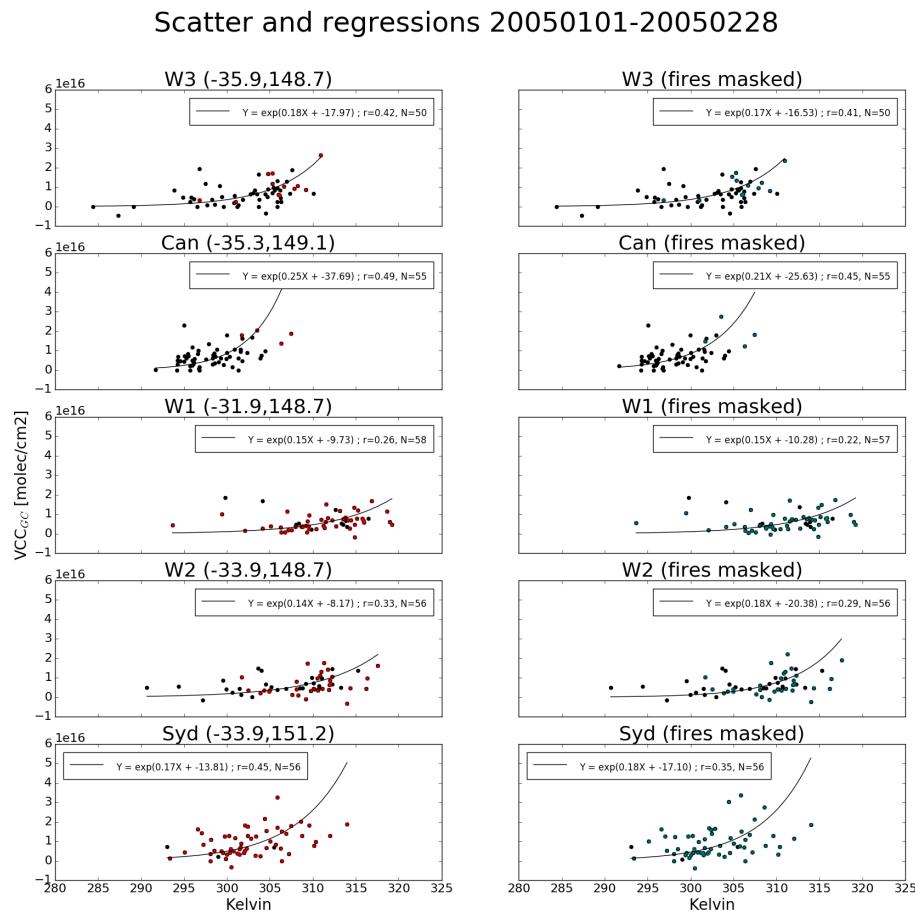


FIGURE 2.35: HCHO vs GEOS-Chem daily midday temperatures
TODO add longer caption

Scatter and regressions (CPC temperature) 20050101-20050228

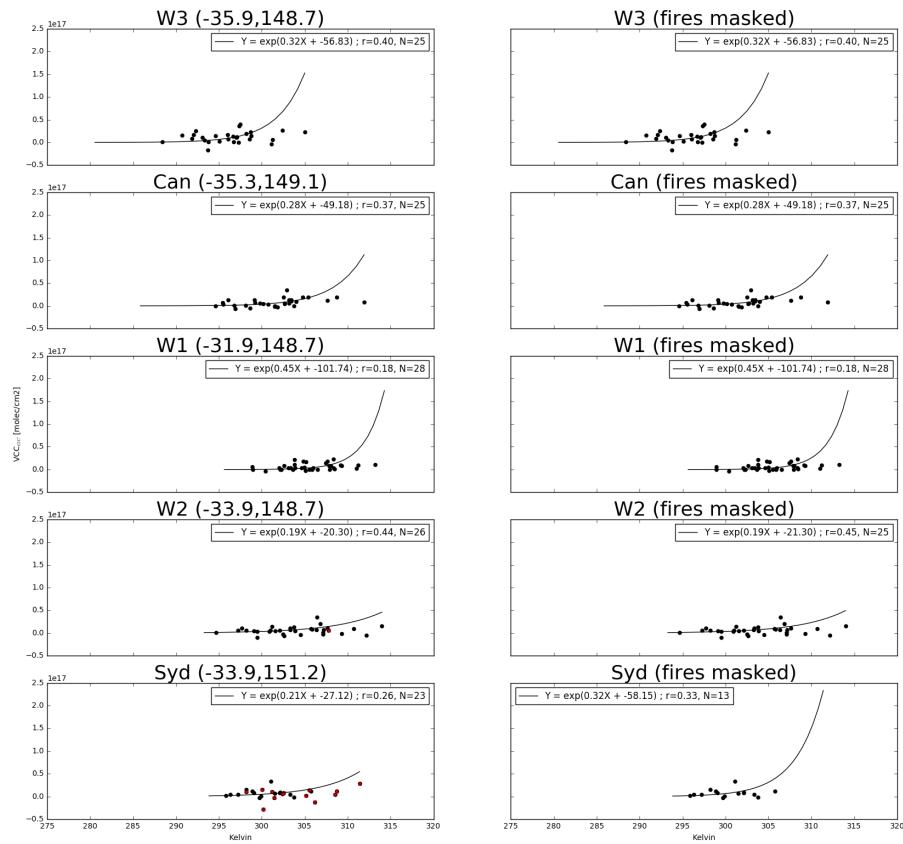


FIGURE 2.36: HCHO vs CPC daily maximum temperatures TODO add longer caption

Temperature 20050101-20050228

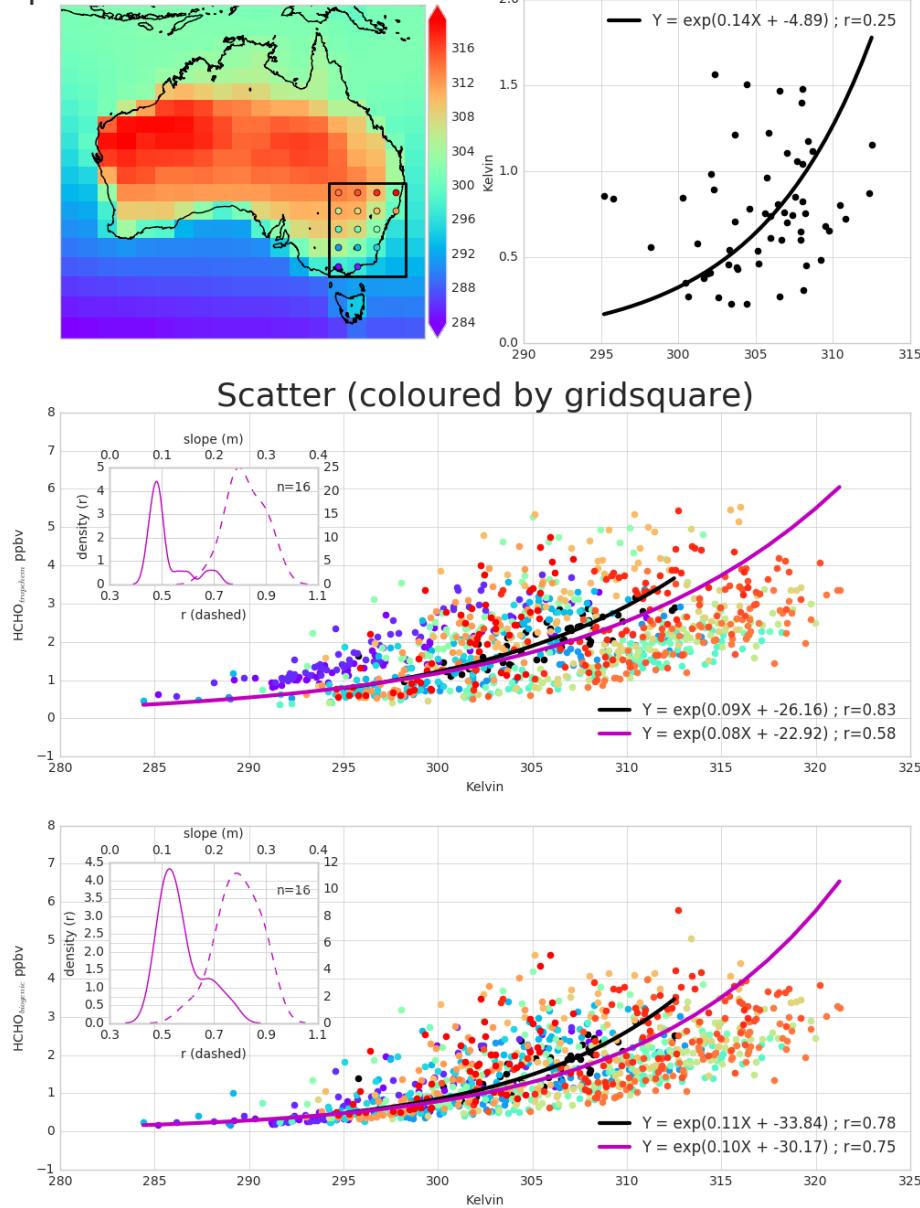


FIGURE 2.37: Top panel: surface temperature averaged over January and February 2005. Bottom panel: surface temperature correlated against temperature over, with different colours for each gridbox, and the combined correlation. A reduced major axis regression is used within each gridbox (shown in top panel) using daily overpass time surface temperature and HCHO amounts (ppbv). The distribution of slopes and regression correlation coefficients (one datapoint per gridbox) for the exponential regression is shown in the embedded plot.

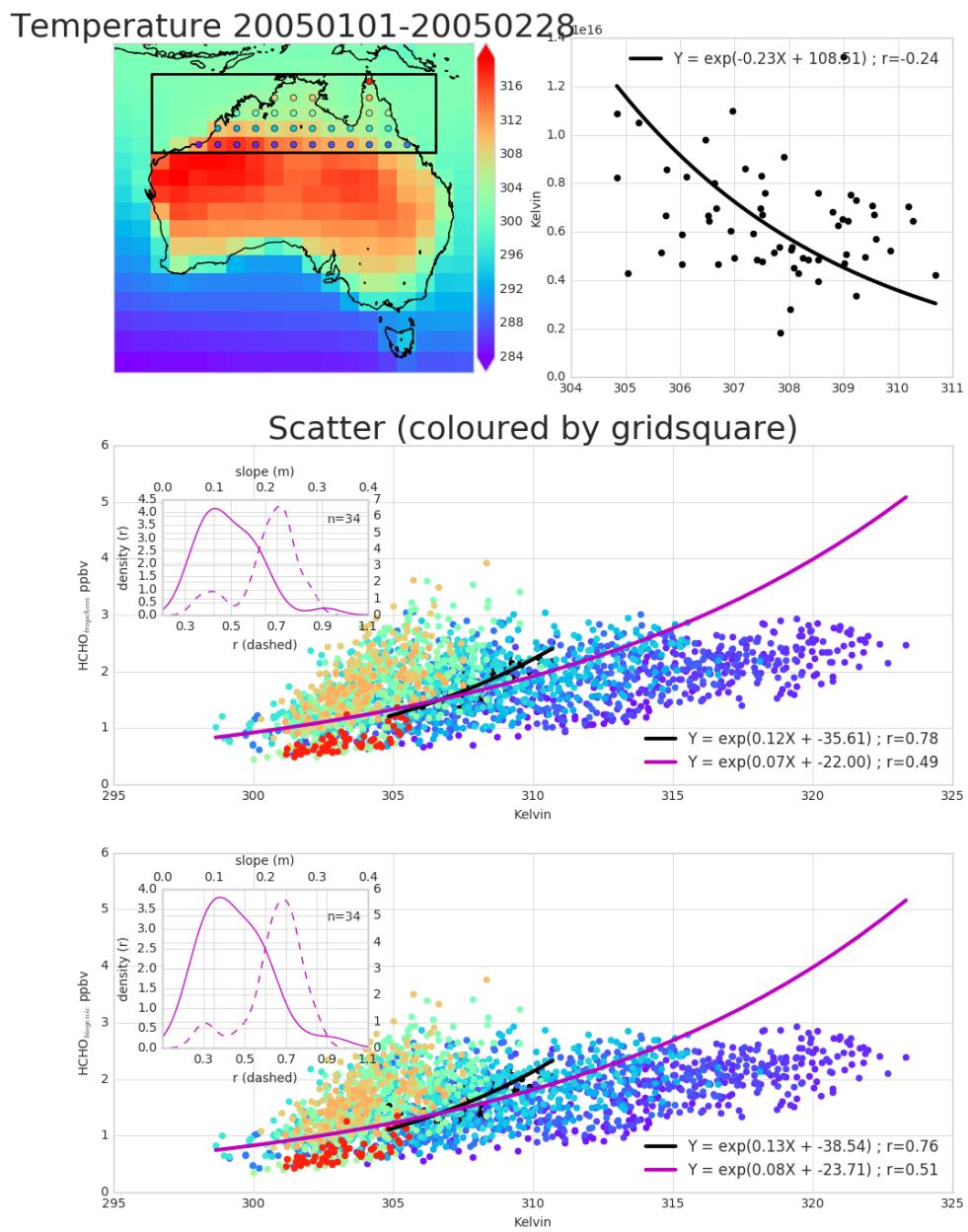


FIGURE 2.38: As figure 2.37 but for northern Australia.

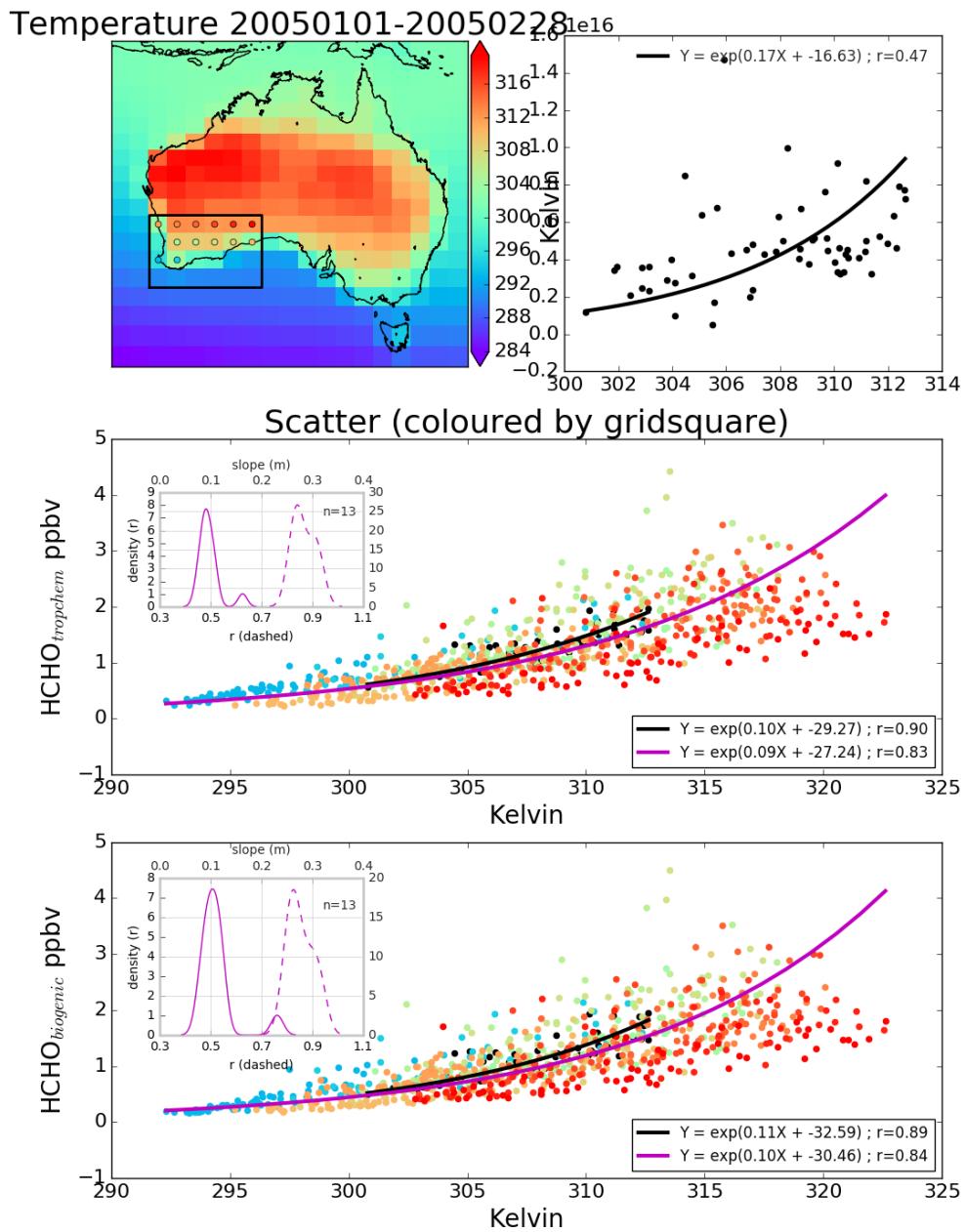


FIGURE 2.39: As figure 2.37 but for south-western Australia.

TABLE 2.4: NO₂ averages by region before and after filtering for anthropogenic emissions using 2005 data from the OMNO2d product.

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

tropospheric NO₂ seen by Aura is shown in figure 2.40. The average tropospheric NO₂ column averaged within all of Australia and then each region shown in this figure is listed in table TODO 2.4.

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

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

This removes both the gridsquares close enough to cities to be affected by their emissions year round, as well as effects from transported pollution plumes. The affects of applying this filter to the OMNO2d product itself can be seen in figure 2.41

The same regions as in figure 2.40 are shown again in figure 2.42, with NO₂ pixels densities for each region shown, along with the threshold of 1×10^{15} molec cm⁻². This led to a reduction of TODO gridsquares from the total available measurement space over Australia. The removal of gridsquares which went above the yearly averaged limit of 1.5×10^{15} molec cm⁻² further reduced the available data by TODO gridsquares.

2.8 Data Access

TODO: ADD MORE HERE

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

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

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

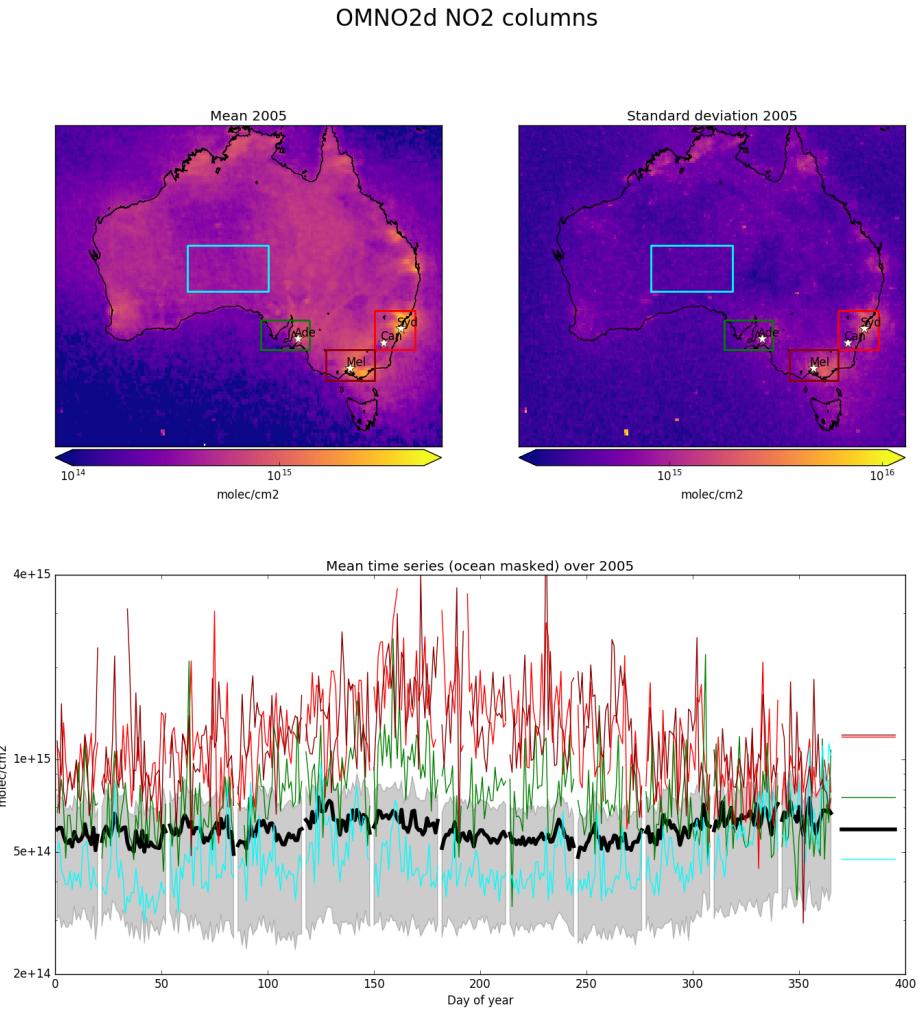


FIGURE 2.40: Mean (top left) and standard deviation (top right) of OMNO2d daily $0.25 \times 0.25^\circ$ tropospheric cloud filtered NO₂ columns. Time series for Australia, and each region (by colour) shown in the bottom panel, with mean for that region shown on the right. A grey shaded area depicts the 25th to 75th percentiles of Australia averaged NO₂ columns for each day in the time series, with a thicker black line showing the Australia-wide mean value.

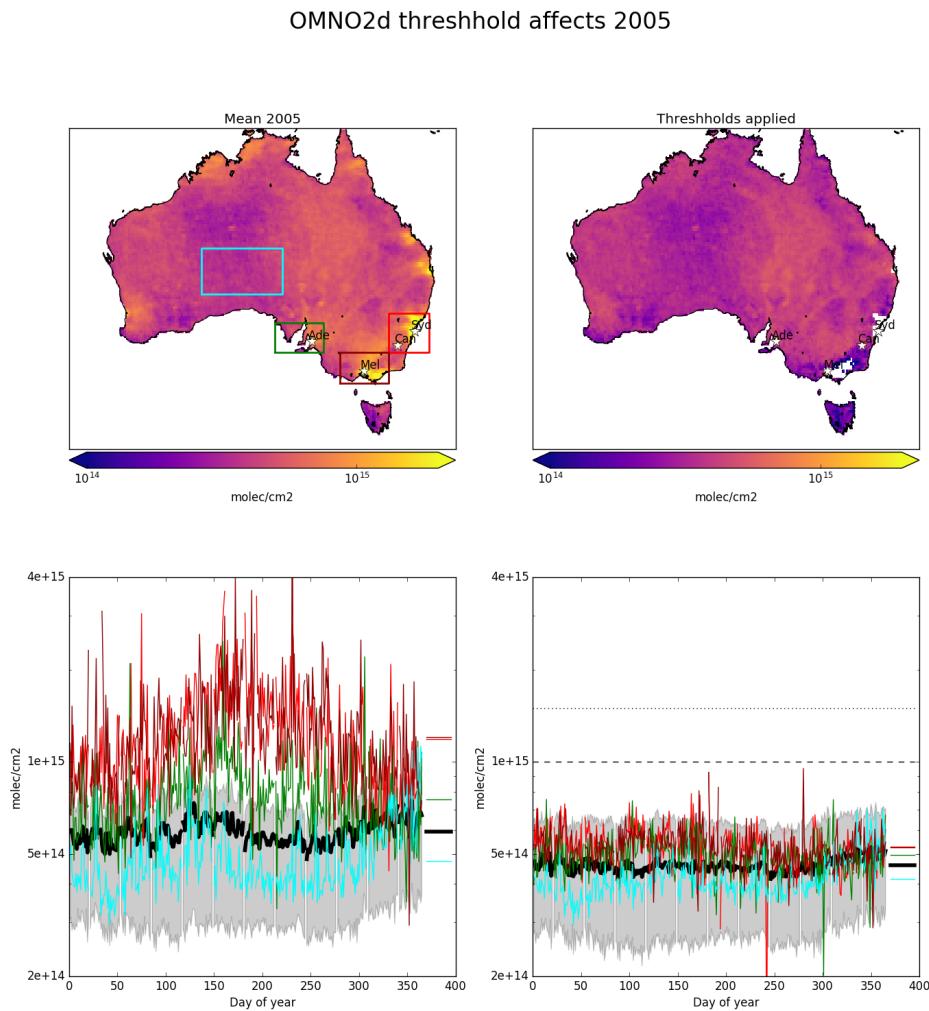


FIGURE 2.41: 2005 OMNO2d NO₂ column mean before (left) and after (right) applying the threshhold filters as described in the text. Time series for Australia, and each region (by colour) shown in the bottom panel, with mean for that region shown on the right.

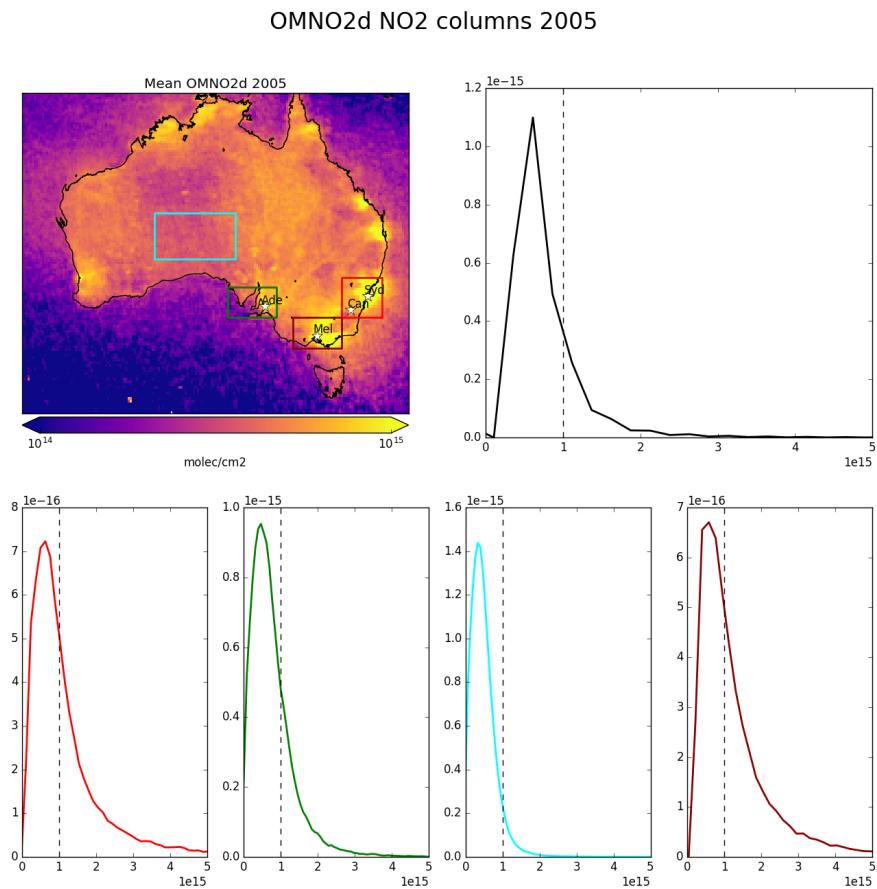


FIGURE 2.42: 2005 OMNO2d NO₂ column means (top left), along with column amount distributions for Australia (top right) and each region shown in the area map (by colour)